



**T.C.
HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**KUMAŞ HATA TESPİTİ İÇİN YAPAY ÖĞRENME TABANLI BİR YAKLAŞIM
ÖNERİSİ: BÜYÜK ÖLÇEKLİ BİR DENİM KONFEKSİYON FİRMASINDA
UYGULANMASI VE HATA KAYNAKLI KARBON AYAK İZİ HESAPLAMA**

**Yasemin EMRE UÇAR
ENDÜSTRİ MÜHENDİSLİĞİ**

**ŞANLIURFA
2026**



**T.C.
HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**KUMAŞ HATA TESPİTİ İÇİN YAPAY ÖĞRENME TABANLI BİR YAKLAŞIM
ÖNERİSİ: BÜYÜK ÖLÇEKLİ BİR DENİM KONFEKSİYON FİRMASINDA
UYGULANMASI VE HATA KAYNAKLI KARBON AYAK İZİ HESAPLAMA**

**Yasemin EMRE UÇAR
ENDÜSTRİ MÜHENDİSLİĞİ**

DANIŞMAN: Doç. Dr. Ahmet Sabri ÖĞÜTLÜ

**ŞANLIURFA
2026**

TEŞEKKÜR

Bu çalışmanın hazırlanması sürecinde bilgi, birikim ve deneyimleriyle bana yol gösteren, her aşamada değerli görüş ve önerileriyle destek olan tez danışmanım Sayın Doç. Dr. Ahmet Sabri Öğütli'ye en içten teşekkürlerimi sunarım.

Hayatımın her döneminde yanımda olan, sevgileri, anlayışları ve destekleriyle bana güç veren aileme; başta annem Senem Emre'ye ve ağabeyim Okan Emre'ye sonsuz teşekkürlerimi sunarım.

Ayrıca, bu zorlu ve yoğun süreç boyunca sabrı, sevgisi ve desteğiyle her zaman yanımda olan kıymetli eşim Ufuk Uçar'a ve varlığıyla bana ilham ve mutluluk kaynağı olan sevgili oğlum Aslan Uçar'a en derin sevgi ve teşekkürlerimi sunarım.

ONAY

İÇİNDEKİLER

ÖZET	i
ABSTRACT	ii
ÇİZELGELER DİZİNİ	iv
ŞEKİLLER DİZİNİ	v
KISALTMALAR	vi
SİMGELER.....	vii
1. GİRİŞ	1
2. ÖNCEKİ ÇALIŞMALAR	7
2.1. Tekstil Sektöründe Kumaş Hata Tespiti Çalışmaları	7
2.2. Yapay Öğrenme Tabanlı Kalite ve Hata Tahmin Yaklaşımları	11
2.3. Makine Öğrenmesi Tabanlı Sınıflandırma Çalışmaları	17
2.4. Karbon Ayak İzi ve Sürdürülebilir Üretim Çalışmaları	23
3. GEREÇ ve YÖNTEM	29
3.1. İşletmenin Tanıtımı	29
3.2. Veri Setinin Tanıtılması	31
3.3. Veri Hazırlama Süreci	35
3.4. Makine Öğrenmesi Modelleri	40
3.4.1. İki Sınıflı Sınıflandırma Yaklaşımı	44
3.4.2. Üç Sınıflı Sınıflandırma Yaklaşımı	46
3.5. Model Performans Ölçütleri	47
3.5.1. Accuracy	49
3.5.2. Precision	50
3.5.3. Recall	51
3.5.4. F1-Score	52
3.6. Karbon Ayak İzi Hesaplama Yöntemi	53
3.6.1. Toplam Hata Adedinin Hesaplanması	55
3.6.2. Rework Süresinin Hesaplanması	56
3.6.3. Ek Elektrik Tüketiminin Hesaplanması	56
3.6.4. Karbon Ayak İzi Hesabı	57
4. BULGULAR	59
4.1. Veri Setine İlişkin Bulgular	59
4.2. Lif Oranlarının Kumaş Hatalarıyla İlişkisi	63
4.2.1. Normallik Testi Bulguları	63
4.2.2. Spearman Korelasyon Analizi Bulguları	64
4.2.3. Değişken Önem Analizi Bulguları	65
4.2.4. Korelasyon ve Değişken Önem Bulgularının Birlikte Sunumu	67
4.3. İki Sınıflı Sınıflandırma Bulguları	68
4.3.1. Model Performans Sonuçları	68
4.3.2. Accuracy Sonuçları	70
4.3.3. Recall Sonuçları	71
4.3.4. F1-Score Sonuçları	72
4.3.5. Genel Değerlendirme	73
4.4. Üç Sınıflı Sınıflandırma Bulguları	73
4.4.1. Model Performans Sonuçları	74

4.4.2. Accuracy Sonuları	76
4.4.3. Recall Sonuları	76
4.4.4. F1-Score Sonuları	77
4.4.5. Genel Deęerlendirme	78
4.5. İki ve Ü Sınıflı Yaklaşımların Karşılaştırılması	78
4.6. Karbon Ayak İzi Bulguları	82
4.6.1. Toplam Hata Kaynaklı Rework Süresi	82
4.6.2. Ek Elektrik Tüketimi Bulguları	83
4.6.3. Karbon Ayak İzi Hesaplama Bulguları	84
4.6.4. Sürdürülebilirlik Açısından Deęerlendirme	84
5. TARTIŞMA	86
5.1. Makine Öğrenmesi Modellerinin Deęerlendirilmesi	86
5.2. Lif Oranlarının Hata Oluşumuna Etkisinin Deęerlendirilmesi	89
5.3. Sürdürülebilirlik Açısından Genel Deęerlendirme	94
5.4. Genel Tartışma Özeti	98
6. SONULAR	102
6.1. Genel Sonular	102
6.2. Çalışmanın Üretim Sürecine Katkıları	104
6.3. Çalışmanın Sürdürülebilirlik Açısından Katkıları	104
7. ÖNERİLER	106
KAYNAKLAR	110
ÖZ GEÇMİŞ	116
EKLER	117

ÖZET

YÜKSEK LİSANS TEZİ

KUMAŞ HATA TESPİTİ İÇİN YAPAY ÖĞRENME TABANLI BİR YAKLAŞIM ÖNERİSİ: BÜYÜK ÖLÇEKLİ BİR DENİM KONFEKSİYON FİRMASINDA UYGULANMASI VE HATA KAYNAKLI KARBON AYAK İZİ HESAPLAMA

Yasemin EMRE UÇAR

HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
ENDÜSTRİ MÜHENDİSLİĞİ

Tez Danışmanı: Doç. Dr. Ahmet Sabri ÖĞÜTLÜ

Yıl: 2026, Sayfa: 119

Bu çalışmada, büyük ölçekli bir denim konfeksiyon işletmesine ait üretim ve kalite kontrol kayıtları kullanılarak kumaş hata düzeylerinin makine öğrenmesi yöntemleriyle sınıflandırılması ve hata kaynaklı yeniden işlem faaliyetlerinin karbon ayak izi üzerindeki etkisinin hesaplanması amaçlanmıştır. Çalışma kapsamında 652 gözlemden oluşan veri seti kullanılmış; pamuk oranı, geri dönüştürülmüş lif oranı, polyester oranı, elastan oranı ve diğer lifler oranı bağımsız değişkenler olarak değerlendirilmiştir. Kumaş hata adedi iki farklı sınıflandırma yapısına dönüştürülmüş; iki sınıflı yaklaşımda gözlemler “Normal” ve “Yüksek”, üç sınıflı yaklaşımda ise “Düşük”, “Orta” ve “Yüksek” hata düzeyleri olarak ele alınmıştır. Sınıflandırma sürecinde KNN, Random Forest, Cost-Sensitive Random Forest, RUSBoost, Bagged Trees, Decision Tree, Naive Bayes, Boosting, SVM-RBF, LDA, Logistic Regression, Linear SVM, ANN ve alt uzay tabanlı modeller karşılaştırılmıştır. Model performansları accuracy, precision, recall ve F1-score ölçütleriyle değerlendirilmiştir. Ayrıca Shapiro-Wilk normallik testiyle korelasyon analizinde kullanılacak değişkenlerin dağılım özellikleri incelenmiş; normal dağılım varsayımının sağlanmaması nedeniyle lif oranları ile kumaş hata adedi arasındaki ilişki Spearman korelasyon analiziyle değerlendirilmiştir. Random Forest temelli değişken önem analiziyle lif oranlarının model içindeki göreceli katkıları belirlenmiştir. Bulgulara göre iki sınıflı yaklaşımda en yüksek F1-score değeri RUSBoost modelinde 0,4883 olarak hesaplanmıştır. Bu modelde accuracy 0,6979, precision 0,4215 ve recall 0,5802 olarak belirlenmiştir. Üç sınıflı yaklaşımda da en başarılı model RUSBoost olmuş, bu modelde accuracy 0,4586, precision 0,4660, recall 0,4584 ve F1-score 0,4589 olarak hesaplanmıştır. Spearman korelasyon analizinde lif oranları ile kumaş hata adedi arasında istatistiksel olarak anlamlı ve güçlü bir ilişki bulunmamıştır. Buna karşılık değişken önem analizinde en yüksek katkı COT oranında elde edilmiş; değişken önem sıralaması COT oranı, EA oranı, Diğer lifler oranı, Recycle oran ve PES oranı şeklinde gerçekleşmiştir. Karbon ayak izi hesabında toplam 31.943 kumaş hatasının 559 saat ek rework süresi, 5.590 kWh ek elektrik tüketimi ve yaklaşık 2,34 ton CO₂e/yıl düzeyinde karbon etkisi oluşturduğu belirlenmiştir. Elde edilen sonuçlar, denim konfeksiyon üretiminde kalite hatalarının yalnızca üretim verimliliği açısından değil, enerji tüketimi ve karbon emisyonu açısından da ölçülebilir etkiler oluşturduğunu göstermektedir.

ANAHTAR KELİMELE: Makine öğrenmesi, Yeniden çalışma, Denim, Kumaş hatası, Karbon ayak izi

ABSTRACT

MASTER’S THESIS

A MACHINE LEARNING BASED APPROACH FOR FABRIC DEFECT DETECTION: IMPLEMENTATION IN A LARGE SCALE DENIM APPAREL COMPANY AND ERROR-BASED CARBON FOOTPRINT CALCULATION

Yasemin EMRE UÇAR

**HARRAN UNIVERSITY
GRADUATE SCHOOL
INDUSTRIAL ENGINEERING**

Thesis Advisor: Assoc. Prof. Ahmet Sabri ÖĞÜTLÜ

Year: 2026, Page: 119

In this study, production and quality control records obtained from a large-scale denim garment manufacturing company were used to classify fabric defect levels through machine learning methods and to calculate the carbon footprint impact of defect-related rework activities. The dataset consisted of 652 observations, and cotton ratio, recycled fiber ratio, polyester ratio, elastane ratio, and other fiber ratio were used as independent variables. Fabric defect count was transformed into two different classification structures. In the binary classification approach, observations were classified as “Normal” and “High”, whereas in the three-class classification approach, they were grouped as “Low”, “Medium”, and “High” defect levels. Several machine learning models, including KNN, Random Forest, Cost-Sensitive Random Forest, RUSBoost, Bagged Trees, Decision Tree, Naive Bayes, Boosting, SVM-RBF, LDA, Logistic Regression, Linear SVM, ANN, and subspace-based models, were compared. Model performances were evaluated using accuracy, precision, recall, and F1-score. In addition, Shapiro-Wilk normality test was used to examine the distributional characteristics of the variables included in the correlation analysis. Since the normality assumption was not met, the relationship between fiber ratios and fabric defect count was evaluated using Spearman correlation analysis. Random Forest-based variable importance analysis was then applied to determine the relative contribution of fiber ratios within the model. According to the findings, the highest F1-score in the binary classification approach was obtained from the RUSBoost model with a value of 0.4883. In this model, accuracy, precision, and recall were calculated as 0.6979, 0.4215, and 0.5802, respectively. In the three-class classification approach, the RUSBoost model also showed the best performance, with accuracy, precision, recall, and F1-score values of 0.4586, 0.4660, 0.4584, and 0.4589, respectively. Spearman correlation analysis did not indicate statistically significant or strong relationships between fiber ratios and fabric defect count. However, variable importance analysis showed that the COT ratio had the highest contribution, followed by EA ratio, other fiber ratio, recycled fiber ratio, and PES ratio. In the carbon footprint calculation, a total of 31,943 fabric defects corresponded to 559 hours of additional rework time, 5,590 kWh of additional electricity consumption, and approximately 2.34 tons CO₂e/year. The results indicate that quality defects in denim garment production create measurable effects not only in terms of production efficiency but also in terms of energy consumption and carbon emissions.

KEYWORDS: Machine Learning, Rework, Denim, Fabric defect, Carbon footprint

ÇİZELGELER DİZİNİ

Çizelge 3.1.	İşletmede denim üretim sürecinin temel aşamaları	30
Çizelge 3.2.	Veri seti değişkenleri ve analizdeki rolleri	33
Çizelge 3.3.	Veri hazırlama sürecinde uygulanan işlemler	38
Çizelge 3.4.	Çalışmada kullanılan makine öğrenmesi modelleri ve seçilme gerekçeleri	43
Çizelge 3.5.	İki sınıflı sınıflandırma yaklaşımında sınıf dağılımı	45
Çizelge 3.6.	Üç sınıflı sınıflandırma yaklaşımında dağılım	47
Çizelge 3.7.	Model performans ölçütleri ve yorumlanma biçimleri	53
Çizelge 3.8.	Karbon ayak izi hesabında kullanılan temel girdiler	55
Çizelge 3.9.	Karbon ayak izi hesaplama adımları	58
Çizelge 4.1.	Veri setinin genel yapısı	59
Çizelge 4.2.	Sayısal değişkenlere ilişkin tanımlayıcı istatistikler	60
Çizelge 4.3.	Değişkenlere ilişkin Shapiro-Wilk normallik testi sonuçları . . .	64
Çizelge 4.4.	. Lif oranları ile kumaş hata adedi arasındaki Spearman korelasyon katsayıları ve anlamlılık değerleri	64
Çizelge 4.5.	Random Forest temelli değişken önem analizi sonuçları	66
Çizelge 4.6.	İki sınıflı sınıflandırma yaklaşımında sınıf dağılımı	68
Çizelge 4.7.	İki sınıflı sınıflandırma modellerinin performans sonuçları . . .	69
Çizelge 4.8.	İki sınıflı modellerin confusion matrix sonuçları	70
Çizelge 4.9.	Üç sınıflı sınıflandırma yaklaşımında sınıf dağılımı	74
Çizelge 4.10.	Üç sınıflı sınıflandırma modellerinin performans sonuçları . . .	75
Çizelge 4.11.	İki ve üç sınıflı sınıflandırma yaklaşımlarının karşılaştırılması .	79
Çizelge 4.12.	Karbon ayak izi hesabında kullanılan temel değerler	82
Çizelge 4.13.	Toplam hata kaynaklı rework süresi hesabı	83
Çizelge 4.14.	Hata kaynaklı ek elektrik tüketimi hesabı	84
Çizelge 4.15.	Hata kaynaklı karbon ayak izi hesabı	84
Çizelge 4.16.	Hata kaynaklı rework, elektrik tüketimi ve karbon ayak izi sonuçları	85

ŞEKİLLER DİZİNİ

Şekil 4.1.	Kumaş hata adedi dağılımı	61
Şekil 4.2.	Lif oranlarına ait kutu grafikleri	62
Şekil 4.3.	Random Forest temelli değişken önem skorları	66
Şekil 4.4.	İki sınıflı sınıflandırma modellerinin LOOCV F1-score değerleri	72
Şekil 4.5.	Üç sınıflı sınıflandırma modellerinin LOOCV F1-score değerleri	77
Şekil 4.6.	İki ve üç sınıflı yaklaşımların LOOCV en iyi performans değerleri	80
Şekil 4.7.	İki ve üç sınıflı yaklaşımların LOOCV en iyi performans değerleri	81

KISALTMALAR

LCA	Yaşam Döngüsü Değerlendirmesi
LDA	Doğrusal Diskriminant Analizi
ML	Makine Öğrenmesi
PES	Polyester oranı
RF	Random Forest
RUSBoost	Rastgele Alt Örneklemeli Boosting
ANN	Yapay Sinir Ağı
CO₂e	Karbondioksit eşdeğeri
COT	Pamuk oranı
EA	Elastan oranı
F1-score	Precision ve recall değerlerinin harmonik ortalaması
KNN	K-En Yakın Komşu
kWh	Kilowatt-saat
SVM	Destek Vektör Makineleri
SVM-RBF	Radyal Tabanlı Fonksiyon Çekirdekli Destek Vektör Makinesi

SİMGELER

dk	Dakika
$\leq$	Küçük Eşit
kg	Kilogram
CO₂	Karbondioksit
%	Yüzde
>	Büyük
p33	33. yüzdellik değer
p66	66. yüzdellik değer
p75	75. yüzdellik değer
kW	Kilowatt
saat	Zaman birimi
ton	Kütle birimi
CO₂e/yıl	Yıllık karbondioksit eşdeğeri emisyonu
kg	Elektrik tüketimine bağlı emisyon faktörü
CO₂e/kWh	
TN	Gerçek negatif
TP	Gerçek pozitif
FN	Yanlış negatif
FP	Yanlış pozitif
n	Gözlem sayısı

1. GİRİŞ

Tekstil ve hazır giyim sektörü, yüksek üretim hacmi, kısa teslim süreleri, değişken müşteri talepleri ve yoğun maliyet baskısı nedeniyle kalite yönetiminin doğrudan üretim performansına yansıdığı sanayi alanlarından biridir. Bu yapı içinde kalite kontrol, yalnızca ürünün kabul edilebilirliğini belirleyen son aşama faaliyeti olarak değerlendirilemez. Üretim hattında ortaya çıkan her hata; yeniden işlem, fire, ek işçilik, teslim süresinde gecikme, enerji tüketiminde artış ve müşteri memnuniyetinde azalma gibi birbirine bağlı sonuçlar doğurmaktadır. Özellikle denim konfeksiyon üretiminde ürün kalitesini etkileyen değişkenler yalnızca dikim ya da son kontrol aşamasıyla sınırlı değildir. Hammadde bileşimi, kumaş yapısı, lif kompozisyonu, kesim özellikleri, dikim operasyonları, yıkama ve son işlem uygulamaları birlikte ilerleyen çok değişkenli bir üretim sistemi oluşturmaktadır.

Denim ürünleri, yaygın kullanımı ve yüksek ticari değeri nedeniyle hazır giyim sektörü içinde özel bir konuma sahiptir. Bununla birlikte denim üretimi, çevresel yükü yüksek süreçlerden biri olarak da tartışılmaktadır. Literatürde denim üretiminin su, enerji ve kimyasal kullanımı bakımından önemli çevresel etkiler oluşturduğu; sürdürülebilir üretim arayışlarının da büyük ölçüde bu etkileri azaltmaya yöneldiği belirtilmektedir (Periyasamy vd., 2023). Tekstil değer zincirinin sera gazı emisyonları, su kullanımı, atık oluşumu ve kaynak tüketimi üzerindeki etkileri de uluslararası raporlarda açık biçimde vurgulanmaktadır. UNEP'in tekstil değer zincirine ilişkin raporunda sektörün çevresel yükünün yalnızca üretim aşamasından ibaret olmadığı, tüketim kalıpları, üretim uygulamaları ve altyapı yatırımlarıyla birlikte ele alınması gerektiği ifade edilmektedir (UNEP, 2023). Bu nedenle tekstil işletmelerinde kalite iyileştirme, yalnızca maliyet azaltıcı bir faaliyet değil, aynı zamanda kaynak verimliliği ve sürdürülebilirlik hedefleriyle bağlantılı bir yönetim alanı olarak görülmelidir.

Kumaş hataları, tekstil üretiminde ürün görünümünü, kullanım performansını ve müşteri kabulünü etkileyen temel kalite problemleri arasında yer almaktadır. Bu hatalar üretimin farklı aşamalarında ortaya çıkabilmekte; zamanında belirlenmediğinde sonraki operasyonlara taşınarak daha yüksek maliyetli düzeltme faaliyetlerine neden olabilmektedir. Geleneksel kalite kontrol uygulamalarında kumaş hatalarının belirlenmesi büyük ölçüde

operatör deneyimine ve görsel kontrole dayanmaktadır. Deneyimli personelin üretim sürecindeki rolü önemli olmakla birlikte, yüksek hız ve yüksek hacimli üretim koşullarında manuel kontrolün sınırları belirginleşmektedir. Yorgunluk, dikkat azalması, yorum farklılığı ve kontrol standartlarındaki değişkenlik, manuel hata değerlendirmesinin tutarlılığını etkileyebilmektedir. Bu nedenle otomatik veya veri destekli kalite kontrol yaklaşımları, tekstil üretiminde uzun süredir önemli bir araştırma alanı olarak gelişmektedir. Kumaş hata kontrolünde bilgisayarlı görme ve sınıflandırıcı tabanlı yaklaşımların manuel kontrole göre daha hızlı ve standartlaştırılabilir çözümler sunduğu farklı çalışmalarda belirtilmiştir (Carrilho vd., 2024; Habib vd., 2014).

Son yıllarda kumaş hata tespiti çalışmalarında yapay zekâ, makine öğrenmesi ve derin öğrenme tabanlı yöntemler öne çıkmaktadır. Bu alandaki çalışmaların önemli bir bölümü kumaş yüzey görüntülerinden hata tespiti yapılmasına odaklanmaktadır. Görüntü tabanlı yaklaşımlarda delik, leke, çizgi, renk farklılığı, kırışıklık veya dokusal bozulma gibi kusurların tespiti amaçlanmakta; CNN, YOLO ve benzeri derin öğrenme mimarileri bu doğrultuda yaygın biçimde kullanılmaktadır. Güncel literatürde otomatik kumaş hata tespit sistemlerinin gerçek zamanlı kalite kontrol için önemli bir potansiyel taşıdığı, ancak farklı kumaş türleri, hata çeşitliliği, veri kalitesi ve model genellenebilirliği gibi konularda hâlâ sınırlılıklar bulunduğu görülmektedir (Carrilho vd., 2024; Machado vd., 2025). Bu noktada dikkat edilmesi gereken husus, her kalite probleminin mutlaka görüntü verisiyle çözümlenmek zorunda olmamasıdır. Üretim işletmelerinin önemli bir kısmında ürün bileşimi, üretim kayıtları, kalite kontrol sonuçları ve hata adetleri gibi tabular veriler düzenli biçimde tutulmakta; bu veriler kalite problemlerinin tahmini ve sınıflandırılması için değerli bir kaynak oluşturmaktadır.

Bu çalışma, görüntü tabanlı bir kumaş yüzey hata tespit sistemi geliştirme iddiası taşımamaktadır. Çalışmanın odağı, büyük ölçekli bir denim konfeksiyon işletmesinden elde edilen üretim ve kalite kontrol verileri üzerinden kumaş hata düzeylerinin makine öğrenmesi yöntemleriyle değerlendirilmesidir. Bu yönüyle çalışma, doğrudan hata görüntüsünü bulmaya değil, lif kompozisyonu değişkenleri kullanılarak hata adedinin sınıfsal düzeyini tahmin etmeye yönelmektedir. Veri setinde pamuk oranı, geri dönüştürülmüş lif oranı,

polyester oranı, elastan oranı ve diğer lif oranları bağımsız değişkenler olarak ele alınmış; kumaş hata adedi ise temel çıktı değişkeni olarak kullanılmıştır. Bu çerçevede, kalite problemini görüntü işleme ekseninden çıkararak üretim verisi temelli bir sınıflandırma problemi olarak ele almaktadır.

Üretim süreçlerinde kalite tahmini, Endüstri 4.0 yaklaşımının veri odaklı karar destek mantığıyla doğrudan ilişkilidir. Makine öğrenmesi tabanlı kestirimci kalite uygulamalarında temel düşünce, üretim sürecinden elde edilen verilerde tekrarlayan örüntülerin belirlenmesi ve bu örüntülerin kalite çıktılarıyla ilişkilendirilmesidir. (Tercan ve Meisen, 2022), kestirimci kalite çalışmalarının üretim verilerinden öğrenerek ürün kalitesinin tahmin edilmesine ve kaliteyi iyileştirecek kararların desteklenmesine olanak sağladığını belirtmektedir. Bu yaklaşım, geleneksel kalite kontrolün reaktif yapısından farklı olarak, hata oluşmadan önce riskli durumların fark edilmesine katkı sağlayabilecek bir analitik zemin sunmaktadır.

Kumaş hata oluşumu doğrusal ve tek nedenli bir süreç değildir. Lif oranları arasında doğal bağımlılıklar bulunmakta, bir değişkenin etkisi diğer bileşenlerin düzeyine göre değişebilmekte ve üretim koşulları bu ilişkileri daha karmaşık hale getirebilmektedir. Bu nedenle yalnızca basit korelasyonlara dayalı değerlendirmeler, hata oluşumunu açıklamakta yetersiz kalabilir. Makine öğrenmesi modelleri ise doğrusal olmayan ilişkileri, değişken etkileşimlerini ve karmaşık sınıf ayrımlarını yakalama kapasitesi nedeniyle kalite tahmini problemlerinde işlevsel araçlar sunmaktadır. Bununla birlikte bu modellerin çıktıları dikkatli yorumlanmalıdır. Bir modelin yüksek genel doğruluk üretmesi, üretim açısından en kritik sınıfı doğru tanıdığı anlamına gelmeyebilir. Özellikle kalite verilerinde yüksek hata sınıfı genellikle daha az sayıda gözlem içerdiğinden, accuracy tek başına yeterli bir performans ölçütü değildir. Bu nedenle bu çalışmada modeller yalnızca accuracy ile değil, precision, recall ve F1-score değerleriyle birlikte değerlendirilmiştir.

Çalışma kapsamında iki farklı sınıflandırma yaklaşımı benimsenmiştir. İlk yaklaşımda kumaş hata adetleri iki sınıfa ayrılarak normal ve yüksek hata düzeyleri oluşturulmuştur. Bu yapı, üretim açısından riskli kumaş kompozisyonlarının daha pratik biçimde ayırt edilmesine olanak sağlamaktadır. İkinci yaklaşımda ise hata adetleri düşük, orta ve yüksek olmak üzere üç sınıfta değerlendirilmiştir. Üç sınıflı yaklaşım, üretim gerçekliğini daha ayrıntılı

temsil etmekte; buna karşılık sınıflandırma problemini daha zor hale getirmektedir. Bu iki yapının birlikte ele alınması, model performansının yalnızca istatistiksel başarı açısından değil, üretim ortamında kullanılabilirlik bakımından da tartışılmasına imkân vermektedir. Çalışmada KNN, Random Forest, Bagged Trees, Decision Tree, Naive Bayes, SVM, LDA, Logistic Regression, ANN, RUSBoost, Cost-Sensitive Random Forest ve alt uzay tabanlı yöntemler gibi farklı algoritmalar karşılaştırılmıştır. Böylece doğrusal, doğrusal olmayan, ensemble tabanlı ve sınıf dengesizliğiyle baş etmeye yönelik modeller aynı veri seti üzerinde değerlendirilmiştir.

Bu çalışmanın bir diğer boyutu, kumaş hatalarının çevresel etkisinin lokal süreç düzeyinde hesaplanmasıdır. Tekstil sektöründe sürdürülebilirlik çoğu zaman hammadde seçimi, su tüketimi, kimyasal azaltımı veya atık yönetimi üzerinden tartışılmaktadır. Ancak kalite hatalarının neden olduğu yeniden işlem faaliyetleri de ek enerji tüketimi ve dolayısıyla ek karbon emisyonu üretmektedir. Hatalı bir ürün yeniden işlendiğinde yalnızca işçilik süresi artmaz; aynı zamanda makine ve ekipman kullanımı uzar, elektrik tüketimi yükselir ve üretim hattının kaynak verimliliği düşer. Bu nedenle kalite problemlerinin karbon ayak iziyle ilişkilendirilmesi, sürdürülebilir üretim açısından göz ardı edilmemesi gereken bir değerlendirme alanıdır.

Karbon ayak izi, bir ürün, süreç veya faaliyetten kaynaklanan sera gazı emisyonlarının karbondioksit eşdeğeri cinsinden ifade edilmesini sağlayan bir göstergedir. ISO 14067 standardı, ürün karbon ayak izinin ve kısmi ürün karbon ayak izinin hesaplanmasına ilişkin ilkeleri ve raporlama esaslarını tanımlamaktadır (ISO, 2018). Bu çalışmada yürütülen hesaplama, ürünün tüm yaşam döngüsünü kapsayan geniş ölçekli bir yaşam döngüsü analizi değildir. Çalışmada yalnızca kumaş hatalarından kaynaklanan yeniden işlem süresi, bu sürenin neden olduğu ek elektrik tüketimi ve elektrik tüketimine bağlı CO₂e emisyonu dikkate alınmıştır. Elektrik tüketimine bağlı emisyonların hesaplanmasında ise GHG Protocol Scope 2 rehberinde tanımlanan satın alınan elektrik kaynaklı emisyon yaklaşımı temel alınabilir (WRI & WBCSD, 2015). Bu sınırlandırma, çalışmanın veri kapsamı ve işletme düzeyindeki ölçülebilirliği açısından daha tutarlı bir yöntemsel çerçeve sunmaktadır.

Bu çalışmada kullanılan veri seti, büyük ölçekli bir denim konfeksiyon işletmesinden

elde edilen 652 gözlemden oluşmaktadır. Veri setinde beş temel lif kompozisyonu değişkeni ve kumaş hata adedi yer almaktadır. Çalışmada öncelikle veri setinin genel yapısı incelenmiş, lif oranlarının hata adediyle ilişkisi Spearman korelasyon analizi ve değişken önem analiziyle ele alınmış, sonrasında iki sınıflı ve üç sınıflı makine öğrenmesi modelleri leave-one-out çapraz doğrulama yaklaşımıyla karşılaştırılmıştır. Son aşamada ise toplam hata adedi üzerinden yeniden işlem süresi, ek elektrik tüketimi ve karbon ayak izi hesaplanmıştır. Bu yapı, kalite kontrol verilerinin yalnızca geçmiş performansı raporlayan kayıtlar olarak değil, üretim riskini ve çevresel yükü değerlendirmeye yarayan analitik girdiler olarak kullanılabileceğini göstermektedir.

Çalışmanın kapsamı bilinçli biçimde sınırlandırılmıştır. Veri setinde makine ayarları, üretim hızı, operatör deneyimi, ortam sıcaklığı ve nem, iplik kalite parametreleri, tedarikçi bilgisi, bakım kayıtları veya proses bazlı detay değişkenleri bulunmamaktadır. Bu nedenle çalışma, denim üretiminde kumaş hatalarının tüm nedenlerini açıklama iddiası taşımamaktadır. Mevcut analiz, lif kompozisyonu ile kumaş hata adedi arasındaki ilişkinin belirli bir işletme verisi üzerinden incelenmesine dayanmaktadır. Bu sınırlılık, çalışmanın değerini azaltmamakta; aksine bulguların hangi sınırlar içinde yorumlanması gerektiğini açık hale getirmektedir. Gerçek işletme verisiyle yürütülen bu tür çalışmalar, akademik modelleme ile üretim sahasındaki karar ihtiyaçları arasında köprü kurması bakımından önem taşımaktadır.

Bu çalışmanın özgün değeri, denim konfeksiyon üretiminden elde edilen gerçek kalite kontrol verilerini makine öğrenmesi yöntemleriyle analiz etmesi ve hata kaynaklı yeniden işlem yükünü lokal karbon ayak izi hesabıyla ilişkilendirmesidir. Literatürde kumaş hata tespiti çoğunlukla görüntü işleme ve derin öğrenme tabanlı otomatik kontrol sistemleri üzerinden ele alınırken, bu çalışma üretim kayıtlarına dayalı tabular veri setiyle hata düzeyi sınıflandırmasına odaklanmaktadır. Ayrıca kalite çıktılarının karbon ayak iziyle birlikte değerlendirilmesi, çalışmayı yalnızca model performansı sunan bir sınıflandırma uygulaması olmaktan çıkarmakta; kalite yönetimi, veri analitiği ve sürdürülebilir üretim arasında daha bütüncül bir ilişki kurmaktadır.

Bu çerçevede çalışma, denim konfeksiyon üretiminde kalite kontrol verilerinin

yalnızca geçmiş hataları raporlayan kayıtlar olarak değil, üretim riski ve çevresel etkiyi birlikte değerlendirmeye olanak sağlayan analitik girdiler olarak kullanılabileceğini göstermektedir. Böylece lif kompozisyonuna dayalı kalite riski sınıflandırması ile hata kaynaklı enerji ve karbon yükünün aynı değerlendirme sistemi içinde ele alınmasına yönelik uygulamalı bir zemin oluşturulmaktadır.

2. ÖNCEKİ ÇALIŞMALAR

2.1. Tekstil Sektöründe Kumaş Hata Tespiti Çalışmaları

Tekstil üretiminde kumaş hatalarının belirlenmesi, kalite kontrol sürecinin en eski fakat güncelliğini koruyan başlıklarından biridir. Kumaş, hazır giyim üretiminin temel girdisi olduğundan, kumaş yüzeyinde ya da yapısında ortaya çıkan hatalar yalnızca ara ürün kalitesini değil, kesim, dikim, yıkama, son kontrol ve müşteri kabul süreçlerini de doğrudan etkilemektedir. Bu nedenle kumaş hatalarının zamanında saptanması, tekstil işletmeleri için hem teknik kalite hem de üretim ekonomisi açısından belirleyici bir konuma sahiptir. Hatanın erken aşamada fark edilmemesi, sonraki işlemlerde fire, ikinci kalite ürün, yeniden işlem ve teslim süresi kaybı gibi sonuçlara yol açmaktadır. Ala ve İkiz, (2015), dokuma üretimi süresince oluşan kumaş hatalarının konfeksiyon aşamasında firelere neden olduğunu, bu hataların üretim sürecinde fark edilmemesi durumunda bitmiş ürünün ikinci kaliteye ayrılabilceğini belirtmiştir. Bu tespit, kumaş hata kontrolünün yalnızca dokuma veya kumaş üretim aşamasına ait teknik bir sorun olmadığını, konfeksiyon üretiminin genel verimliliğiyle de doğrudan ilişkili olduğunu göstermektedir.

Kumaş kalitesinin oluşumu tek bir işlemle açıklanabilecek kadar basit değildir. İplik özellikleri, dokuma hazırlık işlemleri, tezgâh ayarları, üretim hızı, kumaş konstrüksiyonu, operatör müdahalesi ve son işlem koşulları kalite sonucunu birlikte şekillendirmektedir. Kısaoğlu, (2006), dokuma kumaş kalitesinin iplik özelliklerinden başlayarak dokuma hazırlık ve dokuma işlemleri boyunca şekillendiğini, ham veya bitmiş kumaşların hata kontrolünün manuel ya da otomatik sistemlerle yapılabildiğini ifade etmektedir. Bu yaklaşım, kumaş hatasının yalnızca son kontrolde görülen bir kusur olarak değil, üretim sisteminin önceki aşamalarından taşınan bir sonuç olarak ele alınması gerektiğini ortaya koymaktadır. Bu nedenle kumaş hata tespiti çalışmalarında hata türlerinin sayılması kadar, bu hataların hangi üretim koşullarıyla ilişkili olabileceğinin araştırılması da önem taşımaktadır.

Geleneksel kumaş kontrol uygulamaları büyük ölçüde operatör deneyimine ve görsel incelemeye dayanmaktadır. Işıklı kontrol masaları, kalite kontrol panoları ve manuel değerlendirme formları, uzun süre tekstil işletmelerinde temel kontrol araçları olarak kullanılmıştır. Bu yöntemler, özellikle deneyimli personelin bulunduğu üretim

ortamlarında belirli ölçüde işlevsel olsa da yüksek hacimli ve hızlı üretim koşullarında sınırlılıklar taşımaktadır. Kontrol personelinin dikkat düzeyi, yorgunluk, subjektif yorum farkı ve kontrol standardındaki değişkenlik, manuel değerlendirmelerin tekrarlanabilirliğini zayıflatabilmektedir. Kısaoğlu, (2006), manuel kumaş kontrol sistemleri ile otomatik ve çevrim içi kumaş kontrol sistemlerini karşılaştırırken, manuel sistemlerin işletme açısından uygulanabilir olmakla birlikte insan faktörüne bağlı sınırlılıklar içerdiğini vurgulamaktadır. Bu durum, kumaş hata tespiti alanında otomatik ve bilgisayar destekli sistemlere yönelimin temel gerekçelerinden birini oluşturmaktadır.

Kumaş hata tespiti literatüründe ilk dönem çalışmaların önemli bir bölümü istatistiksel kalite kontrol yöntemlerine dayanmaktadır. Ala ve İkiz, (2015), bir dokuma işletmesinde üç hafta boyunca ışıklı panoda yapılan ham kumaş kontrollerinden elde edilen hata kayıtlarını incelemiş, kumaş hatalarını sınıflandırarak Pareto analizi ve p kontrol grafikleri yardımıyla değerlendirmiştir. Bu çalışma, kumaş hata analizinde klasik istatistiksel proses kontrol araçlarının hâlâ işlevsel olduğunu göstermektedir. Pareto analizi, hangi hata türlerinin toplam hata yapısı içinde daha baskın olduğunu belirlemek açısından pratik bir araç sunarken, kontrol grafikleri üretim sürecinin istatistiksel olarak izlenmesine olanak sağlamaktadır. Bununla birlikte bu tür yöntemler çoğunlukla geçmiş hata kayıtlarının değerlendirilmesine dayandığından, hata oluşumunu önceden tahmin etme veya çok değişkenli karmaşık ilişkileri modelleme kapasitesi sınırlı kalmaktadır.

Denim kumaş özelinde yürütülen çalışmalar, kumaş hatalarının üretim süreci ve mamul ürün kalitesiyle ilişkisini daha özgül biçimde ele almaktadır. Acar, (2005), denim kumaş hatalarının optimizasyonuna yönelik çalışmada denim üretim sürecinde ortaya çıkan hata kaynaklarını ve bu hatalara ilişkin çözüm önerilerini incelemiştir. Bu tür çalışmalar, denim üretiminde hata yapısının genel dokuma kumaş üretiminden bazı yönleriyle ayrıştığını göstermektedir. Çünkü denim kumaşlarda iplik yapısı, çözgü boyama, dokuma konstrüksiyonu, elastan veya farklı lif içerikleri, yıkama işlemleri ve son görünüm beklentileri hata algısını doğrudan etkileyebilmektedir. Dolayısıyla denim kumaş hataları yalnızca teknik kusur olarak değil, mamul ürünün estetik ve ticari kabulü açısından da değerlendirilmelidir.

Kapucuoğlu, (2023), denim konfeksiyon işletmelerinde farklı kumaş tiplerine ait hata puanlarını kontrol kartları yardımıyla incelemiş, kumaş özelliklerine bağlı olarak alt ve üst kontrol limitlerinin belirlenmesini amaçlamıştır. Çalışmada üç yıllık işletme verileri kullanılmış ve kumaş hata puanlarının bitmiş ürüne etkisi istatistiksel olarak değerlendirilmiştir. Ayrıca hata puanlarına ilişkin veriler kullanılarak kumaş tiplerinin gruplandırılması amacıyla kümeleme analizi uygulanmıştır. Bu çalışma, denim üretiminde hata analizinin yalnızca hata tespitiyle sınırlı kalmadığını, süreç izleme, kontrol limitlerinin belirlenmesi ve kumaş tiplerinin gruplandırılması gibi karar destek amaçlarıyla da kullanılabileceğini göstermesi bakımından önemlidir. Mevcut çalışma da benzer biçimde gerçek işletme verileri üzerinden ilerlemekte, ancak hata puanı yerine kumaş hata adedini lif kompozisyonu değişkenleriyle birlikte makine öğrenmesi temelli bir sınıflandırma problemi olarak ele almaktadır.

Hazır giyim üretiminde kumaş kalitesini etkileyen faktörlerin çok boyutlu yapısı, yalnızca hata sayımıyla açıklanamayacak kadar karmaşıktır. Tekez ve Özçelik, (2018), konfeksiyon üretimi yapan bir işletmede kumaş kalitesini etkileyen faktörleri bulanık DEMATEL yöntemiyle incelemiş ve uzman görüşlerine dayalı olarak faktörler arasındaki etkileşimleri analiz etmiştir. Çalışmada boyama işleminde kullanılan suyun sertliği önemli bir faktör olarak öne çıkmıştır. Bu sonuç, kumaş kalitesinin yalnızca üretim hattındaki görünür hata kayıtlarıyla değil, proses koşulları, yardımcı malzeme özellikleri ve işletme deneyimiyle birlikte değerlendirilmesi gerektiğini göstermektedir. Mevcut çalışma açısından bu bulgu önemlidir, çünkü eldeki veri seti lif kompozisyonu ve kumaş hata adediyle sınırlıdır. Bu sınır, hata oluşumunun tüm nedenlerini açıklamaya yetmese de lif bileşimi ile hata düzeyi arasındaki ilişkinin veri temelli olarak incelenmesine olanak sağlamaktadır.

Kumaş hata tespitinde otomatik sistemlere yönelim, manuel kontrolün sınırlılıklarına verilen teknolojik bir yanıt olarak gelişmiştir. Bilgisayar destekli kumaş kontrol sistemlerinde kumaş yüzeyinden alınan görüntüler işlenmekte, normal yüzey dokusundan sapan bölgeler hata adayı olarak belirlenmekte ve sınıflandırma algoritmaları yardımıyla hata türleri ayrıştırılmaktadır. Kısaoğlu, (2006), görüntü analizi, sinir ağları, bulanık mantık ve dalgacık temelli modellerin otomatik kumaş kalite kontrol sistemlerinde kullanılabildiğini belirtmiştir.

Bu çerçeve, tekstil sektöründe görüntü işleme tabanlı kalite kontrol uygulamalarının yeni bir eğilim olmadığını, fakat son yıllarda yapay zekâ ve derin öğrenme teknikleriyle daha güçlü hale geldiğini göstermektedir.

Yıldız ve Buldu, (2017), kaşe ve denim kumaş görüntüleri üzerinde kumaş hata tespiti ve sınıflandırması yapmış, termal görüntüleme, dalgacık dönüşümü ve temel bileşen analizi gibi yöntemleri birlikte kullanmıştır. Çalışmada iki farklı kumaş türünde ortalama %95 sınıflandırma başarısı elde edilmiştir. Bu sonuç, görüntü tabanlı yöntemlerin kumaş hata tespitinde yüksek başarı sağlayabileceğini göstermektedir. Ancak burada dikkat edilmesi gereken nokta, bu tür çalışmaların veri yapısının görüntü temelli olmasıdır. Görüntü tabanlı hata tespiti, hatanın yüzeydeki konumunu, biçimini veya görsel karakterini belirlemeye yönelirken, üretim kayıtlarına dayalı tabular veri analizleri daha çok hata düzeyini, risk grubunu veya kalite sonucunu tahmin etmeye odaklanmaktadır. Bu ayrım, mevcut çalışmanın literatürdeki yerini doğru tanımlamak açısından kritik öneme sahiptir.

Pınar vd., (2020)' nin denim kumaşlarda görüntü işleme ile hata tespitine yönelik çalışmaları da bu bağlamda değerlendirilebilir. İlgili çalışma, denim kumaş yüzeylerinin görüntü işleme teknikleriyle incelenmesi ve hatalı bölgelerin belirlenmesi üzerine kurulmuştur. Bu tür araştırmalar, denim özelinde otomatik hata tespit sistemlerinin uygulanabilirliğini göstermesi bakımından değerlidir. Buna karşılık mevcut çalışmada kullanılan veri seti doğrudan görüntü içermemektedir. Bu nedenle çalışmanın yöntemi, görüntü işleme tabanlı hata tespiti değil, lif oranları ve hata adedi üzerinden hata düzeyinin sınıflandırılmasıdır.

Uluslararası literatürde de kumaş hata tespiti alanında benzer bir gelişim çizgisi görülmektedir. Habib vd., (2014), otomatik kumaş hata denetimine ilişkin çalışmalarında manuel kontrolün zaman alıcı olduğunu ve beklenen doğruluk düzeyini her zaman sağlayamadığını belirtmiş, bilgisayarlı görme tabanlı sistemlerin bu sınırlılıkları azaltmak amacıyla geliştirildiğini ifade etmiştir. Hanbay vd., (2016) kumaş hata tespit yöntemlerini yapısal, istatistiksel, spektral, model tabanlı, öğrenme tabanlı, hibrit ve karşılaştırmalı yaklaşımlar altında sınıflandırmıştır. Bu sınıflandırma, kumaş hata tespiti literatürünün yalnızca tek bir yöneme indirgenemeyeceğini göstermektedir. Geleneksel yöntemler belirli kumaş ve hata tiplerinde başarılı sonuçlar verebilirken, öğrenme tabanlı yöntemler daha

karmaşık örüntüleri yakalama kapasitesi nedeniyle son yıllarda daha fazla öne çıkmaktadır.

Güncel çalışmalarda derin öğrenme tabanlı yaklaşımlar, kumaş hata tespiti alanında belirgin bir ağırlık kazanmıştır. (Carrilho vd., 2024), son yıllardaki otomatik kumaş hata tespiti çalışmalarını değerlendirdikleri derlemede, geleneksel yöntemlerden makine öğrenmesi ve derin öğrenme tabanlı modellere doğru belirgin bir yönelim olduğunu belirtmiştir. Bu eğilim, özellikle üretim ortamlarında daha hızlı, daha nesnel ve daha tekrarlanabilir kalite kontrol sistemlerine duyulan ihtiyaçla ilişkilidir. Bununla birlikte derin öğrenme tabanlı yaklaşımların başarılı biçimde uygulanabilmesi için yeterli sayıda etiketlenmiş görüntü verisine, uygun görüntüleme altyapısına ve modelin farklı kumaş/hata türlerine genellenebilirliğini sağlayacak veri çeşitliliğine ihtiyaç vardır. Bu gereklilikler, her işletme için kısa vadede karşılanabilir olmayabilir.

Bu noktada üretim kayıtlarına dayalı kalite analizi ayrı bir önem kazanmaktadır. İşletmelerde kumaş bileşimi, hata adedi, parti bilgisi, kalite kontrol sonucu veya proses kayıtları gibi veriler çoğu zaman görüntü verilerine göre daha erişilebilir durumdadır. Bu veriler, kumaş üzerindeki hatanın görsel konumunu doğrudan göstermese de hata düzeyinin sınıflandırılması, riskli kumaş gruplarının belirlenmesi ve kalite kontrol kararlarının desteklenmesi açısından kullanılabilir. Mevcut çalışma, bu yönüyle görüntü tabanlı hata tespit literatüründen ayrılmakta; kumaş hata problemini üretim kayıtlarına dayalı tabular veri üzerinden ele alan tamamlayıcı bir yaklaşım sunmaktadır.

Genel olarak değerlendirildiğinde, tekstil sektöründe kumaş hata tespiti çalışmaları manuel kontrol, istatistiksel kalite kontrol, görüntü işleme ve yapay öğrenme tabanlı sınıflandırma yaklaşımları etrafında gelişmiştir. Mevcut çalışma bu literatürün doğrudan görüntü işleme kolunda yer almamakta; lif kompozisyonu, hata adedi ve üretim kayıtları üzerinden kalite riskinin veri temelli olarak incelenmesine odaklanmaktadır. Bu nedenle çalışma, kumaş hatasının yalnızca görsel olarak tespit edilmesini değil, hata düzeyine eşlik eden üretim özelliklerinin ve bu hataların operasyonel sonuçlarının birlikte değerlendirilmesini önceleyen bir çerçeveye sahiptir.

2.2. Yapay Öğrenme Tabanlı Kalite ve Hata Tahmin Yaklaşımları

Üretim sistemlerinde kalite kontrol uzun yıllar boyunca ölçüm, gözlem, numune alma,

kontrol kartları ve deneyimli personel değerlendirmesine dayalı biçimde yürütülmüştür. Bu yaklaşım hâlâ birçok işletmede kullanılmakla birlikte, üretim süreçlerinden elde edilen veri miktarının artması kalite yönetiminde daha analitik yöntemlerin kullanımını gündeme getirmiştir. Özellikle sensör verileri, proses kayıtları, ürün özellikleri, kalite kontrol sonuçları ve bakım verileri gibi farklı kaynaklardan elde edilen veriler, kalite problemlerinin yalnızca meydana geldikten sonra belirlenmesini değil, belirli ölçüde önceden tahmin edilmesini de mümkün kılmaktadır. Bu dönüşüm, üretimde kestirimci kalite yaklaşımının temelini oluşturmaktadır.

Kestirimci kalite, üretim sürecinden veya üründen elde edilen veriler kullanılarak kalite çıktılarının tahmin edilmesi ve bu tahminlerin kalite iyileştirme kararlarını desteklemesi olarak değerlendirilmektedir. Tercan ve Meisen, (2022), kestirimci kaliteyi ürün ve proses verilerine dayalı olarak ürünle ilişkili kalite özelliklerinin makine öğrenmesi ve derin öğrenme yöntemleriyle tahmin edilmesi şeklinde tanımlamaktadır. Aynı çalışmada kestirimci kalite uygulamalarının yalnızca gelecekteki kalite değerlerini tahmin etmekle sınırlı olmadığı, kalite sınıflandırması, hata tiplerinin belirlenmesi ve kaliteyi etkileyen süreç değişkenlerinin anlaşılması gibi amaçlara da hizmet edebildiği belirtilmektedir. Bu yaklaşım, mevcut çalışma açısından önemlidir. Çünkü bu çalışmada kumaş yüzeyindeki hatanın görüntü üzerinden bulunması değil, lif kompozisyonu verilerinden hareketle kumaş hata düzeyinin sınıflandırılması amaçlanmaktadır.

Makine öğrenmesi yöntemleri, kalite tahmini problemlerinde özellikle çok değişkenli ve doğrusal olmayan ilişkileri değerlendirebilme kapasitesi nedeniyle öne çıkmaktadır. Üretim ortamlarında kalite çoğu zaman tek bir değişkene bağlı olarak şekillenmez. Hammadde özellikleri, proses parametreleri, makine ayarları, çevresel koşullar ve operatör etkisi bir arada sonuç üretmektedir. Bu tür yapılarda klasik doğrusal modeller bazı ilişkileri açıklayabilse de değişkenler arasındaki etkileşimleri yakalamakta sınırlı kalabilir. Makine öğrenmesi algoritmaları ise veri içindeki örüntüleri doğrudan gözlemleyerek sınıflandırma veya tahmin modelleri kurabilmektedir. Tercan ve Meisen, (2022), 2012–2021 yılları arasında üretimde kestirimci kalite üzerine yayımlanan çalışmaları inceledikleri sistematik derlemede, yapay sinir ağları, derin öğrenme modelleri ve farklı denetimli öğrenme

algoritmalarının kalite tahmininde yaygın biçimde kullanıldığını, buna rağmen gerçek üretim süreçlerine entegrasyon, sınırlı veri setleri ve model genellenebilirliği konularında hâlâ sorunlar bulunduğunu vurgulamaktadır.

Tekstil sektörü, makine öğrenmesi ve yapay sinir ağı uygulamalarının erken ilgi gördüğü üretim alanlarından biridir. Bunun temel nedeni, tekstil üretiminde kalite çıktılarının birçok girdiye bağlı olarak değişmesidir. Lif özellikleri, iplik parametreleri, kumaş konstrüksiyonu, terbiye işlemleri, proses ayarları ve çevresel koşullar ürün performansını birlikte etkilemektedir. Çörekcioğlu vd., (2021), yapay sinir ağlarının tekstil sektöründeki kullanım uygulamalarını değerlendirdikleri çalışmalarında YSA yöntemlerinin iplik parametrelerinin tahmini, dokuma aşaması, terbiye işlemleri, kumaş konfor özellikleri ve çeşitli tekstil performans göstergelerinin modellenmesinde kullanılabildiğini belirtmiştir. Bu derleme, tekstilde yapay öğrenme yöntemlerinin yalnızca hata tespitinde değil, ürün özelliklerinin tahmini ve proses çıktılarının modellenmesinde de kullanılabileceğini göstermektedir.

Ersöz ve Çınar, (2021), pantolon üreten bir tekstil firmasına ait veriler üzerinde veri madenciliği ve makine öğrenimi yaklaşımlarını karşılaştırmış, çoklu doğrusal regresyon, yapay sinir ağları ve karar ağaçları gibi yöntemleri RStudio, Python ve Knime ortamlarında uygulamıştır. Çalışmada model performansları karşılaştırılmış ve tahmin başarısı bakımından yapay sinir ağlarının öne çıktığı ifade edilmiştir. Bu çalışma, mevcut çalışmamızla doğrudan aynı problemi ele almamakla birlikte iki açıdan önemlidir. Birincisi, tekstil işletmelerinde mevcut üretim verilerinin makine öğrenmesiyle modellenebileceğini göstermektedir. İkincisi, farklı algoritmaların aynı veri seti üzerinde karşılaştırılmasının model seçimi açısından gerekli olduğunu ortaya koymaktadır. Mevcut çalışmamızda da benzer biçimde tek bir algoritmanın sonucu üzerinden değerlendirme yapılmamakta, KNN, Random Forest, Bagged Trees, Decision Tree, Naive Bayes, SVM, LDA, Logistic Regression, ANN, RUSBoost, Cost-Sensitive Random Forest ve alt uzay tabanlı yöntemler birlikte ele alınmaktadır.

Yapay sinir ağlarının tekstil alanında kullanımı, özellikle doğrusal olmayan ilişkilerin bulunduğu tahmin problemlerinde dikkat çekmektedir. Uzun vd., (2017), tekstil sektöründe yapay sinir ağının genetik algoritmayla eğitilmesine yönelik çalışmalarında, genetik algoritma destekli eğitimin model performansını iyileştirebildiğini belirtmiştir. Bu tür

çalışmalar, model başarısının yalnızca seçilen algoritmaya değil, eğitim yöntemi, veri ön işleme, parametre ayarı ve optimizasyon stratejilerine de bağlı olduğunu göstermektedir. Bu nokta, mevcut çalışma açısından da geçerlidir. Kumaş hata adedi gibi üretimden gelen kalite çıktıları çoğu zaman düzensiz dağılım gösterebilir, bu nedenle modelin kurulması kadar sınıfların nasıl oluşturulduğu, eğitim-test ayrımının nasıl yapıldığı ve performans ölçütlerinin nasıl yorumlandığı da önem taşımaktadır.

Tekstil ve hazır giyim üretiminde kumaş kalitesini etkileyen faktörler yalnızca sayısal değişkenlerle sınırlı değildir. Uzman deneyimi, proses bilgisi ve üretim koşulları da kalite değerlendirmesinde rol oynamaktadır. (Tekez ve Özçelik, 2018), hazır giyim üretiminde kumaş kalitesini etkileyen faktörleri bulanık DEMATEL yöntemiyle incelemiş ve kaliteyi etkileyen faktörler arasındaki neden-sonuç ilişkilerini uzman görüşleriyle değerlendirmiştir. Bu çalışma, kumaş kalitesinin çok faktörlü yapısını göstermesi bakımından önemlidir. Ancak uzman görüşüne dayalı çok kriterli karar verme yaklaşımları ile makine öğrenmesi tabanlı yaklaşımlar aynı noktada durmamaktadır. Birincisi daha çok uzman bilgisini yapılandırırken, ikincisi veriden örüntü öğrenmeye yönelmektedir. Mevcut çalışmamızda uzman görüşü yerine işletmeden elde edilen üretim verileri kullanılmış, lif kompozisyonu ile kumaş hata adedi arasındaki ilişki veri temelli olarak incelenmiştir.

Kumaş hata tespiti literatürünün önemli bir bölümü görüntü işleme ve derin öğrenme tabanlı yaklaşımlara dayanmaktadır. Carrilho vd., (2024), son yıllarda kumaş hata tespiti alanında makine öğrenmesi ve özellikle derin öğrenme tabanlı yöntemlerin baskın hale geldiğini belirtmektedir. Aynı derlemede kumaş hata tespitinin genellikle görüntü üzerinden yürütüldüğü, geleneksel yöntemlerin istatistiksel, spektral, yapısal ve model tabanlı yaklaşımlar olarak sınıflandırıldığı, güncel çalışmaların ise klasik makine öğrenmesi ve derin öğrenme yöntemleri etrafında yoğunlaştığı aktarılmaktadır. Machado vd., (2025) ise YOLOv11 tabanlı gerçek zamanlı bir sistemle kumaş yüzeyindeki delik, renk akması ve kırışıklık gibi hataları kamera görüntüleri üzerinden tespit etmeye odaklanmıştır. Bu çalışmalar, görüntü tabanlı kumaş hata tespitinin güncel ve güçlü bir araştırma alanı olduğunu göstermektedir. Buna karşılık bu çalışmada kullanılan veri yapısı görüntü değil, üretim kayıtlarından elde edilen tabular verilerdir. Dolayısıyla bu çalışma, kumaş yüzeyindeki

hatanın konumunu veya görsel biçimini belirlememekte, kumaş bileşimine ait sayısal değişkenlerden hareketle hata düzeyini sınıflandırmaktadır.

Bu ayırım, çalışmanın literatürdeki yerini doğru kurmak açısından kritik öneme sahiptir. Görüntü tabanlı hata tespiti sistemlerinde amaç, çoğunlukla kumaş yüzeyinde hatalı bölgenin belirlenmesi, hata türünün sınıflandırılması veya gerçek zamanlı kontrol yapılmasıdır. Bu tür çalışmalarda veri genellikle piksel, görüntü matrisi veya görsel özelliklerden oluşmaktadır. Mevcut çalışmada ise pamuk oranı, geri dönüştürülmüş lif oranı, polyester oranı, elastan oranı ve diğer lif oranları bağımsız değişken olarak kullanılmakta, kumaş hata adedi ise sınıflandırma hedefi olarak değerlendirilmektedir. Bu nedenle çalışma, bilgisayarlı görü tabanlı otomatik hata tespitinden çok, üretim verisi tabanlı kalite riski sınıflandırmasına karşılık gelmektedir. Tercan ve Meisen, (2022) tarafından tanımlanan kestirimci kalite çerçevesi, bu ayrımı açıklamak için daha uygun bir teorik zemin sunmaktadır, çünkü söz konusu çerçeve hem ürün/proses verilerine dayalı kalite tahminini hem de kalite sınıflandırmasını kapsamaktadır.

Makine öğrenmesi tabanlı kalite tahmininde model performansının değerlendirilmesi ayrı bir dikkat gerektirir. Üretim verilerinde sınıflar çoğu zaman dengeli dağılmaz. Normal üretim sınıfı daha fazla gözlem içerirken, yüksek hata sınıfı daha az sayıda gözlemde oluşabilir. Bu durumda model çoğunluk sınıfını iyi öğrenerek yüksek accuracy değeri üretebilir, ancak kalite kontrol açısından asıl kritik olan yüksek hata sınıfını yeterince yakalayamayabilir. Bu nedenle accuracy, tek başına güvenilir bir değerlendirme ölçütü olarak görülmemelidir. Precision, recall ve F1-score gibi ölçütlerin birlikte kullanılması, özellikle azınlık sınıfının önemli olduğu kalite problemlerinde daha anlamlı bir değerlendirme sağlar. Nitekim imalat alanındaki kestirimci kalite çalışmalarında model başarısının yalnızca genel doğrulukla değil, kalite odaklı performans göstergeleri ve gerçek üretim sürecine katkısı üzerinden değerlendirilmesi gerektiği vurgulanmaktadır (Tercan ve Meisen, 2022).

Mevcut çalışmada iki sınıflı ve üç sınıflı sınıflandırma yaklaşımlarının birlikte kullanılması bu bakımdan önemlidir. İki sınıflı yaklaşım, kumaş hata adedini normal ve yüksek hata düzeyleri biçiminde ayırarak üretim açısından riskli grupların pratik şekilde belirlenmesine olanak tanımaktadır. Üç sınıflı yaklaşım ise düşük, orta ve yüksek hata

düzeyleri üzerinden daha ayrıntılı bir kalite sınıflandırması sağlamaktadır. Ancak sınıf sayısı arttığında modelin ayırt etmesi gereken karar sınırları da karmaşılaşmaktadır. Bu nedenle üç sınıflı modellerin daha düşük genel doğruluk üretmesi beklenebilir. Burada temel mesele, en yüksek doğruluğu veren modeli seçmekten ibaret değildir. Üretim uygulaması açısından hangi modelin yüksek hata riskini daha iyi yakaladığı, hangi modelin daha dengeli sonuç verdiği ve hangi sınıflandırma yapısının işletme kararlarına daha uygun olduğu birlikte değerlendirilmelidir.

Tabular üretim verisiyle çalışmanın bazı avantajları ve sınırlılıkları bulunmaktadır. Avantajı, bu tür verilerin birçok işletmede görüntü verilerine göre daha kolay erişilebilir olmasıdır. Üretim reçeteleri, kumaş bileşimi, hata sayıları ve kalite kayıtları çoğu zaman geçmişe dönük olarak tutulmaktadır. Bu veriler kullanılarak işletmenin mevcut dijital altyapısı içinde uygulanabilir karar destek modelleri geliştirilebilir. Buna karşılık tabular veri, hatanın kumaş üzerindeki fiziksel yerini, görsel karakterini veya hata tipini doğrudan göstermez. Dolayısıyla bu çalışmada önerilen yaklaşım, görüntü tabanlı hata tespit sistemlerinin alternatifi değil, kalite riskinin üretim verileri üzerinden önceden değerlendirilmesini sağlayan tamamlayıcı bir yaklaşımdır. Bu konumlandırma, çalışmanın bilimsel tutarlılığını güçlendirmektedir.

Genel olarak değerlendirildiğinde, yapay öğrenme tabanlı kalite ve hata tahmin yaklaşımları tekstil sektöründe iki ana hatta ilerlemektedir. İlk hat, kumaş yüzey görüntülerinden hata belirlemeye yönelik bilgisayarlı görü ve derin öğrenme çalışmalarıdır. İkinci hat ise üretim kayıtları, proses değişkenleri ve ürün özellikleri kullanılarak kalite çıktılarının tahmin edilmesine dayanan veri tabanlı kalite modelleme çalışmalarıdır. Mevcut çalışma ikinci hatta yer almakta, ancak birinci hattaki literatürden tamamen kopmamaktadır. Görüntü tabanlı hata tespiti çalışmaları kumaş hatalarının otomatik belirlenmesi açısından güçlü bir zemin sunarken, bu çalışma hata düzeyinin lif kompozisyonu değişkenleriyle ilişkisini ve üretim kararlarına dönüştürülebilir sınıflandırma sonuçlarını ele almaktadır. Bu yönüyle çalışma, kumaş kalite kontrol literatürüne üretim verisi odaklı, uygulanabilir ve sürdürülebilirlik hesabıyla genişletilebilir bir katkı sunmaktadır.

2.3. Makine Öğrenmesi Tabanlı Sınıflandırma Çalışmaları

Makine öğrenmesi tabanlı sınıflandırma, gözlemlerin belirli özelliklerinden hareketle önceden tanımlanmış sınıflara atanmasını amaçlayan denetimli öğrenme problemlerinden biridir. Üretim sistemlerinde bu yaklaşım, ürün uygunluğu, hata riski, kalite seviyesi, arıza durumu veya proses çıktısının sınıflandırılması gibi farklı amaçlarla kullanılmaktadır. Kestirimci kalite çalışmalarında sınıflandırma modellerinin temel işlevi, geçmiş üretim verilerinden öğrenilen örüntüler yardımıyla yeni ürün veya proses kayıtlarının kalite açısından hangi sınıfa ait olabileceğini tahmin etmektir. Tercan ve Meisen, (2022), üretimde kestirimci kalite uygulamalarının son yıllarda makine öğrenmesi ve derin öğrenme yöntemleriyle genişlediğini, ancak model seçimi, veri kalitesi, sınıf dağılımı ve gerçek üretim süreçlerine aktarılabilirlik konularının hâlâ dikkatle ele alınması gerektiğini belirtmektedir.

Sınıflandırma modelleri, üretim verilerinin yapısına göre farklı avantajlar ve sınırlılıklar göstermektedir. K-En Yakın Komşu algoritması, sınıflandırılacak bir gözlemin eğitim verisindeki en yakın gözlemlerle olan benzerliğine göre karar veren örnek tabanlı bir yöntemdir. Bu algoritma, karar sınırını önceden katı biçimde varsaymadığı için esnek bir yapı sunmaktadır. Buna karşılık değişkenlerin ölçeklerinden ve uzaklık ölçütünden doğrudan etkilenmektedir. Bu nedenle KNN kullanılan çalışmalarda veri normalizasyonu ve uygun komşu sayısının belirlenmesi model performansı açısından kritik hale gelmektedir. Tekstil üretimi gibi sayısal değişkenlerin farklı ölçeklerde bulunabildiği veri yapılarında bu nokta ayrıca önem taşımaktadır.

Karar ağaçları, sınıflandırma problemlerinde yorumlanabilirliği yüksek yöntemler arasında yer almaktadır. Değişkenler belirli eşik değerleri üzerinden ayrılarak ağaç yapısı oluşturulmakta ve sınıflandırma bu dallanma yapısı üzerinden yapılmaktadır. Karar ağacı yaklaşımının en güçlü yönü, model kararlarının görsel ve kural tabanlı biçimde izlenebilmesidir. Bu özellik, üretim sahasında çalışan uzmanların model çıktısını anlamasını kolaylaştırmaktadır. Ancak tekil karar ağaçları, özellikle veri setinde gürültü veya karmaşık değişken etkileşimleri bulunduğunda aşırı öğrenmeye yatkın olabilir. Blockeel vd., (2023), karar ağaçlarının açıklanabilirlik bakımından önemli avantajlar taşıdığını, buna karşılık sorumlu ve güvenilir yapay zekâ uygulamalarında model karmaşıklığı, genellenebilirlik ve

karar kurallarının dikkatle değerlendirilmesi gerektiğini vurgulamaktadır.

Random Forest yöntemi, karar ağaçlarının tekil yapısından kaynaklanan kararsızlıkları azaltmak amacıyla çok sayıda ağacın birlikte kullanıldığı ensemble tabanlı bir yaklaşımdır. Breiman, (2001), Random Forest algoritmasını, her biri rastgele seçilmiş veri ve değişken alt kümeleriyle eğitilen karar ağaçlarının birleşimi olarak tanımlamıştır. Bu yapı, tek bir karar ağacına kıyasla daha kararlı ve genellenebilir sonuçlar üretebilmektedir. Random Forest algoritmasının üretim verilerinde sık tercih edilmesinin nedenlerinden biri, doğrusal olmayan ilişkileri yakalayabilmesi ve değişken önemini değerlendirmeye olanak sağlamasıdır. Mevcut araştırmada lif kompozisyonu değişkenlerinin kumaş hata düzeyine katkısını değerlendirmek açısından bu özellik özellikle anlamlıdır. Ancak Random Forest modelinin yüksek genel doğruluk üretmesi, her zaman azınlık sınıfını yeterli düzeyde öğrendiği anlamına gelmemektedir. Bu nedenle modelin yalnızca accuracy değeriyle değil, sınıf düzeyindeki performans ölçütleriyle birlikte yorumlanması gerekmektedir.

Destek Vektör Makineleri, sınıflar arasında en geniş marjı sağlayan ayırıcı hiper düzlemi bulmaya çalışan güçlü sınıflandırma yöntemlerinden biridir. Cortes vd., (1995), destek vektör yaklaşımının özellikle iki sınıflı sınıflandırma problemlerinde yüksek genelleme kapasitesi sunduğunu belirtmiştir. SVM, doğrusal olarak ayrılmayan veri yapılarında çekirdek fonksiyonları yardımıyla daha karmaşık karar sınırları oluşturabilmektedir. Bu özellik, üretim verilerinde değişkenler arasındaki ilişkilerin doğrusal olmadığı durumlarda avantaj sağlayabilir. Bununla birlikte SVM modellerinin performansı çekirdek seçimi, ceza parametresi, veri ölçeklendirme ve sınıf dağılımından etkilenmektedir. Dengesiz veri yapılarında standart SVM, çoğunluk sınıfına daha yakın karar sınırları üretebilir ve azınlık sınıfına ait kritik gözlemleri kaçırabilir. Bu nedenle SVM tabanlı modellerde sınıf ağırlıkları, yeniden örnekleme veya maliyet duyarlı yaklaşımlar önemli hale gelmektedir.

Naive Bayes algoritması, Bayes teoremine dayanan ve değişkenler arasında koşullu bağımsızlık varsayımı yapan olasılıksal bir sınıflandırma yöntemidir. Domingos ve Pazzani, (1997), basit Bayes sınıflandırıcısının güçlü bağımsızlık varsayımına rağmen birçok sınıflandırma probleminde beklenenden iyi sonuçlar üretebildiğini göstermiştir. Naive Bayes yönteminin avantajı, hızlı çalışması ve sınırlı veri durumlarında uygulanabilir olmasıdır.

Ancak tekstil üretiminde lif oranları gibi değişkenler çoğu zaman birbirinden bağımsız değildir. Bir lif oranındaki artış, toplam kompozisyon yapısı nedeniyle diğer oranların azalmasıyla ilişkili olabilir. Bu nedenle Naive Bayes modelinin çıktıları, bu varsayım dikkate alınarak yorumlanmalıdır. Model düşük hesaplama maliyetiyle karşılaştırmalı bir referans sağlayabilir, fakat karmaşık değişken etkileşimlerini yakalama kapasitesi sınırlı kalabilir.

Lojistik regresyon, sınıflandırma problemlerinde uzun süredir kullanılan yorumlanabilir yöntemlerden biridir. Özellikle iki sınıflı problemlerde bağımsız değişkenlerin sınıf olasılığı üzerindeki etkisini değerlendirmek için tercih edilmektedir. Menard, (2010), lojistik regresyonun kategorik bağımlı değişkenlerin analizinde hem tahmin hem de açıklama amacıyla kullanılabilirliğini belirtmektedir. Üretim kalitesi gibi uygulamalarda lojistik regresyon, karmaşık modellere geçmeden önce temel bir karşılaştırma modeli olarak değerlendirilebilir. Bununla birlikte lif kompozisyonu ile hata düzeyi arasındaki ilişkinin doğrusal olmayan yapıda olması durumunda lojistik regresyonun sınıflandırma başarısı sınırlanabilir. Bu nedenle mevcut araştırmada lojistik regresyon, daha karmaşık modellerin performansını değerlendirmek için anlamlı bir temel model niteliği taşımaktadır.

Yapay sinir ağları, doğrusal olmayan ilişkileri öğrenebilme kapasiteleri nedeniyle tekstil alanında sık kullanılan yöntemlerden biridir. Çörekcioğlu vd., (2021), yapay sinir ağlarının tekstil sektöründe iplik parametrelerinin tahmini, dokuma aşaması, terbiye işlemleri ve kumaş konfor özelliklerinin modellenmesi gibi farklı alanlarda kullanıldığını belirtmiştir. Uzun vd., (2017) ise yapay sinir ağlarının genetik algoritmalarla eğitilmesinin model performansını iyileştirebileceğini göstermiştir. Bu tür modeller, karmaşık veri yapılarında güçlü sonuçlar verebilmekle birlikte, veri miktarı, ağ mimarisi, eğitim algoritması ve aşırı öğrenme riski açısından dikkatli biçimde kurulmalıdır. Ayrıca yapay sinir ağlarının karar mekanizması, karar ağaçları veya lojistik regresyon kadar kolay yorumlanamayabilir. Üretim kararlarında modelin yalnızca doğru sonuç üretmesi değil, uzmanlar tarafından anlaşılabilir olması da çoğu zaman önem taşımaktadır.

Ensemble yöntemleri, birden fazla sınıflandırıcının birlikte kullanılması yoluyla daha güçlü ve kararlı tahminler üretmeyi amaçlamaktadır. Dietterich, (2000), ensemble yöntemlerini birden fazla sınıflandırıcının tahminlerini ağırlıklı ya da çoğunluk oyu ile

birleştiren yaklaşımlar olarak açıklamaktadır. Bagging, boosting, Random Forest ve alt uzay tabanlı yöntemler bu kapsamda değerlendirilebilir. Bagging yaklaşımında modeller farklı örneklem alt kümeleri üzerinde eğitilerek varyansın azaltılması amaçlanırken, boosting yöntemlerinde önceki modellerin hatalı sınıflandırdığı gözlemlere daha fazla ağırlık verilerek ardışık öğrenme sağlanmaktadır. Bu yapı, üretim verilerinde tekil modellerin yakalayamadığı örüntülerin daha iyi öğrenilmesine katkı sağlayabilir. Ancak ensemble modellerinin daha yüksek performans üretmesi, yorumlanabilirlik açısından her zaman avantajlı değildir. Bu nedenle model seçiminde performans, yorumlanabilirlik ve üretimde uygulanabilirlik birlikte değerlendirilmelidir.

Sınıf dengesizliği, makine öğrenmesi tabanlı sınıflandırma çalışmalarında en kritik sorunlardan biridir. Dengesiz veri yapılarında bir sınıf diğer sınıflara göre çok daha fazla gözlem içerdiğinde, standart sınıflandırıcılar çoğunluk sınıfını öğrenmeye eğilim gösterebilir. He ve Garcia, (2009), dengesiz veri setlerinde geleneksel sınıflandırıcıların çoğunluk sınıfına yanlı sonuçlar üretebildiğini ve azınlık sınıfına ait örnekleri yeterli düzeyde ayırt edemeyebildiğini belirtmiştir. Bu sorun üretim kalitesi verilerinde sık görülmektedir. Normal ürün veya düşük hata sınıfı çoğu zaman veri setinde daha baskın durumdayken, yüksek hata sınıfı daha az sayıda gözlemde oluşmaktadır. Oysa üretim yönetimi açısından asıl kritik olan sınıf çoğu zaman bu azınlık sınıfıdır.

Dengesiz veri sorunu, yalnızca veri dağılımına ait teknik bir problem değildir, model çıktılarının işletme kararlarına etkisini doğrudan değiştiren bir durumdur. Bir model çok sayıda normal gözlemi doğru sınıflandırarak yüksek accuracy değerine ulaşabilir. Ancak aynı model yüksek hata sınıfını yakalayamıyorsa, kalite kontrol açısından güvenilir bir karar destek aracı olarak görülemez. Maraş ve Selçukcan Erol, (2022), dengesiz veri kümelerinde sınıflandırma probleminin son yıllarda yoğun biçimde çalışıldığını, örnekleme yöntemleri, maliyet duyarlı öğrenme ve değerlendirme metriklerinin bu alandaki temel çözüm başlıkları arasında yer aldığını belirtmiştir. Bu tespit, mevcut araştırmanın iki sınıflı model sonuçlarının yorumlanması açısından doğrudan önemlidir. Çünkü bazı modellerin yüksek genel doğruluk üretmesine rağmen yüksek hata sınıfını yeterli düzeyde öğrenememesi, dengesiz veri yapısının pratik bir sonucudur.

Dengesiz veri kümeleriyle baş etmek için yeniden örnekleme ve maliyet duyarlı öğrenme yaklaşımları geliştirilmiştir. Yeniden örnekleme yöntemlerinde çoğunluk sınıfı azaltılabilir, azınlık sınıfı artırılabilir veya iki yaklaşım birlikte kullanılabilir. Chawla vd., (2002), SMOTE yönteminde azınlık sınıfına ait sentetik örnekler üretilerek sınıflandırıcının azınlık sınıfını daha iyi öğrenmesinin hedeflendiğini belirtmiştir. Daha güncel çalışmalarda ise ADASYN, hibrit örnekleme ve SVM tabanlı örnek azaltma gibi yaklaşımlar geliştirilmiştir. (Yılmaz Eroğlu, 2025), dengesiz veri setlerinde ADASYN ile örneklem artırma ve SVM tabanlı örneklem azaltmayı birleştiren iki aşamalı bir dengeleme stratejisi önermiş, bu yaklaşımı SVM, Random Forest, XGBoost ve KNN sınıflandırıcılarıyla farklı veri setleri üzerinde test etmiştir. Çalışmada gerçek üretim ortamından elde edilen tekstil veri setinin de kullanılması, dengesiz veri probleminin tekstil üretimi açısından teorik değil, sahada karşılaşılan gerçek bir sorun olduğunu göstermektedir.

Maliyet duyarlı sınıflandırma, dengesiz veri problemini yalnızca gözlem sayıları üzerinden değil, hatalı sınıflandırmanın sonuçları üzerinden ele almaktadır. Üretim kalitesi açısından normal bir ürünü yüksek hata sınıfına atamak ile yüksek hata riski taşıyan bir ürünü normal kabul etmek aynı maliyete sahip değildir. İkinci durumda hatalı ürün üretim sürecinde ilerleyebilir, yeniden işlem yükü artabilir veya müşteri şikâyeti oluşabilir. Bu nedenle bazı modellerde azınlık sınıfına daha yüksek hata maliyeti atanarak sınıflandırıcının kritik sınıfı daha fazla dikkate alması sağlanmaktadır. (Liu ve Guo, 2024), dengesiz çok sınıflı veri kümelerinde doğrusal ayırma analizine maliyet duyarlı ağırlıklandırma ekleyerek sınıf dengesizliğinin etkisini azaltmayı amaçlamıştır. Bu yaklaşım, sınıf dengesizliği sorununa yalnızca örneklem sayısı açısından değil, modelin öğrenme mekanizması içinde de müdahale edilebileceğini göstermektedir.

Model performansının değerlendirilmesinde kullanılan ölçütler, sınıflandırma probleminin amacına uygun seçilmelidir. Accuracy, tüm doğru sınıflandırmaların toplam gözlem sayısına oranını verdiği için genel başarı hakkında hızlı bir fikir sunmaktadır. Ancak sınıf dengesizliği bulunduğu yanıtıcı olabilir. Örneğin gözlemlerin büyük bölümü normal sınıfta yer alıyorsa, model neredeyse tüm gözlemleri normal olarak sınıflandırarak yüksek accuracy elde edebilir. Buna rağmen yüksek hata sınıfını yakalayamadığı için üretim

açısından başarısız sayılabilir. Bu nedenle precision, recall ve F1-score ölçütlerinin birlikte kullanılması gerekmektedir.

Precision, modelin pozitif olarak sınıflandırdığı gözlemler içinde gerçekten pozitif olanların oranını göstermektedir. Yüksek precision, modelin yüksek hata tahmini yaptığında daha güvenilir olduğunu gösterir. Recall ise gerçekte yüksek hata sınıfında bulunan gözlemlerin ne kadarının model tarafından yakalanabildiğini ifade etmektedir. Kalite kontrol bağlamında recall ölçütü özel bir öneme sahiptir, çünkü yüksek hata sınıfının kaçırılması üretim açısından daha riskli sonuçlar doğurabilir. F1-score ise precision ve recall değerlerinin harmonik ortalamasıdır. Bu nedenle modelin yalnızca yüksek hata tahmini yapmadaki doğruluğunu değil, aynı zamanda yüksek hata sınıfını yakalama kapasitesini de birlikte değerlendirmeye olanak tanır.

Mevcut araştırmada iki sınıflı model sonuçlarının yorumlanmasında bu nokta belirleyicidir. Bazı modeller yüksek accuracy değerleri üretse bile yüksek hata sınıfını yeterince öğrenemediğinde kalite kontrol açısından sınırlı bir katkı sunmaktadır. Buna karşılık recall veya F1-score değeri daha dengeli olan modeller, genel doğrulukları bir miktar düşük olsa da üretim riski açısından daha anlamlı sonuçlar verebilir. Bu nedenle model karşılaştırmasında yalnızca en yüksek accuracy değerine odaklanmak doğru değildir. Yüksek hata sınıfının doğru belirlenmesi, üretim sürecinde erken uyarı, yeniden işlem yükünün azaltılması ve kalite maliyetlerinin kontrol altına alınması açısından daha kritik bir ölçüttür.

Bu çerçevede kullanılan sınıflandırma algoritmaları, yalnızca teknik performans üretmek amacıyla değil, farklı öğrenme mantıklarını aynı veri seti üzerinde karşılaştırmak amacıyla ele alınmaktadır. KNN benzerlik temelli bir yaklaşım sunarken, karar ağacı kural tabanlı açıklanabilirlik sağlamaktadır. Random Forest ve Bagged Trees gibi ensemble yöntemleri daha kararlı tahminler üretmeyi hedeflemekte, SVM daha güçlü karar sınırları oluşturabilmekte, Naive Bayes olasılıksal bir referans sunmakta, lojistik regresyon yorumlanabilir temel model işlevi görmekte ve yapay sinir ağları doğrusal olmayan örüntüleri öğrenmeye çalışmaktadır. RUSBoost ve maliyet duyarlı Random Forest gibi yaklaşımlar ise özellikle dengesiz sınıf dağılımı altında yüksek hata sınıfının daha iyi yakalanmasına

yönelik tamamlayıcı modeller olarak değerlendirilmektedir.

Genel olarak bakıldığında, makine öğrenmesi tabanlı sınıflandırma çalışmaları üretim kalitesi problemlerinde güçlü olanaklar sunmaktadır. Ancak bu olanakların anlamlı hale gelebilmesi için modelin veri yapısına uygun seçilmesi, sınıf dengesizliğinin dikkate alınması ve performans ölçütlerinin üretim amacıyla uyumlu yorumlanması gerekir. Bu çalışma, kumaş hata adedini iki ve üç sınıflı yapılar üzerinden değerlendirerek farklı sınıflandırma algoritmalarının üretim verisi üzerindeki davranışını karşılaştırmaktadır. Böylece yalnızca hangi modelin daha yüksek doğruluk verdiği değil, hangi modelin yüksek hata sınıfını daha etkili yakalayabildiği de tartışmaya açılmaktadır.

2.4. Karbon Ayak İzi ve Sürdürülebilir Üretim Çalışmaları

Tekstil ve hazır giyim sektörü, hammadde kullanımı, enerji tüketimi, su ihtiyacı, kimyasal kullanımı, atık oluşumu ve sera gazı emisyonları bakımından çevresel etkileri yüksek sektörler arasında değerlendirilmektedir. Bu etkinin yalnızca nihai ürün üretiminden kaynaklanmadığı; lif üretimi, iplik, dokuma veya örme, boyama, baskı, terbiye, konfeksiyon, dağıtım, kullanım ve bertaraf aşamalarına yayıldığı görülmektedir. Bu nedenle tekstil sektöründe sürdürülebilirlik çalışmaları giderek daha geniş bir çerçevede ele alınmakta; yalnızca çevre dostu hammadde seçimi değil, üretim verimliliği, kalite iyileştirme, yeniden işlem oranlarının azaltılması, enerji yönetimi ve döngüsel ekonomi uygulamaları da bu kapsamda değerlendirilmektedir.

Karbon ayak izi, bir ürün, süreç, faaliyet veya kuruluş nedeniyle ortaya çıkan sera gazı emisyonlarının karbondioksit eşdeğeri cinsinden ifade edilmesini sağlayan temel çevresel göstergelerden biridir. Bu gösterge, üretim faaliyetlerinin iklim değişikliği üzerindeki etkisini sayısallaştırmaya olanak vermektedir. ISO 14067 standardında ürün karbon ayak izinin yaşam döngüsü değerlendirmesi ilkeleriyle uyumlu biçimde hesaplanması gerektiği belirtilmekte; aynı zamanda ürünün tamamını değil, belirli bir bölümünü kapsayan kısmi karbon ayak izi çalışmalarının da yapılabileceği ifade edilmektedir (ISO, 2018). Bu ayırım, tekstil üretimi gibi çok aşamalı ve karmaşık değer zincirlerinde önemlidir. Çünkü her çalışma hammadde üretiminden bertarafa kadar tüm yaşam döngüsünü kapsamak zorunda değildir; bazı çalışmalar belirli bir işletme, proses veya faaliyet düzeyinde daha sınırlı fakat ölçülebilir

bir değerlendirme yapabilir.

Tekstil sektöründe karbon ayak izi ve sürdürülebilirlik literatürü genel olarak dört ana yaklaşım altında değerlendirilebilir. İlk yaklaşım, yaşam döngüsü analizi çalışmalarından oluşmaktadır. Bu çalışmalar, ürünün çevresel etkisini hammadde temininden nihai bertarafına kadar uzanan geniş bir sistem sınırı içinde değerlendirmektedir. İkinci yaklaşım, üretim süreci odaklı karbon ayak izi hesaplamalarına dayanmaktadır. Bu gruptaki çalışmalar belirli bir fabrika, bölüm, proses veya üretim hattı üzerinde yoğunlaşmaktadır. Üçüncü yaklaşım, enerji tüketimi temelli emisyon hesaplamalarını içermektedir. Bu çalışmalarda elektrik, doğal gaz, kömür, motorin veya buhar gibi enerji girdileri üzerinden CO₂e hesaplaması yapılmaktadır. Dördüncü yaklaşım ise kalite problemleri, yeniden işlem, fire, kaynak verimliliği ve sürdürülebilirlik arasındaki ilişkiye odaklanmaktadır. Mevcut çalışma, bu dördüncü yaklaşıma daha yakın konumlanmaktadır. Çünkü burada amaç denim ürününün tüm yaşam döngüsünü hesaplamak değil, kumaş hatalarının neden olduğu yeniden işlem yükünü ve bu yükün ek elektrik tüketimi üzerinden karbon ayak izine etkisini görünür hale getirmektir.

Yaşam döngüsü analizi çalışmaları, tekstil ve denim ürünlerinin çevresel etkilerini geniş bir sistem sınırı içinde değerlendirmesi bakımından literatürde güçlü bir yer tutmaktadır. Khatun, (2024), organik ve konvansiyonel pamuk kullanılarak üretilen jean ürünlerini yaşam döngüsü analiziyle karşılaştırmış; pamuk üretimi, iplik, kumaş üretimi, konfeksiyon, kullanım ve atık aşamalarının çevresel etkiye katkısını değerlendirmiştir. Bu tür çalışmalar, denim ürününün çevresel yükünün yalnızca konfeksiyon aşamasında oluşmadığını, özellikle pamuk üretimi ve ıslak işlemler gibi aşamaların toplam çevresel etki içinde önemli paya sahip olduğunu göstermektedir. Hossain ve Rahman, (2025) da yeşil ve geleneksel denim üretim uygulamalarını karşılaştırdıkları çalışmada, sürdürülebilir üretim uygulamalarının sera gazı emisyonu ve su tüketimi üzerinde belirgin azaltım potansiyeli taşıdığını ortaya koymuştur. Benzer şekilde Semba vd., (2024), bagas esaslı jean üretimi için döngüsel ekonomi modelini karbon ayak izi açısından değerlendirmiş; boyama, ağartma ve terbiye gibi ıslak işlemlerin enerji tüketimi nedeniyle emisyonlarda önemli paya sahip olduğunu belirtmiştir.

Yaşam döngüsü analizi çalışmalarının en güçlü yönü, ürünün çevresel etkisini bütüncül

bir bakışla ele almasıdır. Ancak bu çalışmalar yüksek düzeyde veri gerektirmektedir. Hammadde üretimi, enerji kaynakları, kimyasal kullanımı, taşıma, tüketici kullanımı ve bertaraf aşamalarına ilişkin ayrıntılı veriler olmadan güvenilir bir yaşam döngüsü analizi yürütmek güçleşmektedir. Denim üretimi gibi karmaşık bir ürün grubunda bu durum daha da belirgindir. Bu nedenle her işletme verisiyle tam kapsamlı bir yaşam döngüsü analizi yapmak mümkün olmayabilir. Mevcut çalışma bu nedenle ürün yaşam döngüsünün tamamını değerlendirmek yerine, işletme içindeki kalite hatalarından kaynaklanan yeniden işlem sürecine odaklanmaktadır. Bu tercih, çalışmanın veri kapsamı ve ölçülebilirlik düzeyiyle daha uyumlu bir yöntemsel çerçeve oluşturmaktadır.

İkinci yaklaşım, üretim süreci odaklı karbon ayak izi hesaplamalarıdır. Coşkun ve Doğan, (2021), bir tekstil fabrikasının konfeksiyon, kumaş boyama, baskı ve iplik boyama bölümlerinde 2018 yılı üretim faaliyetlerinden kaynaklanan karbon ayak izini hesaplamıştır. Çalışmada konfeksiyon bölümünün 24,39 kg CO₂e/kg ürün ile en yüksek karbon ayak izi değerine sahip olduğu; kumaş boyama, baskı ve iplik boyama bölümlerinin ise sırasıyla daha düşük değerlere sahip olduğu bildirilmiştir. Araştırmacılar, karbon ayak izinin önemli ölçüde doğal gaz, kömür ve motorin kullanımından kaynaklandığını belirtmiştir. Bu çalışma, tekstil işletmelerinde bölüm bazlı karbon ayak izi hesaplamalarının yapılabileceğini ve üretim bölümleri arasında çevresel yük bakımından farklılıklar bulunabileceğini göstermesi açısından önemlidir.

Üretim süreci odaklı çalışmalar, işletme yönetimi açısından yaşam döngüsü analizlerine göre daha uygulanabilir sonuçlar üretebilir. Çünkü işletmeler çoğu zaman kendi üretim bölümlerine ait enerji tüketimi, üretim miktarı ve yakıt kullanımı verilerine daha kolay ulaşabilmektedir. Bu tür çalışmalar, hangi bölümün daha yüksek emisyon ürettiğini, hangi enerji kaynağının daha baskın olduğunu ve karbon azaltımının nereden başlatılması gerektiğini gösterebilir. Ancak bu yaklaşımın da sınırları vardır. Bölüm bazlı karbon ayak izi hesaplaması, kalite hatalarının neden olduğu ek kaynak kullanımını her zaman ayırtmaz. Örneğin konfeksiyon bölümünün toplam enerji tüketimi hesaplanabilir; fakat bu tüketimin ne kadarının normal üretimden ne kadarının yeniden işlemden kaynaklandığı ayrıca analiz edilmezse kalite problemlerinin çevresel yükü görünmez kalabilir.

Üçüncü yaklaşım, enerji tüketimi temelli karbon emisyonu hesaplamalarına dayanmaktadır. Bu yaklaşımda belirli bir faaliyet, bina, proses veya üretim birimi için tüketilen enerji miktarı belirlenmekte ve uygun emisyon faktörleri kullanılarak CO₂e değeri hesaplanmaktadır. Tosun ve Dede, (2024), enerji tüketiminin karbon ayak izi hesaplamalarındaki rolünü bir mühendislik fakültesi binası üzerinden incelemiş; elektrik ve diğer enerji tüketimlerinden kaynaklanan emisyonları Greenhouse Gas Protocol ve ISO 14064 çerçevesinde değerlendirmiştir. Çalışmada karbon ayak izinin azaltılabilmesi için öncelikle enerji tüketiminin belirlenmesi gerektiği vurgulanmıştır. Bu yaklaşım, mevcut çalışma açısından doğrudan önemlidir. Çünkü yeniden işlemden kaynaklanan karbon ayak izi hesabı da ek enerji tüketiminin sayısallaştırılmasına ve bu tüketimin emisyon faktörüyle ilişkilendirilmesine dayanmaktadır.

Elektrik tüketimine bağlı emisyon hesaplamalarında kullanılan yöntemsel çerçeve ayrıca önem taşımaktadır. GHG Protocol Scope 2 rehberi, satın alınan veya tüketilen elektrik, buhar, ısı ve soğutmadan kaynaklanan dolaylı emisyonların nasıl raporlanacağını açıklamaktadır (WRI & WBCSD, 2015). Bu rehberde konum temelli ve piyasa temelli iki yaklaşım tanımlanmaktadır. Konum temelli yaklaşım, tüketimin gerçekleştiği elektrik şebekesinin ortalama emisyon yoğunluğunu dikkate alırken; piyasa temelli yaklaşım, elektrik tedarik sözleşmeleri ve sertifikalı enerji kaynakları gibi unsurları hesaba katmaktadır. Mevcut çalışma, işletme düzeyinde lokal yeniden işlem etkisini değerlendirdiği için elektrik tüketimi ile emisyon faktörü arasındaki ilişkiyi esas alan sınırlı fakat ölçülebilir bir hesaplama mantığına dayanmaktadır.

Tekstil sektöründe enerji tüketimi ve karbon emisyonu ilişkisi yalnızca enerji kaynağı tercihiyle açıklanamaz. Üretim verimliliği, fire oranı, tekrar işlem gereksinimi, makine duruşları ve hatalı üretim gibi kalite ve süreç faktörleri de toplam kaynak kullanımını artırmaktadır. Nobil vd., (2023) düşük kaliteli ürün sayısının azalmasının yalnızca fireyi azaltmakla kalmayacağını, aynı zamanda yeniden işlem için gereken enerji ve iş yükünü de düşüreceğini ifade etmektedir. Bu bakış, kalite problemleri ile sürdürülebilirlik arasındaki ilişkiyi anlamak açısından değerlidir. Çünkü üretim sisteminde hata oranı arttığında yalnızca üretim maliyeti yükselmez; aynı zamanda aynı ürün miktarına ulaşmak için daha fazla

zaman, daha fazla enerji ve daha fazla kaynak kullanılır.

Kalite problemleri ile sürdürülebilir üretim arasındaki ilişki tekstil işletmeleri açısından özel bir öneme sahiptir. Kumaş hataları veya konfeksiyon hataları yeniden işleme gönderildiğinde, hatanın niteliğine bağlı olarak ek kontrol, sökme, dikim, düzeltme, ütölme, taşıma veya yeniden paketleme işlemleri gerekebilir. Bu işlemlerin her biri belirli bir enerji ve işçilik karşılığına sahiptir. Büyük ölçekli üretim yapan işletmelerde birim ürün başına küçük görünen bu ek yükler, toplam üretim hacmi içinde dikkate değer bir çevresel etkiye dönüşebilir. Bu nedenle kalite hatalarının karbon ayak izi üzerindeki etkisinin hesaplanması, sürdürülebilir üretim çalışmalarında ihmal edilmemesi gereken bir alan olarak görülmelidir.

Mevcut çalışmanın karbon ayak izi yaklaşımı tam yaşam döngüsü analizi değildir. Çalışmada hammadde üretimi, kumaş üretimi, boyama, yıkama, taşıma, kullanım ve bertaraf aşamaları hesaplama sınırına dahil edilmemektedir. Değerlendirme, denim konfeksiyon işletmesinde kumaş hatalarından kaynaklanan yeniden işlem süreciyle sınırlıdır. Toplam hata adedi, hata başına ortalama yeniden işlem süresi, bu sürenin neden olduğu ek elektrik tüketimi ve elektrik tüketimine bağlı CO₂e emisyonu hesaplanmaktadır. Bu sınırlandırma, çalışmanın karbon ayak izi iddiasını abartmadan, eldeki üretim verisiyle doğrudan ilişkilendirilebilir bir çevresel gösterge üretmesini sağlamaktadır.

Bu konumlandırma, literatürdeki dördüncü yaklaşım olan kalite problemleri ile sürdürülebilirlik ilişkisine daha yakındır. Yaşam döngüsü analizi çalışmaları ürünün bütün çevresel yükünü ortaya koyarken, üretim süreci odaklı çalışmalar fabrika veya bölüm düzeyinde emisyonları incelemektedir. Enerji tüketimi temelli hesaplamalar ise tüketilen enerji miktarını karbon emisyonuna dönüştürmektedir. Mevcut çalışma bu yaklaşımlardan yararlanmakla birlikte, asıl olarak kalite hatasının neden olduğu ek kaynak kullanımını görünür hale getirmeye yönelmektedir. Başka bir ifadeyle burada karbon hesabı genel üretim emisyonunu değil, hatadan kaynaklanan ilave çevresel yükü temsil etmektedir.

Bu ayırım bilimsel açıdan önemlidir. Çünkü kalite hataları çoğu zaman maliyet, fire veya müşteri memnuniyeti bağlamında tartışılmakta; çevresel etkileri ikincil planda kalmaktadır. Oysa yeniden işlem gerektiren her hata, normal üretim akışına eklenen fazladan bir faaliyet anlamına gelmektedir. Bu faaliyet enerji tükettiği ölçüde karbon emisyonuna

da katkıda bulunmaktadır. Dolayısıyla kalite iyileştirme faaliyetleri yalnızca verimlilik ve maliyet azaltımı açısından değil, karbon azaltımı ve kaynak verimliliği açısından da değerlendirilmelidir. Bu bakış açısı, sürdürülebilir üretim yaklaşımının kalite yönetimiyle daha doğrudan ilişkilendirilmesine olanak sağlamaktadır.

Özetle, tekstil sektöründe karbon ayak izi ve sürdürülebilir üretim literatürü geniş bir yöntemsel çeşitlilik göstermektedir. Yaşam döngüsü analizi çalışmaları denim ürünlerinin çevresel etkisini bütüncül biçimde ortaya koyarken, süreç odaklı çalışmalar işletme içindeki emisyon kaynaklarını belirlemeye yardımcı olmaktadır. Enerji tüketimi temelli hesaplamalar, karbon emisyonunun ölçülebilir bir girdiye bağlanmasını sağlamaktadır. Kalite problemleriyle sürdürülebilirlik arasındaki ilişkiyi inceleyen çalışmalar ise üretim hatalarının yalnızca kalite maliyeti değil, aynı zamanda ek çevresel yük oluşturduğunu göstermektedir. Bu çalışma, kumaş hata adedi ile yeniden işlem kaynaklı elektrik tüketimini ilişkilendirerek söz konusu dördüncü yaklaşımı denim konfeksiyon üretimi bağlamında ele almaktadır. Böylece kalite kontrol verileri, sürdürülebilir üretim kararlarını destekleyebilecek çevresel bir göstergeye dönüştürülmektedir.

3. GEREÇ ve YÖNTEM

3.1. İşletmenin Tanıtımı

Bu çalışma, tekstil sektöründe faaliyet gösteren ve ağırlıklı olarak denim konfeksiyon üretimi gerçekleştiren büyük ölçekli bir işletmeden elde edilen üretim ve kalite kontrol kayıtları kullanılarak yürütülmüştür. İşletmenin açık adı, kurumsal gizlilik ve veri mahremiyeti nedeniyle belirtilmemiştir. Çalışmada kullanılan veriler, işletmenin gerçek üretim süreçlerinde oluşan kumaş kompozisyonu ve kumaş hata adedi kayıtlarından alınmıştır. Bu yönüyle çalışma, deneysel olarak oluşturulmuş yapay bir veri setine değil, doğrudan saha koşullarında oluşmuş işletme kayıtlarına dayanmaktadır.

İşletmenin idari ve operasyonel yapısı farklı lokasyonlarda örgütlenmiştir. Yönetimsel faaliyetlerin yürütüldüğü merkez ofiste satış ve pazarlama, üretim planlama, tedarik zinciri yönetimi, tasarım, yıkama süreçlerine yönelik Ar-Ge çalışmaları, numune geliştirme, insan kaynakları, kurumsal sosyal sorumluluk, finans ve diğer yönetsel birimler yer almaktadır. Doğrudan üretim faaliyetleri ise Organize Sanayi Bölgesi içerisinde konumlanan üretim tesisinde gerçekleştirilmektedir.

Üretim tesisinde kumaş depolama alanları, kalite ve test laboratuvarları, üretim alanları ve son işlem birimleri bulunmaktadır. Tesis birden fazla üretim yapısından oluşmakta, bu yapılar arasında hammadde, yarı mamul ve mamul akışını sağlayan iç lojistik sistemleri kullanılmaktadır. Bu yapı, üretim sürecinde malzeme akışının kontrollü biçimde sürdürülmesine ve farklı üretim aşamalarının birbirine bağlı olarak ilerlemesine olanak sağlamaktadır.

İşletmede temel üretim grubu denim pantolondur. Üretim sistemi hücresel üretim yaklaşımına göre düzenlenmiş olup üretim hattı “U” biçiminde yapılandırılmış hücrelerden oluşmaktadır. Her hücrede ilgili operasyonları gerçekleştiren birden fazla operatör görev almakta, üretim akışı hammadde girişinden nihai ürün aşamasına kadar kademeli biçimde ilerlemektedir. Bu yapı, yüksek üretim hacimlerinde operasyonların daha düzenli yürütülmesini ve iş akışının hücre bazında izlenebilmesini sağlamaktadır.

Üretim süreci müşteri talepleri doğrultusunda hazırlanan numunelerin değerlendiril-

mesiyle başlamaktadır. Numune geliştirme sürecinde ürün modeli, kumaş yapısı, yıkama görünümü, ölçü özellikleri ve müşteri beklentileri dikkate alınmaktadır. Gerekli numune onaylarının alınmasının ardından seri üretim sürecine geçilmektedir. Seri üretim; kesim, dikim, yıkama, son işlem, kalite kontrol ve paketlenme aşamalarını içeren çok aşamalı bir üretim akışı içinde yürütülmektedir. İşletmede denim üretim sürecinin temel aşamaları Çizelge 3.1’de verilmiştir.

Çizelge 3.1. İşletmede denim üretim sürecinin temel aşamaları

Ana süreç	Alt süreç / operasyon
Kesim	Serim
Kesim	Kesim
Kesim	Meto
Dikim	Ön hazırlık
Dikim	Ön grup
Dikim	Arka hazırlık
Dikim	Arka grup
Dikim	Montaj giriş
Dikim	Montaj çıkış
Yıkama	Yaş işlemler
Yıkama	Kuru işlemler
Yıkama	Finish yıkama
Yıkama	Finish kalite
Son işlemler	Hazırlık
Son işlemler	Temizleme
Son işlemler	Ütü
Son işlemler	Kalite kontrol
Son işlemler	Paketleme

Çizelge 3.1’de görüldüğü gibi denim konfeksiyon üretimi yalnızca dikim aşamasından oluşmamakta; kumaş hazırlığından başlayarak yıkama, son işlem, kalite kontrol ve paketlenme aşamalarına kadar uzanan bütünleşik bir üretim yapısı göstermektedir. Bu çok aşamalı

yapı içinde kumaş kalitesi, üretim sürecinin farklı basamaklarını etkileyebilen temel unsurlardan biridir. Kumaşta bulunan hatalar kesim, dikim, yıkama ve son işlem süreçlerine taşınabildiğinden, kumaş hata adedinin üretim kayıtları üzerinden incelenmesi işletme açısından önem taşımaktadır.

Bu çalışmada kullanılan veri seti işletmenin tüm üretim değişkenlerini kapsamaktadır. Makine ayarları, üretim hızı, operatör bilgisi, vardiya, ortam sıcaklığı ve nem, bakım kayıtları, iplik kalite parametreleri, tedarikçi bilgisi ve hata türleri veri setine dâhil değildir. Çalışmanın veri kapsamı, kumaş kompozisyonundan türetilen lif oranları ile kumaş hata adedi kayıtlarıyla sınırlıdır. Bu nedenle elde edilen sonuçlar işletmenin tüm üretim performansını açıklayan kapsamlı bir proses modeli olarak değil, mevcut üretim ve kalite kontrol kayıtları üzerinden kumaş hata düzeyinin sınıflandırılmasına yönelik uygulamalı bir kalite analizi olarak değerlendirilmelidir.

Gerçek işletme verilerinin kullanılması, çalışmanın saha koşullarıyla ilişkisini güçlendirmektedir. Bu tür veriler laboratuvar verilerine göre daha fazla değişkenlik ve kayıt farklılığı içerebilir. Buna karşılık gerçek üretim kayıtları, işletme ortamında karşılaşılan kalite problemlerinin daha doğrudan değerlendirilmesine olanak sağlar. Bu nedenle işletmenin üretim yapısı, veri kayıt sistemi ve kalite kontrol akışı çalışmanın yönetsel çerçevesinin önemli bir parçasını oluşturmaktadır.

3.2. Veri Setinin Tanıtılması

Bu çalışmada kullanılan veri seti, büyük ölçekli bir denim konfeksiyon işletmesinden elde edilen üretim ve kalite kontrol kayıtlarından oluşmaktadır. Veri seti, işletmede kullanılan kumaş kompozisyon bilgileri ile bu kumaşlara ilişkin hata adetlerini bir araya getirmektedir. Analizlerde temel amaç, kumaş lif kompozisyonu değişkenlerinden hareketle kumaş hata adedinin sınıflandırılmasına yönelik makine öğrenmesi modelleri geliştirmek ve hata düzeyinin üretim verileri üzerinden tahmin edilebilirliğini değerlendirmektir. Bu nedenle veri seti, görüntü tabanlı kumaş hata tespiti için değil, tabular üretim verisiyle hata düzeyi sınıflandırması yapmak için kullanılmıştır.

Veri seti toplam 652 gözlemden, 8 ana değişkenden ve lif oranlarının toplamını kontrol etmek amacıyla kullanılan bir yardımcı sütundan oluşmaktadır. Her gözlem, işletme

kayıtlarında yer alan bir sipariş/kumaş kompozisyon kaydını temsil etmektedir. Veri setinde “Order No” sipariş numarasını, “Kumaş Kompozisyon” ise işletme kayıtlarında yer alan orijinal kumaş bileşimi bilgisini göstermektedir. Bu iki değişken veri setinin tanımlanması ve kayıtların izlenmesi açısından kullanılmış; sınıflandırma modellerinde doğrudan tahmin değişkeni olarak değerlendirilmemiştir. Modelleme sürecinde bağımsız değişkenler olarak pamuk oranı, geri dönüştürülmüş lif oranı, polyester oranı, elastan oranı ve diğer lif oranları kullanılmıştır. Bağımlı değişken ise “Kumaş Hata Adedi” olarak tanımlanmıştır.

Üretim verilerine dayalı makine öğrenmesi çalışmalarında veri setinin açık biçimde tanımlanması, model sonuçlarının güvenilir biçimde yorumlanması açısından zorunludur. Kestirimci kalite çalışmalarında kullanılan verilerin ürün, proses veya kalite çıktılarıyla ilişkili olması ve tahmin edilecek kalite değişkeninin net biçimde belirlenmesi gerekmektedir (Tercan ve Meisen, 2022). Benzer şekilde veri madenciliği ve tahmine dayalı modelleme süreçlerinde gözlem birimi, hedef değişken, açıklayıcı değişkenler ve veri hazırlama adımları açıkça ortaya konulmadan model çıktılarının anlamlı biçimde değerlendirilmesi güçleşmektedir (Han vd., 2012; Kuhn ve Johnson, 2013). Bu nedenle çalışmada kullanılan değişkenler, analizdeki rolleri ve veri türleri Çizelge 3.2’ de ayrıntılı olarak verilmiştir.

Çizelge 3.2. Veri seti değişkenleri ve analizdeki rolleri

Değişken adı	Açıklama	Veri türü	Analizdeki rolü
Order No	Sipariş veya kayıt numarası	Tanımlayıcı	Modelleme dışında bırakılmıştır
Kumas Kompozisyon	İşletme kayıtlarında yer alan orijinal kumaş kompozisyon bilgisi	Metinsel / kategorik tanımlayıcı	Lif oranlarının türetilmesinde ve veri kontrolünde kullanılmıştır
Kumas Hata Adedi	İlgili kumaş kaydına ait hata adedi	Sayısal	Bağımlı değişken / sınıflandırma hedefi
COT oran	Kumaş kompozisyonundaki pamuk oranı	Sayısal	Bağımsız değişken
Recycle oran	Kumaş kompozisyonundaki geri dönüştürülmüş lif oranı	Sayısal	Bağımsız değişken
PES oranı	Kumaş kompozisyonundaki polyester oranı	Sayısal	Bağımsız değişken
EA oranı	Kumaş kompozisyonundaki elastan oranı	Sayısal	Bağımsız değişken
Diğer lifler oranı	Pamuk, geri dönüştürülmüş lif, polyester ve elastan dışında kalan lif oranları	Sayısal	Bağımsız değişken

Veri setinde yer alan lif oranları, işletme kayıtlarında bulunan kumaş kompozisyon bilgilerinden analiz edilebilir sayısal değişkenlere dönüştürülmüştür. Pamuk, geri dönüştürülmüş lif, polyester, elastan ve diğer lif oranları çalışmanın temel açıklayıcı değişkenlerini oluşturmaktadır. Bu değişkenlerin seçilmesinin nedeni, kumaş yapısının ve lif bileşiminin kumaşın üretim davranışı, mekanik özellikleri ve hata oluşum eğilimi üzerinde etkili olabilecek temel faktörler arasında yer almasıdır. Bununla birlikte çalışma, kumaş hata oluşumunu yalnızca lif kompozisyonu ile açıklama iddiası taşımamaktadır. Makine ayarları, üretim hızı, operatör deneyimi, ortam sıcaklığı ve nem, iplik kalite parametreleri, tedarikçi bilgisi, bakım kayıtları ve hata türleri gibi değişkenler veri setinde bulunmadığından analiz kapsamına dahil edilmemiştir. Bu sınırlama, bulguların yorumlanmasında dikkate alınması gereken temel noktalardan biridir.

Veri setinde “Kumas Hata Adedi” değişkeni, modelleme sürecinin temel çıktı

değişkeni olarak kullanılmıştır. Bu değişken ham haliyle sayısal niteliktedir; ancak makine öğrenmesi modellerinde sınıflandırma problemi kurulabilmesi için daha sonraki aşamada hata düzeylerine dönüştürülmüştür. İki sınıflı yaklaşımda kumaş hata adedi normal ve yüksek hata düzeyi olarak ayrılmış, üç sınıflı yaklaşımda ise düşük, orta ve yüksek hata düzeyleri oluşturulmuştur. Böylece hata adedi doğrudan regresyon problemi olarak değil, kalite riskinin sınıflandırılması problemi olarak ele alınmıştır. Bu tercih, üretim ortamında kullanılabilir karar kategorileri oluşturmak bakımından daha uygun görülmüştür.

Veri seti kontrol edildiğinde eksik gözlem bulunmadığı belirlenmiştir. Bu durum, modelleme öncesinde eksik veri tamamlama veya eksik gözlem çıkarma işlemi yapılmasını gerektirmemiştir. Veri setinde lif oranlarının toplamını kontrol etmek amacıyla kullanılan yardımcı sütun tüm gözlemler için %100 değerini vermektedir. Bu nedenle pamuk, geri dönüştürülmüş lif, polyester, elastan ve diğer lif oranlarından oluşan kompozisyon değişkenlerinin veri setinde tutarlı bir yapı gösterdiği kabul edilmiştir. Veri kalitesi açısından bu tür kontroller, üretim verilerine dayalı analizlerde önem taşımaktadır, çünkü verinin eksiksizliği, tutarlılığı ve analiz amacına uygunluğu model sonuçlarının güvenilirliğini doğrudan etkilemektedir (Han vd., 2012; ISO, 2022).

Veri setinde yer alan “Order No” değişkeni her gözlem için benzersiz kayıt numarası niteliğindedir. Bu değişken, gözlemlerin izlenebilirliği açısından korunmuş; ancak modelleme aşamasında bağımsız değişken olarak kullanılmamıştır. Sipariş numarasının modele dahil edilmesi, üretim özelliği taşımayan tanımlayıcı bir bilginin model tarafından öğrenilmesine ve yanıltıcı sonuçlara yol açabileceğinden uygun görülmemiştir. Benzer biçimde Kumaş Kompozisyon değişkeni de orijinal metinsel kompozisyon bilgisini temsil etmekte olup, modelleme sürecinde doğrudan kullanılmamıştır. Bunun yerine bu değişkenden türetilmiş olan sayısal lif oranları analizlerde kullanılmıştır. Tahmine dayalı modelleme çalışmalarında hedef değişkenle doğrudan ilişkili olmayan tanımlayıcı alanların modele alınmaması, modelin genellenebilirliği açısından daha sağlıklı bir yaklaşımdır (Kuhn ve Johnson, 2013).

Bu çalışmada kullanılan veri setinin temel özelliği, gerçek üretim ortamından elde edilmiş olmasıdır. Gerçek işletme verileri, laboratuvar veya simülasyon verilerine göre daha

fazla değişkenlik ve düzensizlik içerebilir; ancak üretim kararlarına aktarılabilirlik açısından daha yüksek uygulama değeri taşımaktadır. Kestirimci kalite literatüründe de gerçek üretim verileriyle yürütülen çalışmaların, kalite tahmin modellerinin saha koşullarındaki kullanılabilirliğini değerlendirmek açısından önemli olduğu belirtilmektedir (Tercan ve Meisen, 2022). Bu kapsamda veri seti, denim konfeksiyon üretiminde lif kompozisyonu ile kumaş hata düzeyi arasındaki ilişkinin incelenmesi ve makine öğrenmesi tabanlı sınıflandırma modellerinin karşılaştırılması için uygun bir temel sağlamaktadır.

Kullanılan veri seti, sipariş numarası, orijinal kumaş kompozisyon bilgisi, kumaş hata adedi ve beş temel lif oranı değişkeninden oluşan tabular bir üretim veri setidir. Analizlerde model girdisi olarak pamuk, geri dönüştürülmüş lif, polyester, elastan ve diğer lif oranları kullanılmış; model çıktısı ise kumaş hata adedinden türetilen hata düzeyi sınıfları olarak belirlenmiştir. Veri setinin kapsamı lif kompozisyonu ve hata adediyle sınırlı olduğundan, elde edilen sonuçlar kumaş hatalarının tüm nedenlerini açıklayan kapsamlı bir proses modeli olarak değil, mevcut üretim kayıtları üzerinden hata düzeyinin sınıflandırılmasına yönelik uygulamalı bir kalite analizi olarak değerlendirilmelidir.

3.3. Veri Hazırlama Süreci

Veri hazırlama süreci, üretim verilerine dayalı makine öğrenmesi uygulamalarında model başarısını doğrudan etkileyen temel aşamalardan biridir. Ham veri setinin modele doğrudan aktarılması çoğu zaman uygun değildir; çünkü veri yapısında eksik kayıtlar, tanımlayıcı fakat modelleme açısından anlamsız değişkenler, ölçek farklılıkları, sınıf oluşturma gereksinimleri veya tutarlılık sorunları bulunabilir. Bu nedenle çalışmada modelleme aşamasına geçilmeden önce veri seti gözlem sayısı, değişken yapısı, eksik veri durumu, değişken rolleri ve sınıflandırma hedeflerinin oluşturulması bakımından kontrol edilmiştir. Tahmine dayalı modelleme çalışmalarında veri ön işleme, veri bölme, model kurma ve model performansını değerlendirme adımlarının bir bütün olarak ele alınması gerektiği belirtilmektedir (Kuhn ve Johnson, 2013). Benzer biçimde veri madenciliği sürecinde veri temizleme, veri dönüşümü ve uygun değişken seçimi, modelleme çıktılarının güvenilirliği açısından temel hazırlık adımları arasında görülmektedir (Han vd., 2012).

Çalışmada kullanılan veri seti Excel formatında düzenlenmiş ve MATLAB ortamına

aktarılmıştır. Veri aktarımı sonrasında ilk olarak satır ve sütun sayıları kontrol edilmiş, veri setinin 652 gözleminden, 8 ana değişkenden ve lif oranlarının toplamını gösteren bir yardımcı kontrol sütunundan oluştuğu doğrulanmıştır. Ardından her değişkenin veri türü ve analizdeki rolü değerlendirilmiştir. “Order No” değişkeni sipariş/kayıt numarasını temsil ettiği için yalnızca izlenebilirlik amacıyla korunmuş; makine öğrenmesi modellerine bağımsız değişken olarak dahil edilmemiştir. Benzer şekilde “Kumaş Kompozisyon” değişkeni orijinal kumaş kompozisyon bilgisini metinsel biçimde içerdiğinden doğrudan model girdisi olarak kullanılmamıştır. Bu değişken yerine, kompozisyon bilgisinden türetilmiş sayısal lif oranları modelleme sürecine alınmıştır. Böylece modellerin üretim açısından açıklayıcı nitelik taşıyan sayısal değişkenler üzerinden kurulması sağlanmıştır.

Veri setinde eksik gözlem bulunup bulunmadığı kontrol edilmiş ve eksik veri olmadığı belirlenmiştir. Bu nedenle eksik değer tamamlama, ortalama ile doldurma, kayıt silme veya benzeri bir işlem uygulanmamıştır. Lif oranlarının tutarlılığı ayrıca değerlendirilmiş ve veri setinde toplam kontrol sütununun tüm kayıtlarda %100 değerini aldığı belirlenmiştir. Bu bulgu, pamuk, geri dönüştürülmüş lif, polyester, elastan ve diğer lif oranlarının veri setinde toplam kompozisyon yapısı bakımından tutarlı olduğunu göstermektedir. Veri kalitesine ilişkin bu tür kontroller, üretim verilerinde modelleme öncesi kritik görülmektedir. ISO 8000-1 standardı, veri kalitesinin iyileştirilmesi ve verinin belirli amaçlara uygunluğunun değerlendirilmesi için genel bir çerçeve sunmaktadır (ISO, 2022).

Modelleme sürecinde bağımsız değişkenler olarak COT oranı, Recycle oran, PES oranı, EA oranı ve Diğer lifler oranı kullanılmıştır. Bağımlı değişken ise Kuma Hata Adedi olarak belirlenmiştir. Ancak Kumaş Hata Adedi değişkeni ham haliyle sürekli/sayısal bir değişken olduğundan, sınıflandırma modellerinde kullanılmak üzere kategorik hata düzeylerine dönüştürülmüştür. Bu dönüşüm, üretim ortamında doğrudan yorumlanabilir kalite risk grupları oluşturmak amacıyla yapılmıştır. Kestirimci kalite çalışmalarında kalite çıktısının açıkça tanımlanması ve modelin tahmin edeceği hedef değişkenin üretim amacıyla uyumlu olması gerekmektedir (Tercan ve Meisen, 2022). Bu doğrultuda hata adedi hem iki sınıflı hem de üç sınıflı sınıflandırma yapısına dönüştürülmüştür.

İki sınıflı sınıflandırma yaklaşımında kumaş hata adedi için 75. yüzdelerik değer eşik

olarak alınmıştır. Veri setinde bu eşik değeri 63 olarak hesaplanmıştır. Buna göre hata adedi 63 ve altında olan gözlemler “Normal”, 63’ün üzerinde olan gözlemler ise “Yüksek” hata sınıfı olarak etiketlenmiştir. Bu yaklaşımın amacı, üretim açısından daha riskli kabul edilebilecek yüksek hata grubunu diğer gözlemlerden ayırmaktır. Sınıf etiketlerinin yüzdelik değerler üzerinden oluşturulması, veri setinin kendi dağılımına dayalı bir sınıflandırma yapısı kurulmasına olanak sağlamaktadır. Özellikle kalite verilerinde mutlak eşik değerlerin işletmeden işletmeye değişebilmesi nedeniyle, dağılım temelli eşikler pratik bir başlangıç noktası sunmaktadır.

Üç sınıflı sınıflandırma yaklaşımında ise kumaş hata adedi 33. ve 66. yüzdelik değerlere göre üç gruba ayrılmıştır. Veri setinde 33. yüzdelik değeri 18, 66. yüzdelik değeri ise 48 olarak hesaplanmıştır. Buna göre hata adedi 18 ve altında olan gözlemler “Düşük”, 19 ile 48 arasında olan gözlemler “Orta”, 48’in üzerinde olan gözlemler ise “Yüksek” hata sınıfı olarak tanımlanmıştır. Üç sınıflı yapı, üretim kalite düzeyini iki sınıflı yaklaşıma göre daha ayrıntılı biçimde temsil etmektedir. Ancak sınıf sayısının artması, modellerin ayırt etmesi gereken karar sınırlarını da karmaşık hale getirdiğinden performans değerlendirmesinde yalnızca genel doğruluk oranına odaklanılmamıştır. Çalışmada kullanılan veri hazırlama adımları Çizelge 3.3’ de özetlenmiştir.

Çizelge 3.3. Veri hazırlama sürecinde uygulanan işlemler

Aşama	Uygulanan işlem	Açıklama
Veri aktarımı	Excel veri setinin MATLAB ortamına aktarılması	Ana veri dosyası analiz ortamına alınmıştır
Veri yapısı kontrolü	Gözlem ve değişken sayısının kontrol edilmesi	Veri setinin 652 gözlem, 8 ana değişken ve bir yardımcı kontrol sütunundan oluştuğu doğrulanmıştır
Eksik veri kontrolü	Eksik değer taraması	Eksik veri bulunmadığından tamamlama işlemi yapılmamıştır
Değişken rolü belirleme	Tanımlayıcı ve model değişkenlerinin ayrılması	OrderNo ve KumasKompozisyon modelleme dışında bırakılmıştır
Bağımsız değişken seçimi	Lif oranlarının model girdisi olarak seçilmesi	Beş sayısal lif oranı bağımsız değişken olarak kullanılmıştır
Bağımlı değişken belirleme	Kumaş Hata Adedi değişkeninin hedef değişken olarak tanımlanması	Hata adedi sınıflandırma hedefi olarak kullanılmıştır
İki sınıflı etiketleme	p75 eşik değerine göre sınıf oluşturma	p75 = 63; Normal ve Yüksek sınıfları oluşturulmuştur
Üç sınıflı etiketleme	p33 ve p66 eşiklerine göre sınıf oluşturma	p33 = 18, p66 = 48; Düşük, Orta ve Yüksek sınıfları oluşturulmuştur
Normallik testi	Shapiro-Wilk testi	Korelasyon analizinde kullanılacak sayısal değişkenlerin normal dağılım varsayımı değerlendirilmiştir
Korelasyon yöntemi seçimi	Spearman sıra korelasyonu	Shapiro-Wilk testi sonucunda normal dağılım varsayımı sağlanmadığı için Pearson yerine Spearman korelasyonu kullanılmıştır
Model doğrulama yaklaşımı	Leave-one-out çapraz doğrulama	Her yinelemede bir gözlem test verisi, kalan gözlemler eğitim verisi olarak kullanılmıştır
Rastgelelik kontrolü	Sabit rastgelelik değeri kullanımı	MATLAB ortamında rng(42) kullanılmıştır

Makine öğrenmesi modellerinin eğitimi ve değerlendirilmesinde veri setinin görece sınırlı büyüklüğü dikkate alınarak leave-one-out çapraz doğrulama yaklaşımı kullanılmıştır. Bu yöntemde her yinelemede veri setindeki bir gözlem test verisi olarak ayrılmış, kalan

gözlemler modelin eğitimi için kullanılmıştır. İşlem tüm gözlemler bir kez test gözlemi olacak şekilde tekrarlanmış ve model performansı tüm yinelemelerden elde edilen tahminler birleştirilerek hesaplanmıştır. Böylece tek bir rastgele eğitim-test ayırımına bağlı kalmadan, her gözlemin model değerlendirme sürecine test gözlemi olarak katkı vermesi sağlanmıştır.

Korelasyon analizine geçmeden önce, analizde kullanılan sayısal değişkenlerin normal dağılım varsayımını sağlayıp sağlamadığı Shapiro-Wilk testi ile değerlendirilmiştir. Shapiro-Wilk testinde sıfır hipotezi, ilgili değişkenin normal dağılım gösterdiği yönündedir. Test sonucunda p değerinin 0,05'in altında olması durumunda normal dağılım varsayımı reddedilmiştir. Bu çalışmada Kumaş Hata Adedi, COT oranı, Recycle oran, PES oranı, EA oranı ve Diğer lifler oranı değişkenleri için normallik testi uygulanmıştır. Normallik testi sonucunda değişkenlerin normal dağılım varsayımını sağlamaması nedeniyle, lif oranları ile kumaş hata adedi arasındaki ilişki Pearson korelasyonu yerine Spearman sıra korelasyonu ile değerlendirilmiştir. Spearman korelasyonu, değişkenler arasındaki doğrusal ilişkiyi değil, sıralı/monoton ilişkiyi değerlendirdiğinden normal dağılım varsayımının sağlanmadığı veri yapılarında daha uygun bir korelasyon yaklaşımı sunmaktadır.

Uzaklık temelli sınıflandırma yöntemleri için değişken ölçeklerinin dikkate alınması gerekmektedir. KNN gibi yöntemlerde gözlemler arasındaki uzaklıklar model sonucunu doğrudan etkileyebilmektedir. Bu nedenle değişkenlerin farklı dağılım özellikleri göstermesi durumunda normalizasyon veya standardizasyon işlemleri model sonuçlarının daha dengeli oluşmasına katkı sağlar. Veri setindeki lif oranları yüzde değerleri olarak benzer ölçekte bulunsa da dağılım genişlikleri ve değişkenlik düzeyleri farklılık gösterebilmektedir. Bu nedenle özellikle uzaklık temelli analizlerde değişkenlerin ölçek etkisi dikkate alınmıştır.

Veri hazırlama sürecinde herhangi bir gözlem veri setinden çıkarılmamıştır. Bunun nedeni, üretim ortamından elde edilen gerçek kayıtların mevcut değişken yapısı içinde korunmak istenmesidir. Aykırı görülebilecek hata adetleri, üretim sürecindeki yüksek hata yükünü temsil edebileceğinden doğrudan silinmemiştir. Kalite problemlerinde yüksek hata değerleri çoğu zaman modelin öğrenmesi gereken kritik üretim durumlarını göstermektedir. Bu tür gözlemlerin gerekçesiz biçimde çıkarılması, yüksek hata sınıfının daha zayıf temsil edilmesine neden olabilir. Bu nedenle hata adedi dağılımındaki yüksek değerler, veri

temizleme aşamasında dışlanmamış; sınıflandırma aşamasında hata düzeyi etiketlerinin oluşturulmasında dikkate alınmıştır. Veri hazırlama süreci; veri aktarımı, eksik veri kontrolü, tanımlayıcı değişkenlerin ayrılması, lif oranlarının model girdisi olarak seçilmesi, kumaş hata adedinin hedef değişken olarak belirlenmesi, iki ve üç sınıflı hata düzeyi etiketlerinin oluşturulması, Shapiro-Wilk normallik testinin uygulanması, korelasyon yönteminin belirlenmesi ve leave-one-out çapraz doğrulama yapısının tanımlanması aşamalarından oluşmaktadır. Bu hazırlık süreci, sonraki aşamalarda uygulanacak makine öğrenmesi modelleri, Spearman korelasyon analizi, değişken önem analizi ve performans değerlendirmeleri için ortak ve tutarlı bir veri yapısı sağlamaktadır.

3.4. Makine Öğrenmesi Modelleri

Bu çalışmada kumaş hata adedinden türetilen hata düzeylerinin sınıflandırılması amacıyla farklı öğrenme yaklaşımlarını temsil eden 15 makine öğrenmesi modeli kullanılmıştır. Model seçiminde tek bir algoritmanın başarısını göstermekten çok, farklı karar yapılarının aynı veri seti üzerindeki davranışını karşılaştırmak amaçlanmıştır. Bu nedenle benzerlik temelli, ağaç tabanlı, olasılıksal, doğrusal, doğrusal olmayan, ensemble ve sınıf dengesizliğine duyarlı modeller birlikte değerlendirilmiştir. Üretim verilerine dayalı kestirimci kalite çalışmalarında model seçiminin veri yapısına, hedef değişkenin niteliğine ve kalite probleminin karar bağlamına göre yapılması gerektiği belirtilmektedir (Tercan ve Meisen, 2022). Bu çalışmada da model çeşitliliği, kumaş lif kompozisyonu ile hata düzeyi arasındaki ilişkinin tek bir öğrenme yaklaşımıyla sınırlı kalmadan incelenmesine olanak sağlamaktadır.

Makine öğrenmesi modellerinin kurulmasında bağımsız değişkenler olarak COT oranı, Recycle oran, PES oranı, EA oranı ve Diğer lifler oranı kullanılmıştır. Hedef değişken ise Kumaş Hata Adedi değişkeninden türetilen hata sınıflarıdır. Çalışmada iki ayrı sınıflandırma yapısı oluşturulmuştur. İlk yapıda hata adedi iki sınıfa ayrılarak “Normal” ve “Yüksek” hata sınıfları belirlenmiştir. İkinci yapıda ise hata adedi “Düşük”, “Orta” ve “Yüksek” olmak üzere üç sınıfta değerlendirilmiştir. Böylece hem üretim açısından pratik bir risk ayrımı sunan iki sınıflı yapı hem de hata düzeylerini daha ayrıntılı temsil eden üç sınıflı yapı karşılaştırmalı olarak incelenmiştir.

Kullanılan modellerin seçilmesinde dört temel gerekçe bulunmaktadır. İlk olarak KNN, Decision Tree, Naive Bayes, Logistic Regression, LDA ve SVM gibi temel sınıflandırıcılar, farklı karar mantıklarını temsil ettikleri için modele alınmıştır. KNN benzerlik temelli bir sınıflandırma yaklaşımı sunarken, Decision Tree değişken eşiklerine dayalı kural benzeri bir yapı kurmaktadır. Naive Bayes olasılıksal bir sınıflandırıcıdır; Logistic Regression ve LDA daha yorumlanabilir doğrusal sınıflandırma yaklaşımlarını temsil etmektedir. SVM modelleri ise sınıflar arasında ayırıcı karar sınırı oluşturma kapasitesi nedeniyle özellikle doğrusal ve doğrusal olmayan sınıflandırma problemlerinde kullanılmaktadır (Cortes vd., 1995; Domingos ve Pazzani, 1997; Kuhn ve Johnson, 2013).

İkinci olarak Random Forest, Bagged Trees, Boosting, Subspace KNN ve Subspace Discriminant Analysis gibi ensemble tabanlı veya çoklu öğrenici kullanan yöntemler modele dahil edilmiştir. Ensemble yöntemleri, tek bir sınıflandırıcının kararına bağlı kalmak yerine birden fazla modelin çıktısını birleştirerek daha kararlı sonuçlar üretmeyi amaçlamaktadır. Dietterich (2000), ensemble yöntemlerinin birden fazla sınıflandırıcının tahminlerini birleştirerek yeni gözlemleri sınıflandırdığını ve bu yaklaşımın çoğu durumda tekil modellere göre daha güçlü performans sağlayabildiğini belirtmektedir. Random Forest yöntemi de çok sayıda karar ağacının birlikte kullanıldığı bir ensemble yaklaşımıdır ve doğrusal olmayan ilişkilerin bulunduğu veri setlerinde yaygın biçimde tercih edilmektedir (Breiman, 2001).

Üçüncü olarak RUSBoost ve Cost-Sensitive Random Forest modelleri, sınıf dengesizliği sorununu dikkate alabilecek yaklaşımlar olarak kullanılmıştır. Üretim kalite verilerinde yüksek hata sınıfı çoğu zaman normal üretim sınıfına göre daha az sayıda gözlem içermektedir. Bu durumda standart sınıflandırıcılar çoğunluk sınıfını daha iyi öğrenme eğilimi gösterebilir ve üretim açısından kritik olan yüksek hata sınıfını yeterli düzeyde yakalayamayabilir. RUSBoost, rastgele alt örnekleme ile boosting yaklaşımını birleştirerek dengesiz veri yapılarında azınlık sınıfının daha görünür hale gelmesini amaçlayan bir yöntemdir (Seiffert vd., 2010). Maliyet duyarlı sınıflandırma yaklaşımlarında ise hatalı sınıflandırma maliyetleri sınıflar arasında farklılaştırılarak kritik sınıfın model tarafından daha fazla dikkate alınması sağlanmaktadır. Bu yaklaşım, yüksek hata sınıfının

kaçırılmasının üretim açısından daha ciddi sonuçlar doğurabileceği kalite kontrol problemleri için önemlidir.

Dördüncü olarak ANN modeli, lif oranları ile hata düzeyi arasındaki olası doğrusal olmayan ilişkileri öğrenebilme kapasitesi nedeniyle modele eklenmiştir. Yapay sinir ağları, girdi değişkenleri ile çıktı sınıfları arasındaki karmaşık örüntüleri katmanlı bir yapı içinde öğrenmeye çalışmaktadır. Bununla birlikte yapay sinir ağlarının yorumlanabilirliği karar ağaçları veya lojistik regresyon gibi modellere göre daha sınırlıdır. Bu nedenle ANN modeli çalışmada tek başına nihai karar modeli olarak değil, diğer sınıflandırıcılarla karşılaştırmalı biçimde değerlendirilmiştir. Çalışmada kullanılan makine öğrenmesi modelleri ve temsil ettikleri öğrenme yaklaşımları Çizelge 3.4’de verilmiştir.

Çizelge 3.4. Çalışmada kullanılan makine öğrenmesi modelleri ve seçilme gerekçeleri

Model	Temel öğrenme yaklaşımı	Çalışmada kullanılma gerekçesi
KNN	Benzerlik / uzaklık temelli sınıflandırma	Benzer lif kompozisyonuna sahip gözlemlerin benzer hata sınıflarında toplanıp toplanmadığını değerlendirmek
RUSBoost	Alt örnekleme destekli boosting	Sınıf dengesizliği durumunda yüksek hata sınıfının yakalanabilirliğini artırmak
Cost-Sensitive Random Forest	Maliyet duyarlı ensemble öğrenme	Hatalı sınıflandırma maliyetini dikkate alarak kritik sınıfın öğrenilmesini desteklemek
Bagged Trees	Bootstrap tabanlı ağaç topluluğu	Tekil karar ağaçlarının kararsızlığını azaltmak ve daha dengeli tahmin üretmek
Random Forest	Rastgele ağaç topluluğu	Doğrusal olmayan ilişkileri yakalamak ve değişken katkısını değerlendirmek
Decision Tree	Kural / eşik tabanlı sınıflandırma	Yorumlanabilir karar yapısı elde etmek
Naive Bayes	Olasılıksal sınıflandırma	Basit ve hızlı çalışan olasılıksal bir referans model oluşturmak
Boosting	Ardışık ensemble öğrenme	Önceki hatalardan öğrenen daha güçlü bir sınıflandırıcı yapısı kurmak
Subspace KNN	Alt uzay tabanlı benzerlik öğrenmesi	Değişken alt kümeleri üzerinden KNN performansını değerlendirmek
SVM-RBF	Doğrusal olmayan destek vektör sınıflandırması	Karmaşık karar sınırlarını yakalayabilmek
LDA	Doğrusal ayırma analizi	Doğrusal sınıf ayrımının yeterliliğini test etmek
Logistic Regression	Doğrusal olasılık temelli sınıflandırma	Yorumlanabilir temel sınıflandırma modeli sağlamak
Linear SVM	Doğrusal destek vektör sınıflandırması	Doğrusal karar sınırıyla sınıflandırma başarısını değerlendirmek
ANN	Yapay sinir ağı	Doğrusal olmayan örüntüleri öğrenme kapasitesini sınamak
Subspace Discriminant Analysis	Alt uzay tabanlı ayırma analizi	Değişken alt kümeleriyle doğrusal ayırma yaklaşımını güçlendirmek

Modellerin tamamı aynı bağımsız değişken seti üzerinde değerlendirilmiştir. Bu tercih, model performanslarındaki farklılıkların değişken setinden değil, algoritmaların

öğrenme yapısından kaynaklanmasına olanak sağlamıştır. Modelleme sürecinde performans değerlendirmesi leave-one-out çapraz doğrulama yaklaşımıyla yapılmıştır. Bu kapsamda her yinelemede bir gözlem test verisi olarak ayrılmış, kalan 651 gözlem modelin eğitimi için kullanılmıştır. Tüm gözlemler bir kez test gözlemi olarak değerlendirildikten sonra elde edilen tahminler birleştirilmiş ve accuracy, precision, recall ve F1-score ölçütleri hesaplanmıştır.

Leave-one-out çapraz doğrulama sürecinde toplam yineleme sayısı gözlem sayısına eşittir. Bu çalışmada $n=652$ olduğundan her model için 652 ayrı eğitim-test döngüsü gerçekleştirilmiştir. Her döngüde 651 gözlem eğitim, 1 gözlem test amacıyla kullanılmıştır. Sonuçta her gözlem için bir tahmin değeri elde edilmiş ve model performans ölçütleri bu 652 tahmin üzerinden hesaplanmıştır.

3.4.1. İki Sınıflı Sınıflandırma Yaklaşımı

İki sınıflı sınıflandırma yaklaşımında kumaş hata adedi $p75$ eşik değeri temel alınarak “Normal” ve “Yüksek” hata sınıflarına ayrılmıştır. Veri setinde 75. yüzdelerik değeri 63 olarak hesaplanmıştır. Buna göre kumaş hata adedi 63 ve altında olan gözlemler “Normal”, 63’ün üzerinde olan gözlemler ise “Yüksek” hata sınıfı olarak etiketlenmiştir. Bu yaklaşımın amacı, üretim sürecinde yüksek hata riski taşıyan kumaş kompozisyonlarının ayırt edilebilirliğini değerlendirmektir.

İki sınıflı yapı, üretim ortamında pratik karar verme açısından kullanışlıdır. Çünkü işletmeler açısından her hata düzeyini ayrıntılı olarak sınıflandırmaktan önce, yüksek hata riski taşıyan grupların erken aşamada belirlenmesi önem taşımaktadır. Bu nedenle $p75$ eşiği, veri setinin üst hata bölgesini diğer gözlemlerden ayıran dağılım temelli bir sınır olarak kullanılmıştır. Böylece mutlak ve işletmeden bağımsız bir hata sınırı belirlemek yerine, veri setinin kendi hata dağılımına dayalı bir sınıflandırma yapısı oluşturulmuştur.

Veri setinde iki sınıflı yapı sonucunda 490 gözlem “Normal”, 162 gözlem ise “Yüksek” sınıfında yer almıştır. Bu dağılım, sınıflar arasında belirli bir dengesizlik bulunduğunu göstermektedir. Normal sınıf veri setinin daha büyük bölümünü oluştururken, yüksek hata sınıfı daha az sayıda gözlemle temsil edilmektedir. Bu durum, kalite kontrol bağlamında dikkatle değerlendirilmelidir. Çünkü standart sınıflandırıcılar çoğunluk sınıfını daha iyi

öğrenerek yüksek genel doğruluk değerleri üretebilir; buna karşılık üretim açısından kritik olan yüksek hata sınıfını yeterli düzeyde yakalayamayabilir. He & Garcia, (2009), dengesiz veri kümelerinde geleneksel sınıflandırıcıların çoğunluk sınıfına yanlı sonuçlar üretebildiğini ve azınlık sınıfının öğrenilmesinde zayıf kalabildiğini belirtmektedir.

Bu nedenle iki sınıflı model sonuçları yorumlanırken yalnızca accuracy değerine odaklanılmamıştır. Yüksek hata sınıfının doğru belirlenmesi, üretim sürecinde yeniden işlem yükünün, kalite maliyetlerinin ve potansiyel kaynak kayıplarının azaltılması açısından daha kritik bir göstergedir. Bu nedenle precision, recall ve F1-score ölçütleri model karşılaştırmasında özellikle dikkate alınmıştır. Recall değeri, gerçekte yüksek hata sınıfında bulunan gözlemlerin ne kadarının model tarafından doğru yakalandığını göstermektedir. Precision, modelin yüksek hata olarak işaretlediği gözlemler içinde gerçekten yüksek hata sınıfına ait olanların oranını ifade etmektedir. F1-score ise precision ve recall değerlerini birlikte değerlendiren daha dengeli bir ölçüt sunmaktadır. Dengesiz veri yapılarında bu ölçütlerin birlikte kullanılması, model performansının daha sağlıklı yorumlanmasına katkı sağlamaktadır (Chawla vd., 2002; He ve Garcia, 2009). İki sınıflı sınıflandırma yapısına ilişkin sınıf dağılımı Çizelge 3.5’ de verilmiştir.

Çizelge 3.5. İki sınıflı sınıflandırma yaklaşımında sınıf dağılımı

Sınıf	Eşik tanımı	Gözlem sayısı	Veri setindeki payı (%)
Normal	Kumaş hata adedi ≤ 63	490	75,15
Yüksek	Kumaş hata adedi > 63	162	24,85
Toplam	—	652	100,00

Bu sınıflandırma yapısı, yüksek hata grubunun üretim açısından ayırt edilebilir olup olmadığını sınamak için kullanılmıştır. Ancak yüksek hata sınıfının daha az sayıda gözlem içermesi nedeniyle, model sonuçlarının değerlendirilmesinde sınıf dengesizliği göz önünde bulundurulmuştur. Bu nedenle RUSBoost ve Cost-Sensitive Random Forest gibi sınıf dengesizliğine duyarlı modellerin sonuçları ayrıca önem taşımaktadır.

3.4.2. Üç Sınıflı Sınıflandırma Yaklaşımı

Üç sınıflı sınıflandırma yaklaşımında kumaş hata adedi p33 ve p66 eşik değerleri kullanılarak “Düşük”, “Orta” ve “Yüksek” hata sınıflarına ayrılmıştır. Veri setinde 33. yüzdelik değer 18, 66. yüzdelik değer ise 48 olarak hesaplanmıştır. Buna göre kumaş hata adedi 18 ve altında olan gözlemler “Düşük”, 19–48 aralığında olan gözlemler “Orta”, 48’in üzerinde olan gözlemler ise “Yüksek” hata sınıfı olarak tanımlanmıştır.

Üç sınıflı yaklaşım, üretim kalite düzeyini iki sınıflı yaklaşıma göre daha ayrıntılı temsil etmektedir. İki sınıflı yapıda veri yalnızca normal ve yüksek hata gruplarına ayrılırken, üç sınıflı yapıda düşük ve orta hata düzeyleri de birbirinden ayrılmaktadır. Bu ayrım, üretim sürecinde kalite davranışının daha kademeli biçimde izlenmesine olanak sağlamaktadır. Özellikle yalnızca yüksek hata riskinin değil, orta düzeyde hata eğilimi gösteren kumaş kompozisyonlarının da belirlenmesi istendiğinde üç sınıflı yapı daha açıklayıcı bir çerçeve sunmaktadır.

Veri setinde üç sınıflı yapı sonucunda 216 gözlem “Düşük”, 215 gözlem “Orta” ve 221 gözlem “Yüksek” sınıfında yer almıştır. Bu dağılım, iki sınıflı yapıya göre daha dengeli bir sınıf yapısı ortaya koymaktadır. Bununla birlikte sınıf sayısının üçe çıkması, modellerin ayırt etmesi gereken karar sınırlarını artırmaktadır. Bu nedenle üç sınıflı yaklaşımda modelleme problemi daha ayrıntılı fakat daha karmaşık bir yapıya dönüşmektedir. Bu noktada model performansının iki sınıflı yapıdaki sonuçlarla doğrudan aynı düzeyde olması beklenmemelidir. Sınıf sayısının artması, özellikle sınıflar arasında geçişkenlik bulunan veri yapılarında modelin karar verme sürecini zorlaştırabilmektedir. Üç sınıflı sınıflandırma yapısına ilişkin sınıf dağılımı Çizelge 3.6’ da verilmiştir.

Çizelge 3.6. Üç sınıflı sınıflandırma yaklaşımında dağılım

Sınıf	Eşik tanımı	Gözlem sayısı	Veri setindeki payı (%)
Düşük	Kumaş hata adedi ≤ 18	216	33,13
Orta	$19 \leq$ Kumaş hata adedi ≤ 48	215	32,98
Yüksek	Kumaş hata adedi > 48	221	33,90
Toplam	—	652	100,00

Üç sınıflı yaklaşımda kullanılan model seti, iki sınıflı yaklaşımla aynı tutulmuştur. Böylece model performanslarındaki farklılıkların sınıf yapısından mı yoksa algoritmaların öğrenme karakterinden mi kaynaklandığı daha açık biçimde değerlendirilebilmiştir. Bu yaklaşım, hata düzeylerinin hem sadeleştirilmiş risk ayrımı hem de daha ayrıntılı kalite sınıflandırması bakımından karşılaştırılmasına olanak sağlamaktadır.

Üç sınıflı model sonuçları yorumlanırken de accuracy tek başına yeterli görülmemiştir. Her ne kadar sınıflar iki sınıflı yapıya göre daha dengeli dağılmış olsa da düşük, orta ve yüksek hata sınıflarının ayrı ayrı yakalanma başarısı önem taşımaktadır. Özellikle yüksek hata sınıfının doğru belirlenmesi, kalite kontrol ve yeniden işlem yükünün azaltılması açısından kritik olmaya devam etmektedir. Bu nedenle üç sınıflı modellerin değerlendirilmesinde de precision, recall ve F1-score ölçütleri dikkate alınmıştır. Bu ölçütler, modelin her sınıfı ne ölçüde ayırt edebildiğini daha ayrıntılı biçimde incelemeye olanak sağlamaktadır.

Özetle, bu çalışmada iki ve üç sınıflı sınıflandırma yaklaşımları birlikte kullanılarak kumaş hata adedinin farklı karar düzeylerinde modellenebilirliği değerlendirilmiştir. İki sınıflı yapı yüksek hata riskinin pratik biçimde ayrıştırılmasına odaklanırken, üç sınıflı yapı kalite düzeylerini daha ayrıntılı biçimde temsil etmektedir. Kullanılan 15 model, farklı öğrenme yaklaşımlarının bu iki sınıflandırma yapısındaki performansını karşılaştırmak ve üretim verisiyle hata düzeyi sınıflandırmasının hangi model aileleriyle daha etkili yapılabileceğini belirlemek amacıyla uygulanmıştır.

3.5. Model Performans Ölçütleri

Makine öğrenmesi modellerinin değerlendirilmesinde kullanılan performans ölçütleri,

model sonuçlarının doğru yorumlanması açısından temel öneme sahiptir. Bir sınıflandırma modelinin başarılı olup olmadığı yalnızca kaç gözlemi doğru sınıflandırdığına göre değerlendirilirse, özellikle dengesiz sınıf dağılımına sahip veri setlerinde yanıltıcı sonuçlara ulaşılabilir. Bu nedenle çalışmada model performansları yalnızca genel doğruluk oranı üzerinden değil, accuracy, precision, recall ve F1-score ölçütleri birlikte dikkate alınarak değerlendirilmiştir.

Sınıflandırma performans ölçütleri genellikle karmaşıklık matrisi üzerinden hesaplanmaktadır. Karmaşıklık matrisi, modelin doğru ve hatalı sınıflandırmalarını sınıf düzeyinde gösteren temel değerlendirme aracıdır. İki sınıflı sınıflandırma problemlerinde bu matris; doğru pozitif, yanlış pozitif, doğru negatif ve yanlış negatif değerlerinden oluşmaktadır. Çok sınıflı problemlerde ise aynı mantık her sınıf için ayrı ayrı değerlendirilebilir. Sokolova ve Lapalme, (2009), sınıflandırma görevlerinde performans ölçütlerinin problem türüne ve sınıf dağılımına göre seçilmesi gerektiğini belirtmektedir. Benzer şekilde Tharwat, (2021), sınıflandırma değerlendirme ölçütlerinin yalnızca matematiksel göstergeler olarak değil, modelin karar verme davranışını açıklayan araçlar olarak yorumlanması gerektiğini vurgulamaktadır.

Bu çalışmada iki sınıflı ve üç sınıflı sınıflandırma yaklaşımları kullanıldığından, performans ölçütlerinin yorumlanmasında sınıf yapısı dikkate alınmıştır. İki sınıflı yaklaşımda veri seti “Normal” ve “Yüksek” hata sınıflarına ayrılmıştır. Veri setinde “Normal” sınıf 490 gözlemden, “Yüksek” sınıf ise 162 gözlemden oluşmaktadır. Bu dağılım, yüksek hata sınıfının normal sınıfa göre daha az temsil edildiğini göstermektedir. Sınıf dengesizliği bulunan veri setlerinde model, çoğunluk sınıfını daha iyi öğrenerek yüksek accuracy değeri üretebilir; ancak üretim açısından kritik olan azınlık sınıfını yeterli düzeyde yakalayamayabilir. Saito ve Rehmsmeier, (2015), dengesiz veri setlerinde model performansının değerlendirilmesinde pozitif sınıfa odaklanan ölçütlerin daha bilgilendirici olabileceğini belirtmiştir. Bu nedenle bu çalışmada özellikle iki sınıflı modellerin değerlendirilmesinde recall ve F1-score ölçütleri daha dikkatli yorumlanmıştır.

Üretim kalite kontrolü açısından yüksek hata sınıfının doğru belirlenmesi, yalnızca istatistiksel bir başarı göstergesi değildir. Yüksek hata sınıfının kaçırılması, hatalı veya riskli

kumaş kompozisyonlarının normal kabul edilmesine neden olabilir. Bu durum yeniden işlem yükünün artması, kalite maliyetlerinin yükselmesi ve ek kaynak tüketimi gibi sonuçlar doğurabilir. Bu nedenle modelin genel doğruluğu kadar, yüksek hata sınıfını ne ölçüde yakalayabildiği de önemlidir. Goutte ve Gaussier, (2005), precision, recall ve F-score ölçütlerinin model karşılaştırmalarında birlikte ele alınmasının değerlendirmeyi daha dengeli hale getirdiğini ifade etmektedir. Bu çalışmada da model başarısı, özellikle yüksek hata sınıfının yakalanabilirliği açısından değerlendirilmiştir.

Model performans ölçütlerinin hesaplanmasında kullanılan temel kavramlar aşağıda verilmiştir:

- TP (True Positive / Doğru Pozitif): Modelin pozitif sınıf olarak tahmin ettiği ve gerçekte de pozitif sınıfta olan gözlemler.
- FP (False Positive / Yanlış Pozitif): Modelin pozitif sınıf olarak tahmin ettiği, ancak gerçekte pozitif sınıfta olmayan gözlemler.
- TN (True Negative / Doğru Negatif): Modelin negatif sınıf olarak tahmin ettiği ve gerçekte de negatif sınıfta olan gözlemler.
- FN (False Negative / Yanlış Negatif): Modelin negatif sınıf olarak tahmin ettiği, ancak gerçekte pozitif sınıfta olan gözlemler.

Bu çalışmada iki sınıflı analizlerde pozitif sınıf, üretim açısından kritik olması nedeniyle “Yüksek” hata sınıfı olarak ele alınmıştır. Böylece recall, precision ve F1-score değerleri özellikle yüksek hata riskinin model tarafından ne ölçüde yakalanabildiğini değerlendirmek için kullanılmıştır.

3.5.1. Accuracy

Accuracy, modelin toplam doğru sınıflandırma oranını ifade etmektedir. Başka bir ifadeyle doğru sınıflandırılan gözlemlerin toplam gözlem sayısına oranı olarak hesaplanmaktadır. Bu ölçüt, modelin genel sınıflandırma başarısı hakkında ilk bakışta anlaşılır bir gösterge sunmaktadır. Accuracy değeri aşağıdaki formülle hesaplanmaktadır:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

Accuracy, sınıfların dengeli dağıldığı veri setlerinde yararlı bir ölçüt olabilir. Ancak sınıf dağılımı dengesiz olduğunda tek başına yeterli değildir. Örneğin gözlemlerin büyük çoğunluğu normal sınıfta yer alıyorsa, model yüksek hata sınıfını zayıf öğrense bile normal sınıfı doğru tahmin ederek yüksek accuracy değeri üretebilir. Bu durumda modelin genel doğruluğu yüksek görünür; fakat üretim açısından asıl kritik olan yüksek hata sınıfı yeterince yakalanmamış olabilir. Bu nedenle accuracy, çalışmada temel performans göstergelerinden biri olarak verilmiş; ancak nihai model değerlendirmesi yalnızca bu ölçüte dayandırılmamıştır.

Dengesiz sınıf yapılarında accuracy'nin yanıltıcı olabileceği farklı çalışmalarda vurgulanmıştır. Saito ve (Rehmsmeier, 2015), özellikle pozitif sınıfın az sayıda temsil edildiği veri setlerinde precision-recall temelli değerlendirmenin daha açıklayıcı olabileceğini belirtmiştir. Bu çalışmada da iki sınıflı model sonuçları yorumlanırken accuracy değerleri, recall ve F1-score sonuçlarıyla birlikte ele alınmıştır.

3.5.2. Precision

Precision, modelin pozitif olarak tahmin ettiği gözlemler içinde gerçekten pozitif olanların oranını göstermektedir. İki sınıflı analizlerde pozitif sınıf “Yüksek” hata sınıfı olarak kabul edildiğinde precision, modelin yüksek hata olarak işaretlediği gözlemlerin ne kadarının gerçekten yüksek hata sınıfına ait olduğunu ifade etmektedir. Precision aşağıdaki formülle hesaplanmaktadır:

$$Precision = TP / (TP + FP) \quad (2)$$

Precision değerinin yüksek olması, modelin yüksek hata tahmini yaptığında daha güvenilir olduğunu gösterir. Ancak precision tek başına yeterli değildir. Çünkü model çok az sayıda gözlemi yüksek hata olarak tahmin edip bunların çoğunu doğru yakalarsa yüksek precision değeri elde edebilir; buna karşılık gerçekte yüksek hata sınıfında bulunan çok sayıda gözlemi kaçırmış olabilir. Bu nedenle precision, recall ile değerlendirilmelidir.

Üretim kalite kontrolü açısından precision, yanlış alarm düzeyini anlamaya yardımcı olur. Modelin gereğinden fazla gözlemi yüksek hata olarak işaretlemesi, işletmede gereksiz kontrol veya müdahale yükü oluşturabilir. Buna karşılık yüksek hata sınıfını fazla

kaçırmamak da önemlidir. Bu nedenle precision, modelin yüksek hata tahminlerindeki güvenilirliğini gösterirken; recall modelin yüksek hata sınıfını yakalama kapasitesini göstermektedir. Goutte ve Gaussier, (2005), precision ve recall ölçütlerinin birlikte değerlendirilmesinin sınıflandırıcı performansını daha anlamlı biçimde yorumlamaya katkı sağladığını belirtmektedir.

3.5.3. Recall

Recall, gerçekte pozitif sınıfta bulunan gözlemlerin model tarafından ne kadarının doğru yakalandığını göstermektedir. İki sınıflı analizlerde pozitif sınıf “Yüksek” hata sınıfı olarak ele alındığında recall, gerçekte yüksek hata sınıfında bulunan kumaş kayıtlarının ne kadarının model tarafından yüksek hata olarak sınıflandırıldığını ifade etmektedir. Recall aşağıdaki formülle hesaplanmaktadır:

$$Recall = TP / (TP + FN) \quad (3)$$

Bu çalışmada recall ölçütü özel bir öneme sahiptir. Çünkü yüksek hata sınıfının kaçırılması, üretim sürecinde riskli kumaş kompozisyonlarının normal sınıf içinde değerlendirilmesine neden olabilir. Böyle bir durumda model, genel doğruluk bakımından başarılı görünse bile kalite kontrol kararları açısından yetersiz kalabilir. Bu nedenle iki sınıflı model sonuçlarında yüksek hata sınıfına ait recall değeri, modelin pratik kullanılabilirliğini değerlendirmek için kritik bir gösterge olarak ele alınmıştır.

Recall değerinin düşük olması, modelin yüksek hata sınıfını yeterince tanıyamadığını gösterir. Bu durum özellikle dengesiz veri setlerinde sık karşılaşılan bir sorundur. Model çoğunluk sınıfı olan normal gözlemleri doğru tahmin ederek yüksek accuracy değerine ulaşabilir; ancak yüksek hata sınıfına ait gözlemleri düşük oranda yakalıyorsa üretim açısından beklenen erken uyarı işlevini yerine getiremez. Davis ve Goadrich, (2006), precision-recall analizinin özellikle pozitif sınıfın önem taşıdığı sınıflandırma problemlerinde model davranışını daha açık biçimde gösterdiğini belirtmiştir. Bu bağlamda recall, çalışmada yüksek hata sınıfının yakalanma başarısını değerlendirmek için temel ölçütlerden biri olarak kullanılmıştır.

3.5.4. F1-Score

F1-score, precision ve recall ölçütlerinin harmonik ortalamasıdır. Bu ölçüt, modelin pozitif sınıfı tahmin etmedeki doğruluğu ile pozitif sınıfı yakalama başarısını birlikte değerlendirmektedir. F1-score aşağıdaki formülle hesaplanmaktadır:

$$F1 - Score = 2x(Precision \times Recall) / (Precision + Recall) \quad (4)$$

F1-score değerinin yüksek olması, modelin precision ve recall bakımından dengeli bir performans gösterdiğine işaret eder. Harmonik ortalama kullanıldığı için precision veya recall değerlerinden biri çok düşük olduğunda F1-score da düşmektedir. Bu özellik, F1-score'u dengesiz sınıf dağılımına sahip sınıflandırma problemlerinde accuracy'ye göre daha dikkatli bir değerlendirme ölçütü haline getirmektedir. (Goutte ve Gaussier, 2005), F-score'un precision ve recall değerlerini birlikte değerlendiren bir gösterge olarak sınıflandırıcı karşılaştırmalarında kullanılabileceğini belirtmiştir. Powers (2011) ise precision, recall ve F-measure gibi ölçütlerin sınıflandırma değerlendirmelerinde yaygın kullanıldığını, ancak bu ölçütlerin yorumlanmasında temel sınıf oranlarının dikkate alınması gerektiğini vurgulamıştır.

Bu çalışmada F1-score, özellikle iki sınıflı modellerin karşılaştırılmasında önemli bir değerlendirme ölçütü olarak kullanılmıştır. Bunun nedeni, iki sınıflı yapıda "Yüksek" hata sınıfının "Normal" sınıfa göre daha az gözlem içermesidir. Böyle bir durumda model yüksek accuracy değerine ulaşsa bile yüksek hata sınıfını yeterli düzeyde öğrenemeyebilir. F1-score, precision ve recall değerlerini birlikte dikkate aldığı için bu tür durumlarda model performansını daha dengeli biçimde değerlendirmeye yardımcı olmaktadır. Bununla birlikte F1-score'un da tek başına mutlak bir başarı ölçütü olarak görülmemesi gerekir. Chicco ve Jurman (2020), ikili sınıflandırma problemlerinde bazı durumlarda Matthews korelasyon katsayısının accuracy ve F1-score'a göre daha bilgilendirici olabileceğini belirtmiştir. Bu değerlendirme, herhangi bir ölçütün tek başına tüm model davranışını açıklayamayacağını göstermektedir. Mevcut çalışma kapsamında ise model sonuçlarının yorumlanmasında accuracy, precision, recall ve F1-score birlikte değerlendirilmiş; ancak yüksek hata sınıfının yakalanması açısından recall ve F1-score daha belirleyici ölçütler olarak ele alınmıştır.

Çalışmada kullanılan performans ölçütleri Çizelge 3.7’de özetlenmiştir.

Çizelge 3.7. Model performans ölçütleri ve yorumlanma biçimleri

Performans ölçütü	Formül	Temel yorum
Accuracy	$((TP + TN) / (TP + TN + FP + FN))$	Modelin toplam doğru sınıflandırma oranını gösterir
Precision	$(TP / (TP + FP))$	Modelin pozitif olarak tahmin ettiği gözlemler içinde gerçekten pozitif olanların oranını gösterir
Recall	$(TP / (TP + FN))$	Gerçekte pozitif sınıfta olan gözlemlerin model tarafından ne kadarının yakalandığını gösterir
F1-score	$(2 \times (Precision \times Recall) / (Precision + Recall))$	Precision ve recall değerlerini birlikte değerlendiren dengeli bir ölçüttür

Çalışmada model performansı çok ölçütlü bir yaklaşımla değerlendirilmiştir. Accuracy, modelin genel doğru sınıflandırma düzeyini göstermesi bakımından raporlanmıştır. Ancak iki sınıflı yapıda yüksek hata sınıfının daha az sayıda gözlem içermesi nedeniyle accuracy tek başına yeterli görülmemiştir. Precision, modelin yüksek hata tahminlerindeki güvenilirliğini; recall, yüksek hata sınıfını yakalama başarısını; F1-score ise bu iki ölçüt arasındaki dengeyi göstermektedir. Bu nedenle özellikle yüksek hata riskinin üretim açısından kritik olduğu durumlarda recall ve F1-score değerleri, model karşılaştırmalarında daha öncelikli biçimde değerlendirilmiştir.

3.6. Karbon Ayak İzi Hesaplama Yöntemi

Bu çalışmada karbon ayak izi hesabı, kumaş hatalarından kaynaklanan yeniden işlem faaliyetlerinin neden olduğu ek elektrik tüketimi üzerinden yapılmıştır. Hesaplama kapsamı bilinçli olarak sınırlandırılmıştır. Ürünün hammadde temini, kumaş üretimi, boyama, yıkama, konfeksiyon, taşıma, kullanım ve bertaraf aşamalarını kapsayan bütüncül bir yaşam döngüsü analizi yapılmamıştır. Bu nedenle burada sunulan hesaplama, ürünün toplam karbon ayak izini değil, kalite hatalarına bağlı lokal süreç etkisini göstermektedir.

Karbon ayak izi hesaplamalarında sistem sınırının açıkça tanımlanması gerekir. ISO 14067 standardı, ürün karbon ayak izinin yaşam döngüsü değerlendirmesi ilkeleriyle uyumlu

şekilde hesaplanmasını öngörmekte; bunun yanında ürünün yalnızca belirli bir bölümünü kapsayan kısmi karbon ayak izi çalışmalarına da olanak tanımaktadır (ISO, 2018). Bu çalışmadaki yaklaşım, kısmi ve süreç odaklı bir hesaplama niteliğindedir. Amaç, denim konfeksiyon üretiminde kumaş hatalarının doğurduğu ek yeniden işlem süresini ve bu sürenin elektrik tüketimi üzerinden oluşturduğu CO₂e emisyonunu hesaplamaktır. Hesaplama süreci dört temel adımdan oluşmaktadır. İlk adımda veri setindeki toplam kumaş hata adedi belirlenmiştir. İkinci adımda toplam hata adedi, hata başına varsayılan ortalama yeniden işlem süresiyle çarpılarak toplam ek rework süresi hesaplanmıştır. Üçüncü adımda bu süre, yeniden işlem faaliyetinde kullanılan ekipmanın ortalama güç tüketimiyle ilişkilendirilerek ek elektrik tüketimine dönüştürülmüştür. Son aşamada ise ek elektrik tüketimi, Türkiye elektrik şebekesi için kullanılan emisyon faktörü ile çarpılarak CO₂e cinsinden karbon ayak izi hesaplanmıştır.

Elektrik tüketimine bağlı emisyonların hesaplanmasında, satın alınan elektrikten kaynaklanan dolaylı emisyonların dikkate alınması gerekmektedir. GHG Protocol Scope 2 rehberi, satın alınan veya tüketilen elektrik, buhar, ısı ve soğutmadan kaynaklanan dolaylı emisyonların hesaplanması için konum temelli ve piyasa temelli yaklaşımları tanımlamaktadır (WRI & WBCSD, 2015). Bu çalışmada işletmenin yenilenebilir enerji sertifikası veya tedarik sözleşmesine dayalı özel bir piyasa temelli emisyon faktörü kullanılmadığından, hesaplama şebeke emisyon faktörü üzerinden ele alınmıştır. IEA Emission Factors veri seti, ülkeler bazında elektrik üretimi ve elektrik tüketimine bağlı emisyon faktörlerinin kullanılabileceği güncel kaynaklardan biridir (IEA, 2025).

Bu hesaplama, kumaş hatalarının üretim sürecinde yalnızca zaman ve işçilik kaybı oluşturmadığını, aynı zamanda ölçülebilir bir enerji tüketimi ve karbon emisyonu etkisi doğurduğunu ortaya koymak için yapılmıştır. Kalite problemleri nedeniyle ortaya çıkan yeniden işlem faaliyetleri, normal üretim akışına eklenen fazladan iş yükü anlamına gelmektedir. Yeniden işlem süresi arttıkça makine ve ekipman kullanım süresi uzamakta, buna bağlı olarak elektrik tüketimi ve dolaylı karbon emisyonu artmaktadır. Sürdürülebilir üretim literatüründe düşük kaliteli ürün sayısının azaltılmasının fire ve yeniden işlem ihtiyacını düşürerek enerji ve iş yükünü azaltılabileceği belirtilmektedir (Nobil vd., 2023). Bu

nedenle kalite hatalarının karbon ayak iziyle ilişkilendirilmesi, üretim verimliliği ve çevresel sürdürülebilirlik arasındaki bağı sayısal olarak görünür hale getirmektedir. Çalışmada kullanılan hesaplama girdileri Çizelge 3.8’de verilmiştir.

Çizelge 3.8. Karbon ayak izi hesabında kullanılan temel girdiler

Parametre	Sembol	Değer	Birim / Açıklama
Toplam kumaş hata adedi	(H_{toplam})	31.943	adet
Hata başına ortalama ek rework süresi	(t_{hata})	1,05	dakika/hata
Ortalama güç tüketimi	(P)	10	kW
Elektrik emisyon faktörü	(EF_{elektrik})	0,4183	kg CO e/kWh
Hesaplama kapsamı	—	Lokal süreç	Hata kaynaklı ek rework etkisi

Bu girdiler kullanılarak toplam hata adedi, ek rework süresi, ek elektrik tüketimi ve karbon ayak izi sırasıyla hesaplanmıştır. Kullanılan eşitlikler aşağıdaki alt başlıklarda verilmiştir.

3.6.1. Toplam Hata Adedinin Hesaplanması

Karbon ayak izi hesabının ilk aşamasında, veri setindeki tüm gözlemlere ait kumaş hata adetleri toplanmıştır. Toplam hata adedi, hata kaynaklı ek yeniden işlem yükünün başlangıç girdisini oluşturmaktadır. Bu değer ne kadar yüksekse, varsayılan hata başına işlem süresi sabit kabul edildiğinde toplam rework süresi de aynı yönde artmaktadır.

Toplam hata adedi aşağıdaki eşitlikle hesaplanmıştır:

$$H_{toplam} = H_i \quad (5)$$

Bu eşitlikte;

H_{toplam} : veri setindeki toplam kumaş hata adedini,

H_i : i’inci gözleme ait kumaş hata adedini,

Σ : tüm gözlemler üzerinden toplam alma işlemini ifade etmektedir.

Bu değer, çalışma kapsamındaki kumaş kayıtlarında tespit edilen toplam hata yükünü temsil etmektedir. Hesaplamanın sonraki aşamalarında bu toplam hata adedi, hata başına ortalama ek yeniden işlem süresi ile ilişkilendirilmiştir.

3.6.2. Rework Süresinin Hesaplanması

Toplam hata adedi belirlendikten sonra, bu hata yükünün üretim sürecinde ne kadar ek yeniden işlem süresi oluşturduğu hesaplanmıştır. Yeniden işlem süresi, kalite hataları nedeniyle normal üretim akışına eklenen kontrol, düzeltme ve işlem faaliyetlerini temsil etmektedir. Bu çalışmada her bir kumaş hatasının ortalama 1,05 dakika ek rework süresi oluşturduğu varsayılmıştır. Bu değer, işletme kayıtlarında kullanılan hesaplama yaklaşımı ve süreç değerlendirmesi temel alınarak sabit kabul edilmiştir.

Toplam ek rework süresi aşağıdaki eşitlikle hesaplanmıştır:

$$Tek = H_{toplam} \cdot x_{thata} / 60 \quad (6)$$

Bu eşitlikte;

Tek: toplam ek rework süresini (saat),

H_{toplam}: toplam kumaş hata adedini,

thata: hata başına ortalama ek işlem süresini (dakika/hata),

60: dakika değerini saate dönüştürmek için kullanılan dönüşüm katsayısını ifade etmektedir.

Bu süre, normal üretim faaliyetinden bağımsız olarak kalite hataları nedeniyle ortaya çıkan ilave üretim yükünü temsil etmektedir.

3.6.3. Ek Elektrik Tüketiminin Hesaplanması

Toplam ek rework süresi hesaplandıktan sonra, bu sürenin neden olduğu ek elektrik tüketimi belirlenmiştir. Yeniden işlem faaliyetlerinde kullanılan ekipmanın ortalama güç tüketimi 10 kW olarak kabul edilmiştir. Ek elektrik tüketimi, toplam ek rework süresi ile ortalama güç tüketiminin çarpılmasıyla hesaplanmıştır.

Ek elektrik tüketimi aşağıdaki eşitlikle hesaplanmıştır:

$$Eek = TekxP \quad (7)$$

Bu eşitlikte;

Eek: hata kaynaklı ek elektrik tüketimini (kWh),

Tek: toplam ek rework süresini (saat),

P: yeniden işlem sürecinde kullanılan ekipmanın ortalama güç tüketimini (kW) ifade etmektedir.

Bu değer, yalnızca hata kaynaklı ilave üretim süresinden doğan elektrik tüketimini temsil etmektedir. Normal üretim sürecinde kullanılan elektrik tüketimi bu hesaplama dahil edilmemiştir.

3.6.4. Karbon Ayak İzi Hesabı

Son aşamada, hata kaynaklı ek elektrik tüketimi CO₂e cinsinden karbon ayak izine dönüştürülmüştür. Bu amaçla ek elektrik tüketimi, Türkiye elektrik şebekesi için kullanılan elektrik emisyon faktörüyle çarpılmıştır. Çalışmada elektrik emisyon faktörü 0,4183 kg CO₂e/kWh olarak alınmıştır. Elektrik tüketimine bağlı emisyon hesaplamalarında ülke veya şebeke düzeyinde emisyon faktörü kullanımı, tüketilen elektriğin karbon yoğunluğunu hesaplama olanak sağlamaktadır.

Karbon ayak izi aşağıdaki eşitlikle hesaplanmıştır:

$$CFek = EekxEFelektrik \quad (8)$$

Bu eşitlikte;

CFek: hata kaynaklı ek karbon ayak izini (kg CO₂e),

Eek: hata kaynaklı ek elektrik tüketimini (kWh),

EFelektrik: elektrik tüketimine bağlı emisyon faktörünü (kg CO₂e/kWh) ifade etmektedir.

Bu değer, çalışma kapsamındaki hata kaynaklı ilave elektrik tüketiminin karbon

karşılığıdır. Dolayısıyla sonuç, denim ürünlerinin toplam yaşam döngüsü karbon ayak izini değil, yalnızca kalite hatalarından kaynaklanan lokal yeniden işlem etkisini temsil etmektedir. Hesaplama süreci Çizelge 3.9'da özetlenmiştir.

Çizelge 3.9. Karbon ayak izi hesaplama adımları

Hesaplama adımı	Eşitlik kodları
Toplam hata adedi	$(H_{\text{toplam}} = \sum H_i)$
Toplam ek rework süresi	$(T_{\text{ek}} = H_{\text{toplam}} \times t_{\text{hata}} / 60)$
Ek elektrik tüketimi	$(E_{\text{ek}} = T_{\text{ek}} \times P)$
Ek karbon ayak izi	$(CF_{\text{ek}} = E_{\text{ek}} \times EF_{\text{elektrik}})$
Ton cinsinden ek karbon ayak izi	$(CF_{\text{ek}} / 1000)$

Bu yöntemsel çerçeve, kalite problemlerinin çevresel etkisini üretim verileriyle ilişkilendirmektedir. Hesaplamanın en güçlü yönü, kumaş hata adedi gibi doğrudan işletme kayıtlarından elde edilen bir kalite göstergesini enerji tüketimi ve karbon emisyonu hesabına bağlamasıdır. Buna karşılık yöntemin sınırları da açık biçimde belirtilmelidir. Hata başına rework süresi ve ortalama güç tüketimi sabit kabul edilmiştir. Ayrıca hesaplama yalnızca elektrik tüketimine bağlı dolaylı emisyonları içermekte; işçilik, yardımcı malzeme, bakım, taşıma, atık, su tüketimi veya proses bazlı diğer enerji kaynakları hesaplama kapsamına alınmamaktadır. Bu nedenle sonuçlar, kalite hatalarının tüm çevresel etkisini değil, elektrik tüketimi üzerinden hesaplanan sınırlı fakat ölçülebilir bir karbon ayak izi etkisini göstermektedir.

4. BULGULAR

4.1. Veri Setine İlişkin Bulgular

Çalışmada kullanılan veri seti, büyük ölçekli bir denim konfeksiyon işletmesine ait üretim ve kalite kontrol kayıtlarından oluşmaktadır. Veri setinde toplam 652 gözlem yer almaktadır. Her gözlem, işletme kayıtlarında bulunan bir sipariş/kumaş kompozisyon kaydını ve bu kayda karşılık gelen kumaş hata adedini temsil etmektedir. Analizlerde model girdisi olarak kumaş kompozisyonundan türetilen lif oranları kullanılmış; model çıktısı ise Kumaş Hata Adedi değişkeninden türetilen hata düzeyi sınıfları olarak değerlendirilmiştir.

Veri setinde yer alan temel değişkenler; sipariş numarası, kumaş kompozisyon bilgisi, kumaş hata adedi, pamuk oranı, geri dönüştürülmüş lif oranı, polyester oranı, elastan oranı ve diğer lifler oranından oluşmaktadır. Sipariş numarası ve orijinal kumaş kompozisyon bilgisi tanımlayıcı nitelikte olduğundan modelleme sürecinde doğrudan bağımsız değişken olarak kullanılmamıştır. Buna karşılık COT oranı, Recycle oran, PES oranı, EA oranı ve Diğer lifler oranı değişkenleri sayısal model girdileri olarak ele alınmıştır.

Veri setinde eksik gözlem bulunmamaktadır. Ayrıca lif oranlarının toplamını kontrol etmek amacıyla yapılan incelemede tüm gözlemlerde toplam kompozisyon değerinin %100 olduğu belirlenmiştir. Bu durum, veri setinin lif oranları toplam kontrolü bakımından tutarlı bir yapıya sahip olduğunu göstermektedir. Veri setinin genel yapısı Çizelge 4.1’de verilmiştir.

Çizelge 4.1. Veri setinin genel yapısı

Özellik	Değer
Toplam gözlem sayısı	652
Ana değişken sayısı	8
Tanımlayıcı değişkenler	Order No, Kumaş Kompozisyon
Bağımsız değişkenler	COT oranı, Recycle oran, PES oranı, EA oranı, Diğer lifler oranı
Hedef değişken	Kumaş Hata Adedi
Eksik veri durumu	Yok
Lif oranları toplam kontrolü	Tüm gözlemlerde %100

Sayısal değişkenlere ilişkin tanımlayıcı istatistikler Çizelge 4.2’de sunulmuştur. Kumaş hata adedi değişkeni 0 ile 582 arasında değişmektedir. Ortalama hata adedi 48,99, medyan

değer ise 31 olarak hesaplanmıştır. Ortalama değer in medyandan yüksek olması, hata adedi dağılımında yüksek değerlere sahip bazı gözlemlerin bulunduğunu göstermektedir. Bu yapı, kumaş hata adedinin tüm gözlemler arasında dengeli biçimde dağılmadığını ve yüksek hata yüküne sahip kayıtların ayrıca dikkate alınması gerektiğini ortaya koymaktadır.

Çizelge 4.2. Sayısal değişkenlere ilişkin tanımlayıcı istatistikler

Değişken	Minimum	Maksimum	Ortalama	Medyan	Standart sapma
Kumaş Hata Adedi	0	582	48,99	31	57,96
COT oranı	0	100	77,55	79	20,04
Recycle oran	0	50	10,38	0	14,38
PES oranı	0	80	2,48	0	6,94
EA oranı	0	5	1,39	1	1,43
Diğer lifler oranı	0	99	8,19	0	17,06

Pamuk oranı değişkeni incelendiğinde, ortalama değer in %77,55, medyan değer in ise %79 olduğu görülmektedir. Bu sonuç, veri setindeki kumaş kayıtlarının büyük bölümünde pamuk oranının yüksek olduğunu göstermektedir. Denim kumaş yapısında pamuk oranının baskın olması beklenen bir durumdur. Bununla birlikte pamuk oranının 0 ile 100 arasında geniş bir aralıkta değişmesi, veri setinde farklı kumaş kompozisyonlarının bulunduğunu göstermektedir.

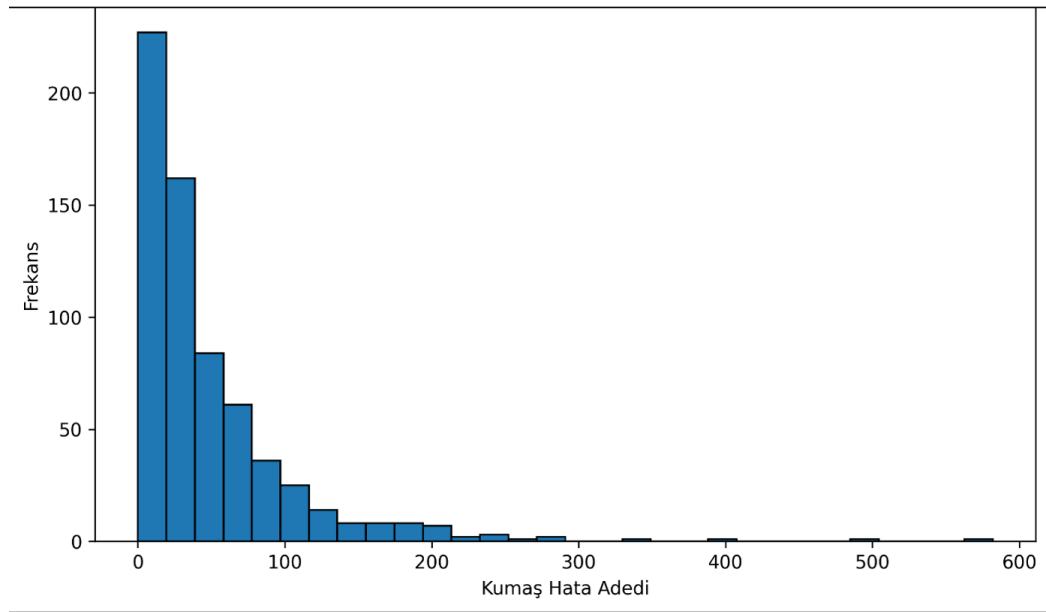
Geri dönüştürülmüş lif oranı 0 ile 50 arasında değişmektedir. Ortalama değer in %10,38, medyan değer ise 0 olarak hesaplanmıştır. Medyan değer in 0 olması, veri setindeki gözlemlerin önemli bir bölümünde geri dönüştürülmüş lif oranının bulunmadığını veya sıfır olarak kaydedildiğini göstermektedir. Buna karşılık maksimum değer in %50'ye ulaşması, bazı kumaş kompozisyonlarında geri dönüştürülmüş lif kullanımının belirgin düzeyde olduğunu ortaya koymaktadır.

Polyester oranı için ortalama değer in %2,48, medyan değer ise 0'dır. Bu durum, polyesterin veri setindeki birçok kumaş kaydında yer almadığını ya da düşük oranlarda bulunduğunu göstermektedir. Ancak maksimum değer in %80 olması, bazı kumaş kompozisyonlarında polyester oranının oldukça yüksek düzeye çıkabildiğini göstermektedir. Elastan oranı ise 0

ile 5 arasında değişmekte olup ortalama %1,39 olarak hesaplanmıştır. Bu bulgu, elastanın denim kumaşlarda genellikle esneklik sağlamak amacıyla sınırlı oranlarda kullanılan bir bileşen olduğunu düşündürmektedir.

Diğer lifler oranı değişkeni 0 ile 99 arasında değişmektedir. Ortalama değer %8,19, medyan değer ise 0 olarak hesaplanmıştır. Medyan değer 0 olması, birçok kumaş kompozisyonunda pamuk, geri dönüştürülmüş lif, polyester ve elastan dışındaki liflerin yer almadığını ya da sınırlı düzeyde bulunduğunu göstermektedir. Buna karşılık maksimum değer %99'a ulaşması, bazı kayıtlarda diğer lif gruplarının kompozisyon içinde baskın hale gelebildiğini ortaya koymaktadır.

Şekil 4.1'de veri setinde yer alan Kumaş Hata Adedi değişkeninin dağılımı gösterilmektedir. Histogram incelendiğinde, gözlemlerin büyük bölümünün düşük hata adedi aralıklarında yoğunlaştığı görülmektedir. Buna karşılık daha yüksek hata değerlerine sahip gözlem sayısı giderek azalmakta ve dağılım sağa çarpık bir yapı göstermektedir.

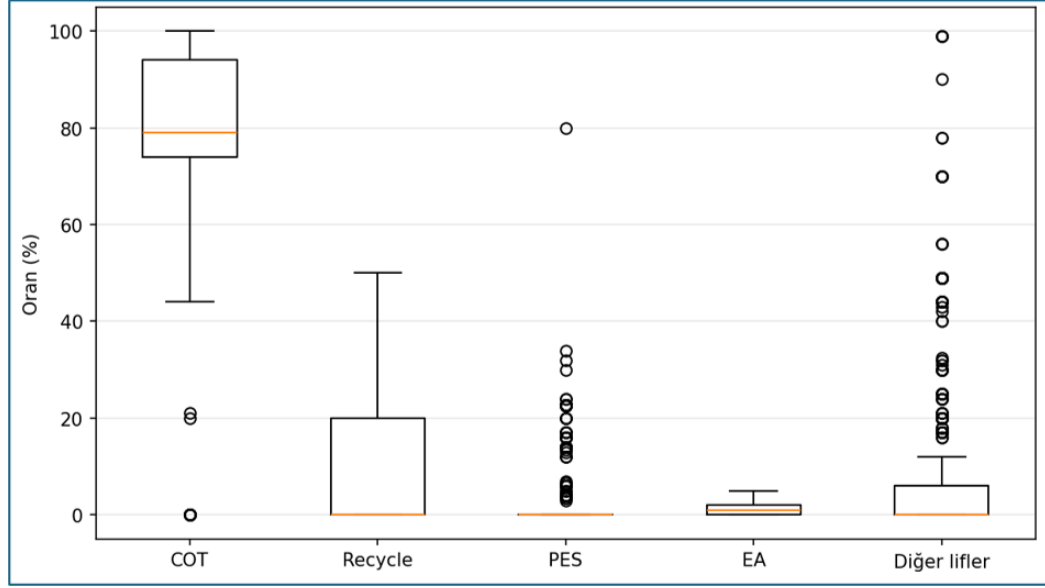


Şekil 4.1. Kumaş hata adedi dağılımı

Bu dağılım yapısı, kumaş hata adedinin veri setinde homojen biçimde dağılmadığını ortaya koymaktadır. Özellikle yüksek hata adedine sahip gözlemlerin daha sınırlı sayıda olması, sınıflandırma modellerinde yüksek hata sınıfının ayrıca dikkatle değerlendirilmesini

gerektirmektedir. Bu nedenle sonraki aşamalarda kumaş hata adedi, iki sınıflı ve üç sınıflı sınıflandırma yapılarına dönüştürülmüştür.

Şekil 4.2’de veri setinde yer alan COT oranı, Recycle oran, PES oranı, EA oranı ve Diğer lifler oranı değişkenlerinin dağılımları kutu grafikleri ile gösterilmektedir. Kutu grafikleri, her bir değişken için medyan değeri, alt ve üst çeyrekleri, yayılım genişliğini ve uç gözlemleri birlikte değerlendirmeye imkân vermektedir.



Şekil 4.2. Lif oranlarına ait kutu grafikleri

Grafik incelendiğinde, COT oranı değişkeninin diğer lif oranlarına göre daha yüksek medyan değere sahip olduğu ve veri setinde baskın lif bileşeni olarak öne çıktığı görülmektedir. Recycle oran değişkeninde medyan değerin 0 olması, geri dönüştürülmüş lif kullanımının tüm kayıtlarda düzenli biçimde yer almadığını göstermektedir. PES oranı büyük ölçüde düşük değerlerde toplanmakta, ancak bazı gözlemlerde daha yüksek değerlere ulaşmaktadır. EA oranı ise dar bir aralıkta değişmekte ve denim kumaş yapısında sınırlı oranlarda yer alan bir bileşen olarak görünmektedir. Diğer lifler oranında ise medyan değerin 0 olmasına rağmen geniş bir yayılım bulunması, bazı kumaş kompozisyonlarında pamuk, geri dönüştürülmüş lif, polyester ve elastan dışındaki liflerin daha belirgin oranlarda yer alabildiğini göstermektedir.

Genel olarak değerlendirildiğinde, veri seti pamuk oranının baskın olduğu, geri

dönüştürülmüş lif ve diğer lif oranlarının bazı kumaş kayıtlarında belirginleştiği, polyester ve elastan oranlarının ise çoğu gözlemde daha sınırlı düzeyde kaldığı bir kompozisyon yapısı göstermektedir. Kumaş hata adedinin sağa çarpık dağılım göstermesi ve yüksek hata değerlerinin daha az sayıda gözlemde yoğunlaşması, hata düzeylerinin sınıflandırma problemi olarak ele alınmasını desteklemektedir. Bu nedenle sonraki analizlerde hem lif oranlarının hata adediyle ilişkisi hem de hata düzeylerinin makine öğrenmesi modelleriyle sınıflandırılabilirliği ayrı başlıklar altında değerlendirilmiştir.

4.2. Lif Oranlarının Kumaş Hatalarıyla İlişkisi

Bu bölümde, kumaş kompozisyonundan türetilen lif oranları ile kumaş hata adedi arasındaki ilişki iki farklı analiz üzerinden değerlendirilmiştir. İlk aşamada, korelasyon analizinde kullanılacak değişkenlerin normal dağılım varsayımını sağlayıp sağlamadığı Shapiro-Wilk testi ile incelenmiştir. Normallik testi sonuçlarına göre değişkenlerin normal dağılım varsayımını karşılamadığı belirlendiğinden, lif oranları ile Kumaş Hata Adedi değişkeni arasındaki ilişki Spearman sıra korelasyonu ile değerlendirilmiştir. Ardından Random Forest temelli değişken önem analizi uygulanarak lif oranlarının model içindeki görece katkıları belirlenmiştir. Böylece lif oranlarının hata adediyle doğrudan ilişkisi ile sınıflandırma modeli içindeki ayırt edici katkısı ayrı ayrı değerlendirilmiştir.

4.2.1. Normallik Testi Bulguları

Korelasyon analizinde kullanılacak değişkenlerin normal dağılım varsayımını sağlayıp sağlamadığı Shapiro-Wilk testi ile değerlendirilmiştir. Test sonuçları Çizelge 4.3'te verilmiştir. Shapiro-Wilk testinde sıfır hipotezi, ilgili değişkenin normal dağılım gösterdiği yönündedir. Elde edilen p değerlerinin 0,05'in altında olması durumunda normal dağılım varsayımı reddedilmektedir.

Çizelge 4.3. Değişkenlere ilişkin Shapiro-Wilk normallik testi sonuçları

Değişken	Shapiro-Wilk W	p değeri	Değerlendirme
Kumaş Hata Adedi	0,6878	$5,70 \times 10^{-33}$	Normal dağılım göstermemektedir
COT oranı	0,8429	$7,04 \times 10^{-25}$	Normal dağılım göstermemektedir
Recycle oran	0,7153	$8,15 \times 10^{-32}$	Normal dağılım göstermemektedir
PES oranı	0,4056	$1,19 \times 10^{-41}$	Normal dağılım göstermemektedir
EA oranı	0,8195	$2,05 \times 10^{-26}$	Normal dağılım göstermemektedir
Diğer lifler oranı	0,5536	$1,16 \times 10^{-37}$	Normal dağılım göstermemektedir

Çizelge 4.3 incelendiğinde, korelasyon analizinde kullanılan tüm değişkenlerde p değerlerinin 0,05'in altında olduğu görülmektedir. Bu nedenle Kumaş Hata Adedi, COT oranı, Recycle oran, PES oranı, EA oranı ve Diğer lifler oranı değişkenleri için normal dağılım varsayımı reddedilmiştir. Değişkenlerin normal dağılım göstermemesi nedeniyle, lif oranları ile kumaş hata adedi arasındaki ilişkinin değerlendirilmesinde Pearson korelasyonu yerine Spearman sıra korelasyonu kullanılmıştır.

4.2.2. Spearman Korelasyon Analizi Bulguları

Lif oranları ile kumaş hata adedi arasındaki ilişki Spearman sıra korelasyonu ile değerlendirilmiştir. Elde edilen sonuçlar Çizelge 4.4'te verilmiştir. Spearman korelasyon katsayıları incelendiğinde, tüm değişkenlerde ilişki düzeyinin oldukça düşük olduğu görülmektedir. COT oranı ile kumaş hata adedi arasındaki Spearman rho değeri -0,0148, Recycle oran için 0,0440, PES oranı için 0,0146, EA oranı için -0,0523 ve Diğer lifler oranı için -0,0655 olarak hesaplanmıştır.

Çizelge 4.4. . Lif oranları ile kumaş hata adedi arasındaki Spearman korelasyon katsayıları ve anlamlılık değerleri

Değişken	Spearman rho	P değeri
Cot oranı	-0,0148	0,7056
Recycle oran	0,0440	0,2614
Pes oranı	0,0146	0,7107
EA oranı	-0,0523	0,1826
Diğer lifler oranı	-0,0655	0,0945

Spearman korelasyon sonuçlarına göre lif oranları ile kumaş hata adedi arasında istatistiksel olarak anlamlı bir ilişki belirlenmemiştir ($p>0,05$). Bu bulgu, mevcut veri setinde COT oranı, Recycle oran, PES oranı, EA oranı ve Diğer lifler oranının kumaş hata adediyle güçlü ve anlamlı bir monoton ilişki göstermediğini ortaya koymaktadır. Diğer lifler oranı değişkeninde p değeri diğer değişkenlere göre daha düşük olmakla birlikte, bu değer de 0,05 anlamlılık düzeyinin üzerinde kalmıştır. Bu nedenle korelasyon bulgularına dayanarak herhangi bir lif oranının kumaş hata adedini doğrudan artırdığı veya azalttığı söylenemez.

Bu sonuç, kumaş hata oluşumunun yalnızca lif oranlarıyla açıklanamayacak çok faktörlü bir yapıya sahip olduğunu göstermektedir. Kumaş hata adedi; lif kompozisyonunun yanı sıra iplik kalitesi, üretim hızı, makine ayarları, operatör etkisi, bakım durumu, yıkama ve son işlem koşulları gibi birçok üretim değişkeninden etkilenebilir. Bu değişkenler veri setinde yer almadığından, Spearman korelasyon analizi yalnızca mevcut lif oranları ile hata adedi arasındaki monoton ilişkiyi sınırlı biçimde göstermektedir.

4.2.3. Değişken Önem Analizi Bulguları

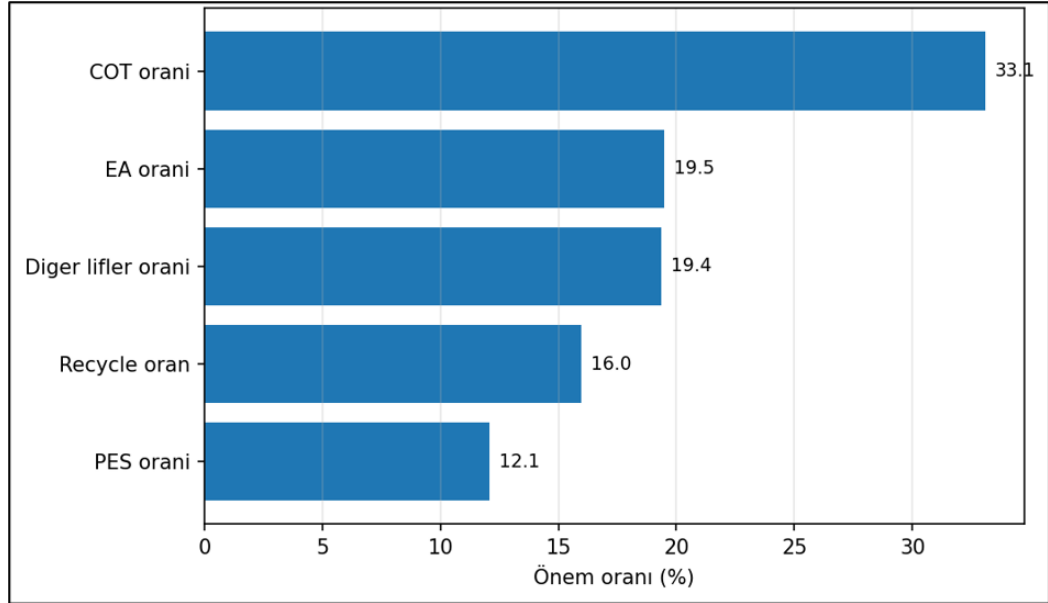
Lif oranlarının sınıflandırma modeli içindeki göreceli katkısını değerlendirmek amacıyla Random Forest temelli değişken önem analizi uygulanmıştır. Elde edilen değişken önem skorları Çizelge 4.5’de verilmiştir. Sonuçlara göre model kararlarına en yüksek katkıyı COT oranı sağlamıştır. COT oranının önem skoru 0,3308, önem oranı ise %33,08 olarak hesaplanmıştır.

COT oranını sırasıyla EA oranı, Diğer lifler oranı, Recycle oran ve PES oranı izlemiştir. EA oranının önem oranı %19,49, Diğer lifler oranının %19,37, Recycle oranının %15,96 ve PES oranının %12,10 olduğu belirlenmiştir. Bu sonuç, Random Forest modelinin sınıflandırma sürecinde tüm lif oranlarından belirli düzeyde yararlandığını; ancak COT oranının diğer değişkenlere göre daha belirgin bir ayırt edici katkı sunduğunu göstermektedir.

Çizelge 4.5. Random Forest temelli değişken önem analizi sonuçları

Sıra	Değişken	Önem skoru	Önem oranı (%)
1	Cot oranı	0,3308	33,08
2	EA oranı	0,1949	19,49
3	Diğer lifler oranı	0,1937	19,37
4	Recycle oran	0,1596	15,96
5	Pes oranı	0,1210	12,10

Şekil 4.3'te Random Forest temelli değişken önem skorları grafiksel olarak sunulmuştur. Grafik incelendiğinde COT oranının diğer değişkenlerden daha yüksek önem skoruna sahip olduğu görülmektedir. EA oranı ile Diğer lifler oranı birbirine oldukça yakın önem düzeyleri göstermiştir. Recycle oran orta düzeyde katkı sunarken, PES oranı model içindeki en düşük önem skoruna sahip değişken olmuştur.

**Şekil 4.3.** Random Forest temelli değişken önem skorları

Değişken önem analizi, korelasyon analizinden farklı olarak değişkenlerin model içindeki katkısını değerlendirmektedir. Bu nedenle COT oranının en yüksek önem skoruna sahip olması, pamuk oranının kumaş hata adedini tek başına doğrudan açıkladığı anlamına gelmemektedir. Bu bulgu, modelin hata düzeyi sınıflandırmasında COT oranından diğer

lif değişkenlerine göre daha fazla yararlandığını göstermektedir. Benzer şekilde EA oranı ve Diğer lifler oranının ikinci ve üçüncü sırada yer alması, bu değişkenlerin de model kararlarında belirli düzeyde ayırt edici bilgi taşıdığını ortaya koymaktadır.

4.2.4. Korelasyon ve Değişken Önem Bulgularının Birlikte Sunumu

Korelasyon ve değişken önem sonuçları birlikte değerlendirildiğinde, lif oranlarının kumaş hata adediyle güçlü ve anlamlı bir monoton ilişki göstermediği; buna karşılık makine öğrenmesi modeli içinde belirli düzeyde ayırt edici katkı sunduğu görülmektedir. Spearman korelasyon analizinde tüm katsayıların düşük düzeyde olması ve p değerlerinin 0,05'in üzerinde kalması, lif oranları ile hata adedi arasında istatistiksel olarak anlamlı bir ilişki bulunmadığını ortaya koymaktadır. Buna karşılık Random Forest değişken önem analizinde COT oranı başta olmak üzere EA oranı, Diğer lifler oranı, Recycle oran ve PES oranı model kararlarına farklı düzeylerde katkı sağlamıştır.

Bu iki bulgu birbirini dışlamamaktadır. Korelasyon analizi değişkenler arasındaki ikili ve doğrusal ilişkiyi gösterirken, Random Forest değişken önem analizi modelin sınıflandırma sürecinde hangi değişkenlerden daha fazla yararlandığını göstermektedir. Bu nedenle düşük korelasyon katsayılarına rağmen bazı lif oranlarının değişken önem analizinde öne çıkması beklenebilir bir durumdur. Özellikle kumaş hata oluşumu gibi çok faktörlü kalite problemlerinde doğrusal olmayan ilişkiler, değişken etkileşimleri ve üretim sürecinden kaynaklanan farklı etkiler bulunabilir.

Lif oranları kumaş hata adedini tek başına doğrudan açıklayan değişkenler olarak değerlendirilmemelidir. Ancak Random Forest temelli değişken önem analizi, özellikle COT oranının model içinde güçlü bir ayırt edici bilgi taşıdığını göstermektedir. EA oranı ve Diğer lifler oranının birbirine yakın önem düzeylerine sahip olması, kumaş kompozisyonunun yalnızca pamuk oranı üzerinden değil, diğer lif bileşenleriyle birlikte değerlendirilmesi gerektiğini düşündürmektedir. Bu nedenle sonraki sınıflandırma bulgularında model performansları yorumlanırken lif oranlarının sınırlı fakat model açısından anlamlı bilgi taşıdığı dikkate alınmıştır.

4.3. İki Sınıflı Sınıflandırma Bulguları

Bu bölümde, kumaş hata adedinin iki sınıflı sınıflandırma yapısına dönüştürülmesiyle elde edilen model performans sonuçları sunulmuştur. İki sınıflı yaklaşımda kumaş hata adedi için 75. yüzdilik değer eşik olarak alınmış ve bu eşik 63 olarak hesaplanmıştır. Buna göre hata adedi 63 ve altında olan gözlemler “Normal”, 63’ün üzerinde olan gözlemler ise “Yüksek” hata sınıfı olarak değerlendirilmiştir. Bu yaklaşımın temel amacı, üretim açısından daha riskli kabul edilebilecek yüksek hata grubunu diğer kayıtlardan ayırmaktır.

4.3.1. Model Performans Sonuçları

İki sınıflı sınıflandırma yaklaşımında sınıf dağılımı Çizelge 4.6’da verilmiştir. Veri setindeki 652 gözlemin 490’ı “Normal”, 162’si ise “Yüksek” hata sınıfında yer almıştır. Oransal olarak değerlendirildiğinde, gözlemlerin %75,15’i Normal sınıfta, %24,85’i Yüksek sınıfta bulunmaktadır. Bu dağılım, iki sınıflı yapıda sınıfların dengeli olmadığını ve yüksek hata sınıfının veri setinde daha sınırlı temsil edildiğini göstermektedir.

Çizelge 4.6. İki sınıflı sınıflandırma yaklaşımında sınıf dağılımı

Sınıf	Eşik tanımı	Gözlem sayısı	Oran (%)
Normal	Kumaş hata adedi ≤ 63	490	75,15
Yüksek	Kumaş hata adedi > 63	162	24,85
Toplam	—	652	100,00

İki sınıflı sınıflandırma modellerinin performans sonuçları Çizelge 4.7’de verilmiştir. Modeller accuracy, precision, recall ve F1-score ölçütleri kullanılarak karşılaştırılmıştır. Bulgular, bazı modellerin genel doğruluk bakımından yüksek değerlere ulaşmasına rağmen yüksek hata sınıfını yakalama konusunda yetersiz kaldığını göstermektedir. Bu nedenle iki sınıflı sınıflandırma sonuçlarının yalnızca accuracy üzerinden değil, recall ve F1-score değerleriyle birlikte yorumlanması gerekmektedir.

Çizelge 4.7. İki sınıflı sınıflandırma modellerinin performans sonuçları

Model	Accuracy	Precision	Recall	F1-score
KNN	0,7009	0,1765	0,0556	0,0845
RUSBoost	0,6979	0,4215	0,5802	0,4883
Cost-Sensitive Random Forest	0,6534	0,3667	0,5432	0,4378
Bagged Trees	0,7439	0,4074	0,0679	0,1164
Random Forest	0,7285	0,3171	0,0802	0,1281
Decision Tree	0,7423	0,3846	0,0617	0,1064
Naive Bayes	0,7025	0,0000	0,0000	0,0000
Boosting	0,7515	0,0000	0,0000	0,0000
Subspace KNN	0,7316	0,2593	0,0432	0,0741
SVM-RBF	0,7515	0,0000	0,0000	0,0000
LDA	0,7515	0,0000	0,0000	0,0000
Logistic Regression	0,7515	0,0000	0,0000	0,0000
Linear SVM	0,7515	0,0000	0,0000	0,0000
ANN	0,7515	0,0000	0,0000	0,0000
Subspace Discriminant Analysis	0,7515	0,0000	0,0000	0,0000

İki sınıflı modellere ait confusion matrix sonuçları Çizelge 4.8’de verilmiştir. Bu tabloda TN, Normal sınıfa ait gözlemlerin doğru biçimde Normal olarak tahmin edilmesini; FP, Normal sınıfa ait gözlemlerin yanlış biçimde Yüksek olarak tahmin edilmesini; FN, Yüksek sınıfa ait gözlemlerin yanlış biçimde Normal olarak tahmin edilmesini; TP ise Yüksek sınıfa ait gözlemlerin doğru biçimde Yüksek olarak tahmin edilmesini göstermektedir.

Çizelge 4.8. İki sınıflı modellerin confusion matrix sonuçları

Model	TN	FP	FN	TP
KNN	448	42	153	9
RUSBoost	361	129	68	94
Cost-Sensitive Random Forest	338	152	74	88
Bagged Trees	474	16	151	11
Random Forest	462	28	149	13
Decision Tree	474	16	152	10
Naive Bayes	458	32	162	0
Boosting	490	0	162	0
Subspace KNN	470	20	155	7
SVM-RBF	490	0	162	0
LDA	490	0	162	0
Logistic Regression	490	0	162	0
Linear SVM	490	0	162	0
ANN	490	0	162	0
Subspace Discriminant Analysis	490	0	162	0

Confusion matrix sonuçları, iki sınıflı sınıflandırma probleminde modellerin yüksek hata sınıfını yakalama davranışlarının birbirinden belirgin biçimde ayrıldığını göstermektedir. Naive Bayes, SVM-RBF, LDA, Logistic Regression, Linear SVM, ANN ve Subspace Discriminant Analysis modelleri test setindeki Normal gözlemlerin tamamını doğru tahmin etmiş, ancak Yüksek hata sınıfındaki 49 gözlemin hiçbirini doğru sınıflandıramamıştır. Buna karşılık RUSBoost ve Cost-Sensitive Random Forest modelleri yüksek hata sınıfına daha duyarlı sonuçlar üretmiştir.

4.3.2. Accuracy Sonuçları

LOOCV sonuçlarına göre iki sınıflı sınıflandırma yaklaşımında en yüksek accuracy değeri Boosting, SVM-RBF, LDA, Logistic Regression, Linear SVM, ANN ve Subspace

Discriminant Analysis modellerinde 0,7515 olarak hesaplanmıştır. Bagged Trees modelinde accuracy 0,7439, Decision Tree modelinde 0,7423, Subspace KNN modelinde 0,7316 ve Random Forest modelinde 0,7285 olarak belirlenmiştir. RUSBoost modelinde accuracy 0,6979, KNN modelinde 0,7009 ve Cost-Sensitive Random Forest modelinde 0,6534 düzeyindedir.

Bu sonuçlar ilk bakışta accuracy değeri yüksek olan modellerin daha başarılı olduğu izlenimini verebilir. Ancak confusion matrix sonuçları, bu değerlendirmenin tek başına yeterli olmadığını göstermektedir. Boosting, SVM-RBF, LDA, Logistic Regression, Linear SVM, ANN ve Subspace Discriminant Analysis modelleri 0,7515 accuracy değerine ulaşmış olmasına rağmen Yüksek hata sınıfındaki 162 gözlemin hiçbirini doğru tahmin edememiştir. Bu modellerin yüksek accuracy üretmesinin temel nedeni, çoğunluk sınıfı olan Normal sınıfa yönelmeleridir.

Bu nedenle iki sınıflı sınıflandırma yaklaşımında accuracy sonuçları dikkatli yorumlanmalıdır. Veri setinde Normal sınıfın daha fazla sayıda gözlem içermesi, bazı modellerin çoğunluk sınıfına yönelerek yüksek accuracy üretmesine neden olmuştur. Üretim açısından asıl kritik olan durum, yüksek hata sınıfının doğru biçimde yakalanmasıdır. Bu nedenle accuracy değeri yüksek olsa bile yüksek hata sınıfını tanıyamayan modellerin pratik kalite kontrol açısından sınırlı katkı sunacağı değerlendirilmiştir.

4.3.3. Recall Sonuçları

Recall sonuçları, modellerin Yüksek hata sınıfını ne ölçüde yakalayabildiğini göstermesi bakımından önem taşımaktadır. LOOCV sonuçlarına göre iki sınıflı yaklaşımda en yüksek recall değeri RUSBoost modelinde elde edilmiştir. Bu modelin recall değeri 0,5802 olarak hesaplanmıştır. RUSBoost modeli, Yüksek hata sınıfındaki 162 gözlemin 94'ünü doğru biçimde Yüksek olarak sınıflandırmıştır.

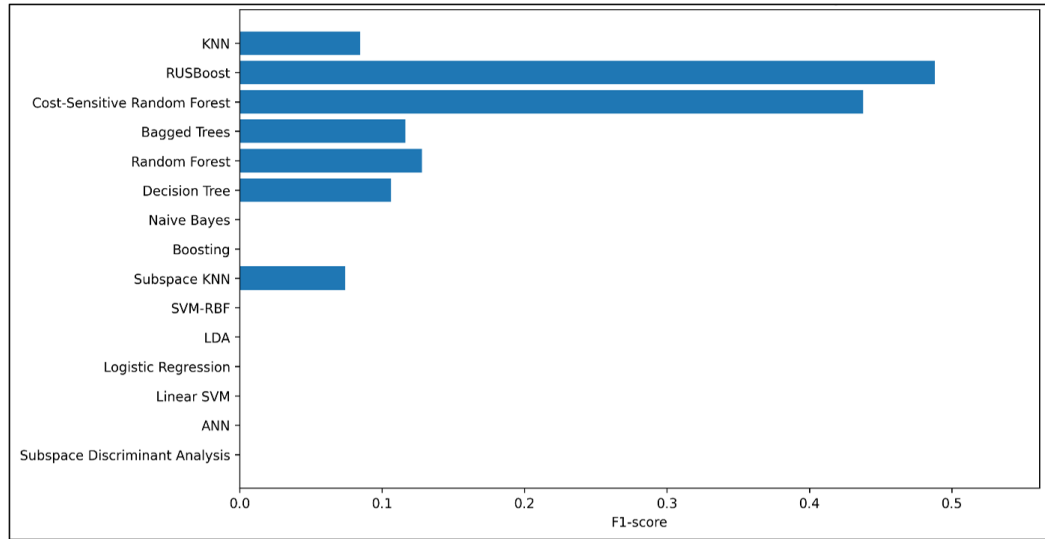
Cost-Sensitive Random Forest modeli recall bakımından ikinci güçlü sonucu üretmiştir. Bu modelde recall değeri 0,5432 olarak hesaplanmış ve Yüksek hata sınıfındaki 162 gözlemin 88'i doğru tahmin edilmiştir. Bu iki model, yüksek hata sınıfını yakalama açısından diğer modellere göre daha etkili sonuçlar vermiştir. Random Forest modelinde recall 0,0802, Bagged Trees modelinde 0,0679, Decision Tree modelinde 0,0617, KNN

modelinde 0,0556 ve Subspace KNN modelinde 0,0432 düzeyinde kalmıştır.

Recall değeri 0 olan modeller, yüksek hata sınıfını hiç yakalayamamıştır. Boosting, SVM-RBF, LDA, Logistic Regression, Linear SVM, ANN ve Subspace Discriminant Analysis modelleri bu grupta yer almaktadır. Bu sonuç, söz konusu modellerin iki sınıflı veri yapısında çoğunluk sınıfı olan Normal sınıfa yöneldiğini göstermektedir. Üretim ortamında yüksek hata sınıfının kaçırılması, hatalı kayıtların normal kabul edilmesine yol açabileceğinden recall değeri bu çalışma açısından kritik bir performans göstergesi olarak değerlendirilmiştir.

4.3.4. F1-Score Sonuçları

F1-score sonuçları, precision ve recall değerlerini birlikte dikkate aldığı için iki sınıflı sınıflandırma performansını daha dengeli biçimde değerlendirmeye olanak sağlamaktadır. Şekil 4.4’de iki sınıflı sınıflandırma modellerine ait F1-score değerleri gösterilmektedir.



Şekil 4.4. İki sınıflı sınıflandırma modellerinin LOOCV F1-score değerleri

LOOCV sonuçlarına göre iki sınıflı sınıflandırma yaklaşımında en yüksek F1-score değeri RUSBoost modelinde elde edilmiştir. RUSBoost modelinin F1-score değeri 0,4883 olarak hesaplanmıştır. Bu modelde accuracy 0,6979, precision 0,4215 ve recall 0,5802 düzeyindedir. RUSBoost modelinin yüksek hata sınıfını yakalama bakımından en güçlü sonucu üretmesi, dengesiz sınıf yapısına sahip iki sınıflı problemde azınlık sınıfının

öğrenilmesi açısından önemli bir bulgudur.

Cost-Sensitive Random Forest modeli ise 0,4378 F1-score değeriyle ikinci sırada yer almıştır. Bu modelde recall 0,5432 olarak hesaplanmış ve yüksek hata sınıfının belirli ölçüde yakalanabildiği görülmüştür. Bununla birlikte RUSBoost modeli hem recall hem de F1-score bakımından daha yüksek değer ürettiği için iki sınıflı yaklaşımda en dengeli model olarak değerlendirilmiştir.

4.3.5. Genel Değerlendirme

İki sınıflı sınıflandırma bulguları, kumaş hata adedinin Normal ve Yüksek hata sınıfları altında belirli ölçüde ayrıştırılabildiğini göstermektedir. Ancak sınıf dağılımındaki dengesizlik model performanslarının yorumlanmasını doğrudan etkilemiştir. Normal sınıfın daha fazla gözlemle temsil edilmesi, bazı modellerin yüksek accuracy değerlerine ulaşmasına rağmen Yüksek hata sınıfını yakalayamamasına neden olmuştur.

LOOCV sonuçlarına göre iki sınıflı yaklaşımda en uygun model RUSBoost olmuştur. RUSBoost modeli F1-score bakımından en yüksek değeri üretmiş ve Yüksek hata sınıfını yakalama açısından diğer modellere göre daha başarılı sonuç vermiştir. Cost-Sensitive Random Forest modeli de yüksek hata sınıfını belirli ölçüde yakalayabilmiş; ancak F1-score bakımından RUSBoost modelinin gerisinde kalmıştır. Bu bulgu, sınıf dengesizliği bulunan kalite kontrol problemlerinde örnekleme destekli boosting yaklaşımlarının etkili olabileceğini göstermektedir.

Genel olarak iki sınıflı yaklaşım, üretim açısından yüksek hata riski taşıyan kayıtların ayrıştırılmasına yönelik pratik bir sınıflandırma yapısı sunmaktadır. Bununla birlikte mevcut veri setinde yalnızca lif oranlarının kullanılması, yüksek hata sınıfının güçlü biçimde ayrıştırılmasını sınırlamıştır. Bu nedenle daha yüksek model başarısı için ilerleyen çalışmalarda lif kompozisyonuna ek olarak hata türleri, üretim parametreleri, makine ayarları, operatör bilgisi, vardiya ve gerçek proses koşullarına ilişkin değişkenlerin de modele dâhil edilmesi gerekmektedir.

4.4. Üç Sınıflı Sınıflandırma Bulguları

Bu bölümde, kumaş hata adedinin üç sınıflı sınıflandırma yapısına dönüştürülmesiyle

elde edilen model performans sonuçları sunulmuştur. Üç sınıflı yaklaşımda kumaş hata adedi 33. ve 66. yüzdellik değerlere göre sınıflandırılmıştır. Buna göre hata adedi 18 ve altında olan gözlemler “Düşük”, 19–48 arasında olan gözlemler “Orta”, 48’in üzerinde olan gözlemler ise “Yüksek” hata sınıfı olarak değerlendirilmiştir. Bu yaklaşım, iki sınıflı yapıya göre hata düzeylerini daha kademeli biçimde incelemeye olanak sağlamaktadır.

4.4.1. Model Performans Sonuçları

Üç sınıflı sınıflandırma yaklaşımında sınıf dağılımı Çizelge 4.9’ da verilmiştir. Veri setindeki 652 gözlemin 216’sı “Düşük”, 215’i “Orta” ve 221’i “Yüksek” hata sınıfında yer almıştır. Sınıf oranları sırasıyla %33,13, %32,98 ve %33,90 olarak hesaplanmıştır. Bu dağılım, üç sınıflı yaklaşımda sınıfların birbirine oldukça yakın sayıda temsil edildiğini göstermektedir.

Çizelge 4.9. Üç sınıflı sınıflandırma yaklaşımında sınıf dağılımı

Sınıf	Eşik tanımı	Gözlem sayısı	Oran (%)
Düşük	Kumaş hata adedi ≤ 18	216	33,13
Orta	$19 \leq$ Kumaş hata adedi ≤ 48	215	32,98
Yüksek	Kumaş hata adedi > 48	221	33,90
Toplam	—	652	100,00

Üç sınıflı sınıflandırma modellerinin performans sonuçları Çizelge 4.10’da verilmiştir. Modeller accuracy, precision, recall ve F1-score ölçütleri kullanılarak karşılaştırılmıştır. Üç sınıflı yapıda sınıf dağılımı dengeli olmasına rağmen, model performanslarının genel olarak orta-düşük düzeyde kaldığı görülmektedir. Bu durum, modelin iki sınıf yerine üç farklı hata düzeyi arasında karar vermesinden ve lif oranlarının hata düzeylerini tek başına güçlü biçimde ayırmakta sınırlı kalmasından kaynaklanmaktadır.

Çizelge 4.10. Üç sınıflı sınıflandırma modellerinin performans sonuçları

Model	Accuracy	Precision	Recall	F1-score
KNN	0,3589	0,3586	0,3600	0,3479
RUSBoost	0,4586	0,4660	0,4584	0,4589
Cost-Sensitive Random Forest	0,4095	0,3972	0,4079	0,3961
Bagged Trees	0,4064	0,3992	0,4053	0,3968
Random Forest	0,4095	0,3961	0,4079	0,3956
Decision Tree	0,4095	0,3995	0,4082	0,3973
Naive Bayes	0,3742	0,3795	0,3709	0,3258
Boosting	0,3911	0,3956	0,3908	0,3795
Subspace KNN	0,4202	0,4197	0,4198	0,4195
SVM-RBF	0,4310	0,4337	0,4312	0,4163
LDA	0,4095	0,4202	0,4080	0,3992
Logistic Regression	0,4049	0,4135	0,4034	0,3944
Linear SVM	0,4049	0,4135	0,4034	0,3944
ANN	0,4034	0,4001	0,4033	0,3963
Subspace Discriminant Analysis	0,3558	0,3514	0,3538	0,3335

Çizelge 4.10 incelendiğinde, üç sınıflı sınıflandırma yaklaşımında en yüksek F1-score değerinin RUSBoost modelinde elde edildiği görülmektedir. RUSBoost modelinde accuracy 0,4586, precision 0,4660, recall 0,4584 ve F1-score 0,4589 olarak hesaplanmıştır. Bu modeli F1-score bakımından Subspace KNN, SVM-RBF ve LDA modelleri izlemiştir. Üç sınıflı yapıda en düşük F1-score değeri ise Naive Bayes ve Subspace Discriminant Analysis modellerinde elde edilmiştir.

Üç sınıflı modellere ait confusion matrix sonuçları, model bazında sınıfların nasıl tahmin edildiğini göstermektedir. RUSBoost modelinde Düşük sınıftaki 216 gözlemin 106'sı, Orta sınıftaki 215 gözlemin 89'u ve Yüksek sınıftaki 221 gözlemin 104'ü doğru tahmin edilmiştir. Bu sonuç, RUSBoost modelinin üç hata düzeyinde de tamamen baskın bir sınıfa

yönelmeden daha dengeli bir tahmin yapısı ürettiğini göstermektedir.

4.4.2. Accuracy Sonuçları

Üç sınıflı sınıflandırma modellerinin accuracy sonuçları incelendiğinde, en yüksek accuracy değerinin RUSBoost modelinde elde edildiği görülmektedir. RUSBoost modelinin accuracy değeri 0,4586'dır. Bu modeli 0,4310 accuracy değeriyle SVM-RBF, 0,4202 değeriyle Subspace KNN, 0,4095 değerleriyle Cost-Sensitive Random Forest, Random Forest ve Decision Tree modelleri izlemiştir. Bagged Trees modelinde accuracy 0,4064, Logistic Regression ve Linear SVM modellerinde 0,4049, ANN modelinde ise 0,4034 olarak hesaplanmıştır.

Üç sınıflı yapıda accuracy değerlerinin iki sınıflı sınıflandırma sonuçlarına göre daha düşük düzeyde kalması beklenen bir durumdur. Çünkü iki sınıflı yaklaşımda model yalnızca Normal ve Yüksek hata ayrımı yaparken, üç sınıflı yapıda Düşük, Orta ve Yüksek hata düzeyleri arasında ayırım yapmak durumundadır. Sınıf sayısının artması, karar sınırlarını daha karmaşık hale getirmektedir.

Accuracy bakımından RUSBoost modelinin öne çıkması, bu modelin üç sınıflı yapıda sınıflar arasında daha dengeli bir tahmin yapısı üretebildiğini göstermektedir. Bununla birlikte en yüksek accuracy değerinin 0,4586 düzeyinde kalması, lif oranlarıyla üç hata düzeyinin güçlü biçimde ayrıştırılamadığını ortaya koymaktadır.

4.4.3. Recall Sonuçları

Recall sonuçları, üç sınıflı yapıda modellerin sınıfları genel olarak ne ölçüde yakalayabildiğini göstermektedir. LOOCV sonuçlarına göre en yüksek recall değeri RUSBoost modelinde elde edilmiştir. RUSBoost modelinin recall değeri 0,4584 olarak hesaplanmıştır. Bu modeli 0,4312 recall değeriyle SVM-RBF, 0,4198 değeriyle Subspace KNN, 0,4082 değeriyle Decision Tree ve 0,4079 değerleriyle Cost-Sensitive Random Forest ile Random Forest modelleri izlemiştir.

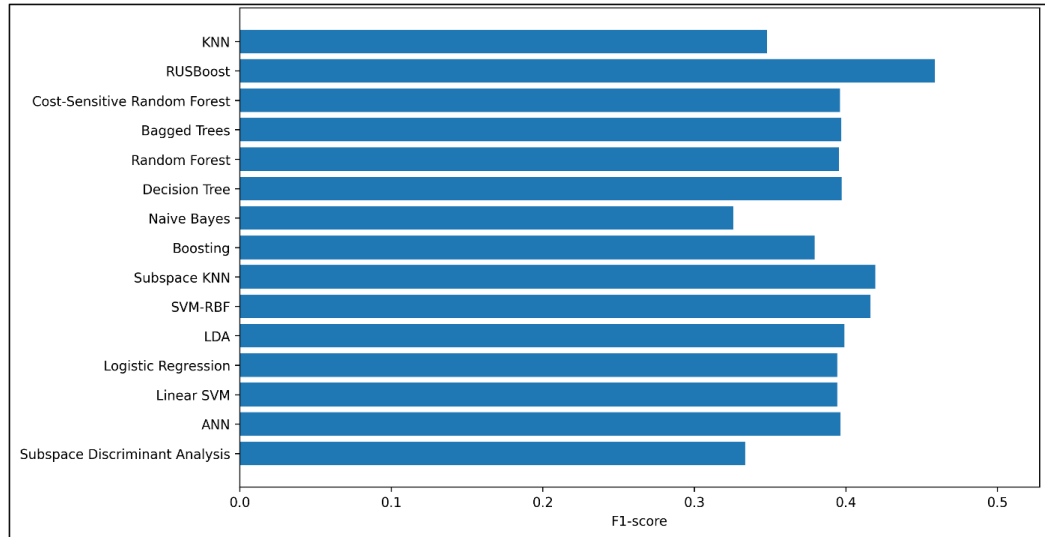
RUSBoost modelinin üç sınıflı yapıda en yüksek recall değerini üretmesi, bu modelin Düşük, Orta ve Yüksek sınıflar arasında daha dengeli bir yakalama performansı sağladığını göstermektedir. Buna karşılık tüm modellerde recall değerlerinin orta düzeyde kalması, hata

düzeylerinin yalnızca lif kompozisyonu değişkenleriyle güçlü biçimde ayrıştırılamadığını ortaya koymaktadır.

Confusion matrix sonuçları sınıf düzeyinde incelendiğinde, RUSBoost modelinde Düşük sınıftaki 216 gözlemin 106'sı, Orta sınıftaki 215 gözlemin 89'u ve Yüksek sınıftaki 221 gözlemin 104'ü doğru tahmin edilmiştir. Bu sonuç, modelin tek bir sınıfa aşırı biçimde yönelmeden üç sınıf için de belirli düzeyde tahmin başarısı ürettiğini göstermektedir.

4.4.4. F1-Score Sonuçları

Üç sınıflı sınıflandırma yaklaşımında F1-score, modellerin sınıflar arasındaki dengeyi ne ölçüde koruyabildiğini değerlendirmek için kullanılmıştır. Bu ölçüt, precision ve recall değerlerini birlikte dikkate aldığı için özellikle sınıflar arası karışmanın bulunduğu durumlarda daha bütüncül bir performans görünümü sunmaktadır. Şekil 4.5'de üç sınıflı sınıflandırma modellerine ait F1-score değerleri gösterilmektedir.



Şekil 4.5. Üç sınıflı sınıflandırma modellerinin LOOCV F1-score değerleri

Üç sınıflı sınıflandırma yaklaşımında en yüksek F1-score değeri RUSBoost modelinde elde edilmiştir. RUSBoost modelinin F1-score değeri 0,4589'dur. Bu modeli 0,4195 F1-score değeriyle Subspace KNN, 0,4163 değeriyle SVM-RBF ve 0,3992 değeriyle LDA modelleri izlemiştir. Random Forest modelinin F1-score değeri ise 0,3956 olarak hesaplanmıştır.

F1-score bakımından RUSBoost modelinin öne çıkması, bu modelin üç sınıflı hata

düzeyi ayırımında sınıflar arasında daha dengeli bir performans üretebildiğini göstermektedir. Bununla birlikte en yüksek F1-score değerinin 0,4589 düzeyinde kalması, lif oranlarının tek başına üç sınıflı hata ayırımında güçlü bir sınıflandırma zemini oluşturmadığını ortaya koymaktadır.

4.4.5. Genel Değerlendirme

Üç sınıflı sınıflandırma bulguları, kumaş hata adedinin Düşük, Orta ve Yüksek hata düzeyleri altında sınıflandırılmasının iki sınıflı yaklaşıma göre daha zor bir problem oluşturduğunu göstermektedir. Sınıf dağılımı dengeli olmasına rağmen model performanslarının orta düzeyde kalması, üç hata düzeyinin yalnızca lif oranları üzerinden net biçimde ayrıştırılamadığını ortaya koymaktadır.

Bu bölümde elde edilen LOOCV sonuçlarına göre, üç sınıflı sınıflandırma yaklaşımında en başarılı model RUSBoost olmuştur. RUSBoost modeli accuracy, precision, recall ve F1-score bakımından en yüksek değerleri üretmiştir. Subspace KNN ve SVM-RBF modelleri de F1-score bakımından RUSBoost modelini izlemiştir; Random Forest modeli ise önceki hold-out sonuçlarına kıyasla daha sınırlı bir performans göstermiştir. Bu durum, model değerlendirme yönteminin sonuçların yorumunu doğrudan etkileyebildiğini göstermektedir.

Üç sınıflı yaklaşım, üretim açısından hata düzeylerinin daha ayrıntılı izlenmesine olanak sağlamaktadır. Düşük, Orta ve Yüksek sınıfların ayrı ayrı değerlendirilmesi, kalite riskinin kademeli biçimde ele alınmasını mümkün kılmaktadır. Ancak mevcut veri setinde yalnızca lif kompozisyonu değişkenlerinin bulunması, modelin ayrıştırma gücünü sınırlamıştır. Özellikle Orta hata sınıfının ayrıştırılmasında görülen zorluk, hata düzeylerinin yalnızca kumaş bileşimiyle değil, üretim sürecine ait ek değişkenlerle birlikte değerlendirilmesi gerektiğini göstermektedir.

4.5. İki ve Üç Sınıflı Yaklaşımların Karşılaştırılması

Bu bölümde, iki sınıflı ve üç sınıflı sınıflandırma yaklaşımlarından elde edilen en iyi model sonuçları birlikte değerlendirilmiştir. İki sınıflı yaklaşımda temel amaç, kumaş hata adedini “Normal” ve “Yüksek” olmak üzere iki sınıfa ayırarak üretim açısından riskli kayıtları belirlemektir. Üç sınıflı yaklaşımda ise hata düzeyleri “Düşük”, “Orta” ve “Yüksek”

olarak ayrılmış, böylece hata yapısının daha kademeli biçimde incelenmesi amaçlanmıştır.

İki yaklaşımın en yüksek performans değerleri Çizelge 4.11’de verilmiştir. LOOCV sonuçlarına göre hem iki sınıflı hem de üç sınıflı sınıflandırma yaklaşımında F1-score bakımından en başarılı model RUSBoost olmuştur. İki sınıflı yaklaşımda RUSBoost modelinde accuracy 0,6979, recall 0,5802 ve F1-score 0,4883 olarak hesaplanmıştır. Üç sınıflı yaklaşımda ise aynı modelde accuracy 0,4586, recall 0,4584 ve F1-score 0,4589 olarak belirlenmiştir.

Çizelge 4.11. İki ve üç sınıflı sınıflandırma yaklaşımlarının karşılaştırılması

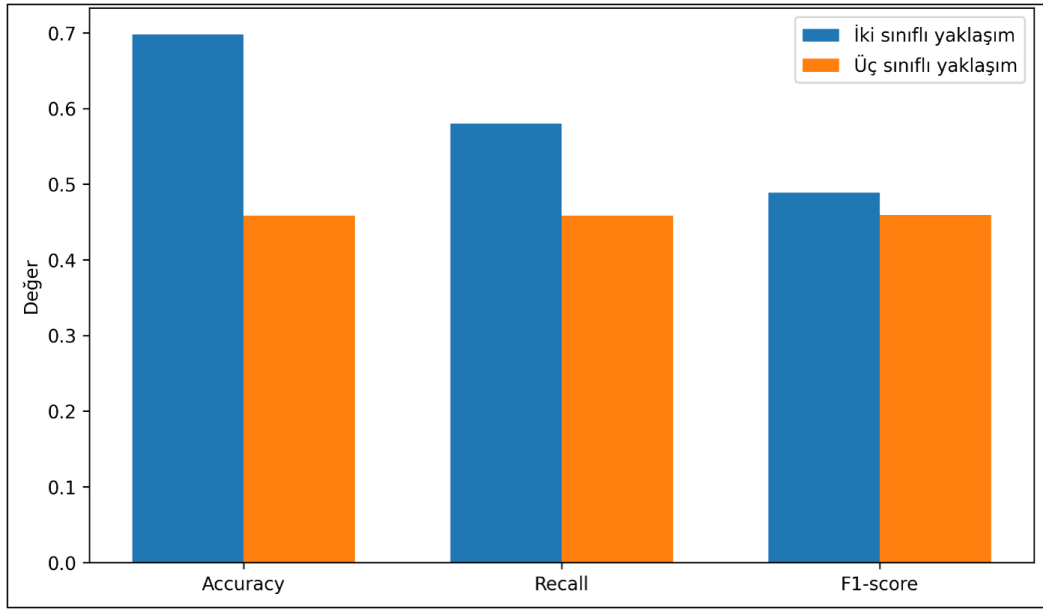
Sınıflandırma yaklaşımı	En başarılı model	Accuracy	Recall	F1-score
İki sınıflı yaklaşım	RUSBoost	0,6979	0,5802	0,4883
Üç sınıflı yaklaşım	RUSBoost	0,4586	0,4584	0,4589

Çizelge 4.11 incelendiğinde, iki sınıflı yaklaşımın accuracy ve recall bakımından üç sınıflı yaklaşıma göre daha yüksek değerler ürettiği görülmektedir. Bu durum beklenen bir sonuçtur; çünkü iki sınıflı yapıda model yalnızca Normal ve Yüksek hata sınıfları arasında ayırım yapmaktadır. Üç sınıflı yapıda ise modelin Düşük, Orta ve Yüksek hata düzeylerini ayrı ayrı tahmin etmesi gerekmektedir. Sınıf sayısının artması, karar sınırlarını daha karmaşık hale getirmiş ve model performanslarını sınırlamıştır.

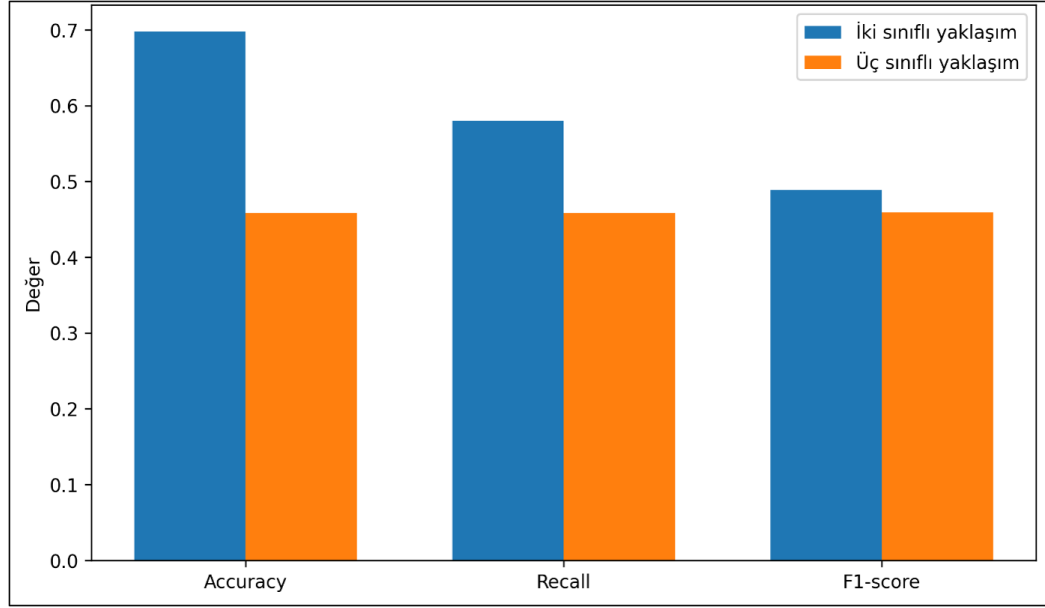
F1-score açısından iki yaklaşım arasındaki fark daha sınırlıdır. İki sınıflı yaklaşımda RUSBoost modelinin F1-score değeri 0,4883 iken, üç sınıflı yaklaşımda RUSBoost modelinin F1-score değeri 0,4589 olarak hesaplanmıştır. Bu sonuç, iki sınıflı yaklaşımın yüksek hata riskini ayırmada daha pratik bir yapı sunduğunu; üç sınıflı yaklaşımın ise daha ayrıntılı fakat daha zor bir sınıflandırma problemi oluşturduğunu göstermektedir.

Şekil 4.6’da iki ve üç sınıflı yaklaşımlarda en başarılı modellerin accuracy, recall ve F1-score değerleri karşılaştırmalı olarak gösterilmektedir. Grafik incelendiğinde, iki sınıflı yaklaşımın özellikle accuracy ve recall bakımından daha yüksek değerlere sahip olduğu görülmektedir. Üç sınıflı yaklaşımda ise performans değerleri birbirine daha yakın ve daha dengeli bir dağılım göstermiştir.

İki sınıflı yaklaşım üretim açısından daha doğrudan yorumlanabilir bir karar yapısı sunmaktadır. Bu yaklaşımda temel amaç, yüksek hata riski taşıyan kayıtların belirlenmesidir. Bu nedenle özellikle kalite kontrol önceliklerinin oluşturulması, riskli kumaş kompozisyonlarının ayrıştırılması ve daha dikkatli kontrol gerektiren kayıtların belirlenmesi açısından iki sınıflı yapı daha uygulanabilir görünmektedir. RUSBoost modelinin bu yapıda öne çıkması, yüksek hata sınıfının üretim açısından kritik kabul edildiği durumlarda örnekleme destekli boosting yaklaşımlarının yararlı olabileceğini göstermektedir.



Şekil 4.6. İki ve üç sınıflı yaklaşımların LOOCV en iyi performans değerleri



Şekil 4.7. İki ve üç sınıflı yaklaşımların LOOCV en iyi performans değerleri

Üç sınıflı yaklaşım ise kalite düzeylerinin daha ayrıntılı izlenmesine olanak sağlamaktadır. Düşük, Orta ve Yüksek hata düzeylerinin ayrı ayrı değerlendirilmesi, üretim sürecinde hata yoğunluğunun kademeli biçimde yorumlanmasına katkı sunabilir. Ancak bu yaklaşımda özellikle Orta sınıfın ayrıştırılması daha zor olmuştur. Orta hata düzeyi, Düşük ve Yüksek sınıflar arasında geçiş niteliği taşıdığından modeller açısından daha belirsiz bir sınıf yapısı oluşturmuştur. Bu nedenle üç sınıflı yaklaşım üretim sürecinin ayrıntılı izlenmesi için yararlı olsa da mevcut veri kapsamıyla güçlü sınıflandırma başarısı üretmekte sınırlı kalmıştır.

İki yaklaşım birlikte değerlendirildiğinde, çalışmada kullanılan lif kompozisyonu değişkenlerinin hata düzeyleri hakkında belirli ölçüde bilgi taşıdığı, ancak tek başına yüksek sınıflandırma başarısı için yeterli olmadığı görülmektedir. İki sınıflı yaklaşımda yüksek hata sınıfının belirli ölçüde ayrıştırılabilmesi, bu yapının üretim karar desteği açısından daha uygun olduğunu göstermektedir. Üç sınıflı yaklaşım ise hata düzeylerini daha ayrıntılı ele almakla birlikte, modelin sınıflar arasındaki sınırları daha zor öğrenmesine neden olmuştur.

İki sınıflı yaklaşım yüksek hata riskinin belirlenmesi amacıyla daha pratik ve uygulanabilir bir yapı sunarken, üç sınıflı yaklaşım hata düzeylerinin kademeli izlenmesi açısından daha açıklayıcıdır. Bu çalışma kapsamında üretim karar desteği açısından iki sınıflı yaklaşımın, kalite düzeylerinin daha ayrıntılı değerlendirilmesi açısından ise üç

sınıflı yaklaşımın tamamlayıcı bir değer taşıdığı değerlendirilmiştir. Her iki yaklaşımda da model performansının daha güçlü hale getirilebilmesi için lif kompozisyonuna ek olarak üretim parametreleri, hata türleri, makine ayarları, operatör bilgisi, vardiya ve gerçek proses koşullarına ilişkin değişkenlerin veri setine eklenmesi gerekmektedir.

4.6. Karbon Ayak İzi Bulguları

Bu bölümde, kumaş hata adedinden kaynaklanan ek rework süresi, bu sürenin neden olduğu ek elektrik tüketimi ve buna bağlı karbon ayak izi hesaplama sonuçları sunulmuştur. Hesaplama, ürünün tüm yaşam döngüsünü kapsayan ayrıntılı bir yaşam döngüsü analizi niteliğinde değildir. Bu bölümde yalnızca kumaş hataları nedeniyle ortaya çıkan yeniden işlem faaliyetlerinin elektrik tüketimine bağlı karbon etkisi değerlendirilmiştir.

Karbon ayak izi hesabında kullanılan temel değerler Çizelge 4.12’de verilmiştir. Veri setinde toplam kumaş hata adedi 31.943 olarak hesaplanmıştır. Hata başına ortalama rework süresi 1,05 dakika olarak kabul edilmiştir. Rework işlemlerinde kullanılan ortalama güç tüketimi 10 kW, elektrik tüketimine bağlı emisyon faktörü ise 0,4183 kg CO₂e/kWh olarak alınmıştır.

Çizelge 4.12. Karbon ayak izi hesabında kullanılan temel değerler

Parametre	Değer	Birim
Toplam hata adedi	31.943	adet
Hata başına ek rework süresi	1,05	dakika/hata
Ortalama güç tüketimi	10	kW
Elektrik emisyon faktörü	0,4183	kg CO ₂ e/kWh

4.6.1. Toplam Hata Kaynaklı Rework Süresi

Veri setinde yer alan tüm gözlemlere ait kumaş hata adetleri toplanmış ve toplam hata adedi 31.943 adet olarak hesaplanmıştır. Hata başına ortalama ek rework süresi 1,05 dakika olarak alınmıştır. Bu iki değer kullanılarak toplam hata kaynaklı rework süresi dakika ve saat cinsinden hesaplanmıştır.

Hesaplama aşağıdaki şekilde yapılmıştır:

· Toplam rework süresi = Toplam hata adedi × Hata başına ek süre

· Toplam rework süresi = $31.943 \times 1,05 = 33.540,15$ dakika

Dakika cinsinden hesaplanan toplam süre saate çevrildiğinde:

· Toplam rework süresi = $33.540,15 / 60 = 559,00$ saat

Bu hesaplama göre kumaş hataları, çalışma kapsamındaki veri setinde yaklaşık 559 saat ek rework süresi oluşturmuştur.

Çizelge 4.13. Toplam hata kaynaklı rework süresi hesabı

Hesaplama adımı	Değer
Toplam hata adedi	31.943 adet
Hata başına ek rework süresi	1,05 dakika
Toplam rework süresi	33.540,15 dakika
Toplam rework süresi	559,00 saat

4.6.2. Ek Elektrik Tüketimi Bulguları

Ortalama güç tüketimi 10 kW olarak alınmıştır. Ek elektrik tüketimi, toplam rework süresi ile ortalama güç tüketiminin çarpılmasıyla hesaplanmıştır.

Hesaplama aşağıdaki şekilde yapılmıştır:

· Ek elektrik tüketimi = Toplam rework süresi × Ortalama güç tüketimi

· Ek elektrik tüketimi = $559,00 \times 10 = 5.590,03$ kWh

Hesaplama sonucunda 559 saatlik rework süresinin 5.590 kWh ek elektrik tüketimine karşılık geldiği belirlenmiştir. Bu değer, kumaş hatalarından kaynaklanan yeniden işlem faaliyetlerinin enerji tüketimi üzerinde doğrudan ölçülebilir bir yük oluşturduğunu göstermektedir.

Çizelge 4.14. Hata kaynaklı ek elektrik tüketimi hesabı

Hesaplama adımı	Değer
Toplam rework süresi	559,00 saat
Ortalama güç tüketimi	10 kW
Ek elektrik tüketimi	5.590,03 kWh
Yuvarlatılmış ek elektrik tüketimi	5.590 kWh

4.6.3. Karbon Ayak İzi Hesaplama Bulguları

Ek elektrik tüketimi hesaplandıktan sonra, bu değer elektrik emisyon faktörü ile çarpılarak CO₂e cinsinden karbon ayak izi belirlenmiştir. Hesaplama ek elektrik tüketimine bağlı emisyon faktörü 0,4183 kg CO₂e/kWh olarak alınmıştır.

Hesaplama aşağıdaki şekilde yapılmıştır:

· Ek karbon ayak izi = Ek elektrik tüketimi × Elektrik emisyon faktörü

· Ek karbon ayak izi = 5.590,03 × 0,4183 = 2.338,31 kg CO₂e

Kilogram cinsinden hesaplanan değer ton cinsine çevrildiğinde:

· Ek karbon ayak izi = 2.338,31 / 1000 = 2,34 ton CO₂e

Bu hesaplama göre kumaş hatalarından kaynaklanan ek rework faaliyetlerinin karbon ayak izi yaklaşık 2,34 ton CO₂e/yıl olarak hesaplanmıştır.

Çizelge 4.15. Hata kaynaklı karbon ayak izi hesabı

Hesaplama adımı	Değer
Ek elektrik tüketimi	5.590,03 kWh
Elektrik emisyon faktörü	0,4183 kg CO e/kWh
Ek karbon ayak izi	2.338,31 kg CO e
Ek karbon ayak izi	2,34 ton CO e/yıl

4.6.4. Sürdürülebilirlik Açısından Değerlendirme

Kumaş hata adedine bağlı rework süresi, ek elektrik tüketimi ve karbon ayak izi

sonuçları Çizelge 4.16’da özetlenmiştir. Buna göre 31.943 kumaş hatası, 559 saat ek rework süresi, 5.590 kWh ek elektrik tüketimi ve yaklaşık 2,34 ton CO₂e/yıl düzeyinde karbon etkisi oluşturmuştur.

Çizelge 4.16. Hata kaynaklı rework, elektrik tüketimi ve karbon ayak izi sonuçları

Gösterge	Sonuç
Toplam hata adedi	31.943 adet
Toplam rework süresi	559,00 saat
Ek elektrik tüketimi	5.590,03 kWh
Ek karbon ayak izi	2.338,31 kg CO ₂ e
Ek karbon ayak izi	2,34 ton CO ₂ e/yıl

Bu bulgular, kumaş hatalarının üretim sürecinde yalnızca kalite ve verimlilik sorunu olarak değil, aynı zamanda kaynak kullanımı ve çevresel etki açısından da değerlendirilmesi gerektiğini göstermektedir. Rework işlemleri, hatalı ürünün tamamen fireye ayrılmasını önleyebilir; ancak aynı zamanda ek zaman, iş gücü, enerji tüketimi ve karbon emisyonu oluşturmaktadır. Bu nedenle sürdürülebilir üretim açısından temel hedef, rework ihtiyacını artırmak değil, hata oluşumunu mümkün olduğunca kaynağında azaltmaktır.

Çalışmada hesaplanan karbon ayak izi değeri sınırlı kapsamlı bir göstergedir. Hesaplama yalnızca hata kaynaklı rework faaliyetinin elektrik tüketimine bağlı etkisini içermektedir. Hammadde üretimi, kumaş üretimi, yıkama, terbiye, lojistik, kullanım ve yaşam sonu aşamaları bu hesaplama dâhil edilmemiştir. Bu nedenle elde edilen 2,34 ton CO₂e/yıl değeri, ürünün toplam çevresel etkisi olarak değil, kalite hatalarından kaynaklanan ek işlem yükünün karbon karşılığı olarak değerlendirilmelidir.

Genel olarak değerlendirildiğinde, kalite hatalarının azaltılması hem üretim verimliliği hem de sürdürülebilirlik açısından önem taşımaktadır. Hata sayısının düşürülmesi, rework süresini azaltarak ek elektrik tüketimini ve buna bağlı karbon emisyonunu sınırlayabilir. Bu yönüyle kumaş hata kontrolü, yalnızca kalite güvence sürecinin değil, aynı zamanda enerji ve karbon yönetiminin de bir parçası olarak ele alınmalıdır.

5. TARTIŞMA

5.1. Makine Öğrenmesi Modellerinin Değerlendirilmesi

Bu çalışmada kumaş hata adedinin iki ve üç sınıflı yapılar üzerinden modellenmesi, üretim verilerine dayalı kalite tahmininin farklı karar düzeylerinde nasıl sonuç verdiğini göstermesi bakımından önem taşımaktadır. Makine öğrenmesi modelleri burada yalnızca genel sınıflandırma başarısını görmek için değil, aynı zamanda hata düzeylerinin hangi koşullarda daha iyi ayrıştırılabildiğini değerlendirmek amacıyla kullanılmıştır. Üretim ortamlarında kalite çıktılarının tahmin edilmesi, çoğu zaman doğrudan ölçülemeyen veya geç fark edilen kalite problemlerinin veri üzerinden erken tanınmasına imkân vermektedir. Tercan ve Meisen, (2022), kestirimci kalite çalışmalarında makine öğrenmesi ve derin öğrenme yöntemlerinin üretim süreçlerinden elde edilen verilerle kalite çıktılarının tahmininde yaygın biçimde kullanıldığını belirtmektedir. Bu çerçevede mevcut çalışma, görüntü tabanlı kumaş hata tespitinden farklı olarak lif kompozisyonu ve hata adedine dayalı tabular üretim verisi üzerinden hata düzeyi sınıflandırmasına odaklanmıştır.

Model performansları leave-one-out çapraz doğrulama yaklaşımıyla değerlendirilmiştir. Bu tercih, veri setinin görece sınırlı büyüklüğü dikkate alındığında model başarısının tek bir rastgele eğitim-test ayırımına bağlı kalmadan incelenmesine olanak sağlamıştır. LOOCV sürecinde her gözlem bir kez test gözlemi olarak kullanılmış ve performans ölçütleri tüm gözlemler üzerinden elde edilen tahminlere göre hesaplanmıştır. Böylece model karşılaştırmaları, veri setinin tamamını değerlendirme sürecine dâhil eden daha kapsamlı bir doğrulama yapısı üzerinden yapılmıştır.

İki sınıflı sınıflandırma sonuçları, model performansının yalnızca accuracy üzerinden değerlendirilmesinin yeterli olmadığını açık biçimde göstermektedir. İki sınıflı yapıda “Normal” sınıf 490 gözlemle veri setinin büyük bölümünü oluştururken, “Yüksek” sınıf 162 gözlemle daha sınırlı temsil edilmiştir. Bu sınıf yapısı nedeniyle bazı modeller genel doğruluk bakımından yüksek değerlere ulaşmış, ancak yüksek hata sınıfını yakalama konusunda zayıf kalmıştır. Özellikle Boosting, SVM-RBF, LDA, Logistic Regression, Linear SVM, ANN ve Subspace Discriminant Analysis modellerinde accuracy 0,7515 olmasına rağmen recall ve F1-score değerlerinin 0 olarak hesaplanması, bu modellerin Yüksek hata sınıfını doğru

biçimde yakalayamadığını göstermektedir. Bu durum, He ve Garcia, (2009) dengesiz veri setlerinde geleneksel sınıflandırıcıların çoğunluk sınıfına yönelme eğilimi gösterebileceği yönündeki değerlendirmesiyle uyumludur.

İki sınıflı yapı içinde RUSBoost modelinin F1-score bakımından en yüksek değeri üretmesi dikkat çekicidir. Bu modelde accuracy 0,6979, precision 0,4215, recall 0,5802 ve F1-score 0,4883 olarak hesaplanmıştır. Cost-Sensitive Random Forest modeli ise 0,4378 F1-score değeriyle ikinci sırada yer almıştır. Bu sonuç, yüksek hata sınıfının yakalanmasının önemli olduğu dengesiz kalite kontrol problemlerinde örnekleme destekli boosting yaklaşımlarının etkili olabileceğini göstermektedir. Seiffert vd., (2010), RUSBoost yaklaşımının rastgele alt örnekleme ile boosting mantığını birleştirerek dengesiz veri setlerinde azınlık sınıfının öğrenilmesini desteklediğini belirtmektedir. Mevcut bulgular da bu genel çerçeveye uyumludur.

Bu noktada iki sınıflı sonuçların en kritik bulgusu, yüksek accuracy değerinin pratik kalite kontrol başarısı anlamına gelmemesidir. Hata sınıflandırması gibi üretim açısından risk taşıyan problemlerde, modelin yalnızca çoğunluk sınıfını doğru tahmin etmesi yeterli değildir. Yüksek hata sınıfının yanlış biçimde “Normal” olarak sınıflandırılması, modelin sahadaki karar desteği değerini zayıflatabilir. Chawla vd., (2002), dengesiz veri setlerinde azınlık sınıfının çoğu zaman daha önemli veya daha yüksek maliyetli sınıf olabileceğini vurgulamaktadır. Bu çalışmada da “Yüksek” sınıf, gözlem sayısı bakımından azınlıkta olmakla birlikte üretim riski bakımından kritik sınıf niteliğindedir. Bu nedenle iki sınıflı model sonuçlarında recall ve F1-score değerleri, accuracy’ye göre daha açıklayıcı ölçütler olarak öne çıkmıştır.

Üç sınıflı sınıflandırma sonuçları iki sınıflı yapıdan farklı bir görünüm sunmuştur. Bu yaklaşımda sınıflar “Düşük”, “Orta” ve “Yüksek” olarak ayrılmış; sınıf dağılımı 216, 215 ve 221 gözlemle oldukça dengeli gerçekleşmiştir. Buna rağmen üç sınıflı modellerde accuracy değerleri iki sınıflı modellere göre daha düşük düzeyde kalmıştır. Bu sonuç, üç sınıflı yapının daha ayrıntılı fakat daha zor bir sınıflandırma problemi oluşturmasıyla ilişkilidir. İki sınıflı yapıda model yalnızca “Normal” ve “Yüksek” ayrımını öğrenirken, üç sınıflı yapıda düşük, orta ve yüksek hata düzeyleri arasında karar vermek durumundadır. Bu nedenle üç sınıflı

sonuçların iki sınıflı sonuçlarla doğrudan aynı başarı düzeyinde olmasını beklemek uygun değildir.

Üç sınıflı yapı içinde RUSBoost modeli accuracy, precision, recall ve F1-score bakımından en yüksek değerleri üretmiştir. RUSBoost modelinde accuracy 0,4586, precision 0,4660, recall 0,4584 ve F1-score 0,4589 olarak hesaplanmıştır. Bu modeli F1-score bakımından Subspace KNN, SVM-RBF ve LDA modelleri izlemiştir. Random Forest modeli ise LOOCV sonuçlarında 0,3956 F1-score değeri üretmiş ve önceki hold-out değerlendirmeye kıyasla daha sınırlı bir konumda kalmıştır. Bu durum, model performansının yalnızca algoritma türüne değil, kullanılan doğrulama yaklaşımına da duyarlı olduğunu göstermektedir.

Üç sınıflı modellerde Orta hata sınıfının ayrıştırılması daha zor görünmektedir. Confusion matrix sonuçlarında Orta sınıfa ait gözlemlerin hem Düşük hem de Yüksek sınıfa kayabildiği görülmüştür. Bu durum, Orta hata düzeyinin model açısından geçiş niteliği taşıdığını ve iki uç sınıfa göre daha belirsiz bir sınıf yapısı oluşturduğunu göstermektedir. Üretim verilerinde kalite sınıfları çoğu zaman keskin sınırlarla ayrılmamaktadır. Hata düzeyi düşükten yükseğe doğru kademeli bir dağılım gösterdiğinde, özellikle orta sınıfın makine öğrenmesi modelleri tarafından net biçimde ayrıştırılması güçleşebilir. Tercan ve Meisen, (2022), kestirimci kalite çalışmalarında kullanılan veri yapılarının, model seçimini ve performansı doğrudan etkilediğini; model başarısının yalnızca algoritma türüne değil, veri kapsamı ve hedef değişkenin tanımlanma biçimine de bağlı olduğunu vurgulamaktadır.

İki ve üç sınıflı yaklaşımlar birlikte değerlendirildiğinde, bu iki yapının aynı amaca hizmet etmediği görülmektedir. İki sınıflı yaklaşım, yüksek hata riskinin ayrıştırılması açısından daha pratik bir karar yapısı sunmaktadır. Buna karşılık üç sınıflı yaklaşım, hata düzeylerini daha kademeli biçimde izlemeye olanak sağlamaktadır. LOOCV sonuçlarına göre her iki sınıflandırma yaklaşımında da RUSBoost modelinin öne çıkması, mevcut veri setinde örnekleme destekli boosting yaklaşımının diğer modellere kıyasla daha dengeli sonuç verdiğini göstermektedir. Bununla birlikte her iki yaklaşımda da model başarısının sınırlı kaldığı noktalar bulunmaktadır. Bu sınırlılık, lif oranlarının kumaş hata adedini tek başına güçlü biçimde açıklamamasıyla ve veri setinde proses düzeyindeki bazı kritik değişkenlerin

bulunmamasıyla ilişkilendirilebilir.

Makine öğrenmesi modelleri, mevcut veri yapısı içinde kumaş hata düzeylerinin belirli ölçüde sınıflandırılabilirliğini göstermiştir. Bununla birlikte modellerin performansı, sınıf yapısına, kullanılan performans ölçütüne ve doğrulama yöntemine göre değişmektedir. İki sınıflı yapıda accuracy değeri yüksek olan bazı modellerin yüksek hata sınıfını yakalayamaması, kalite kontrol bağlamında recall ve F1-score ölçütlerinin önemini artırmıştır. Üç sınıflı yapıda ise RUSBoost modeli daha dengeli sonuç üretmiş, ancak genel performans düzeyi lif kompozisyonu değişkenlerinin tek başına sınıflandırma için sınırlı bilgi taşıdığını göstermiştir. Bu bulgular, makine öğrenmesi yaklaşımlarının denim konfeksiyon üretiminde kalite izleme amacıyla kullanılabileceğini; ancak daha güçlü ve sahaya aktarılabilir modeller için veri kapsamının genişletilmesi gerektiğini ortaya koymaktadır.

5.2. Lif Oranlarının Hata Oluşumuna Etkisinin Değerlendirilmesi

Bu çalışmada lif oranları ile kumaş hata adedi arasındaki ilişki iki farklı düzeyde ele alınmıştır. İlk olarak korelasyon analizinde kullanılacak değişkenlerin normal dağılım varsayımı Shapiro-Wilk testi ile değerlendirilmiş ve değişkenlerin normal dağılım göstermediği belirlenmiştir. Bu nedenle değişkenler arasındaki monoton ilişki Spearman korelasyon analiziyle incelenmiş, daha sonra Random Forest temelli değişken önem analiziyle lif oranlarının sınıflandırma sürecindeki göreceli katkıları değerlendirilmiştir. Bu ayrım önemlidir, çünkü Spearman korelasyonu değişkenler arasındaki sıralı/monoton ilişkiyi gösterirken, değişken önem analizi bir makine öğrenmesi modelinin tahmin sürecinde hangi girdilerden daha fazla yararlandığını göstermektedir. Dolayısıyla korelasyon bulguları ile değişken önem bulguları aynı anlamda yorumlanmamalıdır. Random Forest temelli değişken önem analizleri, modelin karar verme sürecinde değişkenlerden ne ölçüde yararlandığını gösterebilmekle birlikte, bu sonuçlar doğrudan nedensellik kanıtı olarak değerlendirilmemelidir (Breiman, 2001; Genuer vd., 2010; Lazarova-Molnar, 2022).

Spearman korelasyon sonuçları, lif oranları ile kumaş hata adedi arasında istatistiksel olarak anlamlı ve güçlü bir ilişki bulunmadığını göstermiştir. COT oranı için Spearman rho değeri -0,0148, Recycle oran için 0,0440, PES oranı için 0,0146, EA oranı için -0,0523

ve Diğer lifler oranı için -0,0655 olarak hesaplanmıştır. Tüm değişkenlerde p değerleri 0,05'in üzerinde kalmıştır. Bu nedenle mevcut veri yapısı üzerinden lif oranlarının kumaş hata adedini doğrudan ve güçlü biçimde açıkladığını söylemek doğru değildir. Kumaş hatalarının oluşumu çoğu zaman lif kompozisyonundan daha geniş bir üretim bağlamına bağlıdır. Denim üretiminde hata oluşumu; iplik özellikleri, dokuma veya konfeksiyon süreci, makine ayarları, operatör etkisi, bakım koşulları, üretim hızı ve proses parametreleri gibi çok sayıda değişkenin birlikte etkisiyle şekillenebilmektedir (Akter vd., 2021; Eryuruk, 2019; Shahid vd., 2024).

Kumaş hatalarının çok faktörlü yapısı, bu çalışmada elde edilen düşük korelasyon değerlerini anlaşılır kılmaktadır. Lif oranları kumaşın fiziksel davranışını, konforunu, elastikiyetini, mukavemetini veya boyutsal kararlılığını etkileyebilir; ancak hata adedi doğrudan yalnızca bu bileşimden türeyen bir sonuç değildir. Lif kompozisyonu üretim sürecine ilişkin bir bilgi kaynağıdır; fakat makine ayarları, iplik kalitesi, bakım kayıtları, operatör etkisi, vardiya düzeni, üretim hızı, ortam koşulları ve hata türü gibi değişkenler olmadan hata adedini güçlü biçimde açıklaması beklenmemelidir. Bu nedenle kumaş hataları, yalnızca malzeme kompozisyonu üzerinden değil, üretim prosesi ve kalite kontrol sistemiyle birlikte değerlendirilmesi gereken çok boyutlu kalite çıktılarıdır (Tercan ve Meisen, 2022).

Literatürde denim kumaş performansına ilişkin çalışmalar, lif bileşiminin kumaş özellikleri üzerinde etkili olabildiğini göstermektedir. Pamuk, polyester, elastan ve farklı karışım yapıları; kumaşın esneme, geri toplama, kopma mukavemeti, yırtılma mukavemeti, sertlik, tutum ve kullanım performansı gibi özelliklerini etkileyebilmektedir. Akter vd., (2021), pamuk, polyester ve spandeks/elastan içeren denim kumaşlarda lif ve iplik yapısının esneme, mukavemet ve kumaş performansı üzerinde etkili olduğunu belirtmiştir. Benzer şekilde Shahid vd., (2024), farklı denim kumaş yapılarının ağırlık, boyutsal değişim, sertlik, mukavemet, yırtılma dayanımı, esneme ve konfor özellikleri bakımından farklı performanslar gösterebildiğini ortaya koymuştur. Bu çalışmalar lif oranlarının kumaş davranışı açısından anlamlı değişkenler olduğunu göstermektedir. Bununla birlikte söz konusu etkiler daha çok fiziksel ve performans özellikleriyle ilişkilidir. Kumaş hata adedi ise üretim sürecinin farklı aşamalarında oluşan çok boyutlu bir kalite çıktısıdır. Bu nedenle lif bileşiminin kumaş

performansı üzerindeki etkisi ile hata adedini doğrudan açıklama gücü birbirinden ayrı değerlendirilmelidir.

Bu nedenle mevcut çalışmada düşük korelasyon katsayılarının elde edilmesi, lif oranlarının önemsiz olduğu anlamına gelmemektedir. Daha doğru ifade ile bu bulgu, lif oranları ile kumaş hata adedi arasında güçlü ve tek yönlü doğrusal bir bağ bulunmadığını göstermektedir. Hata oluşumu doğrusal olmayan, dolaylı veya etkileşimli ilişkiler içerebilir. Örneğin belirli bir lif oranı tek başına hata adedini artırmayabilir; ancak belirli iplik özellikleri, kumaş konstrüksiyonu, makine ayarları veya üretim koşullarıyla birlikte hata riskine katkıda bulunabilir. Makine öğrenmesi yaklaşımlarının üretim kalitesi çalışmalarında kullanılmasının temel nedenlerinden biri de bu tür doğrusal olmayan ve çok değişkenli ilişkileri belirli ölçüde yakalayabilmesidir (Tercan ve Meisen, 2022).

Random Forest temelli değişken önem analizi, korelasyon sonuçlarından farklı bir görünüm sunmuştur. Bu analizde değişkenlerin model içindeki göreceli katkı sıralaması COT oranı, EA oranı, Diğer lifler oranı, Recycle oran ve PES oranı şeklinde gerçekleşmiştir. COT oranı %33,08 önem oranıyla en yüksek katkıya sahip değişken olmuştur. Bunu %19,49 ile EA oranı, %19,37 ile Diğer lifler oranı, %15,96 ile Recycle oran ve %12,10 ile PES oranı izlemiştir. Bu sıralama, Random Forest modelinin sınıflandırma sürecinde en fazla COT oranından yararlandığını, EA oranı ve Diğer lifler oranının da model kararlarında belirgin düzeyde bilgi taşıdığını göstermektedir. Random Forest modellerinde değişken önem skorlarının, karar ağaçları içinde değişkenlerin ayırıcı katkılarını yansıtmaları nedeniyle yorumlayıcı bir araç olarak kullanılabildiği belirtilmektedir (Breiman, 2001; Genuer vd., 2010).

COT oranının değişken önem analizinde ilk sırada yer alması, denim kumaşın temel yapısı açısından beklenebilir bir bulgudur. Denim kumaşlarda pamuk oranı çoğunlukla ana lif bileşenidir ve kumaşın mekanik davranışı, tutumu, yüzey yapısı ve kullanım özellikleri üzerinde önemli rol oynayabilir. Pamuk ağırlıklı denim kumaş yapısının sektörde yaygın olması, bu değişkenin model içinde daha belirgin ayırt edici bilgi taşımasını anlaşılır kılmaktadır (Akter vd., 2021; Eryuruk, 2019). Ancak bu sonuç, COT oranının kumaş hatalarının doğrudan nedeni olduğu anlamına gelmemektedir. Burada elde edilen bulgu,

pamuk oranının model içi ayırt edicilik düzeyinin diğer lif oranlarına kıyasla daha yüksek olduğunu göstermektedir.

EA oranının ikinci sırada yer alması, elastan içeriğinin model kararlarında dikkate değer bir katkı sunduğunu göstermektedir. Elastan içeren denim kumaşlarda esneme, geri toplama, form koruma ve kullanım performansı gibi özellikler lif bileşiminden etkilenebilmektedir. Akter vd., (2021), elastan/spandeks içeren denim kumaşlarda elastikiyet ve mekanik performansın kumaş yapısına bağlı olarak değiştiğini göstermiştir. Bu özellikler üretim sürecinde kumaşın davranışını ve işlenebilirliğini dolaylı olarak etkileyebilir. Ancak mevcut veri setinde EA oranı ile hata adedi arasındaki korelasyon katsayısı düşük ve negatif düzeydedir. Bu nedenle EA oranının değişken önem analizinde öne çıkması, hata adediyle doğrudan doğrusal bir ilişki kurduğu anlamına gelmemekte; modelin sınıflandırma kararlarında bu değişkenden ayırt edici bilgi olarak yararlandığını göstermektedir.

Diğer lifler oranı, değişken önem sıralamasında üçüncü sırada yer almıştır. Bu değişkenin önem oranı EA oranına oldukça yakındır. Bu durum, pamuk, recycle, polyester ve elastan dışındaki lif bileşenlerinin bazı kumaş kayıtlarında model açısından ayırt edici bilgi taşıdığını göstermektedir. Denim kumaşlarda farklı lif karışımları, yalnızca maliyet veya görünüm açısından değil, aynı zamanda performans, konfor ve fonksiyonel özellikler bakımından da kullanılabilir (Shahid vd., 2024). Bu nedenle Diğer lifler oranının model içinde anlamlı katkı sunması, kumaş kompozisyonundaki farklılaşmanın sınıflandırma sürecinde dikkate alındığını göstermektedir. Ancak bu sonuç “diğer lifler hata oluşturur” şeklinde yorumlanmamalıdır. Daha doğru yorum, farklı lif bileşimlerine sahip bazı kumaş kayıtlarının hata sınıflandırmasında model tarafından ayrıştırılabilir özellikler göstermesidir.

Recycle oran, değişken önem sıralamasında dördüncü sırada yer almıştır. Ancak korelasyon analizinde Recycle oran ile kumaş hata adedi arasındaki ilişki düşük düzeyde kalmıştır. Bu durum, Recycle oranının hata adediyle basit doğrusal bir ilişki göstermediğini, buna karşılık bazı gözlem gruplarında model açısından ayırt edici bilgi sağlayabildiğini düşündürmektedir. Geri dönüştürülmüş lif kullanımının kumaş performansı ve üretim davranışı üzerindeki etkisi, lif kalitesi, karışım homojenliği, iplik özellikleri ve proses koşullarıyla birlikte değerlendirilmelidir. Tek başına Recycle oranı üzerinden hata oluşumuna

ilişkin doğrudan nedensel bir sonuç çıkarmak uygun değildir.

PES oranı ise değişken önem sıralamasında son sırada yer almıştır. Bu durum, mevcut veri setinde polyester oranının model kararlarına diğer lif oranlarına kıyasla daha sınırlı katkı sunduğunu göstermektedir. Polyester ve polyester karışımları denim kumaşlarda mukavemet, boyutsal kararlılık, dayanıklılık ve kullanım performansı gibi özelliklerle ilişkilendirilebilmektedir (Akter vd., 2021; Shahid vd., 2024). Bununla birlikte PES oranının son sırada yer alması, polyester içeriğinin kumaş performansı veya üretim davranışı açısından önemsiz olduğu anlamına gelmemektedir. Yalnızca bu çalışmada kullanılan veri yapısı ve modelleme çerçevesi içinde PES oranının hata sınıflandırmasına katkısı daha düşük bulunmuştur.

Değişken önem sonuçlarının nedensellik olarak okunmaması özellikle önemlidir. Random Forest gibi ağaç temelli yöntemlerde değişken önem skorları, modelin karar süreçlerinde hangi değişkenleri daha fazla kullandığını gösterir; ancak bu değerler doğrudan fiziksel ya da üretimsel neden-sonuç ilişkisi kurmaz. Ayrıca değişken önem skorlarının veri yapısı, değişkenler arası korelasyon, model parametreleri ve kullanılan önem ölçütünden etkilenebileceği bilinmektedir (Genuer vd., 2010; Lazaro-Molnar, 2022). Bu çalışmada COT oranının birinci, EA oranının ikinci ve Diğer lifler oranının üçüncü sırada yer alması, bu değişkenlerin modelin sınıflandırma sürecinde daha fazla ayırt edici bilgi taşıdığını göstermektedir. Buna karşılık bu sıralama, ilgili lif oranlarının kumaş hatalarını doğrudan artırdığı veya azalttığı şeklinde yorumlanmamalıdır.

Genel olarak değerlendirildiğinde, lif oranlarına ilişkin bulgular üç temel sonuca işaret etmektedir. Birincisi, Spearman korelasyon analizi lif oranları ile kumaş hata adedi arasında istatistiksel olarak anlamlı ve güçlü bir monoton ilişki bulunmadığını göstermiştir. İkincisi, Random Forest temelli değişken önem analizi COT oranının model kararlarında en yüksek katkıya sahip olduğunu, bunu EA oranı ve Diğer lifler oranının izlediğini ortaya koymuştur. Üçüncüsü, Recycle oran ve PES oranı da model içinde belirli düzeyde bilgi taşımakla birlikte, göreceli katkıları diğer değişkenlere kıyasla daha düşük kalmıştır.

Bu sonuçlar birlikte değerlendirildiğinde, lif kompozisyonunun hata oluşumuna ilişkin sınırlı fakat tamamen göz ardı edilemeyecek bir bilgi taşıdığı söylenebilir. Mevcut veri yapısı,

lif oranlarının hata adedini tek başına açıklamasına yeterli değildir; buna karşılık makine öğrenmesi modeli içinde bazı lif bileşenlerinin ayırt edici katkı sunduğu görülmektedir. Bu nedenle çalışmanın bulguları, kumaş hata oluşumunun daha güçlü açıklanabilmesi için lif kompozisyonunun yanı sıra iplik özellikleri, proses parametreleri, üretim hattı bilgileri, makine ayarları, operatör bilgisi, vardiya değişkenleri ve hata türlerini içeren daha geniş veri setlerine ihtiyaç olduğunu göstermektedir. Bu değerlendirme, üretim kalitesinin yalnızca malzeme girdileriyle değil, proses verileriyle birlikte modellenmesi gerektiğini vurgulayan kestirimci kalite yaklaşımıyla da uyumludur (Tercan ve Meisen, 2022).

5.3. Sürdürülebilirlik Açısından Genel Değerlendirme

Bu çalışmada kumaş hata adedine bağlı rework süresinin ek elektrik tüketimi ve karbon ayak izi üzerindeki etkisi sınırlı kapsamlı bir hesaplama ile değerlendirilmiştir. Veri setinde toplam kumaş hata adedi 31.943 olarak belirlenmiş, hata başına ortalama rework süresi 1,05 dakika kabul edilmiştir. Bu yaklaşımla toplam rework süresi 33.540,15 dakika, diğer bir ifadeyle yaklaşık 559 saat olarak hesaplanmıştır. Rework sürecinde ortalama güç tüketimi 10 kW olarak alındığında, hata kaynaklı ek elektrik tüketimi 5.590 kWh düzeyine ulaşmıştır. Elektrik tüketimine bağlı emisyon faktörü 0,4183 kg CO₂e/kWh olarak kullanıldığında, kumaş hatalarından kaynaklanan ek karbon etkisi yaklaşık 2.338,31 kg CO₂e, yani 2,34 ton CO₂e/yıl olarak hesaplanmıştır.

Bu sonuç, kalite hatalarının yalnızca üretim verimliliği veya işçilik kaybı açısından değil, enerji tüketimi ve karbon emisyonu açısından da ölçülebilir sonuçlar doğurduğunu göstermektedir. Tekstil ve hazır giyim üretiminde çevresel etkiler çoğunlukla lif üretimi, boyama, terbiye, yıkama, su tüketimi, kimyasal kullanımı ve enerji tüketimi gibi başlıklar üzerinden değerlendirilmektedir (Hasanbeigi ve Price, 2015; Muthu, 2020; Sandin ve Peters, 2018). Bununla birlikte üretim sürecinde ortaya çıkan hata, yeniden işlem ve verimsizlikler de kaynak kullanımını artıran önemli unsurlar arasında yer almaktadır. Rework işlemleri, hatalı ürünün tamamen fireye ayrılmasını önleyebilse de ek iş gücü, makine kullanımı, enerji tüketimi ve zaman kaybı oluşturmaktadır. Bu nedenle kalite yönetimi, yalnızca ürün uygunluğunu sağlayan teknik bir süreç olarak değil, aynı zamanda sürdürülebilir üretim performansını etkileyen bir yönetim alanı olarak değerlendirilmelidir.

Çalışmada hesaplanan 2,34 ton CO₂e/yıl değeri, ürünlerin toplam karbon ayak izini temsil etmemektedir. Bu değer yalnızca kumaş hatalarına bağlı rework faaliyetinin elektrik tüketimi üzerinden hesaplanan sınırlı bir göstergedir. Hammadde üretimi, iplik üretimi, kumaş üretimi, boyama, yıkama, terbiye, konfeksiyon, ambalajlama, lojistik, kullanım ve yaşam sonu aşamaları bu hesaba dâhil edilmemiştir. Yaşam döngüsü değerlendirmesi yaklaşımında sistem sınırlarının, fonksiyonel birimin, veri kaynaklarının ve varsayımların açık biçimde tanımlanması temel bir gerekliliktir (ISO, 2006). Bu nedenle mevcut hesaplama tam kapsamlı bir yaşam döngüsü analizi olarak değil, üretim içi kalite hatalarından kaynaklanan ek işlem yükünün karbon karşılığını görünür kılan dar kapsamlı bir karbon göstergesi olarak ele alınmalıdır.

Karbon ayak izi hesaplamalarında sistem sınırının açık tanımlanması, elde edilen sonucun doğru yorumlanması açısından önemlidir. PAS 2050 ve GHG Protocol Product Standard gibi yaklaşımlar, ürün veya faaliyet kaynaklı sera gazı emisyonlarının hesaplanmasında kapsam, sınır, veri kalitesi ve emisyon faktörü seçiminin açıkça belirtilmesi gerektiğini vurgulamaktadır (Specification, 2008; WRI & WBCSD, 2015). Bu çalışmada kullanılan hesaplama zinciri, toplam hata adedinden rework süresine, rework süresinden elektrik tüketimine ve elektrik tüketiminden CO₂e değerine geçişe dayanmaktadır. Dolayısıyla hesaplanan değer, işletme düzeyinde kalite hatalarının enerji ve karbon etkisini izlemek için kullanılabilecek pratik bir göstergedir; ancak ürünün tüm yaşam döngüsünü temsil eden bütüncül bir karbon ayak izi değeri değildir.

Karbon ayak izi hesabının sınırlı kapsamlı olması, bulgunun önemini azaltmamaktadır. Aksine, yalnızca rework kaynaklı elektrik tüketiminin bile ölçülebilir bir karbon etkisi oluşturması, kalite hatalarının sürdürülebilir üretim açısından dikkate alınması gerektiğini göstermektedir. Tekstil sektöründe enerji kullanımı, sera gazı emisyonlarının önemli kaynaklarından biri olarak değerlendirilmektedir (Hasanbeigi ve Price, 2015). Özellikle denim üretiminde boyama, yıkama, kurutma, ütüleme ve son işlem aşamaları enerji, su ve kimyasal kullanımı bakımından çevresel yük oluşturabilmektedir (Abbate vd., 2023; Muthu, 2020). Bu bağlamda hata kaynaklı yeniden işlem faaliyetlerinin azaltılması, aynı üretim hacmi için daha az enerji tüketilmesine ve daha düşük karbon emisyonu oluşmasına katkı

sağlayabilir.

Kumaş hatalarının azaltılması, karbon etkisini iki yönden sınırlayabilir. Birincisi, rework ihtiyacını düşürerek ek elektrik tüketimini azaltabilir. İkincisi, hata kaynaklı üretim akış bozulmalarını azaltarak üretim hattının daha verimli çalışmasına katkı sağlayabilir. Üretim sistemlerinde kalite iyileştirme, fire ve yeniden işlem miktarının azaltılması yoluyla kaynak verimliliğini artıran önemli bir uygulama alanıdır (Garza-Reyes, 2015; Vinodh vd., 2015). Bu açıdan kalite kontrol sistemleri yalnızca hatalı ürünleri tespit eden bir mekanizma olarak değil, üretim kaynaklarının daha etkin kullanımını destekleyen bir sürdürülebilirlik aracı olarak da değerlendirilmelidir. Özellikle denim konfeksiyon üretiminde hata kaynaklı tekrar işlemlerin azaltılması, işletme düzeyinde enerji yönetimi ve karbon azaltımı hedefleriyle ilişkilendirilebilir.

Bu çalışmada karbon ayak izi hesabı, kumaş hata adedi üzerinden oluşturulan basit ve izlenebilir bir hesaplama zincirine dayanmaktadır. Toplam hata adedinin rework süresine, rework süresinin elektrik tüketimine ve elektrik tüketiminin CO₂e değerine dönüştürülmesi, kalite verisinin çevresel performans göstergesine çevrilmesine olanak sağlamıştır. Bu yaklaşım, üretim işletmeleri için uygulanabilir bir izleme mantığı sunmaktadır. Çünkü birçok işletmede hata kayıtları düzenli olarak tutulmakta; ancak bu kayıtların enerji ve karbon karşılığı çoğu zaman sistematik biçimde hesaplanmamaktadır. Bu çalışma, kalite kayıtlarının çevresel performans göstergeleriyle ilişkilendirilebileceğini göstermesi bakımından uygulamaya dönük bir katkı sunmaktadır.

Bununla birlikte hesaplamada bazı varsayımlar bulunmaktadır. Hata başına rework süresi 1,05 dakika olarak kabul edilmiş, rework sürecinde ortalama güç tüketimi 10 kW alınmış ve elektrik kaynaklı emisyon faktörü 0,4183 kg CO₂e/kWh olarak kullanılmıştır. Bu değerler hesaplamayı mümkün kılmakla birlikte, işletmeye, makine parkına, rework türüne, enerji kaynağına ve üretim koşullarına göre değişebilir. Karbon ayak izi hesaplamalarında kullanılan emisyon faktörleri ve faaliyet verileri sonuçları doğrudan etkilediğinden, bu tür hesaplamalarda veri kaynağı ve varsayımlar açıkça belirtilmelidir (Wiedmann ve Minx, 2008; WRI ve WBCSD, 2015). Bu nedenle hesaplanan 2,34 ton CO₂e/yıl değeri mutlak ve tüm işletmelere genellenebilir bir değer olarak değil, çalışma kapsamındaki veri seti ve

varsayımlar altında elde edilen bir gösterge olarak yorumlanmalıdır.

Karbon ayak izi sonucunun yorumlanmasında hata türlerinin veri setinde yer almaması da önemli bir sınırlılıktır. Her kumaş hatasının aynı rework süresine ve aynı enerji tüketimine neden olduğu varsayılmıştır. Oysa gerçek üretim ortamında bazı hatalar kısa süreli düzeltmelerle giderilebilirken, bazı hatalar daha uzun işlem, yeniden kontrol veya ek ekipman kullanımı gerektirebilir. Benzer şekilde bazı hatalar yalnızca manuel düzeltme gerektirirken, bazıları yıkama, ütüleme veya yeniden kalite kontrol gibi enerji tüketen ek adımlar oluşturabilir. Bu nedenle ilerleyen çalışmalarda hata türü, rework tipi, işlem süresi ve kullanılan ekipman bilgileri birlikte kaydedilirse karbon etkisi daha ayrıntılı biçimde hesaplanabilir. Yaşam döngüsü değerlendirmesi literatüründe veri ayrıntısı ve süreç temsiliyetinin sonuç güvenilirliği üzerinde belirleyici olduğu vurgulanmaktadır (Finnveden vd., 2009; ISO, 2006).

Bu bölümde elde edilen bulgular, çalışmanın makine öğrenmesi sonuçlarıyla birlikte değerlendirildiğinde daha bütünlüklü bir anlam kazanmaktadır. Lif oranları ve hata adedi arasındaki ilişki güçlü doğrusal bir yapı göstermemiş, buna karşılık bazı makine öğrenmesi modelleri yüksek hata riskini belirli ölçüde ayırt edebilmiştir. Eğer üretim sürecinde yüksek hata riski taşıyan kayıtlar daha erken belirlenebilirse, rework ihtiyacı oluşmadan önce düzeltici veya önleyici kontrollerin yapılması mümkün olabilir. Kestirimci kalite yaklaşımları, üretim verilerinden yararlanarak kalite çıktılarının önceden tahmin edilmesini ve süreç müdahalelerinin daha erken yapılmasını desteklemektedir (Tercan ve Meisen, 2022). Bu nedenle makine öğrenmesi destekli hata tahmini, yalnızca kalite güvence açısından değil, enerji tüketimi ve karbon emisyonlarının azaltılması açısından da dolaylı katkı sağlayabilecek bir araç olarak değerlendirilebilir.

Kumaş hatalarına bağlı rework faaliyetleri işletme düzeyinde ölçülebilir bir enerji ve karbon yükü oluşturmaktadır. Bu çalışmada hesaplanan 2,34 ton CO₂e/yıl değeri sınırlı kapsamlı olmakla birlikte, kalite hatalarının çevresel etkisini görünür kılması bakımından önemlidir. Hata azaltımı, yalnızca ürün kalitesini ve üretim verimliliğini artıran bir faaliyet değil, aynı zamanda enerji tüketimini ve karbon emisyonunu azaltabilecek bir sürdürülebilirlik stratejisidir. Bu nedenle denim konfeksiyon üretiminde kalite yönetimi,

makine öğrenmesi destekli hata tahmini ve karbon izleme uygulamalarının birlikte ele alınması, daha verimli ve çevresel açıdan daha sorumlu üretim sistemlerinin geliştirilmesine katkı sağlayabilir.

5.4. Genel Tartışma Özeti

Bu çalışmada denim konfeksiyon üretiminde kumaş hata adedi; lif kompozisyonu, makine öğrenmesi tabanlı sınıflandırma ve hata kaynaklı karbon ayak izi hesaplaması çerçevesinde birlikte değerlendirilmiştir. Çalışmanın temel yaklaşımı, kumaş hata kayıtlarını yalnızca kalite kontrol çıktısı olarak ele almakla sınırlı değildir. Bunun yerine hata adedi, üretim verimliliği, modelleme potansiyeli ve sürdürülebilirlik göstergeleriyle ilişkilendirilmiştir. Bu yönüyle çalışma, üretim kalitesinin veri odaklı biçimde izlenmesi ve yorumlanması açısından bütüncül bir değerlendirme sunmaktadır. Üretim süreçlerinde kalite çıktılarının veri temelli yöntemlerle tahmin edilmesi, kestirimci kalite yaklaşımının temel uygulama alanlarından biridir (Tercan ve Meisen, 2022).

Makine öğrenmesi bulguları, lif kompozisyonu değişkenlerinin kumaş hata düzeylerini belirli ölçüde ayırt edebildiğini; ancak tek başına güçlü bir sınıflandırma başarısı sağlamadığını göstermiştir. LOOCV sonuçlarına göre iki sınıflı sınıflandırma yaklaşımında en dengeli sonuç RUSBoost modelinden elde edilmiştir. Bu modelde accuracy 0,6979, recall 0,5802 ve F1-score 0,4883 olarak hesaplanmıştır. Bazı modeller accuracy bakımından 0,7515 değerine ulaşmış olsa da yüksek hata sınıfını hiç yakalayamamıştır. Bu durum, dengesiz sınıf yapılarında accuracy değerinin tek başına yeterli olmadığını ve özellikle riskli sınıfın yakalanmasında recall ile F1-score ölçütlerinin daha açıklayıcı olduğunu göstermektedir. Dengesiz veri setlerinde çoğunluk sınıfına yönelme, sınıflandırma modellerinde bilinen bir sorundur ve bu tür problemlerde performansın yalnızca genel doğruluk üzerinden yorumlanması yanıltıcı olabilir (He ve Garcia, 2009; Saito ve Rehmsmeier, 2015).

Üç sınıflı sınıflandırma yaklaşımında da en başarılı model RUSBoost olmuştur. RUSBoost modelinde accuracy 0,4586, precision 0,4660, recall 0,4584 ve F1-score 0,4589 olarak hesaplanmıştır. Üç sınıflı yapıda sınıf dağılımı dengeli olmasına rağmen model performanslarının iki sınıflı yaklaşıma göre daha düşük kalması, hata düzeylerinin üç ayrı kategoriye ayrılmasının daha karmaşık bir sınıflandırma problemi oluşturduğunu

göstermektedir. Buna karşılık RUSBoost modelinin üç sınıflı yapıda da en yüksek F1-score değerini üretmesi, bu modelin mevcut veri setinde sınıflar arasında daha dengeli bir tahmin yapısı oluşturabildiğini göstermektedir.

İki ve üç sınıflı yaklaşımlar birlikte değerlendirildiğinde, iki yöntemin farklı karar amaçlarına hizmet ettiği görülmektedir. İki sınıflı yaklaşım, yüksek hata riskinin belirlenmesi açısından daha pratik bir yapı sunmaktadır. Üç sınıflı yaklaşım ise hata düzeylerini daha kademeli biçimde incelemeye olanak sağlamaktadır. LOOCV sonuçlarına göre her iki yaklaşımda da RUSBoost modelinin öne çıkması, örnekleme destekli boosting yaklaşımının mevcut veri yapısında diğer modellere göre daha dengeli sonuç verdiğini göstermektedir. Ancak her iki yaklaşımda da performans değerleri orta düzeyde kalmıştır. Bu durum, lif oranlarının hata sınıflarını tek başına güçlü biçimde ayırmak için yeterli olmadığını ortaya koymaktadır.

Lif oranlarının hata oluşumuna etkisi değerlendirildiğinde, Spearman korelasyon analizi ile değişken önem analizinin farklı sonuçlar sunduğu görülmüştür. Spearman korelasyon analizinde lif oranları ile kumaş hata adedi arasında istatistiksel olarak anlamlı ve güçlü bir ilişki bulunmamıştır. Bu bulgu, kumaş hata adedinin lif oranlarıyla doğrudan ve tek yönlü biçimde açıklanamayacağını göstermektedir. Denim üretiminde hata oluşumu; lif kompozisyonunun yanı sıra iplik özellikleri, kumaş konstrüksiyonu, üretim hattı, makine ayarları, bakım koşulları, operatör etkisi, vardiya düzeni ve proses parametreleri gibi çok sayıda değişkenin birlikte etkisiyle şekillenebilir.

Random Forest temelli değişken önem analizi ise lif oranlarının model içindeki görece katkılarını daha ayrıntılı biçimde göstermiştir. Değişken önem analizi, özellikle COT oranının model kararlarında öne çıktığını, EA oranı ve Diğer lifler oranının da sınıflandırma sürecine yakın düzeyde katkı sunduğunu göstermiştir. Bu sonuçlar, lif kompozisyonunun hata adedini doğrudan açıklamasa da model içinde ayırt edici bilgi taşıdığını ortaya koymaktadır. Bu sonuç, modelin sınıflandırma sürecinde en fazla COT oranından yararlandığını, EA oranı ve Diğer lifler oranının da belirgin düzeyde ayırt edici bilgi taşıdığını göstermektedir. Random Forest modellerinde değişken önem skorları, modelin karar süreçlerinde değişkenlerden nasıl yararlandığını göstermesi açısından yararlı

olmakla birlikte, bu skorların doğrudan nedensellik kanıtı olarak değerlendirilmemesi gerekir (Genuer vd., 2010; Lazarova-Molnar, 2022).

COT oranının değişken önem analizinde ilk sırada yer alması, denim kumaşın temel yapısı açısından anlamlıdır. Pamuk, denim kumaşlarda ana lif bileşeni olarak kumaşın tutumu, mekanik davranışı, yüzey özelliği ve kullanım performansı üzerinde etkili olabilir. EA oranının ikinci sırada bulunması, elastan içeriğinin kumaş davranışı ve işlenebilirlik açısından model kararlarında bilgi taşıdığını düşündürmektedir. Diğer lifler oranının üçüncü sırada yer alması ise pamuk, recycle, polyester ve elastan dışındaki lif bileşenlerinin bazı kumaş kayıtlarında ayırıştırıcı bir özellik gösterdiğini ortaya koymaktadır. Bununla birlikte bu değişken önem sıralaması, söz konusu liflerin hata oluşumunu doğrudan artırdığı veya azalttığı anlamına gelmemektedir. Daha doğru yorum, lif kompozisyonunun model içinde sınırlı fakat dikkate değer bir ayırt edici bilgi taşıdığıdır.

Karbon ayak izi bulguları, kumaş hata adedinin çevresel açıdan da izlenebilir sonuçlar doğurduğunu göstermiştir. Veri setinde toplam kumaş hata adedi 31.943 olarak hesaplanmış; bu hata sayısı 559 saat ek rework süresine, 5.590 kWh ek elektrik tüketimine ve yaklaşık 2,34 ton CO₂e/yıl karbon etkisine karşılık gelmiştir. Bu değer, ürünlerin toplam karbon ayak izini temsil etmemektedir. Hesaplama yalnızca kumaş hatalarına bağlı rework faaliyetinin elektrik tüketimi üzerinden yapılmıştır. Yaşam döngüsü değerlendirmesi yaklaşımında sistem sınırlarının açık biçimde tanımlanması temel bir gereklilik olduğundan, bu sonuç tam kapsamlı bir yaşam döngüsü analizi olarak değil, üretim içi kalite hatalarının oluşturduğu ek karbon yükünü gösteren sınırlı kapsamlı bir gösterge olarak değerlendirilmelidir (ISO, 2006; WRI & WBCSD, 2015).

Karbon ayak izi sonucunun sınırlı kapsamlı olması, bulgunun önemini azaltmamaktadır. Aksine, yalnızca rework kaynaklı elektrik tüketiminin dahi ölçülebilir bir karbon etkisi oluşturması, kalite hatalarının sürdürülebilir üretim açısından da ele alınması gerektiğini göstermektedir. Tekstil ve hazır giyim sektöründe çevresel etkiler çoğunlukla hammadde üretimi, boyama, terbiye, su kullanımı, kimyasal tüketimi ve enerji kullanımı üzerinden tartışılmaktadır (Hasanbeigi ve Price, 2015; Sandin ve Peters, 2018). Ancak üretim içi yeniden işlem, fire, zaman kaybı ve enerji kullanımı da çevresel performans üzerinde etkili

olabilmektedir. Bu nedenle kalite hatalarının azaltılması, yalnızca ürün uygunluğunu artıran bir uygulama değil, aynı zamanda enerji tüketimini ve karbon emisyonunu azaltabilecek bir sürdürülebilirlik stratejisi olarak değerlendirilmelidir.

Çalışmanın bulguları birlikte değerlendirildiğinde, lif kompozisyonu değişkenlerinin kumaş hata adedi hakkında sınırlı ancak tamamen göz ardı edilemeyecek bilgi taşıdığı görülmektedir. Korelasyon katsayılarının düşük olması, basit doğrusal ilişkilerin zayıf olduğunu göstermiştir. Buna karşılık Random Forest değişken önem analizi ve sınıflandırma modelleri, lif oranlarının model kararlarında belirli düzeyde katkı sağladığını ortaya koymuştur. Bu bulgular, kumaş hata oluşumunun tek bir değişkenle açıklanamayacak çok faktörlü bir yapı taşıdığını desteklemektedir.

Bu çalışma aynı zamanda veri kapsamının model başarısı üzerindeki belirleyici rolünü ortaya koymuştur. Kullanılan veri seti lif oranları ve kumaş hata adedi üzerinden yapılandırılmıştır. Buna karşılık hata türü, iplik özellikleri, kumaş konstrüksiyonu, makine ayarları, üretim hızı, operatör bilgisi, vardiya, bakım durumu, ortam koşulları ve proses parametreleri gibi değişkenler veri setinde yer almamaktadır. Bu eksiklik, model performanslarının sınırlı kalmasında önemli bir etken olabilir. Kestirimci kalite çalışmalarında model başarısı, yalnızca algoritmanın seçimine değil, kullanılan veri kapsamına, veri kalitesine ve hedef değişkenin tanımlanma biçimine de bağlıdır (Tercan ve Meisen, 2022). Bu nedenle ilerleyen çalışmalarda daha geniş ve proses düzeyinde zenginleştirilmiş veri setlerinin kullanılması, hata tahmin modellerinin açıklayıcılığını ve uygulanabilirliğini artırabilir.

Genel olarak değerlendirildiğinde, bu çalışma denim konfeksiyon üretiminde kalite verilerinin makine öğrenmesi ve karbon ayak izi hesaplamasıyla birlikte ele alınabileceğini göstermiştir. İki sınıflı yaklaşım, yüksek hata riskini belirleme açısından daha pratik bir karar desteği sunarken, üç sınıflı yaklaşım hata düzeylerinin daha ayrıntılı izlenmesine imkân sağlamaktadır. Değişken önem analizi COT oranı başta olmak üzere bazı lif bileşenlerinin model kararlarında daha belirgin katkı sunduğunu göstermiştir. Karbon ayak izi hesaplaması ise kalite hatalarının enerji ve çevresel etki boyutunu görünür hale getirmiştir. Bu nedenle çalışmanın genel bulgusu, kalite yönetimi, veri analitiği ve sürdürülebilirlik hedeflerinin denim konfeksiyon üretiminde birlikte değerlendirilmesi gerektiğidir.

6. SONUÇLAR

6.1. Genel Sonuçlar

Bu çalışmada büyük ölçekli bir denim konfeksiyon işletmesine ait üretim ve kalite kontrol kayıtları kullanılarak kumaş hata adedi, lif kompozisyonu, makine öğrenmesi tabanlı sınıflandırma ve hata kaynaklı karbon ayak izi hesabı çerçevesinde değerlendirilmiştir. Çalışma kapsamında 652 gözlemden oluşan veri seti kullanılmış; COT oranı, Recycle oran, PES oranı, EA oranı ve Diğer lifler oranı bağımsız değişkenler olarak ele alınmıştır. Kumaş hata adedi ise hem sürekli değişken olarak incelenmiş hem de iki sınıflı ve üç sınıflı sınıflandırma yaklaşımlarına dönüştürülerek modellenmiştir.

Veri setinde toplam kumaş hata adedi 31.943 olarak hesaplanmıştır. Kumaş hata adedi değişkeninin ortalaması 48,99, medyanı 31 ve standart sapması 57,96'dır. Bu sonuç, hata adedinin gözlemler arasında geniş bir dağılım gösterdiğini ve bazı kumaş kayıtlarında hata yoğunluğunun belirgin biçimde arttığını göstermektedir. Lif kompozisyonu açısından değerlendirildiğinde, COT oranı ortalaması %77,55, Recycle oran ortalaması %10,38, PES oranı ortalaması %2,48, EA oranı ortalaması %1,39 ve Diğer lifler oranı ortalaması %8,19 olarak belirlenmiştir.

Shapiro-Wilk normallik testi sonucunda korelasyon analizinde kullanılan değişkenlerin normal dağılım varsayımını sağlamadığı belirlenmiştir. Bu nedenle lif oranları ile kumaş hata adedi arasındaki ilişki Spearman sıra korelasyonu ile değerlendirilmiştir. Spearman korelasyon analizi, lif oranları ile kumaş hata adedi arasında istatistiksel olarak anlamlı ve güçlü bir monoton ilişki bulunmadığını göstermiştir. Bu sonuç, kumaş hata adedinin yalnızca lif oranlarıyla doğrudan ve güçlü biçimde açıklanamayacağını ortaya koymaktadır. COT oranı ve Recycle oran ile kumaş hata adedi arasındaki korelasyon katsayıları pozitif olmakla birlikte oldukça düşük düzeydedir. PES oranı, EA oranı ve Diğer lifler oranı için ise düşük düzeyde negatif korelasyon katsayıları elde edilmiştir. Bu sonuç, kumaş hata adedinin yalnızca lif oranlarıyla doğrudan ve güçlü biçimde açıklanamayacağını ortaya koymaktadır.

Random Forest temelli değişken önem analizi, lif oranlarının model içindeki göreceli katkılarını daha ayrıntılı biçimde göstermiştir. Değişken önem analizi sonucunda COT

oranının model kararlarında en yüksek katkıya sahip değişken olduğu, EA oranı ve Diğer lifler oranının da bunu izlediği belirlenmiştir. Bu bulgu, lif kompozisyonunun hata adedini tek başına güçlü biçimde açıklamadığını, buna karşılık sınıflandırma modelleri içinde belirli düzeyde ayırt edici bilgi taşıdığını göstermektedir.

İki sınıflı sınıflandırma yaklaşımında kumaş hata adedi 75. yüzdelerik değeri olan 63'e göre "Normal" ve "Yüksek" olarak sınıflandırılmıştır. Veri setinde 490 gözlem Normal, 162 gözlem ise Yüksek hata sınıfında yer almıştır. LOOCV sonuçlarına göre iki sınıflı sınıflandırma yaklaşımında F1-score bakımından en başarılı model RUSBoost olmuştur. Bu modelde accuracy 0,6979, precision 0,4215, recall 0,5802 ve F1-score 0,4883 olarak hesaplanmıştır. Cost-Sensitive Random Forest modeli ise 0,4378 F1-score değeriyle ikinci sırada yer almıştır.

İki sınıflı sınıflandırma sonuçları, accuracy değerinin tek başına yeterli bir performans ölçütü olmadığını göstermiştir. Bazı modeller 0,7515 accuracy değerine ulaşmasına rağmen yüksek hata sınıfını hiç doğru tahmin edememiştir. Bu nedenle üretim açısından kritik olan yüksek hata sınıfının yakalanmasında recall ve F1-score değerlerinin daha açıklayıcı olduğu sonucuna ulaşılmıştır.

Üç sınıflı sınıflandırma yaklaşımında kumaş hata adedi 33. ve 66. yüzdelerik değerlere göre "Düşük", "Orta" ve "Yüksek" olarak sınıflandırılmıştır. Düşük sınıfta 216, Orta sınıfta 215 ve Yüksek sınıfta 221 gözlem yer almıştır. LOOCV sonuçlarına göre üç sınıflı sınıflandırma yaklaşımında en başarılı model yine RUSBoost olmuştur. Bu modelde accuracy 0,4586, precision 0,4660, recall 0,4584 ve F1-score 0,4589 olarak hesaplanmıştır. Üç sınıflı yaklaşımda sınıf dağılımı dengeli olmasına rağmen model performanslarının orta düzeyde kalması, hata düzeylerinin üç ayrı kategori altında ayrıştırılmasının daha karmaşık bir problem oluşturduğunu göstermektedir.

İki ve üç sınıflı yaklaşımlar birlikte değerlendirildiğinde, iki sınıflı yaklaşımın yüksek hata riskini belirleme açısından daha pratik bir yapı sunduğu, üç sınıflı yaklaşımın ise hata düzeylerini daha ayrıntılı izlemeye olanak sağladığı sonucuna ulaşılmıştır. LOOCV sonuçlarına göre her iki yaklaşımda da RUSBoost modelinin öne çıkması, mevcut veri setinde örnekleme destekli boosting yaklaşımının daha dengeli sonuç verdiğini göstermektedir.

6.2. Çalışmanın Üretim Sürecine Katkıları

Bu çalışma, denim konfeksiyon üretiminde kalite kontrol verilerinin yalnızca geçmiş hataları gösteren kayıtlar olarak değil, üretim sürecini değerlendirmeye ve riskli kayıtları belirlemeye yardımcı olabilecek analitik girdiler olarak kullanılabileceğini göstermiştir. Lif kompozisyonu değişkenleri, kumaş hata adedini tek başına güçlü biçimde açıklamamış olsa da makine öğrenmesi modelleri içinde belirli düzeyde ayırt edici katkı sağlamıştır. Bu durum, işletmelerde mevcut kalite kayıtlarının veri analitiğiyle birlikte değerlendirildiğinde üretim kararlarına katkı sunabileceğini ortaya koymaktadır.

İki sınıflı sınıflandırma yaklaşımında RUSBoost modelinin en dengeli sonucu vermesi, yüksek hata riski taşıyan kayıtların belirlenmesinde örnekleme destekli boosting yaklaşımlarının yararlı olabileceğini göstermektedir. Üretim ortamında yüksek hata sınıfının kaçırılması, yeniden işlem, zaman kaybı, ek işçilik ve teslim süresi açısından daha yüksek maliyet oluşturabilir. Bu nedenle yüksek hata sınıfını yakalama başarısı, yalnızca istatistiksel bir model başarısı değil, doğrudan üretim yönetimi açısından da önem taşımaktadır.

Üç sınıflı sınıflandırma yaklaşımı, hata düzeylerinin daha kademeli biçimde izlenmesine katkı sağlamıştır. Düşük, Orta ve Yüksek sınıfların ayrı ayrı değerlendirilmesi, kalite kontrol sürecinde yalnızca hatalı/hatasız ayrımı yapılmasından daha ayrıntılı bir izleme yapısı sunmaktadır. Ancak üç sınıflı yaklaşımda model performanslarının daha sınırlı kalması, üretim verilerinde ara kalite düzeylerinin ayrıştırılmasının daha zor olduğunu göstermektedir.

Çalışma, üretim sürecinde kalite tahmini yapılabilmesi için veri kapsamının önemini de ortaya koymuştur. Mevcut veri setinde yalnızca lif oranları ve kumaş hata adedi yer almaktadır. Buna rağmen bazı modeller yüksek hata sınıfını belirli düzeyde ayırt edebilmiştir. Daha kapsamlı üretim verileri kullanıldığında, model başarısının ve üretim sürecine katkının artması beklenmektedir. Bu nedenle çalışma, işletmelerde veri toplama ve kalite kayıtlarının standartlaştırılması açısından da yol gösterici niteliktedir.

6.3. Çalışmanın Sürdürülebilirlik Açısından Katkıları

Çalışmada kumaş hatalarından kaynaklanan rework faaliyetlerinin enerji tüketimi ve karbon emisyonu üzerindeki etkisi de değerlendirilmiştir. Toplam 31.943 kumaş hatasının

559 saat ek rework süresine neden olduğu belirlenmiştir. Ortalama güç tüketimi 10 kW olarak alındığında, bu rework süresi 5.590 kWh ek elektrik tüketimine karşılık gelmiştir. Elektrik emisyon faktörü 0,4183 kg CO₂e/kWh olarak kullanıldığında, hata kaynaklı karbon etkisi yaklaşık 2,34 ton CO₂e/yıl olarak hesaplanmıştır.

Bu değer, ürünlerin toplam karbon ayak izini temsil etmemektedir; yalnızca kumaş hatalarına bağlı rework faaliyetinin elektrik tüketimi üzerinden hesaplanan sınırlı kapsamlı bir göstergedir. Bu nedenle sonuç, ürün bazlı bütüncül karbon ayak izi olarak değil, kalite hatalarının işletme içi ek işlem yükünü gösteren bir sürdürülebilirlik göstergesi olarak değerlendirilmelidir. Bu nedenle elde edilen değer, ürün bazlı toplam karbon ayak izi olarak değil, kalite hatalarının işletme içi ek işlem yükünü gösteren sınırlı kapsamlı bir sürdürülebilirlik göstergesi olarak değerlendirilmelidir.

Buna rağmen hesaplanan değer, kalite hatalarının sürdürülebilirlik açısından ölçülebilir bir karşılığı olduğunu göstermektedir. Rework ihtiyacının artması, yalnızca ek işçilik ve zaman kaybı oluşturmaz; aynı zamanda ek elektrik tüketimine ve buna bağlı karbon emisyonuna da neden olur. Bu nedenle kalite hatalarının azaltılması, üretim verimliliğinin yanında enerji verimliliği ve karbon azaltımı açısından da önem taşımaktadır.

Çalışmanın sürdürülebilirlik açısından temel katkısı, kalite kontrol verileri ile karbon ayak izi göstergeleri arasında bağlantı kurulabileceğini göstermesidir. İşletmelerde hata adedi, rework süresi, kullanılan ekipman ve enerji tüketimi birlikte izlenirse kalite hatalarının çevresel etkisi daha somut biçimde hesaplanabilir. Böylece kalite yönetimi, yalnızca ürün uygunluğu ve müşteri beklentileri açısından değil, aynı zamanda kaynak kullanımı ve emisyon azaltımı açısından da değerlendirilebilir.

Genel olarak bu çalışma, denim konfeksiyon üretiminde kalite yönetimi, veri analitiği ve sürdürülebilirlik hedeflerinin birlikte ele alınabileceğini ortaya koymuştur. Hata azaltımı; üretim süresini, ek enerji tüketimini ve karbon emisyonunu azaltabilecek bir iyileştirme alanı olarak görülmelidir. Bu nedenle çalışma, sürdürülebilir üretim uygulamalarında kalite verilerinin daha aktif kullanılabileceğini göstermesi bakımından katkı sunmaktadır.

7. ÖNERİLER

İşletmelere Yönelik Öneriler

İşletmelerin kumaş hata kayıtlarını yalnızca kalite kontrol amacıyla değil, veri analitiği ve karar destek amacıyla da düzenli biçimde kullanmaları önerilmektedir. Hata adedi, lif kompozisyonu, sipariş bilgisi ve kalite kontrol sonuçları standart bir veri yapısı içinde tutulduğunda makine öğrenmesi modelleri daha güvenilir biçimde uygulanabilir. Bu nedenle kalite kayıtlarının dijital, düzenli ve analiz edilebilir bir formatta saklanması önem taşımaktadır.

Kumaş hata adedinin yanı sıra hata türü bilgisi de veri setine eklenmelidir. Bu çalışmada hata adedi toplam değer olarak ele alınmıştır. Ancak hatanın leke, yırtık, iplik kopuğu, renk farkı, dokuma bozukluğu, yüzey kusuru veya dikim kaynaklı hata gibi türlere ayrılması modelleme gücünü artırabilir. Hata türlerinin kaydedilmesi, hangi hataların daha fazla rework süresi, enerji tüketimi ve karbon etkisi oluşturduğunu belirlemek açısından da önemlidir.

İşletmelerde lif kompozisyonu bilgileri, üretim süreci verileriyle birlikte değerlendirilmelidir. Mevcut çalışmada lif oranları belirli düzeyde ayırt edici bilgi taşımakla birlikte, hata adedini tek başına güçlü biçimde açıklamamıştır. Bu nedenle lif oranlarına ek olarak iplik numarası, iplik tipi, gramaj, kumaş konstrüksiyonu, üretim hattı, makine ayarları, üretim hızı, operatör bilgisi, vardiya, bakım kayıtları, ortam sıcaklığı, nem ve tedarikçi bilgisi gibi değişkenlerin kalite kayıtlarıyla ilişkilendirilmesi önerilmektedir.

Yüksek hata riski taşıyan kayıtların belirlenmesinde yalnızca genel doğruluk değerine dayalı modeller kullanılmamalıdır. Bu çalışmada bazı modeller yüksek accuracy değeri üretmesine rağmen yüksek hata sınıfını yakalayamamıştır. Bu nedenle işletme uygulamalarında recall, precision, F1-score ve confusion matrix sonuçları birlikte değerlendirilmelidir. Özellikle yüksek hata sınıfının kaçırılması işletme açısından daha maliyetli ise recall değeri daha dikkatli izlenmelidir. Yanlış alarm sayısının azaltılması öncelikliyse precision değeri de karar sürecine dâhil edilmelidir.

İşletmelerde iki sınıflı sınıflandırma yaklaşımı, yüksek hata riskinin belirlenmesi

için pratik bir karar destek aracı olarak kullanılabilir. Bu yaklaşımda üretim kayıtları “Normal” ve “Yüksek” hata riski taşıyan gruplar olarak ayrıştırılabilir. Böylece yüksek hata riski taşıyan kumaş kompozisyonları veya üretim kayıtları daha dikkatli kontrol edilebilir. Üç sınıflı yaklaşım ise kalite düzeylerinin daha kademeli izlenmesi gereken durumlarda kullanılabilir.

Özellikle Recycle oranı yüksek kumaş grupları, Diğer lifler oranı yüksek kumaş grupları ve PES/EA oranı farklılaşan kayıtlar kalite kontrol açısından ayrıca takip edilebilir. Ancak bu gruplar doğrudan riskli olarak kabul edilmemeli; hata türü, üretim hattı, tedarikçi, proses parametreleri ve rework sonuçlarıyla birlikte değerlendirilmelidir.

Kalite kayıtları enerji ve karbon göstergeleriyle ilişkilendirilmelidir. İşletmelerde hata adedi, rework süresi, kullanılan ekipman, elektrik tüketimi ve emisyon faktörü birlikte kaydedilirse, kalite hatalarının çevresel etkisi daha doğru biçimde hesaplanabilir. Bu yaklaşım, kalite yönetimi ile enerji ve karbon yönetimi arasında doğrudan bağlantı kurulmasını sağlayabilir.

Rework süreçleri daha ayrıntılı izlenmelidir. Her kumaş hatasının aynı düzeltme süresine ve aynı enerji tüketimine neden olmadığı dikkate alınarak hata türüne göre rework süresi ölçülmelidir. Ayrıca rework sırasında kullanılan makine, ekipman ve enerji kaynağı ayrı ayrı kaydedilmelidir. Böylece hangi hata türlerinin daha fazla enerji tüketimine ve karbon emisyonuna neden olduğu daha net biçimde belirlenebilir.

İşletmelerin kalite kontrol süreçlerinde makine öğrenmesi destekli karar destek sistemlerine geçiş yapmaları önerilmektedir. Bu sistemler, yüksek hata riski taşıyan üretim kayıtlarının erken belirlenmesine katkı sağlayabilir. Ancak model çıktıları doğrudan nihai karar olarak değil, kalite uzmanlarının değerlendirmesini destekleyen yardımcı araçlar olarak kullanılmalıdır.

Gelecek Çalışmalar İçin Öneriler

Gelecek çalışmalarda daha geniş kapsamlı veri setleri kullanılmalıdır. Bu çalışma 652 gözlemden oluşan bir veri setiyle yürütülmüştür. Daha uzun dönemli, farklı üretim partilerini kapsayan ve farklı kumaş tiplerini içeren veri setleri kullanıldığında model

performansları daha güçlü biçimde değerlendirilebilir. Ayrıca farklı işletmelerden elde edilecek verilerle modellerin genellenebilirliği test edilebilir.

Veri setine üretim sürecine ilişkin ek değişkenler dâhil edilmelidir. Lif oranlarının yanı sıra iplik özellikleri, kumaş konstrüksiyonu, gramaj, makine ayarları, üretim hızı, operatör bilgisi, vardiya, bakım kayıtları, ortam koşulları ve tedarikçi bilgisi gibi değişkenlerin modele eklenmesi önerilmektedir. Bu değişkenler, kumaş hata adedinin daha güçlü biçimde açıklanmasına ve tahmin edilmesine katkı sağlayabilir.

Hata türleri ayrı ayrı modellenmelidir. Bu çalışmada kumaş hata adedi toplam değer olarak ele alınmıştır. Gelecek çalışmalarda hata türlerine göre ayrı sınıflandırma modelleri kurulabilir. Böylece hangi lif kompozisyonlarının veya üretim koşullarının hangi hata türleriyle daha fazla ilişkili olduğu daha ayrıntılı biçimde değerlendirilebilir.

Farklı sınıflandırma eşikleri denenmelidir. Bu çalışmada iki sınıflı yaklaşımda 75. yüzdellik değer, üç sınıflı yaklaşımda ise 33. ve 66. yüzdellik değerler kullanılmıştır. Gelecek çalışmalarda işletme hedeflerine göre farklı eşik değerleri belirlenebilir. Örneğin yüksek hata sınıfının daha dar veya daha geniş tanımlandığı alternatif senaryolar oluşturularak model performansları karşılaştırılabilir.

Sınıf dengesizliği sorununa yönelik farklı yöntemler denenmelidir. Bu çalışmada RUSBoost ve Cost-Sensitive Random Forest gibi yaklaşımlar kullanılmıştır. Gelecek çalışmalarda SMOTE, ADASYN, farklı örnekleme yöntemleri, sınıf ağırlıklandırma stratejileri ve maliyet duyarlı öğrenme yaklaşımları birlikte değerlendirilebilir. Böylece özellikle yüksek hata sınıfının daha başarılı biçimde yakalanması sağlanabilir.

Gelişmiş makine öğrenmesi modelleriyle karşılaştırmalı analizler yapılmalıdır. Random Forest ve Cost-Sensitive Random Forest modelleri bu çalışmada dikkat çekici sonuçlar vermiştir. Gelecek çalışmalarda XGBoost, LightGBM, CatBoost, gelişmiş boosting algoritmaları, derin öğrenme yaklaşımları ve hibrit modeller denenebilir. Bu modellerin performansı yalnızca accuracy ile değil, recall, precision, F1-score, confusion matrix ve sınıf bazlı performans ölçütleriyle birlikte değerlendirilmelidir.

Açıklanabilir yapay zekâ yöntemleri kullanılmalıdır. Random Forest temelli

değişken önem analizi bu çalışmada lif oranlarının model içindeki görelî katkılarını göstermiştir. Ancak ilerleyen çalışmalarda SHAP, permutation importance ve partial dependence plot gibi yöntemlerle değişkenlerin model kararlarına nasıl etki ettiğî daha ayrıntılı biçimde incelenebilir. Bu tür yöntemler, model çıktılarının kalite uzmanları tarafından daha kolay yorumlanmasına katkı sağlayabilir.

Karbon ayak izi hesabı daha ayrıntılı hale getirilmelidir. Bu çalışmada yalnızca kumaş hatalarına bağırlı rework faaliyetinin elektrik tüketimi dikkate alınmıştır. Gelecek çalışmalarda hammadde üretimi, iplik üretimi, kumaş üretimi, boyama, yıkama, terbiye, konfeksiyon, ambalajlama, lojistik, kullanım ve yaşam sonu aşamaları da hesaba dâhil edilerek daha kapsamlı bir yaşam döngüsü değerlendirmesi yapılabilir. Böylece hata kaynaklı rework faaliyetlerinin toplam çevresel etki içindeki payı daha net biçimde ortaya konulabilir.

Rework süresi ve enerji tüketimi hata türüne göre ölçülmelidir. Bu çalışmada hata başına ortalama rework süresi kullanılmıştır. Gelecek çalışmalarda her hata türü için gerçek rework süresi, kullanılan ekipman, enerji tüketimi ve işlem tekrarı ayrı ayrı kaydedilmelidir. Bu sayede karbon ayak izi hesabı daha hassas ve işletme gerçekliğine daha uygun hale getirilebilir.

Görüntü verileri ile tabular üretim verileri birlikte kullanılabilir. Kumaş hata tespitinde görüntü işleme ve derin öğrenme yaklaşımları yaygın biçimde kullanılmaktadır. Gelecek çalışmalarda kumaş yüzey görüntüleri, lif kompozisyonu, üretim parametreleri ve kalite kayıtları birlikte modellenenebilir. Bu tür çok kaynaklı veri yapıları, hem hata türünün tespit edilmesine hem de hata düzeyinin tahmin edilmesine katkı sağlayabilir.

Son olarak, kalite yönetimi, veri analitiğî ve sürdürülebilirlik göstergelerini birleştiren karar destek sistemleri geliştirilmelidir. Böyle bir sistem, üretim sırasında yüksek hata riski taşıyan kayıtları erken belirleyebilir, rework ihtiyacını azaltabilir, enerji tüketimini düşürebilir ve karbon emisyonlarının izlenmesine katkı sağlayabilir. Bu nedenle gelecek çalışmalar yalnızca model performansına değıl, modellerin işletme süreçlerine nasıl entegre edileceğine de odaklanmalıdır.

KAYNAKLAR

- Abbate, S., Centobelli, P., Cerchione, R., Nadeem, S. P., & Riccio, E. (2023). Sustainability trends and gaps in the textile, apparel and fashion industries. *Environment, Development and Sustainability*, 26(2), 2837–2864. <https://doi.org/10.1007/s10668-022-02807-8>
- Acar, A. (2005). *Denim kumaş hatalarının optimizasyonuna yönelik çözüm önerileri* [Yüksek Lisans Tezi]. Marmara Üniversitesi, Fen Bilimleri Enstitüsü.
- Ahn, H., & Chang, T. W. (2019). A Similarity-Based Hierarchical Clustering Method for Manufacturing Process Models. *Sustainability*, 11(9), 2560. <https://doi.org/10.3390/su11092560>
- Akter, N., Repon, M. R., Mikučionienė, D., Jalil, M. A., Islam, T., & Karim, M. R. (2021). Fabrication and characterization of stretchable denim fabric using core spun yarn. *Heliyon*, 7(12), e08532. <https://doi.org/10.1016/j.heliyon.2021.e08532>
- Ala, D. M., & İkiz, Y. (2015). Dokuma Üretimi Süresince Oluşan Kumaş Hatalarının Belirlenmesine Yönelik İstatistiksel Bir Araştırma. *Pamukkale University Journal of Engineering Sciences*, 21(7).
- Alam, S. M. M. (2022). Application of k-means clustering to enhance the quality control and the impact of the productivity in the apparel industries. *International Journal of Research Publication and Reviews*, 2582, 7421.
- Blockeel, H., Devos, L., Frénay, B., Nanfack, G., & Nijssen, S. (2023). Decision trees: from efficient prediction to responsible AI. *Frontiers in Artificial Intelligence*, 6, 1124553. <https://doi.org/10.3389/frai.2023.1124553>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- BSI. (2008). *Specification for the Assessment of the Life Cycle Greenhouse Gas Emissions of Goods and Services* (tek. rap. No. PAS 2050:2008). British Standards Institution.
- Carrilho, R., Yaghoubi, E., Lindo, J., Hambarde, K., & Proença, H. (2024). Toward Automated Fabric Defect Detection: A Survey of Recent Computer Vision Approaches. *Electronics*, 13(18), 3728. <https://doi.org/10.3390/electronics13183728>
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>
- Cortes, C., Vapnik, V., & Saitta, L. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- Coşkun, S., & Doğan, N. A. (2021). Tekstil Endüstrisinde Karbon Ayak İzinin Belirlenmesi. *Süleyman Demirel University Journal of Natural and Applied Sciences*, 25(1), 28–35.
- Çörekcioğlu, M., Ercan, E., & Elibüyük, A. (2021). Yapay Sinir Ağı Yöntemlerinin Tekstil Sektöründe Kullanım Uygulamaları. *Teknik Bilimler Dergisi*, 11(2), 14–20.

- Davis, J., & Goadrich, M. (2006). The relationship between precision-recall and ROC curves. *Proceedings of the 23rd International Conference on Machine Learning (ICML)*, 148, 233–240. <https://doi.org/10.1145/1143844.1143874>
- Dietterich, T. G. (2000). Ensemble methods in machine learning. İçinde *Lecture Notes in Computer Science* (ss. 1–15, C. 1857). Springer. https://doi.org/10.1007/3-540-45014-9_1
- Domingos, P., & Pazzani, M. (1997). On the Optimality of the Simple Bayesian Classifier under Zero-One Loss. *Machine Learning*, 29(2–3), 103–130. <https://doi.org/10.1023/A:1007413511361>
- Eroğlu, D. Y. (2025). Dengesiz Veri Setleri İçin İki Aşamalı Dengeleme Stratejisi: Adasyn İle Örneklem Artırma, SVM Tabanlı Örneklem Azaltma. *Uludağ University Journal of The Faculty of Engineering*, 30(3), 825–844. <https://doi.org/10.17482/uumfd.1650000>
- Ersöz, F., & Çınar, Y. (2021). Veri Madenciliği ve Makine Öğrenimi Yaklaşımlarının Karşılaştırılması: Tekstil Sektöründe bir Uygulama. *Avrupa Bilim ve Teknoloji Dergisi*, 29(29), 397–414. <https://doi.org/10.31590/ejosat.990007>
- Eryuruk, S. H. (2019). The effects of elastane and finishing processes on the performance properties of denim fabrics. *International Journal of Clothing Science and Technology*, 31(2), 243–258. <https://doi.org/10.1108/IJCST-05-2018-0061>
- Feng, Y., Zhou, M., & Tong, X. (2021). Imbalanced classification: A paradigm-based review. *Statistical Analysis and Data Mining*, 14(5), 383–406. <https://doi.org/10.1002/sam.11521>
- Finnveden, G., Hauschild, M. Z., Ekvall, T., Guinée, J., Heijungs, R., Hellweg, S., & Suh, S. (2009). Recent developments in Life Cycle Assessment. *Journal of Environmental Management*, 91(1), 1–21. <https://doi.org/10.1016/j.jenvman.2009.06.018>
- Garza-Reyes, J. A. (2015). Green lean and the need for Six Sigma. *International Journal of Lean Six Sigma*, 6(3), 226–248. <https://doi.org/10.1108/IJLSS-04-2014-0010>
- Genuer, R., Poggi, J. M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14), 2225–2236. <https://doi.org/10.1016/j.patrec.2010.03.014>
- Goutte, C., & Gaussier, E. (2005). A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. İçinde *Advances in Information Retrieval* (ss. 345–359, C. 3408). Springer. https://doi.org/10.1007/978-3-540-31865-1_25
- Gupta, A., Jain, P. K., & Kumar, D. (2013). A novel approach for part family formation using K-means algorithm. *Advances in Manufacturing*, 1(3), 241–250. <https://doi.org/10.1007/s40436-013-0028-9>
- Habib, M. T., Faisal, R. H., Rokonzaman, M., & Ahmed, F. (2014). Automated Fabric Defect Inspection: A Survey of Classifiers. *International Journal in Foundations*

- of Computer Science & Technology*, 4(1), 17–25. <https://doi.org/10.5121/ijfcst.2014.4102>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3. bs.). Morgan Kaufmann Publishers.
- Hanbay, K., Talu, M. F., & Özgüven, Ö. F. (2016). Fabric defect detection systems and methods—A systematic literature review. *Optik*, 127(24), 11960–11973. <https://doi.org/10.1016/j.ijleo.2016.09.110>
- Hasanbeigi, A., & Price, L. (2015). A technical review of emerging technologies for energy and water efficiency and pollution reduction in the textile industry. *Journal of Cleaner Production*, 95, 30–44. <https://doi.org/10.1016/j.jclepro.2015.02.079>
- He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284. <https://doi.org/10.1109/TKDE.2008.239>
- Hossain, K. R., & Rahman, M. F. (2025). Reducing the environmental impact of Denim: A comparative study of green and conventional manufacturing practices. *Textile & Leather Review*, 8, 38–71. <https://doi.org/10.31881/TLR.2025.01>
- IEA. (2025). Emissions Factors 2025 [Erişim tarihi: 22 Haziran 2026].
- ISO. (2006). *Environmental management: Life cycle assessment – Principles and framework* (tek. rap. No. ISO 14040:2006). International Organization for Standardization.
- ISO. (2018). ISO 14067:2018 – Greenhouse gases: Carbon footprint of products [Erişim tarihi: 19 Haziran 2026].
- ISO. (2022). *ISO 8000-1:2022 Data quality – Part 1: Overview* (tek. rap. No. ISO 8000-1:2022). International Organization for Standardization.
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- Kapucuoğlu, D. (2023). *Denim kumaş üretiminde hataların analizi ve mamul ürüne etkisinin incelenmesi* [Yüksek Lisans Tezi]. Tekirdağ Namık Kemal Üniversitesi.
- Kaufman, L., & Rousseeuw, P. J. (2009). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley-Interscience.
- Khatun, M. M. (2024). Life Cycle Assessment of Jeans Production Using Organic and Conventional Cotton. *Tekstilec*, 67(2), 139–150. <https://doi.org/10.14502/tekstilec.67.2024139>
- Kısaoglu, Ö. (2006). Kumaş kalite kontrol sistemleri. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 12(2), 233–241.
- Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. Springer. <https://doi.org/10.1007/978-1-4614-6849-3>
- Lazarova-Molnar, S. (2022). Data-driven Simulation. *2022 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, 18. <https://doi.org/10.23919/SPA53010.2022.9927604>

- Liu, Y., & Guo, S. (2024). Weighted Linear Discriminant Analysis: An Effective Feature Extraction Method for Multi-Class Imbalanced Datasets. *Symmetry*, 16(12), 1656. <https://doi.org/10.3390/sym16121656>
- Machado, R., Barros, L. A. M., Vieira, V., da Silva, F. D., Costa, H., & Carvalho, V. (2025). Textile Defect Detection Using Artificial Intelligence and Computer Vision—A Preliminary Deep Learning Approach. *Electronics*, 14(18), 3692. <https://doi.org/10.3390/electronics14183692>
- Maraş, A., & Erol, Ç. S. (2022). Dengesiz Veri Kümeleriyle Sınıflandırmada Gelişen Trendler: İlerlemenin Bibliyometrik Analizi. *International Journal of Informatics Technologies*, 15(3), 275.
- McQueen, J. B. (1967). Some methods of classification and analysis of multivariate observations. *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, 281–297.
- Menard, S. (2010). *Logistic Regression: From Introductory to Advanced Concepts and Applications*. Sage.
- Murtagh, F., & Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification*, 31(3), 274–295. <https://doi.org/10.1007/s00357-014-9161-z>
- Muthu, S. S. (2020). *Assessing the Environmental Impact of Textiles and the Clothing Supply Chain*. Woodhead Publishing.
- Nobil, E., Cárdenas-Barrón, L. E., de J. Loera-Hernández, I., Smith, N. R., Treviño-Garza, G., Céspedes-Mota, A., & Nobil, A. H. (2023). Sustainability Economic Production Quantity with Warm-Up Function for a Defective Production System. *Sustainability*, 15(2), 1397. <https://doi.org/10.3390/su15021397>
- Periyasamy, A. P., Wiener, J., & Militky, J. (2023). Critical Review on Sustainability in Denim: A Step toward Sustainable Production and Consumption of Denim. *ACS Omega*, 8(5), 4472. <https://doi.org/10.1021/acsomega.2c05553>
- Peterson, A. D., Ghosh, A. P., & Maitra, R. (2018). Merging K-means with hierarchical clustering for identifying general-shaped groups. *Stat*, 7(1), e172. <https://doi.org/10.1002/sta4.172>
- Pınar, Z., Gülağız, F. K., Altuncu, M. A., & Şahin, S. (2020). Denim kumaşlarda görüntü işleme ile hata tespiti. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 9(4), 1609–1620. <https://doi.org/10.17798/bitlisfen.713196>
- Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLOS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- Sandin, G., & Peters, G. M. (2018). Environmental impact of textile reuse and recycling – A review. *Journal of Cleaner Production*, 184, 353–365. <https://doi.org/10.1016/j.jclepro.2018.02.266>

- Seiffert, C., Khoshgoftaar, T. M., Hulse, J. V., & Napolitano, A. (2010). RUSBoost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, 40(1), 185–197. <https://doi.org/10.1109/TSMCA.2009.2029559>
- Semba, T., Furukawa, R., & Itsubo, N. (2024). Carbon Footprint for Jeans' Circular Economy Model Using Bagasse. *Sustainability*, 16(14), 6044. <https://doi.org/10.3390/su16146044>
- Shahid, M. A., Okyay, N., & Babaarslan, O. (2024). A Comparative Analysis of Denim Fabric Performances from Cotton/Polyester Blended Rigid and Stretched Yarns. *Fibers*, 12(10), 86. <https://doi.org/10.3390/fib12100086>
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>
- Tekez, E. K., & Özçelik, T. Ö. (2018). Hazır Giyim Üretiminde Kumaş Kalitesine Etki Eden Faktörlerin Analizi. *Academic Platform – Journal of Engineering and Science*, 6(3), 104–112. <https://doi.org/10.21541/apjes.348506>
- Tercan, H., & Meisen, T. (2022). Machine learning and deep learning based predictive quality in manufacturing: a systematic review. *Journal of Intelligent Manufacturing*, 33(7), 1879–1905. <https://doi.org/10.1007/s10845-022-01963-8>
- Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics*, 17(1), 168–192. <https://doi.org/10.1016/j.aci.2018.08.003>
- Tosun, G., & Dede, Ö. T. (2024). Karbon Ayak İzi Hesaplamalarında Enerji Tüketiminin Örnek Bir Çalışma ile İncelenmesi. *The Black Sea Journal of Sciences*, 14(2), 932–943. <https://doi.org/10.31466/kfbd.1449706>
- UNEP. (2023). *Sustainability and Circularity in the Textile Value Chain: A Global Roadmap* (tek. rap.). United Nations Environment Programme. Paris.
- Uzun, E., Söner, A., Bulut, Y., Kökçam, A. H., & Yılmaz, Ü. (2017). Tekstil sektöründe yapay sinir ağının genetik algoritmayla eğitimi uygulaması.
- Vinodh, S., Ruben, R. B., & Asokan, P. (2015). Life cycle assessment integrated value stream mapping framework to ensure sustainable manufacturing: a case study. *Clean Technologies and Environmental Policy*, 18(1), 279–295. <https://doi.org/10.1007/s10098-015-0982-2>
- Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), 236–244. <https://doi.org/10.1080/01621459.1963.10500845>
- Wiedmann, T., & Minx, J. (2008). A definition of 'carbon footprint'. *Ecological Economics Research Trends*, 1, 1–11.
- WRI & WBCSD. (2015). *GHG Protocol Scope 2 Guidance: An Amendment to the GHG Protocol Corporate Standard* (tek. rap.). World Resources Institute.

Yıldız, K., & Buldu, A. (2017). Kumaş hata tespiti ve sınıflandırmada dalgacık dönüşümü ve temel bileşen analizi. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 23(5), 622–627. <https://doi.org/10.5505/pajes.2016.02258>

ÖZ GEÇMİŞ

KİŞİSEL BİLGİLER

Ad Soyadı : YASEMİN EMRE UÇAR
Öğrenci No : 235120003
E-posta : yaseminnemre@gmail.com
Telefon : 05428128385
Doğum Yeri : Malatya
Doğum Tarihi : 1991-09-26
ORCID : 0000-0002-9872-4183
ABD (TR) : Endüstri Mühendisliği
ABD (EN) : Industrial Engineering
:

EĞİTİM BİLGİLERİ

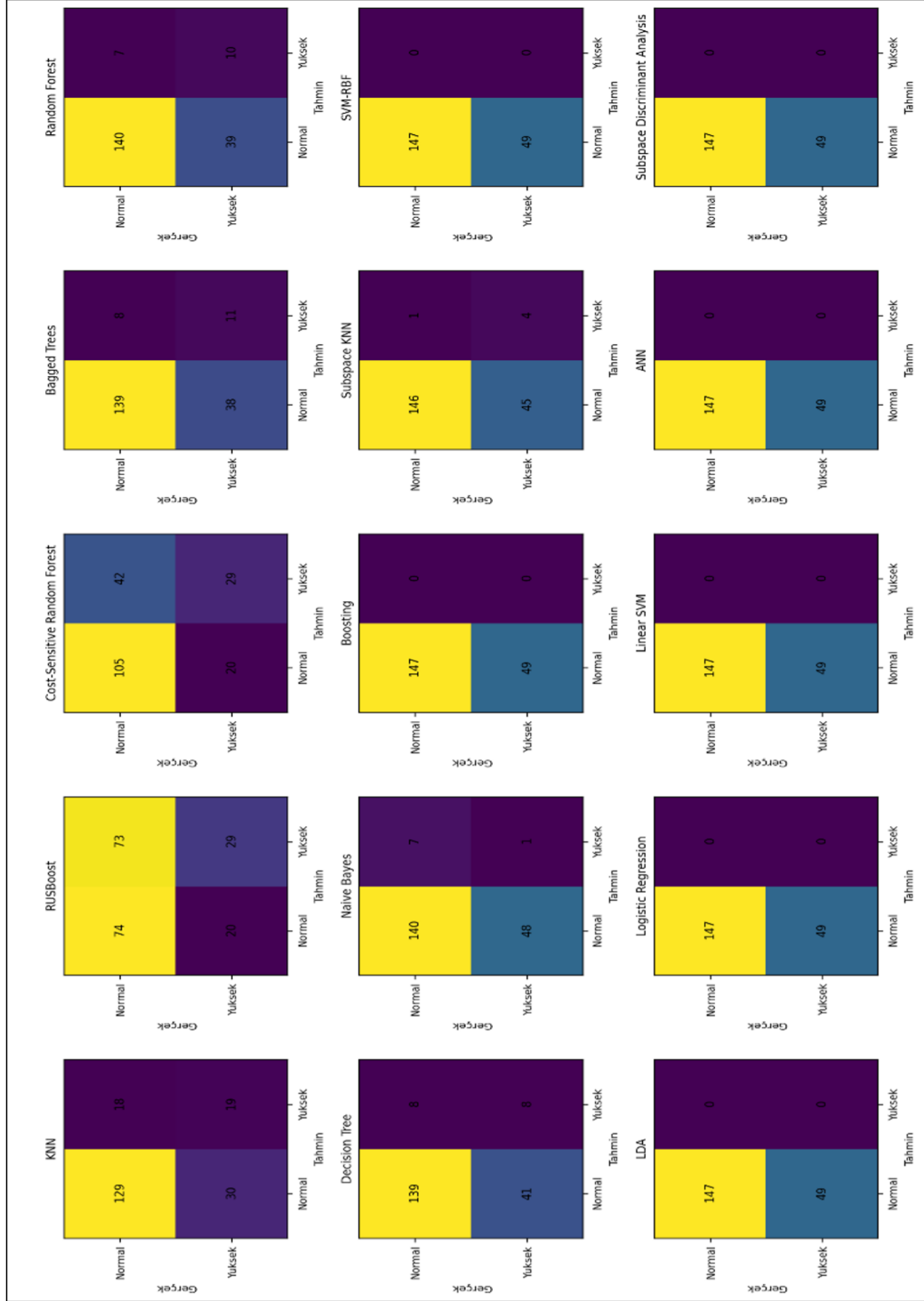
Derece	Adı, Bölüm	Yıl
1.	Erciyes Üniversitesi / Endüstri Mühendisliği	2013-2017

DİĞER BİLGİLER

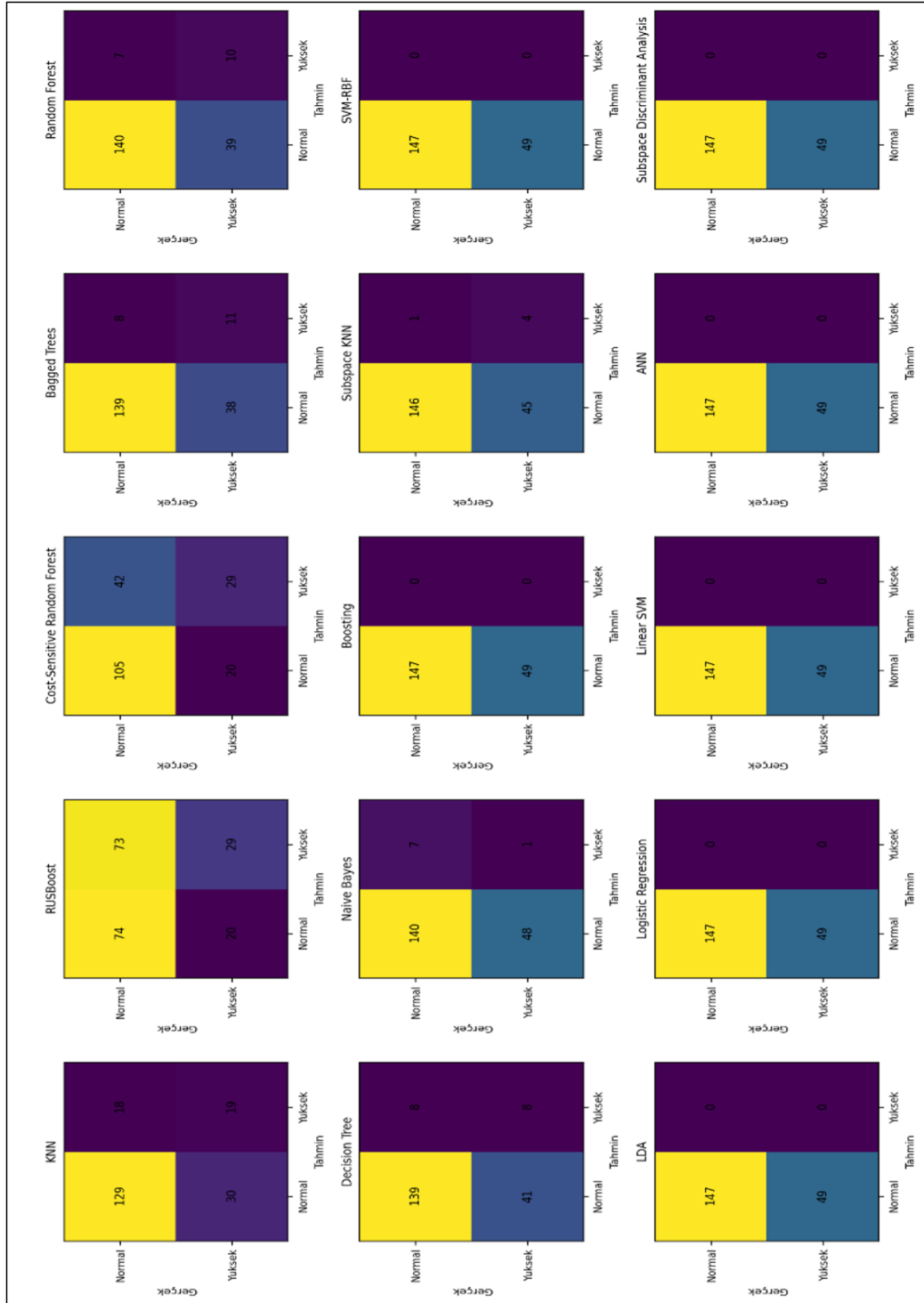
İş Deneyimi;Erzurum Büyükşehir Belediyesi-Endüstri Mühendisi (Erzurum) 2026-
Devam EdiyorBaykan Denim-Üretim Yöneticisi (Malatya) 2020-2025Kaizen PR-Dijital
Operasyonlar ve Süreç Yönetimi Uzmanı (Kayseri) 2018-2020

EKLER

Ek 1. İki sınıflı modeller için confusion matrix çıktıları



Ek 2. İki sınıflı modeller için confusion matrix çıktıları



Ek 3. Üç sınıflı modeller için confusion matrix çıktıları

