

**Tedarik Zinciri Risk Analizinde Metodolojik Bir Anomali ve Optimizasyonu: Kara Kutu Metriğin Sınıflandırma Etkisi ve Özellik Mühendisliği ile Çözümü**

**Meriç Özcan<sup>1</sup>, Mırgavam Gavamı<sup>2\*</sup>, Buket Gürlek<sup>3</sup>**

<sup>1</sup>Ege Üniversitesi, Fen Fakültesi, İstatistik Bölümü

<sup>2</sup>Ege Üniversitesi, Fen Fakültesi, İstatistik Bölümü

<sup>3</sup>Ege Üniversitesi, Fen Fakültesi, İstatistik Bölümü

\*İletişim kurulacak yazar e-mail: mericozcan.edu@gmail.com

**Özet**

*Bu çalışma, kargo taşımacılığındaki kesintileri öngörmek amacıyla geliştirilen yapay zekâ tabanlı bir risk tahmin sisteminin metodolojik analizini sunmaktadır. Yaklaşık 650.000 kayıt içeren dört veri seti birleştirilmiş, 32.065 yüksek riskli gözlem üzerinde modelleme yapılmıştır.*

*İlk modellerde kullanılan “disruption\_likelihood\_score” adlı tescilli kara kutu metriği, modelin  $F1 = 1.00$  düzeyinde yapay bir başarı göstermesine neden olmuştur. Ancak bu skorun, mevcut değişkenlerle açıklanamadığı ( $R^2 < 0$ ) belirlenmiş ve kara kutu bağımlılığı istatistiksel olarak kanıtlanmıştır. Bu bulgu üzerine, **Interaction** ve **Pseudo Index** gibi sentetik öznitelikler üretilmiş; fakat bu yeni değişkenler model performansını artırmamıştır.*

*Söz konusu skor modelden çıkarıldığında, XGBoost sınıflandırma modeli  $F1 = 0.6393$  düzeyinde gerçek performansa ulaşmış ve risk tahminleri yalnızca operasyonel ve çevresel faktörlere dayanmıştır. Özellik önem (feature importance) analizine göre en etkili dört faktör: **hava durumu** (%14,85), **trafik yoğunluğu** (%11,74), **araç tipi** (%11,37) ve **bölge türü** (%8,07) olarak belirlenmiştir.*

*Sonuç olarak, çalışma kara kutu bağımlılığını ortadan kaldırarak **eyleme dönük (actionable)** bir tedarik zinciri optimizasyonunu mümkün kılmıştır. Bu yaklaşım, risk yönetiminde yorumlanabilir yapay zekâyâ dayalı yeni bir metodolojik paradigma önermektedir.*

**Anahtar Kelimeler:** *Tedarik Zinciri Riski, Kara Kutu Metriği, Özellik Mühendisliği, XGBoost, Yorumlanabilirlik, Veri Sızıntısı Önleme*

**A Methodological Anomaly in Supply Chain Risk Analysis and Its Optimization: The Classification Effect and Feature Engineering Resolution of the Black-Box Metric**

**Meriç Özcan<sup>1</sup>, Mırgavam Gavamı<sup>2\*</sup>, Buket Gürlek<sup>3</sup>**

<sup>1</sup>Ege University, Science Faculty, Statistics, Turkey

<sup>2</sup>Ege University, Science Faculty, Statistics, Turkey

<sup>3</sup>Ege University, Science Faculty, Statistics, Turkey

\*Corresponding author e-mail: mericozcan.edu@gmail.com

***Abstract***

*This study presents a methodological analysis of an AI-based risk prediction system developed to anticipate disruptions in cargo transportation. Four datasets comprising approximately 650,000 records were merged, and modeling was performed on 32,065 high-risk observations.*

*The initial models that included a proprietary 'disruption\_likelihood\_score' produced an artificial F1 score of 1.00. However, this metric could not be explained by the existing variables ( $R^2 < 0$ ), statistically proving its black-box dependency. Following this finding, synthetic features such as Interaction and Pseudo Index were engineered, but they did not improve model performance.*

*After removing the score, the XGBoost model achieved a real-world performance level of  $F1 = 0.6393$ , relying solely on operational and environmental factors. Feature importance analysis showed that the most influential factors were Weather (14.85%), Traffic Density (11.74%), Vehicle Type (11.37%), and Region Type (8.07%).*

*Consequently, this study eliminated black-box dependency and enabled actionable supply chain optimization. The proposed approach represents a new methodological paradigm for interpretable AI in risk management.*

***Keywords: Supply Chain Risk, Black-Box Metric, Feature Engineering, XGBoost, Interpretability, Data Leakage Prevention***