



VUCA KOŞULLARINDA TÜRKİYE’NİN SERAMİK SOFRA EŞYALARI İHRACATINDA CİDDİ DARALMA RİSKİ: AÇIKLANABİLİR MAKİNE ÖĞRENMESİ YAKLAŞIMI

Bekir Onur Kınalıbaş^{1*}, Ercan Taşkın²

¹Kütahya Dumlupınar Üniversitesi, Uluslararası Ticaret ve Finansman Anabilim Dalı, Kütahya, Türkiye

²Kütahya Dumlupınar Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Uluslararası Ticaret ve Finansman Bölümü, Kütahya, Türkiye

Corresponding Author: bekir.kinalibas0@ogr.dpu.edu.tr

ÖZET: Bu çalışma, Türkiye’nin seramik sofraya eşyaları ihracatında gelecek yıl ortaya çıkabilecek ciddi daralma riskini incelemektedir. Analiz, 2003–2024 döneminde HS 691110 ve HS 691200 ürün grupları için 168 hedef ülkeyi kapsayan ülke–ürün–yıl verisine dayanmaktadır. Ciddi daralma, bir pazara yapılan ihracatın izleyen yılda en az %30 azalması olarak tanımlanmıştır. Lojistik regresyon, Random Forest, HistGradientBoosting ve XGBoost modelleri zaman sırasını koruyan eğitim, doğrulama ve test dönemlerinde karşılaştırılmıştır. Ortak 975 gözlemlili test örnekleminde geliştirilmiş XGBoost modeli 0,6655 ROC-AUC, 0,4247 PR-AUC ve 0,6260 dengeli doğruluk değerine ulaşmıştır. Yeni model, ilk modele göre daha az yanlış alarm üretmiş; buna karşılık bazı gerçek daralma vakalarını kaçırmıştır. Açıklanabilirlik analizi, geçmiş ihracat eğilimi, oynaklık, birim değer hareketleri, hedef ülkedeki enflasyon ve büyüme ile pazar yoğunlaşmasının tahminlerde öne çıktığını göstermiştir. Bulgular, modelin kesin bir tahmin sistemi olmaktan çok, yakından izlenmesi gereken hedef pazarları belirlemeye yardımcı olan bir erken uyarı aracı olarak kullanılabileceğine işaret etmektedir.

Anahtar Kelimeler: VUCA, seramik sofraya eşyaları, ihracat daralması, pazar riski, XGBoost, erken uyarı

PREDICTING SEVERE EXPORT CONTRACTION RISK IN TÜRKİYE’S CERAMIC TABLEWARE TRADE UNDER VUCA CONDITIONS: AN EXPLAINABLE MACHINE LEARNING APPROACH

ABSTRACT: This study examines the risk of severe export contraction in Türkiye’s ceramic tableware trade. The analysis uses country–product–year data covering 168 destination markets for HS 691110 and HS 691200 during 2003–2024. Severe contractions are defined as a decline of at least 30% in exports to a destination market in the following year. Logistic regression, Random Forest, HistGradientBoosting, and XGBoost models were compared through chronologically separated training, validation, and test periods. On the common test sample of 975 observations, the improved XGBoost model achieved a ROC-AUC of 0.6655, a PR-AUC of 0.4247, and a balanced accuracy of 0.6260. The improved model produced fewer false alarms than the initial model, although it missed some actual contraction cases. The explainability analysis showed that previous export patterns, volatility, unit-value movements, inflation, economic growth, and market concentration were influential in the predictions. The findings

suggest that the model is more suitable as an early-warning and market-prioritization tool than as a precise forecasting system.

Keywords: VUCA, ceramic tableware, export contraction, market risk, XGBoost, early warning

1. GİRİŞ

Uluslararası ticaret, son yıllarda ekonomik dalgalanmaların yanı sıra salgın, enerji maliyetleri, jeopolitik gerilimler, lojistik aksamalar ve ticaret politikalarındaki değişikliklerden etkilenmektedir. Bu gelişmeler ihracatçı işletmeler açısından yalnızca satış hacmini değil, hangi pazarlarda kalınacağına ve hangi pazarlara kaynak ayrılacağına ilişkin kararları da zorlaştırmaktadır. Yönetim yazınında bu çevre; oynaklık, belirsizlik, karmaşıklık ve muğlaklık boyutlarını bir araya getiren VUCA kavramıyla açıklanmaktadır [1].

Dış pazarlarda faaliyet göstermek, müşteri ilişkileri kurma, dağıtım kanalı oluşturma, ürünleri yerel beklentilere uyarılama ve mevzuata uyum sağlama gibi maliyetler doğurur. Gelecekteki talep ve ticaret koşulları belirsizleştğinde firmalar yeni yatırımları erteleyebilir, mevcut pazarlardaki faaliyetlerini daraltabilir veya pazardan çekilebilir. Politika belirsizliğinin ihracat yatırımı ve pazara katılım üzerindeki olumsuz etkileri daha önce ortaya konulmuştur [2]–[4]. Küresel üretim ve dağıtım ağları da bir ülkede başlayan sorunun diğer pazarlara taşınmasını kolaylaştırmaktadır [5].

İhracat riskini yalnızca Türkiye’deki üretim koşullarıyla açıklamak bu nedenle yeterli değildir. Hedef ülkedeki büyüme, enflasyon, ithalat talebi ve kur hareketleri de satışların devamlılığını etkileyebilir. Dünya Belirsizlik Endeksi ülkeler arasında karşılaştırılabilir bir gösterge sunmakla birlikte [6], pazar kırılganlığı tek bir endeksle açıklanamaz. Geçmiş satışların istikrarı, enerji maliyetleri ve ihracatın birkaç ülkeye yoğunlaşması farklı risk kanalları yaratmaktadır.

Türkiye’nin seramik sofa ve mutfak eşyaları ihracatı bu açıdan anlamlı bir inceleme alanıdır. HS 691110 kodlu porselen ürünler ile HS 691200 kodlu porselen dışı seramik ürünler aynı ürün ailesinde yer alsa da hedef pazarları, fiyat yapıları ve daralma eğilimleri bakımından farklılaşabilir. Sektör toplamındaki artış, bazı ülkelerde yaşanan ciddi kayıpları görünmez hâle getirebilir. İşletmeler ve ihracatı destekleyen kurumlar açısından asıl ihtiyaç, hangi ürün–pazar eşleşmelerinin daha yakından izlenmesi gerektiğini önceden görebilmektir.

Bu amaçla kullanılan makine öğrenmesi yöntemleri, çok sayıda göstergenin birlikte değerlendirilmesine ve doğrusal olmayan ilişkilerin yakalanmasına imkân vermektedir. Çalışmada temel model olarak kullanılan XGBoost, ardışık karar ağaçlarını birleştiren ölçeklenebilir bir güçlendirme yöntemidir [7]; bu çalışmada hedef pazarları risk düzeyine göre sınıflandırmak için kullanılmıştır. Modelin hangi bilgilere dayanarak sonuç ürettiğini görmek amacıyla SHAP yöntemi uygulanmıştır [8]. Böylece analiz yalnızca “hangi pazar riskli?” sorusuna değil, “model bu pazarı neden riskli görüyor?” sorusuna da yanıt vermektedir.

Araştırmanın amacı, Türkiye’nin seramik sofa eşyaları ihracatında izleyen yıl en az %30 daralma yaşanması riskini kamuya açık verilerle incelemek ve riskli pazarları ayırmada öne çıkan göstergeleri belirlemektir. Analiz 2003–2024 dönemini, iki ürün grubunu ve 168 hedef ülkeyi kapsamaktadır. Çalışmanın katkısı, toplam ihracat performansından çok pazar düzeyindeki kırılganlığa odaklanması ve elde edilen tahminleri dış ticaret kararları açısından yorumlamasıdır. Bu çerçevede araştırma aşağıdaki sorulara yanıt aramaktadır.

AS1. Kamuya açık ticaret ve makroekonomik veriler kullanılarak izleyen yıldaki ciddi ihracat daralması, pazar düzeyinde anlamlı ölçüde ayırt edilebilir mi?

AS2. Cari dönem bilgileri ile gecikmeli ihracat göstergelerinin birlikte kullanılması modelin ayırt etme performansını geliştirir mi?

AS3. Ciddi daralma riskinin ayırt edilme düzeyi HS 691110 ve HS 691200 ürün grupları arasında farklılaşmakta mıdır?

AS4. Modelin risk sıralamasında hangi ihracat, pazar yapısı ve makroekonomik göstergeler öne çıkmakta; bu göstergelerin SHAP katkıları hangi yönde değişmektedir?

2. LİTERATÜR TARAMASI

VUCA yaklaşımının dört boyutu ihracat kararlarında farklı biçimlerde ortaya çıkar. Talepteki hızlı düşüş oynaklığı, gelecekte uygulanabilecek ticaret önlemlerinin bilinmemesi belirsizliği, maliyet ve pazar koşullarının birbirine bağlı olması karmaşıklığı ifade eder. Yaşanan düşüşün geçici mi kalıcı mı olduğunun anlaşılamaması ise muğlaklık yaratır [1]. Bu ayrım, ihracat riskini tek bir göstergeye indirgemek yerine farklı kaynaklarıyla birlikte değerlendirmeyi gerektirir.

Belirsizlik arttığında firmalar yeni pazarlara girişten kaçınabilir veya mevcut yatırımlarını sınırlandırabilir. Handley ve Limão'nun çalışmaları, politika belirsizliğinin dış pazara giriş ve ticaret hacmi üzerinde caydırıcı olabildiğini göstermektedir [2], [3]. Carballo ve arkadaşları ise ekonomik koşullar ile politika belirsizliğinin birlikte değerlendirilmesi gerektiğini vurgulamaktadır [4]. Bu bulgular, ihracat daralmasının yalnızca fiyat rekabetiyle değil, firmaların geleceğe ilişkin beklentileriyle de ilişkili olduğunu göstermektedir.

Küresel üretim ve tedarik ağları, belirsizliğin etkisini ülkeler arasında taşıyabilir. Salgın dönemine ilişkin araştırmalar, üretim ağlarının ekonomik daralmanın yayılmasında önemli rol oynadığını ortaya koymuştur [5]. Dolayısıyla bir hedef pazardaki risk değerlendirilirken o ülkenin ekonomik koşulları kadar küresel maliyet baskılarının da göz önünde bulundurulması gerekir.

Ürün-ülke düzeyindeki ticaret ilişkileri her zaman uzun süre devam etmez. Ürün niteliğinin ticaret ilişkisinin süresini etkilediği daha önce gösterilmiştir [9]. Bu durum, aynı ürün ailesinde yer alan porselen ve porselen dışı seramik ürünlerin farklı pazar dayanıklılığına sahip olabileceğini düşündürmektedir. Toplam ihracatın artması da bütün pazarlarda olumlu bir görünüm bulunduğu anlamına gelmez. Nitekim dış pazarların iç talep daralmasının etkisini kısmen dengeleyebildiği gösterilmiştir [10].

Pazar çeşitlendirmesi, tek bir ülkeye bağımlılığı azaltarak şoklara karşı hareket alanı sağlayabilir. Buna karşılık çok sayıda küçük pazara dağılmak yönetim ve dağıtım maliyetlerini artırabilir. Bu nedenle yoğunlaşma ile risk arasında her durumda geçerli basit bir ilişki beklenmemelidir. Çalışmada aktif pazar sayısı, pazar payları, ilk beş ülkenin toplam payı ve çeşitlilik göstergeleri birlikte ele alınmıştır.

Kur ve birim değer hareketleri de pazar ilişkilerinin sürdürülmesinde önemlidir. Kur değişimleri firmaların fiyat, miktar ve ürün kapsamı kararlarına farklı biçimlerde

yansıyabilmektedir [11]. Ancak birim değerdeki artış doğrudan daha yüksek kalite veya kârlılık anlamına gelmez. Bu gösterge, fiyatlama ve ürün bileşimindeki değişimleri yaklaşık olarak yansıtan bir işaret olarak değerlendirilmiştir.

İhracat daralması, çok sayıda koşulun birlikte etkili olduğu ve bazı etkilerin belirli eşiklerden sonra güçlendiği bir süreçtir. Bu nedenle yalnızca doğrusal ilişkilere dayanan modeller her örüntüyü yakalayamayabilir. Random Forest ve gradyan artırma yöntemleri, değişkenler arasındaki daha karmaşık ilişkileri incelemeye imkân verir [12], [13]. XGBoost ise bu yaklaşımın yaygın kullanılan uygulamalarından biridir [7].

Daralma vakalarının toplam gözlemler içindeki payı sınırlı olduğundan genel doğruluk tek başına yeterli değildir. Dengesiz sınıflarda kesinlik-duyarlılık ölçütlerinin daha açıklayıcı olabildiği dikkate alınarak [14], modelin gerçek daralmaları yakalama ve yanlış alarm üretme düzeyi birlikte değerlendirilmiştir. Üretilen risk puanlarının gerçekleşen sonuçlarla uyumu da olasılıksal tahminlerin değerlendirilmesinde kullanılan Brier yaklaşımı çerçevesinde incelenmiştir [15].

Son yıllardaki çalışmalar, makine öğrenmesinin dış ticaret ve sınır ötesi risk analizinde yalnızca tahmin aracı olarak değil, karmaşık ilişkilerin yorumlanmasını sağlayan bir çerçeve olarak da kullanılmaya başladığını göstermektedir. Silva, Wilhelm ve Amancio [16], uluslararası ticaret ağlarından türetilen göstergeleri ekonomik büyüme tahminlerine dâhil etmiş; doğrusal olmayan modelleri SHAP ile açıklayarak ticaret ağı yapısının tahmin bilgisini artırabildiğini göstermiştir. Çalışmanın bağımlı değişkeni ihracat daralması değildir; ancak ticaret ağları, doğrusal olmayan tahmin ve açıklanabilirlik arasındaki bağ bakımından mevcut araştırmaya yakın bir yöntemsel örnek sunmaktadır.

Risk ve erken uyarı yazınında da benzer bir yönelim vardır. Shi ve arkadaşları [17], sınır ötesi egemen risk aktarımını XGBoost-SHAP yaklaşımıyla inceleyerek risk kanallarının farklı koşullarda aynı yönde çalışmadığını göstermiştir. Purnell, Etemadi ve Kamp [18] ise finansal ağlarda erken uyarı amacıyla makine öğrenmesi ile Shapley değerlerini bir araya getirmiş ve karar vericiler açısından açıklanabilirliğin önemini özellikle vurgulamıştır. Bu çalışmalar doğrudan ürün-ülke düzeyinde ihracat daralmasını tahmin etmemektedir. Dolayısıyla Türkiye'nin seramik sofa eşyaları ihracatında izleyen yıl yaşanacak belirgin kaybı, zaman sırasını koruyan bir test düzeni ve açıklanabilir makine öğrenmesiyle ele alan sektörel boşluk devam etmektedir.

3. YÖNTEM

3.1. Araştırma Tasarımı ve Veri Kaynakları

Araştırma nicel ve tahmin odaklı bir panel veri tasarımına sahiptir. Gözlem birimi hedef ülke-ürün-yıl birleşimidir. Analiz, Türkiye'nin HS 691110 kodlu porselen sofa ve mutfak eşyaları ile HS 691200 kodlu diğer seramik sofa ve mutfak eşyaları ihracatını kapsamaktadır. Bütün veriler açık erişimli ve toplulaştırılmış ikincil kaynaklardan alınmış, işlemler Python ortamında yürütülmüştür.

Yıllık ihracat verileri UN Comtrade'den alınmıştır [19]. Veri temizliği sonrasında 4.715 pozitif ticaret kaydı kalmış; ticaret yapılmayan ara yılların da izlenebilmesi için yıllık seri

tamamlanmıştır. Böylece 168 hedef ülkeyi ve 5.212 ülke–ürün–yıl gözlemini kapsayan panel oluşturulmuştur. Hedef ülkelere ait büyüme, kişi başına gelir, enflasyon, nüfus, dışa açıklık, döviz kuru ve ithalat göstergeleri Dünya Bankası’ndan elde edilmiştir [20]. Küresel belirsizlik Dünya Belirsizlik Endeksiyle [6], enerji maliyetleri Avrupa doğal gaz fiyatlarıyla temsil edilmiştir [21].

3.2. Daralma Ölçütü ve Kullanılan Göstergeler

Cari yılda en az 1.000 ABD doları ihracat yapılan ülke–ürün gözlemleri aktif pazar kabul edilmiştir. Ciddi daralma etiketi, bilgi yılı t ile izleyen sonuç yılı $t+1$ açık biçimde eşleştirilerek oluşturulmuştur: aktif bir pazarda $t+1$ yılındaki ihracatın t yılına göre en az %30 azalması durumunda ciddi daralma=1, diğer durumlarda 0 olarak kodlanmıştır. İhracatın $t+1$ yılında 1.000 ABD dolarının altına düşmesi ayrıca pazar kaybı göstergesi olarak hesaplanmıştır.

Veri tabanı 2024 yılında sona erdiği için, 2024 yılına ait açıklayıcı değişkenlerle 2025 daralması üretilmemiştir; çünkü 2025 gerçekleşmesi gözlenmemektedir. Test dönemindeki “2024” ifadesi sonuç yılını göstermektedir: bu gözlemlerde model 2023 ve daha önceki bilgilerle kurulmuş açıklayıcı değişkenleri kullanmakta, hedef değişken ise 2023’ten 2024’e gerçekleşen ihracat değişiminden oluşturulmaktadır. Başka bir ifadeyle 2024 yılı, 2025’i tahmin eden bir bilgi yılı olarak değil, 2023 bilgilerinin karşılaştırıldığı gözlenmiş hedef yıl olarak yer almaktadır. Bu eşleştirme, mevcut olmayan 2025 bilgisinin modele dolaylı biçimde girmesini önlemektedir.

Tahminlerde pazarın geçmiş ve cari ihracat düzeyi, büyüme eğilimi, son üç yıldaki oynaklık, birim değer hareketleri, pazar payı, pazarda geçirilen süre, pazar yoğunlaşması ve çeşitlilik dikkate alınmıştır. Hedef ülkelerin ekonomik koşulları büyüme, kişi başına gelir, enflasyon, nüfus, ithalat, dışa açıklık ve kur değişimleriyle temsil edilmiştir. Finansal kriz, salgın ve enerji baskısı dönemlerini gösteren zaman değişkenleri de ortak dönem etkilerini dikkate almak için kullanılmıştır.

3.3. Zaman Ayrımı ve Hedef Yılın Eşleştirilmesi

Zaman ayrımı açıklayıcı değişkenlerin bulunduğu bilgi yılına göre değil, tahmin edilen sonuç yılı ($t+1$) esas alınarak yapılmıştır. Sonuç yılı 2017 ve öncesi olan gözlemler eğitim, 2018–2020 olanlar doğrulama, 2021–2024 olanlar test kümesine ayrılmıştır. Buna karşılık gelen son bilgi yılları sırasıyla 2016 ve öncesi, 2017–2019 ve 2020–2023’tür. Geliştirilmiş modelde eğitim, doğrulama ve test kümeleri 2.196, 643 ve 1.002 gözlemde oluşmaktadır. Rastgele bölme yerine kronolojik ayırım, değerlendirmeyi gerçek ileri dönem tahmin düzenine yaklaştırmak ve zaman sırasını korumak amacıyla tercih edilmiştir [22].

3.4. Model Seçimi, Hiperparametreler ve Test Setinin Korunması

İlk aşamada lojistik regresyon, Random Forest, HistGradientBoosting ve XGBoost modelleri karşılaştırılmıştır. Sınıf dengesizliği nedeniyle model seçiminde genel doğruluktan çok ciddi daralma sınıfını ayırt eden ölçütlere ağırlık verilmiştir. Nihai geliştirme aşamasında XGBoost için 53 açıklayıcı değişken içeren özellik seti kullanılmış ve zaman temelli doğrulama örnekleminde 80 hiperparametre bileşimi değerlendirilmiştir. Seçim puanı $0,60 \times \text{PR-AUC} + 0,25 \times \text{F1} + 0,15 \times \text{ROC-AUC}$ olarak tanımlanmış; böylece azınlık sınıfındaki daralmaları ayırt etme başarısı seçimde belirleyici kılınmıştır.

En yüksek doğrulama puanına ulaşan XGBoost yapılandırmasında maksimum ağaç derinliği 3, minimum çocuk ağırlığı 3, öğrenme oranı 0,02, alt örneklem oranı 0,90, değişken örnekleme oranı 0,75, L1 düzenlileştirme katsayısı 0,70 ve L2 katsayısı 3,00'dır. Sınıflandırma eşiği de test sonuçları görülmeden önce belirlenmiştir: doğrulama örnekleminde kesinliğin en az 0,30 olması koşuluyla F1 skorunu en yüksek yapan 0,4736 eşiği seçilmiş ve testte değiştirilmeden uygulanmıştır.

Nihai modelleme protokolünde test kümesi kilitli tutulmuştur. Test gözlemleri özelliklerin veri bağımlı biçimde seçilmesinde, eksik değer işlemlerinin veya uç değer sınırlarının belirlenmesinde, algoritma tercihinde, hiperparametre aramasında ya da karar eşiğinin saptanmasında kullanılmamıştır. Bu işlemler eğitim verisi ve gerektiğinde doğrulama dönemiyle sınırlandırılmış; 2021–2024 sonuç yıllarını içeren test kümesi yalnızca seçilmiş modelin ileri dönem performansını raporlamak için açılmıştır. Tablo 2'deki test sonuçları bu nedenle model seçme ölçütü değil, seçilmiş modellerin sonradan yapılan karşılaştırmalı performans raporudur.

3.5. Performans Ölçütleri ve Açıklanabilirlik

Model başarısı ROC-AUC, PR-AUC, dengeli doğruluk, kesinlik, duyarlılık, özgüllük, F1 ve Brier skoru üzerinden değerlendirilmiştir [14], [15]. İlk ve geliştirilmiş modeller ortak 975 test gözleminde karşılaştırılmış; performans farklarının örnekleme duyarlılığı 2.000 tekrarlı eşleştirilmiş bootstrap analiziyle incelenmiştir. Ürün grupları için ayrıca ayrı modeller kurulmuştur. Açıklanabilirlik analizinde ortalama mutlak SHAP değerleri değişken katkısının büyüklüğünü, işaretli SHAP değerleri ise katkının risk puanını ortalama olarak hangi yöne taşıdığını göstermek için kullanılmıştır [8]. Bu değerler nedensel katsayı olarak yorumlanmamış; yüksek mutlak katkıya karşın işaretli ortalaması sıfıra yakın değişkenler, katkının gözlemler arasında yön değiştirebildiğine ve doğrusal olmayan eşik veya etkileşim örüntülerinin bulunabileceğine ilişkin tanımlayıcı bir işaret olarak ele alınmıştır.

4. BULGULAR

4.1. Veri Setinin Genel Görünümü

Nihai panel, 2003–2024 döneminde 168 hedef ülke ve iki ürün grubuna ait 5.212 gözlemden oluşmaktadır. İlk modelleme setinde ciddi daralma oranı %32,22, pazar kaybı oranı %7,98'dir. Cari dönem ve geçmiş eğilimi birlikte gösteren yeni değişkenlerin eklenmesiyle geliştirilmiş modelde kullanılabilir gözlem sayısı 3.841'e yükselmiştir.

Tablo 1. Veri seti ve modelleme örneklemi

Gösterge	Değer
Dönem	2003–2024
Ürün grubu	2
Hedef ülke	168
Panel gözlemi	5.212

Gösterge	Değer
Geliştirilmiş model gözlemi	3.841
Eğitim / doğrulama / test	2.196 / 643 / 1.002
Ortak test örnekleme	975
Ortak testte daralma	274 (%28,10)

4.2. Modellerin Karşılaştırılması

İlk aşamada tek bir algoritma bütün ölçütlerde üstün değildir. Random Forest en yüksek ROC-AUC ve duyarlılığı, HistGradientBoosting ise en yüksek dengeli doğruluk ve F1 değerini üretmiştir. Lojistik regresyon iki grubu yeterince ayıramamıştır. Tablo 2'deki test sonuçları nihai XGBoost yapılandırmasının seçimi için kullanılmamış; bu seçim 3.4'te açıklanan doğrulama prosedürüyle yapılmıştır.

Tablo 2. İlk sınıflandırma modellerinin test performansı

Model	ROC-AUC	PR-AUC	Dengeli doğruluk	Kesinlik	Duyarlılık	F1
XGBoost	0,6407	0,3812	0,5897	0,3354	0,7956	0,4719
Random Forest	0,6457	0,3812	0,5519	0,3049	0,9526	0,4619
HistGradientBoosting	0,6335	0,3734	0,5994	0,3393	0,8321	0,4820
Lojistik regresyon	0,5395	0,2797	0,5064	0,2836	1,0000	0,4419

4.3. Geliştirilmiş Modelin Sonuçları

Cari dönem göstergeleri ve pazarın yakın geçmişini daha iyi yansıtan değişkenler eklendiğinde modelin ayırt etme gücü yükselmiştir. Doğrulama ve test dönemlerindeki ROC-AUC değerlerinin birbirine yakın olması, performansın test döneminde belirgin biçimde bozulmadığını göstermektedir.

Tablo 3. Geliştirilmiş modelin doğrulama ve test performansı

Örneklem	N	ROC-AUC	PR-AUC	Dengeli doğruluk	Kesinlik	Duyarlılık	F1
Doğrulama	643	0,6824	0,5618	0,6631	0,5388	0,5946	0,5653
Test	1.002	0,6823	0,4665	0,6354	0,4076	0,6881	0,5120

İlk ve geliştirilmiş modelin adil biçimde karşılaştırılabilmesi için iki model ortak 975 gözlem üzerinde yeniden değerlendirilmiştir. Yeni model daha yüksek kesinlik ve özgüllük üretmiş, yanlış pozitif sayısını 432'den 289'a düşürmüştür. Buna karşılık duyarlılık gerilemiş ve kaçırılan gerçek daralma sayısı 56'dan 92'ye çıkmıştır. Başka bir ifadeyle model daha seçici hâle gelmiş, ancak bu seçiciliğin bir maliyeti olmuştur.

Tablo 4. İlk ve geliştirilmiş XGBoost modellerinin ortak test örneklemindeki performansı

Model	ROC-AUC	PR-AUC	Dengeli doğruluk	Kesinlik	Duyarlılık	Özgüllük	F1
İlk XGBoost	0,6407	0,3812	0,5897	0,3354	0,7956	0,3837	0,4719

Model	ROC-AUC	PR-AUC	Dengeli doğruluk	Kesinlik	Duyarlılık	Özgüllük	F1
Geliştirilmiş XGBoost	0,6655	0,4247	0,6260	0,3864	0,6642	0,5877	0,4886

4.4. Sonuçların Kararlılığı ve Uygulamadaki Anlamı

Eşleştirilmiş bootstrap analizi, yeni modeldeki hangi değişimlerin daha kararlı olduğunu göstermektedir. PR-AUC, dengeli doğruluk, kesinlik ve özgüllükteki iyileşmelerin güven aralıkları sıfırı içermemiştir. ROC-AUC ve F1 değerleri yükselmiş olsa da bu iki ölçütteki farklar daha ihtiyatlı yorumlanmalıdır. Duyarlılıktaki düşüş ise belirgindir.

Tablo 5. Eşleştirilmiş bootstrap ile temel performans farkları

Gösterge	Fark	%95 alt sınır	%95 üst sınır
ROC-AUC	0,0248	-0,0053	0,0573
PR-AUC	0,0435	0,0083	0,0772
Dengeli doğruluk	0,0363	0,0049	0,0688
Kesinlik	0,0510	0,0260	0,0777
Duyarlılık	-0,1314	-0,1843	-0,0784
Özgüllük	0,2040	0,1683	0,2396
F1	0,0167	-0,0142	0,0491

Modelin ortalama risk tahmini, ortak test örnekleminde gözlenen %28,10 daralma oranının üzerinde kalmıştır. Bu nedenle modelin ürettiği değerler “kesin daralma olasılığı” olarak değil, pazarları görel risk düzeyine göre sıralayan işaretler olarak okunmalıdır. Benzer biçimde karar eşliğinin düşürülmesi daha fazla daralmayı yakalamakta, fakat uyarı sayısını aşırı artırmaktadır. Eşğin yükseltilmesi ise yanlış alarmları azaltırken gerçek daralmaların önemli bir bölümünü gözden kaçırmaktadır. Uygun eşik, kurumun yanlış alarm ile kaçırılan risk arasında kurduğu dengeye bağlıdır.

Performans değerlerinin uygulamadaki karşılığı ihtiyatlı biçimde okunmalıdır. Ortak test örneklemindeki 0,6655 ROC-AUC, rastgele seçilen bir daralma gözlemi ile daralma yaşamayan bir gözlem karşılaştırıldığında modelin riskli olanı yaklaşık üçte iki oranında daha üst sıraya yerleştirebildiğini ifade eder. PR-AUC’nin 0,4247 olması, pozitif sınıfın test örneklemindeki %28,10’luk payının üzerinde bir ayırt etme bilgisi bulunduğunu gösterir; ancak bu düzey otomatik karar vermek için güçlü bir ayırım anlamına gelmez. Seçilmiş eşikte kesinlik 0,3864 ve duyarlılık 0,6642’dir. Dolayısıyla modelin daha uygun kullanım alanı, bütün pazarlar hakkında kesin hüküm vermek değil, analist incelemesine alınacak ülke–ürün eşleşmelerini önceliklendirmektir.

Bu çerçevede model; ülke masaları, ihracatçı birlikleri veya firma satış ekipleri için bir ön tarama katmanı olarak kullanılabilir. Yüksek risk puanı alan bir pazarın sipariş akışı, müşteri yoğunluğu, ödeme koşulları, mevzuat değişiklikleri ve yerel talep görünümü ayrıca incelenmeden pazardan çıkış, fiyat değişikliği veya kaynak aktarımı kararı verilmemelidir. Modelin değeri, sınırlı izleme kapasitesini görel olarak daha kırılgan görünen pazarlara yönlendirmesinde ortaya çıkmaktadır.

4.5. Ürün Grupları ve Öne Çıkan Göstergeler

Ürün bazında kurulan modeller aynı performansı göstermemiştir. HS 691110 kodlu porselen ürünlerde daha güçlü bir ayırt etme düzeyi elde edilmiştir. HS 691200 modelinde duyarlılık çok yüksek görünse de bu sonuç çok düşük karar eşiği ve yoğun yanlış alarm üretimiyle birlikte gerçekleşmiştir. Dolayısıyla iki ürün grubuna aynı karar kuralının uygulanması uygun değildir.

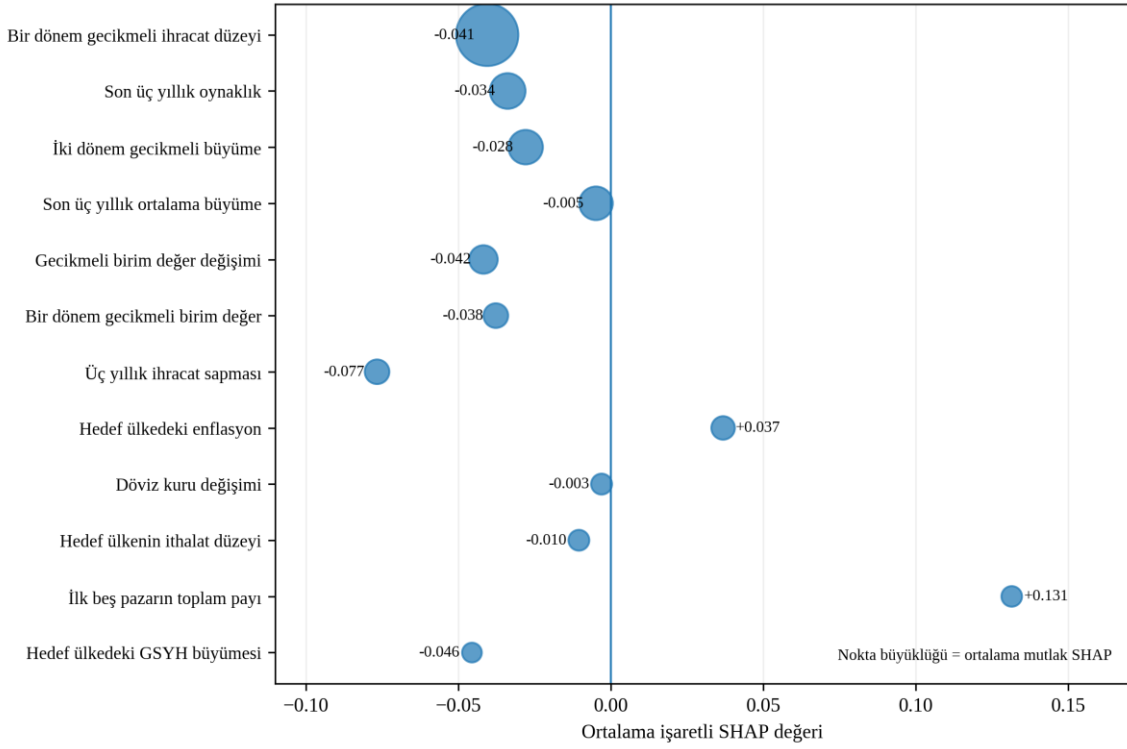
Tablo 6. Ürün grupları için ayrı modellerin performansı

Ürün	Test N	ROC-AUC	PR-AUC	Dengeli doğruluk	Kesinlik	Duyarlılık	F1
HS 691110	524	0,7192	0,5017	0,6658	0,4170	0,6901	0,5199
HS 691200	478	0,6689	0,4601	0,5423	0,3402	0,9739	0,5042

Açıklanabilirlik analizi, modelin yalnızca genel ekonomik göstergelere dayanmadığını göstermiştir. Pazarın geçmiş ihracat düzeyi, son üç yıldaki oynaklık, büyüme eğilimi ve birim değer hareketleri en önemli bilgi kaynakları arasındadır. Hedef ülkedeki enflasyon ve ekonomik büyüme ile Türkiye'nin ihracatının birkaç büyük pazarda yoğunlaşması da risk ayırımına katkı sağlamıştır.

Tablo 7. Modelin tahminlerinde öne çıkan göstergeler

Sıra	Gösterge	Ortalama mutlak SHAP
1	Bir dönem gecikmeli ihracat düzeyi	0,297761
2	Son üç yıllık oynaklık	0,172327
3	İki dönem gecikmeli büyüme	0,168396
4	Son üç yıllık ortalama büyüme	0,165517
5	Gecikmeli birim değer değişimi	0,150178
6	Bir dönem gecikmeli birim değer	0,140435
7	Üç yıllık ihracat sapması	0,140259
8	Hedef ülkedeki enflasyon	0,137756
9	Döviz kuru değişimi	0,132185
10	Hedef ülkenin ithalat düzeyi	0,132031
11	İlk beş pazarın toplam payı	0,131568
12	Hedef ülkedeki GSYH büyümesi	0,129809



Şekil 1. Öne Çıkan Değişkenlerde SHAP Katkısının Ortalama Yönü ve Büyüklüğü

Not: Yatay eksen ortalama işaretli SHAP değerini, nokta büyüklüğü ortalama mutlak SHAP değerini göstermektedir. SHAP değerleri nedensel etki değildir.

Şekil 1, değişken önem sıralamasına ortalama katkı yönünü eklemektedir. İlk beş pazarın toplam payının ortalama işaretli SHAP değeri +0,1315, enflasyonunki +0,0368'dir; bu değişkenlerin SHAP katkıları test örnekleminde ortalama olarak risk puanını yukarı taşımıştır. Gecikmeli ihracat, oynaklık ve birim değer göstergelerinde ortalama işaret negatiftir. Bu değerler sabit marjinal etki değildir ve değişken düzeyiyle risk arasında tekdüze bir ilişki göstermez.

Son üç yıllık ortalama büyüme (ortalama mutlak SHAP=0,1655; ortalama işaretli SHAP=-0,0048) ve döviz kuru değişimi (0,1322; -0,0030), yüksek katkı büyüklüğüne rağmen sıfıra yakın ortalama yön göstermektedir. Bu durum, pozitif ve negatif katkıların gözlemler arasında dengelenebildiğini gösterir. Ortalama işaretli SHAP ilişkisinin fonksiyonel biçimini tek başına ortaya koymadığından, doğrusal olmayanlık veya belirli eşikler bu şekilden çıkarılmamıştır.

4.6. Bulguların Genel Değerlendirmesi

Bulgular, seramik sofraya eşyaları ihracatındaki ciddi daralmaların geçmiş pazar davranışı ile hedef ülke koşulları birlikte değerlendirildiğinde belirli ölçüde ayırt edilebildiğini göstermektedir. Geliştirilmiş model yüksek doğruluklu bir tahmin sistemi değildir. Buna rağmen ilk modele göre daha az yanlış alarm üretmesi, sınırlı izleme kaynaklarının daha dar bir pazar grubuna yöneltilmesi açısından önemlidir. Modelin göreceli risk sıralaması, firma içi satış bilgileri ve saha değerlendirmeleriyle birleştirildiğinde daha işlevsel hâle gelebilir.

5. TARTIŞMA VE SONUÇ

5.1. Bulguların Uygulamadaki Karşılığı

Bu çalışma, Türkiye'nin seramik sofa ve mutfak eşyaları ihracatında gelecek yıl yaşanabilecek ciddi daralma riskini hedef ülke ve ürün düzeyinde incelemiştir. Geliştirilmiş XGBoost modeli ortak test örnekleminde 0,6655 ROC-AUC ve 0,4247 PR-AUC üretmiştir. Bu değerler risk sıralamasının rastgele bir ayırımdan daha fazla bilgi taşıdığını, fakat yüksek doğruluklu ve kendi başına karar verebilecek bir tahmin sistemi düzeyine ulaşmadığını göstermektedir. Son yıllarda uluslararası ticaret ağlarının makine öğrenmesiyle tahmin edilmesi ve SHAP ile açıklanması [16], sınır ötesi risk kanallarının XGBoost-SHAP ile incelenmesi [17] ve açıklanabilir erken uyarı sistemlerinin geliştirilmesi [18], karmaşık ekonomik risklerde performans kadar yorumlanabilirliğin de önem kazandığını göstermektedir. Bu çalışmanın bulguları da aynı doğrultuda, tahmin skorunun tek başına sonuç değil, daha ayrıntılı pazar incelemesini başlatan bir sinyal olarak kullanılmasının daha uygun olduğuna işaret etmektedir.

Modelin asıl kazanımı yanlış alarm yapısındaki iyileşmedir. Aynı 975 test gözleminde yanlış pozitif sayısı 432'den 289'a inerken özgüllük 0,3837'den 0,5877'ye yükselmiştir; buna karşılık duyarlılık 0,7956'dan 0,6642'ye gerilemiştir. Bu değişim, erken uyarı sistemlerinde tek bir performans ölçütünün yeterli olmadığını gösterir. Çok sayıda pazarı izleyebilen bir kurum daha düşük eşikle daha fazla gerçek daralmayı yakalamayı tercih edebilir; analist kapasitesi sınırlı bir kullanıcı ise daha seçici bir uyarı listesine ihtiyaç duyabilir. Ürün gruplarındaki farklı performans da aynı karar eşiğinin HS 691110 ve HS 691200 için mekanik biçimde uygulanmaması gerektiğini göstermektedir.

5.2. Kurumsal ve Politika Çıkarımları

Uygulama açısından en doğrudan kullanım alanı pazar önceliklendirmesidir. Ticaret Bakanlığı düzeyinde ülke-ürün bazlı risk puanları, mevcut dış ticaret istihbaratı ve ticaret müşavirliklerinden gelen nitel bilgilerle birlikte değerlendirilen bir erken uyarı ekranının girdilerinden biri olabilir. Böyle bir kullanımda model, destek verilecek veya terk edilecek pazarları otomatik olarak belirlemek yerine, hangi ülke-ürün eşleşmelerinde daha ayrıntılı inceleme gerektiğini işaretlemelidir. İhracatçı birlikleri aynı yaklaşımı sektör üyelerine yönelik ülke izleme raporlarında, pazar çeşitlendirme çalışmalarında ve dönemsel risk bilgilendirmelerinde kullanabilir. Özellikle aynı ülkeye ilişkin riskin iki ürün grubunda farklılaşması, sektörel uyarıların ürün ayırımını korumasını gerekli kılmaktadır.

Firma düzeyinde ise model çıktısının sipariş portföyü, müşteri yoğunluğu, tahsilat koşulları, dağıtım kanalı ve fiyatlama bilgileriyle birleştirilmesi daha anlamlıdır. Yüksek risk puanı, doğrudan pazardan çıkış sinyali olarak değil; satış ekibinin ilgili pazardaki müşteri kaybı, sipariş erteleme veya fiyat baskısı gibi göstergeleri kontrol etmesini tetikleyen bir öncelik sinyali olarak kullanılabilir. Bu ayırım önemlidir; çünkü kullanılan kamu verileri firma içi sözleşme ve müşteri bilgilerini içermemektedir.

5.3. Sınırlılıklar ve Sonuç

Çalışmanın temel sınırlılığı, verilerin ülke ve ürün düzeyinde toplulaştırılmış olmasıdır. Firma büyüklüğü, müşteri yapısı, sözleşme süresi, marka konumu ve finansal dayanıklılık gibi unsurlar modele dâhil edilememiştir. Birim değer kalite veya kârlılığın doğrudan ölçüsü

değildir. Test dönemi ileri tarihli gözlemleri içerdiği için zaman bütünlüğünü korumakla birlikte, tek bir sektör ve iki HS ürün grubuyla sınırlıdır. Ayrıca SHAP açıklamaları modelin tahmin mekanizmasını betimler; nedensel etki göstermez. Gelecek araştırmalarda firma ve işlem düzeyindeki veriler, aylık ticaret hareketleri, navlun maliyetleri ve jeopolitik risk göstergeleriyle dış doğrulama yapılabilir. Sonuç olarak geliştirilen yaklaşım ciddi ihracat daralmalarını kusursuz biçimde öngören bir sistem değildir. Daha savunulabilir kullanım, kamuya açık verilerden hareketle yakından izlenmesi gereken ürün-pazar eşleşmelerini sıralayan, gerekçesi incelenebilir bir erken uyarı katmanı oluşturmaktır.

KAYNAKLAR

- [1] Bennett, N., & Lemoine, G. J. (2014). What a difference a word makes: Understanding threats to performance in a VUCA world. *Business Horizons*, 57(3), 311–317. <https://doi.org/10.1016/j.bushor.2014.01.001>
- [2] Handley, K., & Limão, N. (2015). Trade and investment under policy uncertainty: Theory and firm evidence. *American Economic Journal: Economic Policy*, 7(4), 189–222. <https://doi.org/10.1257/pol.20140068>
- [3] Handley, K., & Limão, N. (2017). Policy uncertainty, trade, and welfare: Theory and evidence for China and the United States. *American Economic Review*, 107(9), 2731–2783. <https://doi.org/10.1257/aer.20141419>
- [4] Carballo, J., Handley, K., & Limão, N. (2022). Economic and policy uncertainty: Aggregate export dynamics and the value of agreements. *Journal of International Economics*, 139, 103661. <https://doi.org/10.1016/j.jinteco.2022.103661>
- [5] Bonadio, B., Huo, Z., Levchenko, A. A., & Pandalai-Nayar, N. (2021). Global supply chains in the pandemic. *Journal of International Economics*, 133, 103534. <https://doi.org/10.1016/j.jinteco.2021.103534>
- [6] Ahir, H., Bloom, N., & Furceri, D. (2022). The world uncertainty index (NBER Working Paper No. 29763). National Bureau of Economic Research. <https://doi.org/10.3386/w29763>
- [7] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY: Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939785>
- [8] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 30* (pp. 4765–4774). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html
- [9] Besedeş, T., & Prusa, T. J. (2006). Product differentiation and duration of US import trade. *Journal of International Economics*, 70(2), 339–358. <https://doi.org/10.1016/j.jinteco.2005.12.005>
- [10] Almunia, M., Antràs, P., Lopez-Rodriguez, D., & Morales, E. (2021). Venting out: Exports during a domestic slump. *American Economic Review*, 111(11), 3611–3662. <https://doi.org/10.1257/aer.20181853>
- [11] Chatterjee, A., Dix-Carneiro, R., & Vichyanond, J. (2013). Multi-product firms and exchange rate fluctuations. *American Economic Journal: Economic Policy*, 5(2), 77–110. <https://doi.org/10.1257/pol.5.2.77>
- [12] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [13] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- [14] 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>
- [15] Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1), 1–3. [https://doi.org/10.1175/1520-0493\(1950\)078<0001:VOFEIT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2)
- [16] Silva, T. C., Wilhelm, P. V. B., & Amancio, D. R. (2024). Machine learning and economic forecasting: The role of international trade networks. *Physica A: Statistical Mechanics and its Applications*, 649, 129977. <https://doi.org/10.1016/j.physa.2024.129977>
- [17] Shi, G., Chen, Z., Luo, W., & Wei, Z. (2024). Cross-border spillover of imported sovereign risk to China: Key factors identification based on XGBoost-SHAP explainable machine learning algorithm. *Finance Research Letters*, 70, 106307. <https://doi.org/10.1016/j.frl.2024.106307>
- [18] Purnell, D., Jr., Etemadi, A., & Kamp, J. (2024). Developing an early warning system for financial networks: An explainable machine learning approach. *Entropy*, 26(9), 796. <https://doi.org/10.3390/e26090796>

- [19] United Nations Statistics Division. (n.d.). UN Comtrade Database [Data set]. Retrieved July 30, 2026, from <https://comtradeplus.un.org/>
- [20] World Bank. (n.d.). World Development Indicators [Data set]. Retrieved July 30, 2026, from <https://databank.worldbank.org/source/world-development-indicators>
- [21] International Monetary Fund. (n.d.). Global price of Natural gas, EU [PNGASEUUSDM] [Data set]. Retrieved July 30, 2026, from FRED, Federal Reserve Bank of St. Louis: <https://fred.stlouisfed.org/series/PNGASEUUSDM>
- [22] Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis*, 120, 70–83. <https://doi.org/10.1016/j.csda.2017.11.003>