

T.C. KAMU GÖZETİMİ, MUHASEBE VE DENETİM STANDARTLARI KURUMU

YAPAY ZEKÂ DESTEKLİ FİNANSAL ANALİZ PROJESİ

1. Kurum Hakkında Bilgi

Kamu Gözetimi, Muhasebe ve Denetim Standartları Kurumu (KGK), ülkemizde finansal raporlama ve bağımsız denetim sisteminin yüksek bir kalitede standartlaştırılarak güvenilirliğinin sağlanması yoluyla sermaye piyasasının ve ülke ekonomisinin güçlendirilmesi için düzenleme, denetleme ve gözetim ile uygulama yetkilerini haiz statüsüyle faaliyet göstermektedir.

KGK; başta halka açık şirketler, borçlanma araçları ihraç eden şirketler, bankalar, sigorta şirketleri olmak üzere bağımsız denetime tabi olmaya ilişkin cari limitlere göre Türkiye'nin en büyük 15.000'e yakın işletmesinin bağımsız denetim süreçlerini incelemek ve gözetmek amacıyla kurulmuştur. Bu işletmelerin ülkemiz Gayrisafi Millî Hasılasına (GSMH) katkısının %85'ten fazla olduğu tahmin edilmektedir.

2. Problem Tanımı

Kurumumuzun öncelikli hedeflerinden biri finansal risklerin erken aşamada öngörülebilmesidir. Ancak mevcut analiz çalışmalarımızın tüm sektörleri kapsayacak bütüncül bir yapay zekâ mimarisıyla tam olarak desteklenmemesi, bu hedefin önündeki temel gelişim alanını oluşturmaktadır. Yapılan ön çalışmaların henüz sınırlı sayıda sektörü kapsamaması ve sistemin tüm veriye yayılan kapsayıcı bir anomali tespitine imkân tanınamaktadır.

3. Mevcut Durum

Kurumumuza ait Elektronik Finansal Raporlama (EFR) Sistemi, şirketlerin finansal tablolarına ait verileri 2018 yılından itibaren barındırmaktadır. Veriler, Excel formatında belirli şablonlarda Sözleşme Bilgi Giriş Portalı (SBG) kanalıyla temin edilmektedir. 2018-2024 yıllarına ait finansal veriler, kontrollerden geçirilerek, NACE kodlama sistemine göre belirlenmiş olan 22 Ana ve 272 Alt sektöre uygun bir biçimde oluşturulan finansal tablolar (Finansal Durum Tablosu, Kâr veya Zarar Tablosu ve Nakit Akış Tablosu) ve 39 finansal analiz oranları (kârlılık analizi, devir hızı analizi, likidite analizi ve finansal yapı analizine ilişkin ortalama oranlar) "Finansal Analiz Platformu (FinAP)" projesi aracılığıyla "Bağımsız Denetime Tabi Sektör Verileri" başlığıyla KGK internet sitesinde yayımlanmaktadır.

EFR sistemi çerçevesinde temin edilen veriler yıllık (12 aylık) veriler olup kamuya açıklama yapmak zorunda olan şirketlerin (yaklaşık 600 şirket) ara dönem (3'er aylık ve/veya 6 aylık)

verileri de yapılan protokol kapsamında ilgili kurumlardan Excel formatında temin edilmektedir. Şirketlerin, kendi finansal oranlarını, sektördeki durumunu ve risklerini görebileceği; bunları paydaşlarla paylaşabileceği bir raporlama aracı (analiz.kgk.gov.tr) Python dilinde yazılmış olup şu anda test aşamasındadır.

Söz konusu verileri kullanarak, finansal başarısızlık bakımından riskli şirketlerin tespitine yönelik olarak yapılan ön çalışmalarda, finansal başarısızlık göstermeye yönelik 8 kriterden herhangi birini sağlayan şirket finansal başarısız olarak kabul edilmiştir. Çalışmada, 9 farklı alt sektör üzerinde gerçekleştirilen analizler doğrultusunda her bir sektör için özelleştirilmiş lojistik regresyon modelleri geliştirilmiştir. Sektör bazında finansal başarı ve başarısızlık sınıflandırmalarını gerçekleştirebilmek amacıyla belirlenen 22 finansal oran kullanılarak çeşitli değişken kombinasyonları test edilmiş ve bu sektörler için oluşturulan 100 model arasından en yüksek doğru sınıflandırma oranına sahip 9 model seçilmiştir. Bu modellerin doğru sınıflandırma yüzdesi en düşük %82,2 ve F1 skoru %70'dir.

4. Projenin Amacı ve Kapsamı

KGK tarafından yürütülmekte olan FinAP projesinde yapay zekâ (YZ) yardımıyla firmaların finansal başarısızlıklarının önceden tespit edilmesi ve sektörel bazlı finansal anomalilerin tespit edilmesi amaçlanmaktadır.

Proje kapsamında, FinAP sistemine entegre olacak aşağıdaki özelliklere sahip bir sistem geliştirilecektir:

- EFR aracılığıyla yüklenen, KGK veri tabanlarında depolanan ve yıllık olarak elde edilen yaklaşık 30 GB boyutundaki verileri analiz ederek finansal başarısızlık ve belirli finansal oranlar bakımından sektöründeki şirketlerden büyük oranda farklılaşan riskli şirketleri tespit edecektir.
- Şirketlerin finansal durum ve performansları hakkında, KGK tarafından belirlenecek finansal oranlar ve firmanın bu oranlardaki yıllık değişimlerini model sonuçları ile karşılaştırabilen, farklı kullanıcılara yetkileri kapsamında raporlayabilen ekranlar sunacaktır.

5. Teknik İsterler

1. İDARE (KGK) tarafından, finansal başarısızlık bakımından riskli şirketlerin tespitine yönelik olarak yapılan çalışmalarda, finansal başarısızlık göstermeye yönelik 8 kriterden herhangi birini sağlayan şirketlerden “finansal başarısız” olarak etiketlenenler, YÜKLENİCİ'ye (Proje Yürütücüsü Kuruluş) sağlanacaktır.
2. Geliştirilecek çözüm için kullanılacak veriler İDARE'nin veri tabanlarında depolanmakta

olup yıllık yaklaşık 30 GB boyutundadır. Proje için geliştirme sürecinde gerekli veriler İDARE tarafından gerekli gizlilik sözleşmesi imzalandıktan sonra sağlanacaktır.

3. Geliştirilecek çözüm; clustering algoritmaları, olasılık tabanlı algoritmalar, makine öğrenmesi, derin öğrenme ve/veya dönüştürücü (transformers) gibi güncel yapay zekâ algoritmalarını kullanacaktır.
4. Geliştirilecek çözüm, NACE kodlama sistemine göre belirlenen 272 alt sektörün her biri için performans metriği olarak en az %70'lik bir Makro F1 skoru sağlayacaktır. Modelin eğitilmesi ve test edilmesi için istatistiksel olarak yeterli sayıda başarısız şirket verisi barındırmayan alt sektörler, İDARE'nin onayıyla hiyerarşik olarak bir üst NACE grubuna (Ana Sektör) dahil edilerek değerlendirilecektir.
5. Geliştirilecek çözüm, güncel verilerle tekrar eğitilebilecek ve performansını sürekli olarak iyileştirebilecek uçtan uca bir MLOps mimarisine sahip yapıda olacaktır. Çözüm, yıllık verilerin yanı sıra ara dönem verileri de kullanabilecektir.
6. Geliştirilecek çözüm, finansal başarısızlık ve sektörel bazlı finansal anomalilerin tespit edilmesi kapsamında yetkilendirilen kullanıcılara raporlama ekranı sunacaktır.
7. Geliştirilecek çözüm, şirketlerin finansal durum ve performansları hakkında, İDARE tarafından belirlenen finansal oranlar ve bu oranlarda yıllar itibarıyla meydana gelen değişimler bakımından, model sonuçlarını (Örneğin, "... şirketin cari oranı 0,8 hesaplanmış olup bu oranın önceki yılın cari oranı olan 0,9'dan ve sektör ortalaması olan 1,2'den düşük olduğu tespit edilmiştir. Bu seviyedeki bir cari oran, işletmenin mevcut dönen varlıklarının kısa vadeli borçlarını ödemek için yetersiz olduğunu gösterir. Bu durum ise, işletmenin likidite sorunları yaşadığını ve borçlarını zamanında ödeyememe riski olduğunu gösterebilir.") raporlayacaktır.
8. Geliştirilecek sistem İDARE'nin mevcut sistemleri (FinAP, EFR, SBG) ile uygulama katmanı üzerinden (REST API, SOAP, gRPC, ODATA) entegrasyon sağlayabilir yapıda tasarlanacak daha sonra yeni entegrasyonlar eklenmesine olanak sağlayacaktır. Ayrıca, İDARE'nin yazılım altyapı bileşenleri ile entegrasyon sağlayacaktır.
9. Geliştirilecek sistem, İDARE'nin belirlediği teknolojiler hariç herhangi bir ticari lisans bağımlılığı olmaksızın, açık kaynaklı teknolojiler kullanılarak geliştirilecektir.
10. Geliştirilecek sistem için gerekli olan donanım ve altyapı özellikleri, YÜKLENİCİ tarafından belirlenecek ve sözleşmenin imzalanma tarihinden itibaren en geç 3 ay içinde İDARE'ye sunulacaktır. Gerekli donanımlar ve altyapı, İDARE tarafından proje takvimine uygun olarak temin edilecektir.
11. Proje süresince gerekli geliştirme ortamı ve altyapı İDARE tarafından kurum içinde çalışılacak şekilde ve/veya uzaktan erişim (VPN) şeklinde sağlanacaktır.
12. Geliştirilecek sistem, YÜKLENİCİ ve İDARE'nin birlikte karar vereceği yazılım geliştirme

dilleriyle geliştirilecektir.

13. Geliştirilecek sistem kurum içi kapalı devre (on-premise) olarak çalışacaktır.
14. Geliştirilecek sistem, toplamda 25.000 kullanıcıya, eş zamanlı olarak 200 kullanıcıya hizmet verecektir. Performans testleri, test raporunda raporlanacaktır.
15. Sistemdeki tüm kullanıcı işlemleri ve sistem hataları, değiştirilemez bir merkezi loglama yapısında tutulacak; donanım kaynaklarının (CPU/RAM/GPU) kullanımını gerçek zamanlı bir izleme ve arayüz üzerinden izlenebilir olacaktır.
16. Kimlik doğrulama ve yetkilendirme İDARE'nin yetkilendirme servisi veya İDARE ile mutabık kalınacak yöntem üzerinden yapılacaktır. Kurum içi kullanıcılar için Active Directory ile giriş imkânı sağlanacaktır.
17. Geliştirilecek sistem, YÜKLENİCİ tarafından OWASP Top 10 standartlarına uygun olarak hiçbir güvenlik açığı içermeyecek şekilde geliştirilecektir.
18. Geliştirilecek sistem canlıya alınmadan önce bağımsız bir kurum veya İDARE tarafından yapılacak Sızma Testi'nden başarıyla geçecektir.
19. YÜKLENİCİ; veri setinin ve geliştirilen algoritmaların ISO/IEC TR 24027 standardına uygun olarak yanlılık (bias) analizini gerçekleştirecek, sistemin tasarımından yaygınlaştırılmasına kadar olan tüm süreçlerdeki olası riskleri ISO/IEC 23894 rehberliğinde sistematik olarak yönetecek ve genel yönetim yapısını ISO/IEC 42001 Yapay Zeka Yönetim Sistemi prensipleriyle uyumunu gözetecek ve bu analizleri raporlayacaktır.
20. Proje kapsamındaki tüm veri işleme, saklama ve model geliştirme süreçleri TS ISO/IEC 27001 Bilgi Güvenliği Yönetim Sistemi standartlarına uygun olarak yürütülecektir.
21. Uygulamanın kaynak kodları, veri tabanlarına ait tüm elemanları, güvenlik şifreleri ve ilgili tüm dokümanlar İDARE'ye teslim edilecektir.
22. YÜKLENİCİ, İDARE'nin uygun gördüğü personele teorik ve uygulamalı eğitimleri verecektir. Eğitimler için gerekli eğitim dokümanlarını hazırlayıp, İDARE'ye teslim edecektir.

6. Ek (Dokümanlar)

Proje kapsamında aşağıda belirtilen raporlar hazırlanacak ve İDARE ile proje planında belirtilen tarihlerde paylaşılacaktır. İDARE gerekli görmesi durumunda ek raporlar talep edebilecektir.

- Üst Düzey Sistem Tasarımı Dokümanı

Geliştirilecek sistemin genel yapısı ve işleyişi tanımlanmalı, sistem mimarisi detaylı şekilde açıklanmalıdır. Ana bileşenlerin işlevleri ve bunlar arasındaki etkileşimler belirlenmeli, veri akışı ve süreçler netleştirilmelidir. Performans, ölçeklenebilirlik ve güvenlik gibi teknik

gereksinimler ele alınmalı ve riskler ile varsayımlar ortaya konulmalıdır. Bu doküman, sistemin başarılı şekilde tasarlanıp uygulanabilmesi için yol gösterici bir rehber olmalıdır.

- Analiz ve Tasarım Dokümanı

Sistemin mevcut durumu analiz edilmeli ve ihtiyaç duyulan özellikler belirlenmelidir. Teknik gereksinimlerden tasarım ilkelerine kadar olan süreç açık bir şekilde sunulmalı, kullanıcı deneyimini iyileştirmek için arayüz tasarımı yapılmalıdır. Ayrıca, veri tabanı yapısı ve sistemin mimari çözümü detaylandırılmalıdır.

- Veri Analiz Raporu

Bu raporda, analiz edilen veri setleri ve kullanılan yöntemler açıklanmalı, veri setinin temel özellikleri tanımlanmalıdır. Bulgular, grafikler ve tablolarla görselleştirilmeli ve analizden elde edilen sonuçlar yorumlanmalıdır. Stratejik kararlar için öneriler geliştirilmeli ve sonraki adımlar belirlenmelidir. Rapor, veri temelli karar alma süreçlerine rehberlik etmelidir.

- Test Raporları

Test süreçleri ve sonuçları detaylı şekilde sunulmalıdır. Test amaçları ve kapsamı netleştirilmeli, uygulanan test senaryoları ve kullanılan yöntemler, performans ve hedeflenen çıktılarla uyum açıklanmalıdır.

- Kılavuzlar ve Teknik Dokümantasyonlar

Sistemin farklı kullanıcı rollerine göre etkin bir şekilde işletilebilmesi, yönetilebilmesi ve diğer sistemlerle entegre edilebilmesi için gerekli tüm belgeler ve kılavuzlar hazırlanmalıdır. Bu kapsamda; son kullanıcılar için sistemin fonksiyonel kullanımını adım adım açıklayan Kullanıcı Kılavuzu, sistem yöneticileri için kurulum, yapılandırma ve bakım süreçlerini içeren Sistem Kılavuzu sunulmalıdır. Ayrıca, sistemin sunduğu servislerin ve veri yapılarının teknik detaylarını içeren kapsamlı bir API Dokümantasyonu hazırlanmalıdır. Tüm dokümanlar güncel olmalı, görsel öğelerle desteklenmelidir.