

5. Sonuç ve Gelecek Çalışmalar

- Genel bir veri seti hakkında (Takasbank, tefas)
- Problem sorusu hakkında
- Tez çalışmasının önemi(Akademi, sanayi işbirliği, tefas ilk defa kullanılması)
- Algoritma karşılaştırmaları özet(avantaj, dezavantaj)
- Model sonuçlarının genel karşılaştırılması özet
- Tercih edilen modelin açıklanması
- Tez yazarken kısıtlayıcı faktörler
- Gelecek çalışmalar için önerileri(optimizasyon, risk çalışmaları, daha fazla veriyle besleme)

Fonlar, finans sektöründe önemli bir yatırım aracı olmakta ve yatırımcılar için kazançlı yatırımlar yapacakları fonların seçilmesi için yatırımcı tercihlerinin yönetimde, yönlendirilmesinde bir çok aracı kurum ve kuruluştaki bulunmaktadır. Temel olarak, yatırımcılara kar sağlayacak kazançlar elde edilmesi için birden fazla kurum, kuruluş ve finans araştırmacıları fon alım satımında danışmanlık yapmaktadır. Doğru tercihlerin yapılması içinse, piyasa araştırmaları, fonların fiyat dalgalanmaları ve bir çok parametreyi hesaba katarak öngöründe bulunmaya çalışmaktadır. Bu süreçlerin hepsi, geleneksel anlamda, yapılan araştırmalardan elde edilen verilerinden hareketle fon danışmanlarının tecrübeleri ve öngördükleri bağlamında yatırımcıya öneride bulunabilmektedir. Fakat, her ne kadar bu öngörüler belli araştırmalar kapsamında yapılsada, günümüz teknolojik gelişmelerin geldiği noktada sunulan imkanlar ölçüsünde çok düşük kalmaktadırlar. Artık teknolojik imkanlar, yapay zeka çalışmaları ile birlikte bu tür bir çok manuel yapılan işlemleri akıllı sistemlere dönüştürerek, üst düzey hesaplamalar, birden fazla parametreyi hesaplamalar katarak işlemler yapılabilmektedir.. Bu amaçla bu tez çalışmasında, makine öğrenme algoritmaları kullanılarak fonların fiyat tahmini için model geliştirilmesi yapılmıştır.

Fon fiyat tahmin sistemi için geliştirilmesi yapılan model çalışmasında kullanılan veri setinin gerçek bir veri olması, yapılan modelin tahmin başarısında gerçek piyasadaki fon bilgilerine yakın değerler olacağı ve model tutarlılığı ve güvenilirliği açısından çok önemli olacaktır . Bundan dolayı veri seti, Takas İstanbul(İstanbul Takas ve Saklama Bankası A.Ş.-Takasbank)’un platform ve veri kaynağı sağlayıcılığı yaptığı Türkiye Elektronik Fon Dağıtım Platformu (TEFAS) web sitesi üzerinden erişime açık olan fon bilgilerinde elde edilmiştir. Her bir fon için 02.01.2019 –

31.12.2019 tarihleri arasında tipi, türü, toplam değeri, tedavüldeki pay sayısı, pay alan kişi sayısı, fiyatı ve menkul (26 çeşit) oranları ile fonların üzerinde önemli bir etkiye sahip olan her bir tarih için faiz bilgisi, altın fiyatı ve dolar fiyatı baz alınarak veri seti hazırlanmıştır. Veri kümesi, toplam 187,438 veri ve 37 kolondan oluşturulmuştur. Veri seti, tutarlı ve temizlenmiş bir veri için ön işleme adımlardan geçirilerek deneysel çalışmalarda kullanılmak üzere, eğitim ve test setine ayrılmıştır.

Model geliştirme çalışmaları, Kısmi En Küçük Kareler Regresyonu (PLSR), Ridge Regresyonu(RR), Destek Vektör Regresyonu (SVR) ve Yapay Sinir Ağları(YSA) algoritmalarıyla yapılmıştır. Bölüm 4'te her bir algoritmanın bölüm başlığı altında ayrıntılı olarak model geliştirme süreçleri anlatılmıştır. Geliştirilen model başarı değerlendirmelerini; Hata Karelerinin Ortalaması(MSE), Hata Kare Ortalamasının Karekökü(RMSE), Ortalama Mutlak Hata(MAE) ve R Kare Oranı(R2) ölçüm metriklerine göre yapılmış. Tablo 4.13'te model geliştirme çalışmaları yapılmış olan 4 algoritmanın doğrulanmış model ölçüm metrik değerleri; MSE, RMSE, MAE ve R2 sonuçlarına göre kurulan modellerden PLSR, RR ve SVR algoritmaların kabul edilebilirlik oranlarının çok düşük olduğu görülmüştür. YSA ile elde edilen tahmin değerlerinin başarı oranı ise, model önerisi için kabul edilebilir ölçekte olduğu gözlemlenmiştir. Ayrıca, 4 algoritmayla tahmin edilen değerlerinin gerçek değerlerden ne kadar uzakta olduğunu ifade eden artık verileri görselleştirme işlemleri şekil 4.23, 4.24, 4.25 ve 4.26'da yapılmıştır. Bu veri artık görsel grafiklerine göre modellerin tahmin değerlerinin tutarlılık yapısı daha anlaşılır kılmaktadır. Bu veri artık dağılım grafiklerine bakıldığında da, YSA model tahmin veri grafik dağılımın, tahmin edilen değerlerin yayılımının diğer 3 algoritma grafiklerine göre daha dengeli olduğu görülmektedir. Bu değerlendirmelerden hareketle, fonların fiyat tahmini çalışmasında YSA ile geliştirilen model tercih edilmiştir.

Sonuç olarak, fonların fiyat tahmin için geliştirilen model çalışmasında, YSA modeli tahmin edilen fon fiyat değerinin birim pay başına düşen hata oranı, RMSE değerine göre (+/-) 6.2 olarak bulunmuştur. Model başarı yüzdesi ise, R2 ölçüm metrik sonucuna bakılarak %90 üzerinde bir tahmin başarısı olarak elde edilmiştir. Son olarak, 4 modelin tahmin başarısı R2 ölçüm metriğine göre yapılmış olup, şekil 4.27'de de model karşılaştırması yapılmıştır. Ancak nihai model karşılaştırılmasının yüzdelik sonucu R2 ile değerlendirilemez. Diğer ölçüm metrik sonuçları da göz önüne alındığında regresyon problemleri için PLSR, RR ve SVR modellerinin kullanılabilirliklerini yitirmedikleri tespit edilmiştir. Daha doğru sonuçlar için PLSR, RR ve SVR modellerin

optimizasyon bölümlerinde modeller daha geniş deneysel değer aralıkları ile test edilebilir ve veri setlerinden korelasyon ilişkisi zayıf olan değişkenler çıkarılarak modeller tekrar değerlendirilebilir.

Mevcut kısıtlar Bu çalışmada 1 yıllık veri seti kullanılarak model tahmini çalışması yapılmıştır. Fakat takasbankta fonlar için 7-8 yıllık verilerin mevcut olduğu bilgisi alınmış, Model başarı yüzdesini daha güvenilir kılcak büyük verilerle de modeller tekrar açıştıırakbilir.

Ayrıca, kullanılan bilgisayarın teknik kısıtları bağlamında da daha yüksek test setleri ile optimizasyon yapılmamıştır. Mevcut pandamı sürecinden kaynaklı erişilebilir labarotuvary imkanlarıda kullanılmamıştır. Bu kapsamda, teknik özellikleri daha gelişmiş bilgisayardan modeller tekrar çalıştırılabilir, hiper parametre setleri genişletitilrereck.

Tezin önemli noktalarından biri, çalışmada kullanılan veri setinin Takasbanktan alınan gerçek verileden oluşması, elde edilen sonuçların gerçek hayattan karşılaşılan problemlere çözüm olabileceği zemini hazırlamış olmaktadır. Bu bağlamda akademi ve sanayi iş biriliğinin kapsamında bir çok projenin geliştirilemsi için örnek oluşturması. ayrıca, erişime açık olan TEFAS verilerini ilk defa doğrudan alınarak bu çalışma yapılmıştır. literartüer katkısın olacaktır. Tefas verilerin kullanım yaygınşatırarak ekosisitene katkısı olacaktır. Ayrıca, takasbank'raki bir çok projeye de uyaralanabilecek şekllidne örnek kaynak oluşturulmasına katkı yapılmıştır.

gelecek çalışmalar için, tefas platformu için fonların fiyat tahmini için yapılan bu model çalışması birinci adımı sayılacak tahmin fonksiyonu oluşturumuştur. Ancak, fon tipleri ve fon turleri bazında kırımlar yapılarak model daha spesikleştirilebilri. Yatırımcılar için fon kazançlarını nasıl optimize edileceği ve risk getirilerini minimize edilecek model çalışmaları yapılabilir.