

1. Giriş

Günümüz bilgi çağı veya dijital çağın insanlara sunduğu yeni imkanlarla birlikte artık bilginin çok hızlı bir şekilde dolaşıma girmesine ve erişimini kolaylaştırmasıyla beraber son yüzyılda bilgiyi geliştiren, inovatif teknolojilere yatırım yapan ülkeler global pazarlarda rekabet üstünlüğü elde etmekte ve ekonomileri için büyük bir katma değer faktörü oluşturmaktadırlar. Artık ülkelerin gelişmişlik düzeyleri geleneksel yöntemlerle elde ettikleri gelir oranları belirlememekte, tam aksine bilgiye, bilgi teknolojisine, yenilikçi teknolojilere ve bilgi merkezli üretilen sistemlere yaptıkları katkılar, yatırımlar şekillendirmektedir. Dolayısıyla modern çağın getirdiği bu perspektifle güç ve gelişimi elinde tutan ülkelerin, asıl odak merkezleri bilgi ve otomasyon sistemlerine yapılan yatırımlar ön plana çıkarmaktadır. Bu kapsamda dijital çağa beraber bir çok kavram, kurum ve geleneksel kültürü yapı bozuma uğratacak şekilde, yeniden tanımlamaktadır. Ekonomi'den sağlık sektörüne, eğitim, sanayi ve daha benzeri bir çok sektörde yeni ufuklar ve imkanlar sunmaktadır. Bu sektörlerdeki teknolojik yatırımların geldiği noktada artık insan gücünün üretimi yerine uzman sistemler, akıllı yazılımlar, donanımlar, robotik tasarımlar gibi yapay zekanın ve alt dalları olan makine öğrenmesi ve derin öğrenme konularını kapsayan geniş bir yelpazede araştırma ve geliştirme çalışmaları yapılmaktadır. Bu çalışmaların temelinde geleneksel olan yapıların nasıl otomatize edilerek uzman sistemlere dönüştürüleceği üzerinde durulmaktadır. Sektörel bazda yapılan bu otomasyon sistemleri bir çok kolaylıklar getirmektedir. Bu tez çalışmasında, finans sektörüne yönelik teknolojinin sunduğu imkanlar kapsamında fonların fiyatını tahminleyecek bir model uygulaması üzerinde durulacaktır.

Finans sektörüne dair yapılacak fonların fiyat tahminleri üzerinde birçok araştırmacı uzun yıllardır çalışmaktadır. Fon, kısaca açıklaması yapılırsa eğer; yatırım yapmayı düşünen yatırımcılara belli bir orandan katılım pay hakkı verilerek toplanan paralar ile katılma pay sahipleri adına, risk çeşitlendirilmesi ve inançlı mülkiyet esaslarına göre, portföy işletme amacıyla kullanılan mal varılığıdır[1]. Fonların fiyat tahmin çalışmalarındaki hedeflenen amaç, bir yatırımcının doğru yatırım araçlarına yatırım yaparak kazanç elde edilmesini sağlayacak yatırım aracının değerini önceden öngörülüp öngörülemezliği ancak yatırımcıların öngörü de bulunması mevcut bir çok parametrenin dahil olmasıyla öngörü yapmak kolay olmamaktadır. Öngörü kavramı ise, belirli varsayımlar öncülüğünde bir değişkenin gelecekte elde edeceği çıktı değerini önceden yaklaşık olarak belirlenmesi biçiminde tanımlanabilmektedir. Doğru öngörü (tahmin) 'de bulunmanın bir çok başarılı kararları beraberinde getireceği ve bu ulaşılan başarıların en maksimum düzeye

çıkartılabileceği gerçeği, öngörü çalışmalarına olan merakı, ilgiyi artırarak bu alanda sürekli önemli gelişmelerin oluşmasına ön ayak olmaktadır. Daha önceleri geleneksel olarak matematiksel ve istatistiksel metotlarla yapılan tahmin denemeleri, özellikle teknolojik gelişmelerle beraber artık daha başarılı sonuçlar vermiştir.

Bu bağlamdan hareketle son yıllarda, makine öğrenme metotları bir çok farklı alanda kullanımları yaygın bir şekilde giderek artmakta ve uygulama alanları genişlemektedir. Makine öğrenmenin temel uygulama alanlarından biri olan tahmin edebilmektir ve daha doğru tahmin değerleri elde edebilmek için daha optimize tahmin modelleri üzerine kullanılabilir. Ayrıca, Makine öğrenme çalışmalarının finansal tahmin uygulamalarında kullanımı ve alınan başarılı sonuçlar araştırmacıların bu alana dair taleplerini artırmaktadır.

Bu tez çalışmamızda da, makine öğrenme metotları kullanılarak, finans sektörü için ihtiyaçta göre en doğru veya doğruya en yakın sonuçlar elde edebilecek bir fon fiyat tahminleme model önerisi üzerinde durulacaktır. Geliştirilecek olan fonların fiyat tahmin model önerisi, finans sektöründeki tahmin çalışmalarının yatırımcılar için önemli bir konu olduğu göz önüne alındığında, fiyat tahmin modellemesinin bu alana dair önemli bir katkı sağlayacağı aşıkardır. Yatırımcıların alım satımını yapacağı fonlardan kâr merkezeli kazanç elde etmelerini sağlayacak, tahminlenebilir değerler üzerinden çıkarımlar ve analizler yapılarak fon yatırımcılarına kazançlı yatırımlar için doğru yönlendirilmenin oluşmasına zemin oluşturacaktır. Bu kapsamda, yatırımcı fon alırken, makine öğrenme yöntemiyle oluşturulacak tahmin modelinde fon fiyatını belirleyen birden fazla parametre değerleri olan; fonun mevcut fiyatı, alan kişi sayısı, menkul dağılım oranları, toplam piyasa değeri ve bunların dışında piyasalarda çok büyük oranda etkisi olan, faiz oranları, döviz fiyatları, kıymetli maden fiyatları gibi değişkenleri de tahmin sistemi üzerindeki etkisini değerlendirilmeye katarak, fon alım satımını yapacak yatırımcıya belli bir sapma değeri oranında o fonun fiyatını öngörülebilir yaparak, yatırımcı tercihleride çok büyük bir etki yaratacağı kaçınılmazdır. Buraya kadar yapılan bu açıklamalardan da görülmektedir ki, geliştirilmesi yapılacak fon fiyat tahmin model sistemi, finans sektörü için önemli bir yere sahip olduğu ve yatırımcıların da doğru yatırım araçlarına yönlendirilmesini sağlayacak araştırmalarda önemli bir katkı oluşturacağı öngörülmektedir. Ayrıca, geliştirilecek olan fon fiyat tahmin model sistemi için deneysel çalışmalarda kullanılacak veri setinin gerçek verilerden oluşması, elden edilecek sonuçların gerçek piyasalardaki fonların fiyat değerleri ile örtüşür olacağından, tahmin modelinin daha doğru ve güvenilir bir alt yapıya sahip olmasını sağlayacaktır. Kullanılan veri kümesi, Türkiye’de fonların tek platform üzerinde toplanmasını

sağlayarak, fon piyasası için önmeli katkı sağlayan Takas İstanbul(İstanbul Takas ve Saklama Bankası A.Ş.-Takasbank) tarafından platform sahipliği ve sağlayıcılığı yaptığı TEFAS(Türkiye Elektronik Fon Alım Satım Platformu) web sitesi üzerinden alınan fon bilgilerinden oluşturulmaktadır. TEFAS platformu ise; Tüm fonların tek bir sistem üzerinden karşılaştırması ve tek bir yatırım hesabıyla piyasadaki tüm fonlara erişme imkanı oluşturmasını sağlayan elektronik bir fon platformudur[2]. Bu veri setini kullanarak yapılacak model önerisinde, TEFAS platformu üzerinde halka açık olan gerçek fon bilgilerinin kullanılması sağlayarak, bu alanda yapılacak bir çok çalışma için araştırmacılara örnek oluşturulması ve platform verilerinin benzeri araştırma ekosisteminde de yaygınlaştırılmasına katkı sağlayacağı da öngörülmektedir.

Son olarak, bu tez çalışmasının diğer bölümleri hakkında bilgi verilerek giriş bölümünü tamamlamış olacağız. Genel olarak tez çalışması 5 bölümden oluşmaktadır. İlk olarak, giriş bölümünde, teknolojik gelişmelerin gelindiği son nokta ve tüm sektörlerde kullanım olarak sunduğu imkanlar üzerin bir inceleme yapılmış, daha sonra teknolojinin finans sektöründe kullanımında ve sağladığı kolaylıkların değerlendirilmesi gerçekleştirilmiştir. Sonrasında, finans sektöründe önemli bir yere sahip olan tahmin çalışmalarının öneminde bahsedilmiş ve tez çalışmasının temel önerisi üzerinde, tezin amacı, önemi, deneysel bölümde kullanılan veri setinin önemi ve tezin model çalışmasının sektöre katkısı bağlamında bir değerlendirme yapılarak bölüm bitirilmiştir.

İkinci bölüm olan literatür araştırmasında, makine öğrenme çalışmaları ile ilgili ekonomi, sağlık, eğitim ve sanayi gibi bir çok sektörde yapılmış olana araştırmalar incelenmiş ve incelenmesi yapılan araştırmalar genel olarak; çalışmanın amacı, kullanılan makine öğrenme algoritmaları, veri seti ve elde edilen başarı sonuçları kapsamında tasnifleri yapılarak, makine öğrenme çalışmalarının genel kullanım alanları ve sonuçları değerlendirilmiştir. Daha sonrasında, spesifik olarak makine öğrenme çalışmalarının fiyat tahmini uygulamalarında kullanım sıklıkları, en çok kullanılan algoritmalar, başarı sonuçları ve hangi sektörlerde yoğun olarak çalışmaların yapıldığı araştırmaları genel olarak incelerek, sınıflandırılması yapılmıştır. Literatür bölümün son başlığı ise, model öneri çalışmasının geliştirme aşamasında kullanılan algoritmaların, neden tercih edildiği, fiyat tahmin çalışmalarında kullanım sıklıkları ve elde edilen başarı çıktıları gibi sorular kapsamında, yapılmış çalışmalar incelenerek model çalışmasında kullanılan algoritmaların, tercih edilmesini sağlayan yeterli argüman zeminin oluşturulmasını gerçekleştirilerek bölüm çalışması tamamlanmıştır.

Üçüncü bölüm ise kullanılan makine öğrenme algoritmaları yani metod çalışmaları üzerinde durulmuştur. Bu bölümde, tahmin problemlerinin genelde regresyon algoritmalarının konusu olduğu için fonların fiyat tahmin model geliştirilmesi, doğrusal çoklu regresyon algoritmalarından; Kısmi En Küçük Kareler Regresyonu(PLSR) ve Ridge Regresyonu(RR) kullanılmış ve doğrusal olmayan çoklu regresyon algoritmalarından ise; Destek Vektör Regresyonu(SVR) ve Yapay Sinir Ağları(YSA) incelenerek çalışmalar yapılmıştır. Yapılan incelemeler, kullanılan her bir algoritmanın bölüm başlığı altında, teorik yapısı, matematiksel denklem yapıları ve algoritmanın çalışma adımlarından bahsedilerek bölüm incelenmesi bitirilmiştir.

Dördüncü bölüm olan model çalışmasında, ilk olarak veri seti hakkında ayrıntılı bilgiler verilmiş ve veri kümesi üzerinden, veri ön işlem adımları sırasıyla yapılarak deneysel çalışmalarda kullanılması için temizlenmiş bir veri seti haline getirilmiştir. Daha sonra model geliştirme için kullanılacak olan PLSR, RR, SVR ve YSA dört algoritmanın her biri için ayrı olarak adımlar halinde, modelin eğitim ve test veri setleriyle model kurulma aşaması ve model sonuçları elde edilmiştir. Daha sonra kurulan model üzerinde her bir algoritmanın optimizasyon problemine konu olan değişkenleri incelenerek, model optimizasyon çalışmalarının gerçekleştirilmesini sağlayıp nihayi doğrulanmış model sonuçlarına varılmıştır. Ayrıca, her bir model sonuçlarının değerlendirmesini dört algoritma için ortak ölçüm metrikleri; Hata Karelerinin Ortalaması(MSE), Hata Kare Ortalamasının Karekökü(RMSE), Ortalama Mutlak Hata(MAE) ve R Kare Oranı(R²) olan metotlarla ölçerek her biri için, optimize edilmemiş model ile doğrulanmış model sonuçları karşılaştırması yapılmıştır. Son olarak bu bölümde, dört modelin doğrulanmış sonuçları üzerinden karşılaştırma tabloları yapılarak, her bir algoritmayla kurulan model için değerlendirmeler yapılmış ve en yüksek başarı oranını dört ölçüm metriğine göre YSA algoritmasıyla elde edildiği sonucuna ulaşılmıştır.

Tez çalışmasının son bölümü ise sonuç ve gelecek çalışmalar için yapılan değerlendirmelerle bitirmiş olacağız. Sonuç bölümünde, fonların fiyat tahmini için kullanılan PLSR, RR, SVR ve YSA dört algoritmanın avantaj ve dezavantajlarından bahsedilmiş, daha sonra bu algoritmaları kullanarak geliştirilen model sonuçları değerlendirilmesi yapılmış ve tez çalışmasının temel amacı olan model önerisi bağlamında tercih edilen modelin seçilmesini sağlayan faktörlerin, kısıtların kapsamlı değerlendirilmesi gerçekleştirilmiştir. Ayrıca, deneysel çalışmalarda karşılaşılan teknik kısıtlar üzerinden de durulmuştur. Son olarak ise, tez çalışmasının genel değerlendirilmesi yapılarak gelecek çalışmalar içinde önerilerde bulunmuştur.

Referanslar

1. <https://www.spk.gov.tr/Sayfa/Dosya/922>
2. <https://www.spk.gov.tr/Duyuru/Dosya/20150107/1>, <https://www.tefas.gov.tr/>