***DENETİMLİ (SUPERVİSED) MAKİNE ÖĞRENMESİ KULLANARAK***

***STOK PAZARI\BORSA’YI ANALIZE ETME***

SHIRAZ AMADU BELLO

*Mühendislik ve Doğa Bilimleri,Bilgisayar Mühendislık, Ist. Sabahattin Zaim Üniversitesi.*

IMAN AL MOLLA

*Mühendislik ve Doğa Bilimleri,Bilgisayar Mühendislık, Ist. Sabahattin Zaim Üniversitesi.*

MULHAM SALLAH *Mühendislik ve Doğa Bilimleri,Bilgisayar Mühendislık, Ist. Sabahattin Zaim Üniversitesi.*

**GİRİŞ**

**Özet-Stok Pazar ticcaret yapmanın en karmaşık/kompleks yöntemlerinden biridir. Kar elde etmeyi ve riskleri paylaşımında Küçük çaplı mülkiyetler, kooperatif komisyoncular, banka gibi sektörleri stok pazarları dayanırlar. Oldukça kompleks bir modeldir. Bu çalışmada Makine öğrenmesi algoritmalrı kullanarak gelecek stok fiyatları tahmin etmeye çalışacaktır. Açık kaynak kutuphanleri(Open Source Library ) ve var olan algoritmaları kullanarak tahminini oldukça zor olan Stok pazarı tahmin etmeye kolaylaştırmayı hedeflenmektedir. Bu basit uygulamayı kabul edilebilir sonuçları nasıl ortaya çıkartacağını incelenecektir. Sonuçları tamamen sayısal rakmaları ve kuralları dayalı olacaktır ve tahmin esnasında gerçek Stok pazarına uyar veya uymaz[1].**

**Ana kelimeler-** Veri analizi, Uygulama, Denetimli Makine öğrenme, Linear Regrasyon, Detimli (Supervised) Makine öğrenmesi, Stok Pazarı, Borsa, Tahmin, Temeller.

**STOK PAZAR** Insanların borsa hisselerini pazarlamak, yatırım yapmak ve şirketlerin hisselerini satın alarak para kazanmayı ve kar elde etmeyi olan bir sistemdir. Çok eski bir yatırım yöntemdir. Bu platformda doğru kararları alınırsa oldukça kazançlıdır. İşin gerçeği bu platformdaki fiyatların esnekliliği tahmin etmeyi oldukça zor veya tahmin edilemez durumundadır, teknolojininin yardımı burada muhtaçtır. **Makine öğrenmesi** tahmin etmemizde kolaylık sağlayan bir teknoloji araçtır.

Bildiğımız gibi borsa öyle önemli bir platformdur ki herekesi ya bireysel yada ulusal seviyesinde etkiler[3].

**TAHMİN ETME MODELİ**

Tahmin etmek sürecinde birkaç adımları mevcuttur. Bunları;

1)Veri analize adımı(Data analysis)

2)Eğitmek ve test etme süci(Training and Testing )

3)Sonuç adımı(Results)

**VERİ ANALİZE**

Bu adımda elimizdeki ham(işlenmemiş) verileri inceleyerek tahmin etmemize yardımcı olabilecek özellikleri çıkarılacaktır.

Veri Kümenin öznitellikleri şunlardır:

**Open(Açılış):** Borsanın açılış fiyatıdır.

**High (Yüksek):** En yüksek fiyatıdır.

**Low (Düşük):** En düşük fiyatıdır.

**Close (Kapanış):** Borsanın kapanış fiyatıdır.

**Volume (Hacmi):** Gün içindeki toplam hareketlerin miktarıdır.

**Split (Paylaşım Oranı):** Kişinin veya kurumun yatırmına göre elde ettiği oranıdır. Düzeltilmiş de denilebiyor. Bunlar:

1)Düzeltilmiş Açılış

2)Düzeltilmiş Yüksek.

3)Düzeltilmiş Düşük.

4)Düzeltilmiş Kapanış.

5)Düzeltilmiş Hacmi.

Tahmin etmeye çalışılacak öznitellik **KAPANIŞ’tır.** Bu öznitellikler graf hallindedir**.** Bu tür grafları **OHLCV** [2](Open, High, Low, Close,Volume) olarak bilinmektedir. Borsaları bu grafı kullanır.

**EĞİTMEK VE TEST ETMEK**

Elde ettiğimiz verileri makine öğrenme modellemize uygulanılacaktır. Python veya Matlab programlama dilleri kullanılacaktır. Modellimize Denetimli (Supervised) Makine öğrenmesi algoritma kullanarak verilerimizin ilişkilerini tespit etmek ve sınıflandırmayı yapılacaktır**. Linear Regrasyon** basit fakat oldukça verimli bir algoritmadır.

Bu aşamada hem modelimizi eğitiriz hemede yeni verileri test ederiz.

**SONUÇ**

Bu aşamada modelimizin bitmiş hallidir. Artık gelecek borsa’nın tahminlerinde kullanılabilir. Doğruluk oranı bu aşamada çok önemlidir. Modelim doğruluk oranı en az %95 olacak şekilde olmayı hedefleniyor. Doğruluk oranı hesaplamak için

1)RMSE (Root Mean Square Error)

2)Confusion Matrix (Karışıklık Matriksi)

**Kullanılacak Algoritmalar**

1. Supervised Machine Learning
2. Linear Regrasyonu
3. Support Vector Machine Learning
4. Convolutional Neural Network.

**Kullanılacak veri Seti**

Verimiz [www.quandl.com](http://www.quandl.com*) ‘dan indirdik. Bu veri kümesi 15 yıllık Borsa Pazarının hareketlerini kapsar.

Siteden “GOOGL/WIKI “ ile erişilir. Python ile veri seti çekip ‘doya\_adı.csv’ formatında kaydedilir.

**ÖRNEK ÇALIŞMALAR**

\*K. Pahwa and N. Agarwal, "Stock Market Analysis using Supervised Machine Learning," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, 2019, pp. 197-200.

\* Pabuçcu, H. (2019). Borsa Endeksi Hareketlerinin Tahmini: Trend Belirleyici Veri. Selçuk Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi, 22 (1), 246-256.

\* *Dev Shah*; *Haruna Isah* and *Farhana Zulkernine*, (2019), [Stock Market Analysis: A Review and Taxonomy of Prediction Techniques](https://econpapers.repec.org/article/gamjijfss/v_3a7_3ay_3a2019_3ai_3a2_3ap_3a26-_3ad_3a234740.htm), *International Journal of Financial Studies*, **7**, (2), 1-22

**\*** Li, X., Xie, H., Wang, R. *et al.* Empirical analysis: stock market prediction via extreme learning machine. *Neural Comput & Applic* **27,**67–78 (2016). <https://doi.org/10.1007/s00521-014-1550-z>

\*Predictability of machine learning techniques to forecast the trends of market index prices: Hypothesis testing for the Korean stock markets. Pyo S., Lee J., Cha M., Jang H. (2017)  *PLoS ONE*,  12  (11) , art. no. e0188107

**REFERANSLAR**

[1] K. Pahwa and N. Agarwal, "Stock Market Analysis using Supervised Machine Learning," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, 2019, pp. 197-200.

[2] Hurwitz, E. and Marwala, T., 2012. Common mistakes when applying  
computational intelligence and machine learning to stock market  
modelling. arXiv preprint arXiv:1208.4429.

[3] Yadav, Sameer. (2017). STOCK MARKET VOLATILITY – A STUDY OF INDIAN STOCK MARKET. Global Journal for  
Research Analysis. 6. 629-632.