

Ekin Bozyel - İletişim Bilgileri

- Cep tel: 0534 328 27 56
- Mail adresi : ekinbozyel@gmail.com
- Linkedin hesabı : (<https://www.linkedin.com/in/ekin-bozyel-453934269/>)
- Github hesabı : (<https://github.com/john-fante>)
- Kaggle hesabı : (<https://www.kaggle.com/banddaniel>)

Dosyalar

1. **case-study-dataguess-ekin-bozyel.ipynb** ana çözüm dosyası

Aşağıdaki adımları kullanarak bir çözüm geliştirdim

- Veri bulma, (kaggle dataseti -> <https://www.kaggle.com/datasets/benjaminpo/s-and-p-500-with-dividends-and-splits-daily-updated>, version 66, 2025-01-31 e kadar olan veri), veri okuma ve ön işleme (preprocessing)
- Oracle, IBM hisselerinin günlük close ve volume verileri ile Microsoft hissesinin günlük volume ve open price verileri kullanarak **Microsoft hissesinin günlük close price** değeri hesaplamaya çalışıldı.
- Feature engineering adımında öncelikle time series veriye uygun cycle featureları oluşturuldu (weekly, daily, quarterly), Oracle ve IBM hisselerinin rsi, macd, signal_line ve histogram gibi teknik indikatörlerden gelen featureları oluşturuldu. (Not modele Microsoft hissesinin bir gün önceki günlük volume değeri ve open price değeri verildi, o yüzden feature engineering kısmında kullanılmadı.)
- Model olarak optuna ile hyperparameter tuning yapılmış bir **LinearRegression** modeli kullanıldı. Burada eğitim tekniği olarak zaman verisine uygun olan **online/incremental learning** kullanıldı. Bu teknik ile model her yeni örnekte (burada her gün gelen yeni veri ile) modelin eğitilmesi amaçlamaktadır. Bu yüzden klasik '**batch learning**' te olan train test seti ayırımı yerine model eğitim aşamasında tek tek verilerden öğrenmektedir. **Data leakage** problemi olmaması adına tahmin edilen Microsoft close price bir gün geriye kaydırılmıştır, bu şekilde model bugünkü verilerle yarının close price değerini tahmin edebilmektedir. Eğitim için özel pipeline fonksiyonu kullanıldı.
- **Evaluation** için MAE ve RMSE metricleri online learning modeli için rolling ile kullanılmaktadır.
- **Hyperparameter optimizasyonu** (window_size, learning_rate, intercept_lr, l2) için optuna ile 2000 adımlı trial oluşturuldu.
- **Sonuç MAE: 0.000713 , RMSE: 0.000713**