

Prédiction des Prix et Actifs du S&P 500 grâce à du Machine Learning

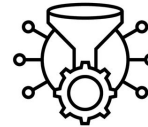
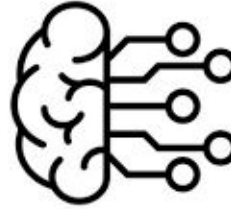
Andrea CHECCHI
Hugo MADOUAS
Colin LEBLANC
Paul ROFIDAL
Dylan SAYARATH

Le projet en bref

Objectif: Prédire les prix du S&P500 à l'aide du machine learning

3 Étapes:

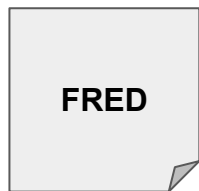
- Collecte et Préparation des Données
- Développement du Modèle Prédictif
- Analyse des Résultats et Interprétation



Collecte et Préparation des Données



5 Years



Prix SP500, Volume, VIX

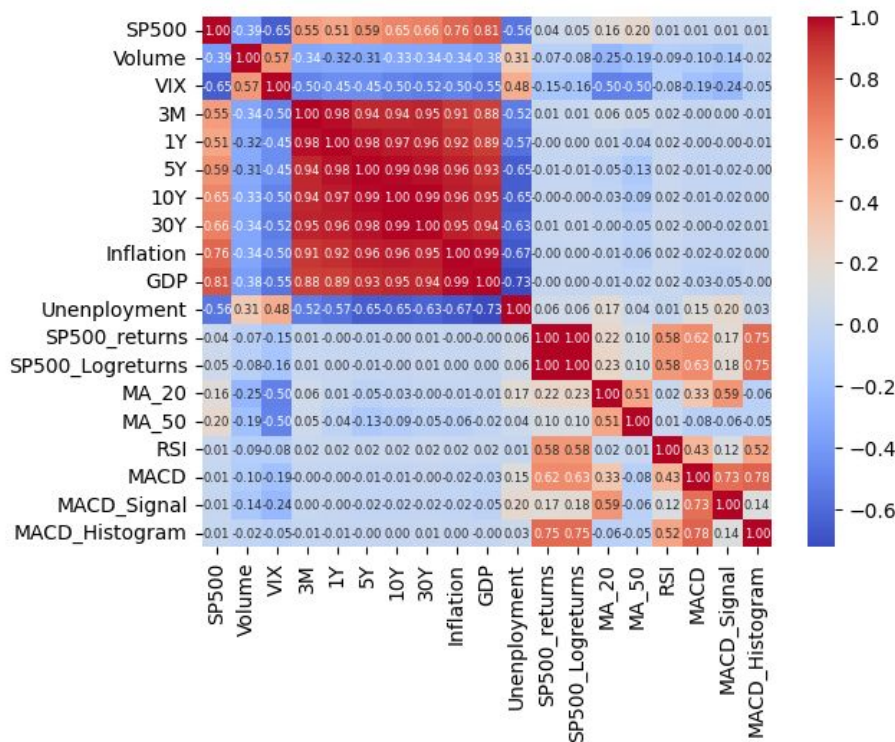
Unemployment, GDP, Rates, CPI

Traitement (NAN,
valeurs aberrantes)



Features Engineering:
MACD, RSI , Return, MA

Collecte et Préparation des Données



Utilisation d'une matrice de corrélations pour identifier les variables à utiliser

Développement du Modèle Prédictif



Modèle utilisés:

- Régression linéaire
- Random Forest Classifier et Gradient Boosting Classifier (Hausse/Baisse)
- Gradient Boosting Regressor (Valeur numériques)

Les Modèles sont entraînés et testés grâce à un test set et train set répartis selon la loi de pareto (80/20)

Analyse des Résultats et Interprétation



Résultats des différents modèles

Modèles	Régression linéaire	Gradient Boosting Classifieur	Random Forest Classifieur	Gradient Boosting Regressor
Scores	$R^2 = 4.67\%$	Accu = 50.21%	Accu = 49,58%	MAE = 0.0057

Analyse des Résultats et Interprétation



Construction d'une stratégie d'investissement à partir du Gradient Boosting Classifier

- Achat si Hausse prédite
- Vente si Baisse prédite

