|  |  |
| --- | --- |
|  |  |
| **Mockup of Machine Learning Model** | Stock Market Analysis |

|  |  |
| --- | --- |
| **PREPARE THE PROBLEM** | There have been two significant downturns in the stock market since March 2020.  Our group will create and train machine learning models that address questions related to stock performance and investor sentiment during this period. |
| **DATA SELECTION** | Data Extraction  Data Transformation  Data Storage – Postgres/AWS RDS |
| **CLEAN DATASET** | Import Dependencies  Load Dataset  Split Dataset into Training and Testing |
| **MACHINE LEARNING MODELS** | **Stock Sector Comparison**  Create Supervised Learning Model (Logistic Regression)  Train  Predict (Classification)  Validate (Accuracy)  **ESG Stock Performance**  Create Unsupervised Learning Model (K-means)  Train (Principal Component Analysis)  Predict (Determine Pattern)  Create a 3D-Scatter with the PCA data and clusters  **Investor Sentiment**  Create Natural Language Processing Model  Train  Predict (positive, negative, and neutral scoring)  Validate |
| **SAVE AND EXPORT** | Hierarchical Data Format (HDF5) file |