I want you to write a readme.md file for GitHub for the following project. The project is prediction of stock returns. The project is developed in python. It’s an emulsification of frameworks in MLops. An aws ec2 instance of windows server is used to host the application. Postgres is used as Database storage. S3 is used to store model. AWS RDS is used to host Postgres. The structure of the project is divided into 2 major parts:

1. Initial Setup (Data Science):
   1. Data is soured from yfinance till date from starting of these stocks – Apple, Tesla, MSFT, NVIDIA, AMZN.
   2. Now quality of data is checked, and regular EDA is performed.
   3. With the cleaned data from previous step, table for each ticker is created in the database. We have used pyspark for this.
      1. The naming format is <ticker>\_daily\_stock\_data.
      2. The columns are open, close, high, low, volume, adj\_close, created\_at, last\_modified\_at.
   4. With source tables, we have created feature tables for each ticker. The technical features are:
      1. Moving Averages (MA): Short-term (e.g., 10-day) and long-term (e.g., 50-day) moving averages.
      2. Exponential Moving Averages (EMA): Like MAs but gives more weight to recent prices.
      3. Relative Strength Index (RSI): Measures the magnitude of recent price changes to evaluate overbought or oversold conditions.
      4. MACD (Moving Average Convergence Divergence): Shows the relationship between two moving averages of a stock’s price.
      5. Bollinger Bands: A volatility indicator using standard deviation.
   5. Now once features are also computed using pyspark we ingested these into feature tables with following format <ticker>\_features.
      1. The columns are the features mentioned plus date column as index as same as in daily\_stock\_data table and created\_at, last\_modifed\_by
   6. Now we have all the data in one place. I have jumped into machine learning part.
   7. First, we have started ingesting data from the tables and processed into windowed dataset suing Tensorflow and keras.
   8. Then we have used forward selection technique to perform
   9. Then we built a base model of linear regression model.
   10. Then we with various parameters (different layers) and different architectures (pure dense layers, CNN, LSTM) we have created some 20 models.
   11. Then we trained these models and selected the best model.
   12. The we applied kerasTuner to finalize the hypermeters and we also used mlflow to record how these models are performing and we have stored all this information back into postgres.
   13. Finally, we have a model that is researched enough for deployment.
2. MLOps (ML Engineering):
   1. Now we shall create the following scripts and use Jenkins to create jobs:
      1. New data sorcing and ingestion
      2. Feature creation for new data
      3. Data drift check
      4. Data ingestion for ML model
      5. Model drift check
      6. New model training
      7. New model storage
      8. Model rollout and deployment
   2. We shall only train new model if there is a data drift or model drift happened and we shall deploy new model and store this in S3 for future analysis or reference.
   3. We also create logging logic for monitoring and any analysis purposes in future.
   4. Finally, we shall create a flask application and host on this EC2 to provide the access to public.
   5. The application has option for user to select any of the 5 tickers and then we shall provide past 60 days data and show next 5 days predicted values in different colors in a dashboard.
   6. The flask application appropriately selects the ticker model since we have model for each ticker and then ingests the past 60 days for inference purposes.
   7. Note this data sorucing and ingestion in point 1 shall happen daily in the moring before market opens and only runs if the previous day has marketopen to avoid unnecessary compute.