**Meeting Summary for Yousaf Abdul Khaliq's Personal Meeting Room**

Mar 31, 2025 07:26 PM Central Time (US and Canada) ID: 644 948 8621

Top of Form

Quick recap

Yousaf outlined the requirements for the reproducibility project for data ethics class, which involves finding a paper, reproducing its results, and analyzing its ethical implications. The team discussed data sources for historical stock prices, with concerns raised about using potentially proprietary Bloomberg data, and found that Yahoo Finance can be used to obtain daily stock price data from 1990 to 2018. The team also explored using publicly available data and a code to reproduce results from a paper and agreed to reconvene and try out some things before starting the second portion of their work.

Next steps

All team members to review the paper "Forecasting Directional Movements of stock Prices for intraday trading, using LSTM and Random Forests".

Yousaf to share the GitHub repository invite with the team for collaboration.

All team members to set up and familiarize themselves with the shared Word document for report writing.

Yousaf to share the Python script for pulling stock price data from Yahoo Finance with the team.

All team members to attempt reproducing the paper's results using the provided code and data.

Team to reconvene and discuss progress on reproducing the paper's results.

Team to start working on the second portion of the project after the next meeting.

Summary

Reproducibility Project for Data Ethics Class

In the meeting, Yousaf outlined the requirements for the reproducibility project for data ethics class. The project involves finding a paper, reproducing its results, and analyzing its ethical implications. The paper must have a publicly available dataset and code, and be from the last five years. The team will document their initial analysis, reproduction attempt, and ethics analysis. The quality of their documentation, thoroughness, and technical implementation quality will affect their grade. The team also discussed the importance of setting a reasonable statistical variance threshold for reproducing results. They found a potential paper on stock price forecasting, but noted that the required stock price data might not be publicly available.

Historical Stock Price Data Sources

The group discusses data sources for historical stock prices, with concerns raised about using potentially proprietary Bloomberg data. Yousaf demonstrates that Yahoo Finance can be used to obtain daily stock price data from 1990 to 2018, using Apple as an example. The team also considers alternative sources like Kaggle and briefly shares their opinions on coding environments such as Google Colab, PyCharm, and Jupyter notebooks.

Days in a Year Discussion

The team discussed the number of days in a year, including leap years, and found it to be approximately 10,000. They also explored using publicly available data and a code to reproduce results from a paper. The team agreed to reconvene and try out some things before starting the second portion of their work.

Bottom of Form