**Project Meeting Minutes #5 (Week 11 and Week 12)**

**GROUP 3 – DAB422-24W-001**

**Project Name: Crime Rate Analysis in Toronto, Canada**

|  |  |  |
| --- | --- | --- |
| **Date and Time** | **Location** | **Attendees** |
| 19th March and 26th March 2024  8:00 AM– 10:00 AM | 1 Riverside Drive  R1012 | - Prof. Abiodun Sodiq Shofoluwe  - Srilakshmi Gummadidala  - Tehsin Shaikh |

**Agenda Items:**

1. Review changes made based on professor's suggestions: feature selection approach, Lazy Predict, and scaling for LSTM.
2. Preparation & presentation of the Interim Presentation.
3. Discuss and confirm final model selection.
4. Discuss deployment strategies for finalized models.

**Meeting Minutes:**

1. The meeting commenced with a thorough review of the changes implemented based on the professor's recommendations from the previous week. These changes primarily centered around refining the modeling approach to address overfitting and enhance model performance. These include proceeding with a different approach for feature selection, opting for a subset of features, we selected the top 10 features deemed most relevant to the analysis. This strategic decision was intended to streamline the modeling process and improve the interpretability of the results. Additionally, the team utilized Lazy Predict, a tool designed to expedite model evaluation by providing quick insights into the performance of various algorithms. Furthermore, to mitigate potential issues related to overfitting, the team implemented scaling techniques specifically tailored to the LSTM model, aimed at optimizing its performance on the dataset.
2. The interim presentation was expertly prepared and delivered by Srilakshmi Gummadidala and Tehsin Shaikh. During the presentation, we comprehensively covered the work accomplished to date, providing an overview of our project from inception to the current stage. This included introducing the project, summarizing the exploratory data analysis (EDA), and process, discussing the identification of feature importance, and detailing our modeling efforts. Through clear and concise communication, we effectively conveyed the progress and insights gleaned from our analysis to date.
3. Through rigorous testing and comparison, the team arrived at the consensus that the most suitable final model comprises Neural Networks and time series models for temporal analysis. This conclusion was drawn based on empirical evidence suggesting that other classification models were prone to overfitting and therefore deemed less reliable for the task at hand. As part of the ongoing efforts to enhance the analysis and interpretate the model better, a visualization graph illustrating the Neural Network model's performance was created.
4. Additionally, during our discussions, the professor suggested exploring various deployment options suitable for our project's needs. We are actively working on investigating these suggestions to identify the most appropriate deployment techniques tailored to our machine learning models. These suggestions will help us ensure the seamless integration and efficient operation of our models in practical applications.
5. We updated the professor about the absence of two team members, Vinod Soloman Santhakumar and Yen Nga Le. Vinod Soloman Santhakumar has been absent since the Mid-Term Interview due to personal matters, while Yen Nga Le had to undertake emergency travel to her home country, Vietnam.

**Key Resources and Repository:**

* Main dataset link – Crime rate:

[Major Crime Indicators Open Data | Major Crime Indicators Open Data | Toronto Police Service Public Safety Data Portal](https://data.torontopolice.on.ca/datasets/TorontoPS::major-crime-indicators-open-data/explore)

* Additional dataset link - Demographic:

[Police Annual Statistical Report - Arrested and Charged Persons - City of Toronto Open Data Portal](https://open.toronto.ca/dataset/police-annual-statistical-report-arrested-and-charged-persons/)

* GitHub link:

<https://github.com/VinodSolomon/Crime-Rate-Analysis-Toronto>

**Next Meeting date and time:** April 2nd, 2024, 8:00 AM - 10:00 AM