Project Charter for DATA ANALYSIS & MODELING TECHS (DAMT)

Project Name: Bot detection in social media

Problem Statement:
This project aims to develop a more effective method for identifying bot accounts on social media, by leveraging data analysis to address issues such as spam, scams, malware, and state-actor propaganda

Project Goal:
By the end of this project, we'll develop a ~95% accurate bot detection model which will be trained by more than 1M data with 15+ features that will help to reduce spams in online social media.

Business Case:
The business case for this project revolves around safeguarding the social media platform's ecosystem. By deploying a more effective method to identify and combat bot accounts, the project aims to reduce spam, scams, and propaganda, fostering a safer online environment. This can attract and retain users, improve platform reputation, and potentially increase advertising revenue as users engage more confidently with authentic content. Additionally, addressing these issues aligns with regulatory compliance and ethical considerations, enhancing the platform's credibility.

Project Scope:
This project involves:

- Gathering data from social media profiles
- Selecting key features and Preprocessing
- Development of a Machine Learning model to distinguish between bot and human profiles

Deliverables

- Bot Detection Model: A trained model capable of identifying bots with high accuracy.
- Documentation: Comprehensive documentation detailing the model development process, including data preprocessing, model selection, training, and evaluation metrics.

Team Members:

1. Tasdiqul Islam
2. Nagmat Nazarov
3. Varsha Bathala
4. Sai Jyothi Yalamuri
5. Aqsa Yousaf

Benefits:

1. Enhanced social media platform security.
2. Reduced spread of spam and scams.
3. Safer online environment for users.
4. Better protection against malicious activities.

Timeline:

1. Proposal (Problem Identification-) – February 5th, 2024
2. Model Development - February 26th, 2024
3. Analysis – March 26th, 2024
4. Final report- April 18th, 2024