**739916Initial Project Planning Template**

|  |  |
| --- | --- |
| Date | 27 September2024 |
| Team ID | 739916 |
| Project Name | Toxic Comment Classification for Social Media using NLP |
| Maximum Marks | 4 Marks |

**Product Backlog, Sprint Schedule, and Estimation (4 Marks)**

Use the below template to create a product backlog and sprint schedule

| **Sprint** | **Functional Requirement (Epic)** | **User Story Number** | **User Story / Task** | **Story Points** | **Priority** | **Team Members** | **Sprint Start Date** | **Sprint End Date (Planned)** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sprint-1 | Data collection | USN-1 | As a user, I can collect a diverse dataset of social media comments, including both toxic and non-toxic comments from various social media platforms | 5 | High | ABBINENI BHUMIKA | 20/10/2024 | 21/10/2024 |
| Sprint-2 | Data Preprocessing | USN-2 | As a user, I can import libraries like numpy, pandas, matplotlib and seaborn. As a user I can read the datasets | 3 | High | DHARLA POOJITHA  ABBINENI BHUMIKA | 22/10/2024 | 23/10/2024 |
| Sprint-2 | Data Preprocessing | USN-4 | As a user, I In data preprocessing for toxic comment classification, vectorizing involves transforming text data into numerical representations for machine learning models | 3 | Medium | ABBINENI BHUMIKA  DADI ABHINAY | 24/10/2024 | 26/10/2024 |
| Sprint-3 | Model Building&  Model Selection | USN-5 | As a user, I can select Logistic regression is chosen for toxic comment classification on social media due to its simplicity, interpretability, and effectiveness in binary classification tasks. | 3 | High | DHARLA POOJITHA  ABBINENI BHUMIKA | 27/10/2024 | 28/10/2024 |
| Sprint-4 | Model Training | USN-6 | As a user, I can Training a logistic regression model on the preprocessed data . Evaluating the model’s performance by using relevant metrics like accuracy\_score, precision etc. Building a multi-label classifier using Logistic Regression. For each target variable, the model is fitted and saved | 3 | Medium | DADI ABHINAY  DIDDI AKSHAY | 29/10/2024 | 1/11/2024 |
| Sprint-5 | Creating and Generating Submission file | USN-7 | As a user, I can create and generate a submission file for toxic comment classification, you compile your model's predictions into the required CSV format and save the trained model using pickle.dump for future use | 3 | High | ABBINENI BHUMIKA  DHARLA POOJITHA | 2/11/2024 | 4/11/2024 |
| Sprint-6 | Application Building | USF-9 | As a user, In building a Flask app for toxic comment classification, you create API endpoints to accept text input, process it through your trained model, and return the prediction results | 4 | High | DADI ABHINAY | 5/11/2024 | 7/11/2024 |
| Sprint-6 | Model Development | USF-10 | To build an HTML page for toxic comment classification on social media, as a user, you design a web interface to input comments, integrate the model for prediction, and display the results | 4 | Medium | DADI ABHINAY | 8/11/2024 | 10/11/2024 |
| Sprint-6 | Model Development | USF-11 | As a user ,I can Execute the python code and after the module is running, open the localhost page and give the text to be predicted or tested. | 3 | Medium | DADI ABHINAY | 11/11/2024 | 13/11/2024 |
| Sprint-6 | Project Report | USN-12 | Reporting the project | 4 | Medium | DADI ABHINAY  ABBINENI BHUMIKA  DHARLA POOJITHA  DIDDI AKSHAY | 14/11/2024 | 15/11/2024 |