Initial Project Planning Template

Date	26 November 2024
Team ID	SWTID1726490119
Project Name	Toxic Comment Classification for Social Media
Maximum Marks	4 Marks

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

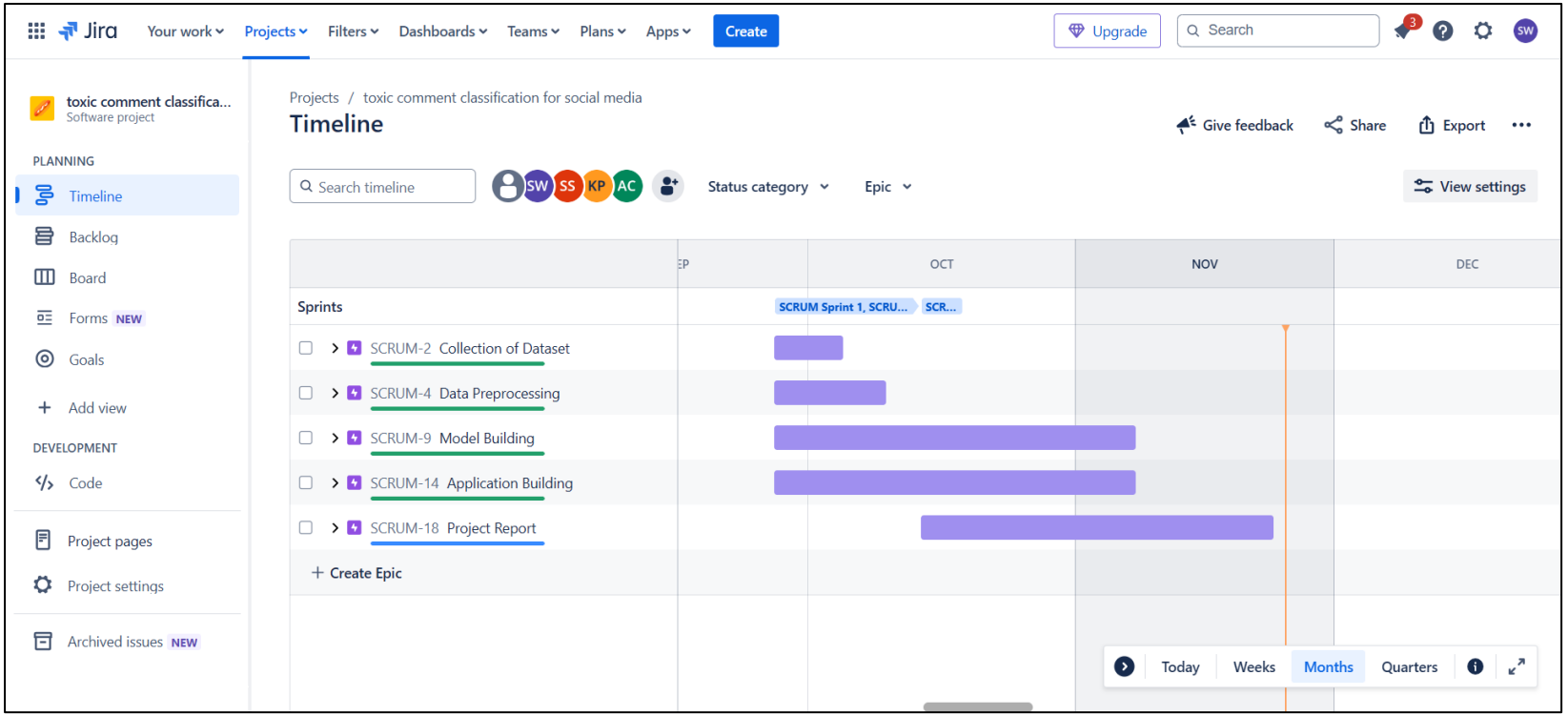
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members	Sprint Start Date	Sprint End Date (Planned)
Sprint-1	Download the dataset	USN-3	Download the dataset from Kaggle. Store it in the project directory for preprocessing.	1	High	Ankita Chavhan	27 September 2024	4 October 2024
Sprint-2	Import required libraries	USN-5	Import necessary libraries such as pandas, numpy, scikit-learn, tensorflow/keras, nltk, and re for text cleaning and processing	1	High	Ankita Chavhan	27 September 2024	9 October 2024
Sprint-2	Read the datasets	USN-6	Load the dataset using pandas.read_csv() to read the data and inspect it using basic head(), info(), and describe() functions	2	High	Ankita Chavhan		
Sprint-2	Text processing	USN-7	Remove special characters, HTML tags, and URLs.	3	High	Saloni Warang		

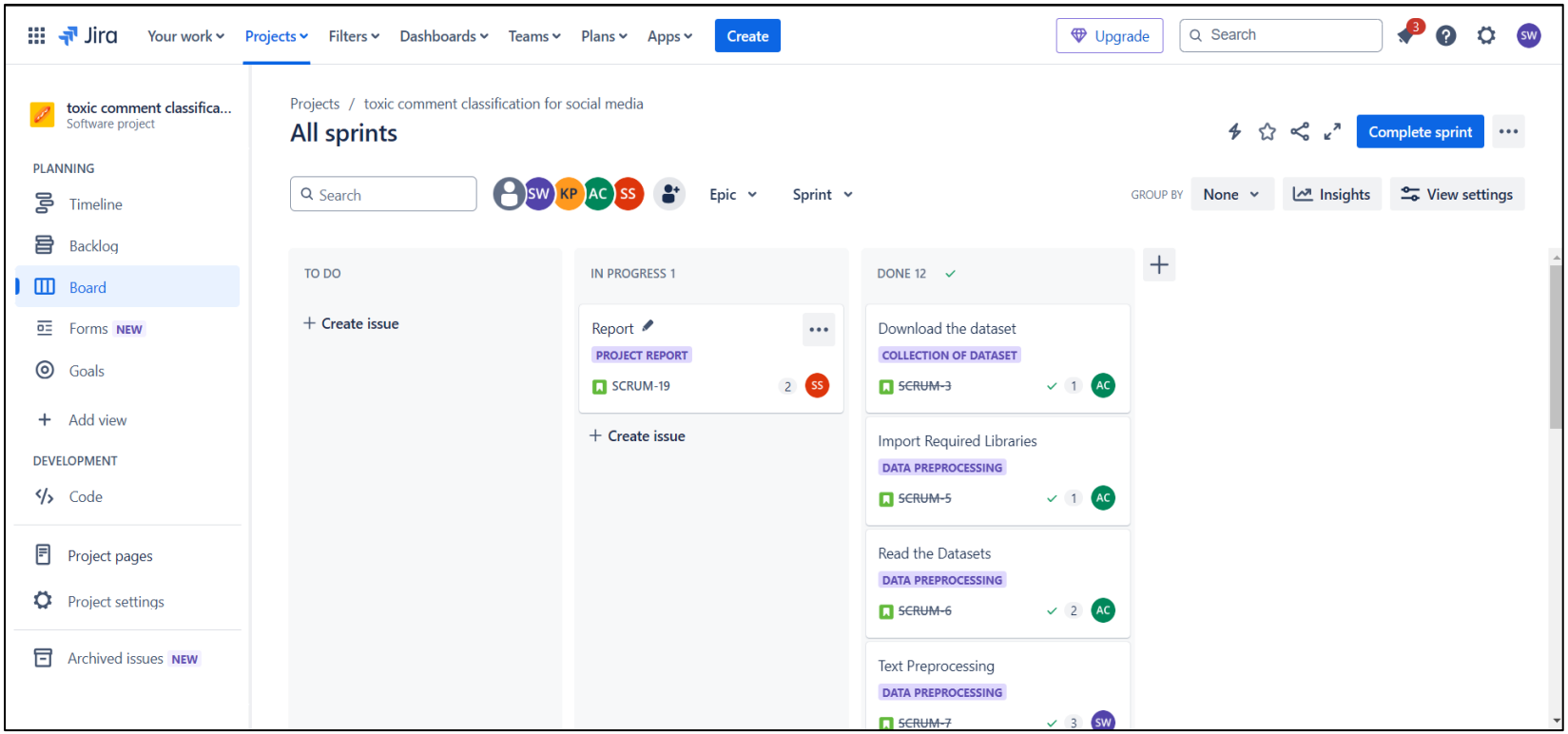
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members	Sprint Start Date	Sprint End Date (Planned)
			Convert all text to lowercase and remove stopwords. Tokenize and lemmatize the text					
Sprint-2	Vectorizing the data	USN-8	TF-IDF (Term Frequency-Inverse Document Frequency) Word2Vec or GloVe embeddings.	3	High	Saloni Warang		
Sprint-2	Exploratory data analysis (EDA) and visualizations	USN-13	Perform exploratory data analysis (EDA) and create visualizations to better understand the dataset, identify patterns, and gain insights that will guide feature engineering and model selection.	5	Medium	Kanishka Pawar		
Sprint-3	Choose the appropriate model	USN-10	CNN (Convolutional Neural Network) for handling text as sequences. RNN (Recurrent Neural Network) or LSTM (Long Short-Term Memory) to capture sequential information. Transformer-based models like BERT if you opt for more advanced models.	3	High	Saloni Warang	27 September 2024	14 October 2024

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members	Sprint Start Date	Sprint End Date (Planned)
Sprint-3	Build and train the model	USN-11	Build the model architecture in TensorFlow/Keras. Train the model using the training dataset, monitor loss, and accuracy metrics	5	High	Kanishka Pawar		
Sprint-3	Hyperparameter tuning	USN-12	Experiment with hyperparameters (learning rate, batch size, epochs, optimizers). Tune parameters to improve model performance.	5	Medium	Kanishka Pawar		
Sprint-4	Flask App	USN-15	Set up a Flask environment and build a simple backend API to load the trained model.	5	Medium	Sonal Solanki	27 September 2024	16 October 2024
Sprint-4	Build an HTML page	USN-16	Design a frontend using HTML/CSS to accept user input (comment). Provide a button to submit the comment for toxic classification.	3	Low	Sonal Solanki		
Sprint-4	Execute and test the model	USN-17	Test the Flask app by sending user comments for classification. Output the result on the HTML page (toxic or non-toxic).	3	High	Kanishka Pawar		

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members	Sprint Start Date	Sprint End Date (Planned)
Sprint-5	Report	USN-19	Report will provide insights into the development process and document key findings for stakeholders.	2	Medium	Sonal Solanki	14 October 2024	18 October 2024

Screenshots:




Jira
Your work
Projects
Filters
Dashboards
Teams
Plans
Apps
Create
Upgrade
Search
toxic comment classifica...
Software project
PLANNING
Timeline
Backlog
Board
Forms NEW
Goals
Add view
DEVELOPMENT
Code
Project pages
Project settings
Archived issues NEW
Projects / toxic comment classification for social media
All sprints
Complete sprint
Search
Epic
Sprint
GROUP BY
None
Insights
View settings
TO DO
Create issue
IN PROGRESS 1
Report
PROJECT REPORT
SCRUM-19
Create issue
DONE 12
Download the dataset
COLLECTION OF DATASET
SCRUM-3
Import Required Libraries
DATA PREPROCESSING
SCRUM-5
Read the Datasets
DATA PREPROCESSING
SCRUM-6
Text Preprocessing
DATA PREPROCESSING
SCRUM-7