Department Of Computer Engineering

**SEM VII & VIII**: WEEK – **Academic Year:** 2021 to 2022

|  |  |
| --- | --- |
| **Date** | **Progress Achieved** |
| 08/08/2021 | 3 topics were shortlisted along with research papers to be referenced. |
| 16/08/2021 | Final topic was decided – **Citizen Problem Detection from Twitter data using Naïve Bayes Classifier** and the workflow was discussed. |
| 02/09/2021 | Project synopsis was updated. |
| 24/9/2021 | Extracted tweets using Tweepy API |
| 7/10/2021 | Preprocessing the data by removing emoji, links and Data cleaning + research about Naive Bayes algorithm |
| 24/10/2021 | Project synopsis was updated with class diagram |
| 8/11/2021 | Trained a Naive Bayes Model with sklearn + research paper and report preparation |
| 15/12/2021 | Research About Realtime Extraction and processing |
| 12/1/2022 | Setup for Apache Kafka and Zookeeper - Realtime Extraction. |
| 22/1/2022 | Realtime Extraction Successful + Feeding the Data to the model  (**Facing Some Issues**) |
| 28/1/2022 | MongoDB setup + saving the tweets in the backend successfully |
| 5/2/2022 | Solving the realtime error and getting last id for saving the tweets from mongoDB collection |
| 8/4/2022 | Finished With Research Paper |
| 10/4/2022 | Saving unique Tweets + adding location , date , time username of the tweets in the mongodb collection |

**Guide reviews :**

**Signature**

Team Member 1: Amruta Koshe

Team Member 2: Nikita Sarode

Team Member 3: Nidhi Vanjare

Team Member 4: Rakshita Tantry

**Project Guide** : Ramya R B Ma’am

**Signature:**

**Date:**