**Major Project-I Progress Report**

**Class – B.E. Sem- VII**

|  |  |
| --- | --- |
| **Project Title** | **A Framework for Social Media Opinion Mining for Low Resource Marathi Text** |
| **Name of the student** | **Dhruv Talati, Naitik Rathod, Nishit Mistry, Manan Parikh** |
| **SAP ID of the student** | **60004180022, 60004180054, 60004180066, 60004180049** |
| **Name of Internal Project Guide** | **Dr. Pratik Kanani** |
| **Name of External Project Supervisor ( If Applicable)** |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Fortnight Milestones** | | **Reporting Dates** | **Brief Description of Work** | **Internal Project Guide Remarks with signature** |
| **Fortnight 1** | **Week 1** | **21/08/21**  **Saturday** | 1. **Project topic approval after presentation.** 2. **Finalized to classify text into 5 classes and not 3 based on the score achieved.** 3. **Model selection for the task of text classification.** 4. **Finalizing XLM-RoBERTa due to its better performance over the other multilingual models.** 5. **Searched for publicly available datasets for marathi tweets.** 6. [**Paper Review: 1 on Experimental Evaluation of Deep Learning models for Marathi Text Classification.**](https://docs.google.com/document/d/1zzuMjG5D7pQU4uklSgXaUVh_1HDbh9R77B5Vhkgz9dc/edit) |  |
| **Week 2** | **28/08/21**  **Saturday** |
| **Fortnight 2** | **Week 3** | **4/09/21**  **Saturday** | 1. **Cleaned the dataset, removed the unwanted English texts, links, and mentions. Kept the marathi hashtags as they could be important.** 2. **Started understanding the architecture and the working on RoBERTa and then building of the XLM-R model for our task.** 3. **Paper review 2-** [**Coarse and Fine-Grained Hostility Detection in Hindi Posts Using Fine Tuned Multilingual Embeddings**](https://docs.google.com/document/d/1cVDtByoaK9ZHVbPPxfIAX2YPId0LRKIcnNAbQYbyJKE/edit?usp=sharing) 4. **Paper review 3:** [**Detection of Low-toxic texts in similar sets using a modified XLM-R Neural networks and toxicity confidence parameters**](https://docs.google.com/document/d/11xipQwG1oEfWOJMaUVT0QgjZq26u9rZ4aWWaIqbVMl0/edit) |  |
| **Week 4** | **11/09/21**  **Saturday** |
| **Fortnight 3** | **Week 5** | **18/09/21**  **Saturday** | 1. **Paper review 4:** [**The Impact of Translating Resource-Rich Datasets to Low-Resource Languages Through Multi-Lingual Text Processing**](https://docs.google.com/document/d/141mEs6vlwi-ZrhGO4PDBq4a8TntAxzFJ2fKn5jEtj30/edit) |  |
| **Week 6** |  |
| **Fortnight 4** | **Week 7** |  |  |  |
| **Week 8** |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Fortnight 5** | **Week 9** |  |  |  |
| **Week 10** |  |
| **Fortnight 6** | **Week 11** |  |  |  |
| **Week 12** |  |
| **Fortnight 7** | **Week 13** |  |  |  |
| **Week 14** |  |

**Student’s Signature with Date**

**Recommendations/Remarks by the External Project Supervisor (If applicable):**

**Signature of Internal Project Guide Signature of External Project Supervisor (If applicable)**