**7PAM2002 Data Science Project Module**

**Choosing a Project Form**

**Semester B 2024/2025**

This form will be used by your supervisor to agree your project topic and dataset. Complete as much of the form as possible then submit the form into ‘Assignments’ on the Project Module Canvas site.

**Student Full Name:** **Ali hamza**

**Student ID: 23037713**

**Email Address: alihamzaamjad006@gmail.com**

**Course: 1 year / 2 year sandwich / 2 year Advanced Research** *(delete as appropriate)*

**Semester intake to the course: A / B** *(delete as appropriate)*

|  |  |
| --- | --- |
| SECTION A |  |
| **What is your proposed project title or topic?** | Sentiment Analysis of Emotions in Twitter and Amazon Reviews |
| **What is the Research Question for your project?** | |  | | --- | | How accurately can machine learning models classify emotions in Twitter and Amazon reviews using sentiment analysis datasets? |  |  | | --- | |  | |
| **Dataset website address (or organisation and person the dataset is from).** | <https://huggingface.co/datasets/carblacac/twitter-sentiment-analysis>  <https://huggingface.co/datasets/hugginglearners/amazon-reviews-sentiment-analysis> |
| **Where was the data originally collected? (who, when, where)** | The collected sample of the Twitter dataset consists of tweets and is expressed of different emotions, by Carb Lacac. The data was produced by social media users worldwide between 2019 and 2021. The product reviews dataset of Amazon’s e commerce platform contributed by customers in general from all over the world in different years until 2023 are curated by HuggingLearners. The two datasets both served a sentiment analysis research purpose where we had labeled positive, negative and neutral sentiments to train and test the machine learning models with such examples. |
| **What type of data are you using? (e.g. image/ tabular/category/continuous etc)** | |  | | --- | | This study uses a tabular dataset that contains textual movie reviews as in one side and sentiments to be labelled on the other side. With this structured format, pre-processing, feature extraction, model training can all be done in efficient ways for sentiment analysis. |  |  | | --- | |  | |
| SECTION B |  |
| Can you attend all 6 supervision sessions on-campus? If not state reason. | Yes I Can |
| What was the subject of your BSc degree and any other Master’s degrees you have taken? | BS In Mathematics |
| What are your career aims and/or the industry sector you would like to get into? | To pursue a career in data science or machine learning, ideally within the e-commerce or technology sector. |
| A brief account of your programming/data science experience (including work or placement). | - Proficient in Python and libraries such as Pandas, NumPy, and Scikit-learn. |
| Any work experience (including non-computing related). | I have Worked healthcare sector in Pakistan. |
| Do you have any hobbies, activities or interests. | Passionate about NLP and sentiment analysis. |