Minutes 1

|  |  |
| --- | --- |
| Subject: | COMP90051 Statistical Machine Learning |
| Group name/ identifier | Group 14 |
| Meeting Location, Date & Time | ERC Library, 31/08/2023 4.00pm |
| Group members present | Yinuo Sun, Yufeng Xie, Xueke Zheng |
| Apologies |  |

# Actions from previous meetings

|  |  |  |
| --- | --- | --- |
| **Task** | **Who is responsible** | **Deadline** |
| Read the project specification | All | 31/08 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

# Items discussed and decisions made

|  |  |
| --- | --- |
| **Item** | **Decision (if applicable)** |
| Brainstorm about the main task of the project |  |
| Look at the datasets and discuss about the strategy of balancing domain 1 and domain 2 data | Try resampling methods, e.g. under resample / over resample |
| Baseline model choice | Try logistic regression/ SVM/ MLPC models etc. |
|  |  |
|  |  |

# Agreed actions from this meeting

|  |  |  |
| --- | --- | --- |
| **Task** | **Who is responsible** | **Deadline** |
| EDA of the raw data | Yufeng | 04/09 |
| Vectorization of the text | Xueke | 05/09 |
| Baseline model training | Yinuo | 07/09 |
|  |  |  |
|  |  |  |

# Next meeting

|  |  |
| --- | --- |
| Location | ERC Library |
| Date & Time | 07/09/2023 |
| Goals | Discuss about solution of any issues coming in the working process;  Evaluate baseline model performance;  Try more complex models |