**16:52 submission | COMP90049 Report for Toxicity classification in online comments**

The goal of the report was to experiment with including unlabelled data and reduce the number of feature dimensions but make them more meaningful and see what effects this produces when classifying online toxicity with different machine learning models. The results evaluated by the machine learning models used accuracy, precision, recall macro F1-score and the area under curve. Then each model was introduced and evaluated, and a selection of the best performing models were chosen to start the experiment.

The author has done well in respects to the logical flow of the report where they have shown their thinking and have utilised meaningful data to illustrate their decision-making process. This stems from elaborating on related literature where anonymous has explained the key concepts of each paper. Additionally having a summary of datasets as a table and referring to the dataset while explaining the question and sets the background necessary for a reader to follow.

The author mentioned key aspects of literature but could’ve gone into more detail regarding the outcomes and why that paper is relevant to the report. Also, the discussion was too short and did not fully justify the decisions made to answer the questions nor did it explicitly mention a hypothesis regarding the research questions. This could’ve been compensated by reducing the explanation (definitions of common terms) as the audience are machine learning students and rationalize the decisions and thought process more on the models used in relation to the question and hypothesis.

**15:29 submission | Toxicity classification in Online Comments**

The experiment uses Naïve Bayes and Logistic regression to reviews the performance of the two algorithms and whether they can classify toxicity in online comments. This extends to investigating labelled and unlabelled data through semi supervised learning and determining whether unlabelled data improves effectiveness of toxicity classification.

Anonymous has done well regarding the literature review where they explain ley aspects of the papers referenced and summarise the reports succinctly and are relevant to the experiment. They also expanded well on the performance model and the theoretical properties and answers their research question well. The conclusion

The report could have been improved by having a logical progression of the decision making and justify the steps taken and why those were relevant to the experiment, so in a sense; reflecting on the methods used. Experimenting with distinctive features of the data and different parameters would have yielded more thorough investigation and a discussion on how this was done would have been appropriate as well as the performance of each individual experiment.