ISY503 Intelligent Systems Individual report

In our group project we focused on the analysing of Amazon product reviews, preprocessing and modelling phases were my main part in this group project. I contributed mostly in data cleaning by developing scripts to remove punctuation, normalize text, and filter out extremely short reviews, ensuring high-quality input for our model. My partner and I we implemented the tokenization process, converting the text data into numerical formats which are suitable for machine learning algorithms. Secondly, I also worked in neural network architecture included in embedding layers and LSTM cells and give machine training at the end. In the machine training I monitored accuracy and loss metrics, making adjustments to improve performance.

Each group member contributed significantly to the project.

In sentiment analysis, several ethical considerations need to pay attention on. The potential for algorithmic bias is very possible; if the training data is imbalanced or reflects prejudiced opinions, the model may inadvertently include these biases. In additional, privacy concerns arise with user generated content, “as obtaining and utilizing such data without consent poses ethical dilemmas (Binns, 2018)”. Transparency in model predictions is crucial to maintain user trust and accountability. Additionally, care should be taken to ensure that the model does not perpetuate harmful stereotypes or misinformation, necessitating the implementation of fairness-aware algorithms (Caliskan et al., 2017).

**References**

* Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency*, 149-159.
* Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora necessarily contain human biases. *Science*, 356(6334), 183-186. DOI: 10.1126/science.aal4230.

Top of Form