Research Proposal Review

**Description**

Considering your thoughts on your chosen area of interest for your project:

* Which of the methods described in this week's reading would you think would suit your purpose?
* Which data collection methods would you consider using?
* Which required skills will you need to have or develop for the chosen project?

**Solution**

**1. Which of the methods described in this week’s reading would suit your purpose?**

For this project, a conclusive, descriptive research design would be most appropriate. My aim is to develop and test a sentiment analysis framework and then evaluate its performance using measurable metrics (accuracy, F1-score, precision, recall). This fits with a quantitative approach, as I will be using structured experiments on multiple datasets to produce replicable and comparable results. While primarily quantitative, the early stages will also include exploratory elements, for example identifying existing tools, preprocessing techniques, and algorithms to inform the final system design.

**2. Which data collection methods would you consider using?**

I plan to use secondary data collection by sourcing publicly available sentiment-labelled datasets such as Twitter/X Sentiment140, IMDB movie reviews, and Amazon product reviews. These will allow structured experimentation without the need for primary data gathering. However, if I decide to test domain adaptability, I may collect primary data via online surveys or questionnaires to create a small, domain-specific sentiment dataset for model fine-tuning. All datasets will be prepared through preprocessing methods such as tokenisation, stop-word removal, and feature extraction.

**3. Which required skills will you need to have or develop for the chosen project?**

To complete this project, I will need to develop or strengthen:

* Machine Learning and NLP skills, implementing algorithms (SVM, Random Forest, CNN, LSTM, BERT) and preprocessing pipelines.
* Programming skills, proficiency in Python and libraries such as Scikit-learn, TensorFlow, PyTorch, and Hugging Face Transformers.
* Data handling and preprocessing skills, cleaning, transforming, and feature engineering for text data.
* Evaluation and statistical analysis, applying and interpreting performance metrics in a quantitative research framework
* Research design and reporting, aligning methodology with research questions, ensuring reproducibility, and applying ethical considerations in dataset use.