Find out the Data source

Expected results! getting a 95 or 98% accuracy etc.

What software will you use! List down

**Research Aims ..**

Objectives 1

Objectives 2

Objectives 3

Objectives 4

# **What software will you use! List down**

1. Python: Key for data processing, sentiment analysis, and model creation.
2. Web scraping tools:
   * Praw
   * Google API Client
   * BeautifulSoup
   * Selenium
3. Machine learning libraries:
   * scikit-learn
   * Naive Bayes
   * Support Vector Machines (SVM)
   * Recurrent Neural Networks (RNN)
4. Natural Language Processing (NLP) libraries:
   * NLTK (Natural Language Toolkit)
   * spaCy
5. Sentiment analysis tools:
   * VADER
   * BERT
   * RoBERTa
   * LexMo
6. Data visualization tools:
   * Matplotlib
   * Seaborn
7. Version control:
   * GitHub
8. Integrated Development Environment (IDE):
   * Jupyter Notebook

These software resources will be used for various tasks including data collection, preprocessing, model development, analysis, and visualization as outlined in the project proposal.

# **Find out the Data source**

1. Social Media Platforms:
   1. YouTube
   2. Reddit
2. E-commerce Platforms:
   1. Amazon
   2. eBay

These platforms will be used to collect user-generated content such as comments, reviews, and discussions related to Apple products for sentiment analysis. The data will be obtained through a combination of APIs (for platforms that provide them) and web scraping (for platforms without APIs).

# **Proposed Model Development**

The proposed model development for this research project involves the following steps:

1. Data Preprocessing and Splitting:
   * The data will undergo preprocessing to ensure it is clean and ready for modelling. This could entail tasks such as eliminating unwanted elements like noise, irrelevant data, and special characters, as well as performing actions like tokenization and lemmatization, along with other text-processing procedures.
   * Following preprocessing, the data will be split into two sets: the training set, used to train the models, and the testing set, employed to evaluate their effectiveness.
2. Model Selection:
   * The choice of algorithms for sentiment analysis will be informed by the literature review. Some of the algorithms mentioned in the proposal are:
     + Naive Bayes
     + Support Vector Machines (SVM)
     + Recurrent Neural Networks (RNN)
     + BERT
     + LexMo
   * These algorithms have been identified from previous research for their effectiveness in sentiment analysis tasks.
3. Model Training:
   * The selected algorithms will be trained on the training set of data. This step involves feeding the algorithm with labelled data (where sentiments are known) to learn the patterns associated with positive, negative, or neutral sentiments.
4. Model Evaluation:
   * After training, the models will be evaluated using the testing set to assess their performance. Common evaluation metrics may include accuracy, precision, recall, F1-score, and others depending on the specific goals of the sentiment analysis.
5. Fine-tuning and Hyperparameter Optimization:
   * Depending on the initial results, models may be fine-tuned or their hyperparameters adjusted to improve performance.

Overall, the goal is to select and train models that accurately predict sentiment from the collected data. The choice of algorithms will be guided by the literature review and experimentation with different methods to find the most effective approach.

# **Research Aims**

The project aims to investigate user sentiments related to Apple products across diverse digital platforms. It seeks to understand how these sentiments change over time and their broader implications for business strategies. Additionally, the project aims to contribute to the development of sentiment analysis methodologies within the digital landscape and enhance academic discussions on language processing and sentiment analysis techniques.

* Compare and contrast different sentiment analysis methods like Naive Bayes, Support Vector Machines, Recurrent Neural Networks, BERT, and LexMo for social media and e-commerce text data, making sure they accurately interpret sentiment while considering platform changes, and product features.
* Perform an in-depth comparative sentiment analysis on discussions surrounding Apple products, encompassing both social media and e-commerce platforms, utilizing advanced data science techniques(natural language processing, topic modeling, time series analysis, text analytics, data mining, etc.).
* Investigate and analyse the dominant emotions people feel in separate platforms about Apple products over time, considering factors like product updates (release of iPhone 15), to understand sentiment trends.
* Discover common sentiment trends, key themes, important keywords, and significant topics to understand what influences people's opinions about Apple products on social media and e-commerce platforms.
* Assess how sentiment analysis can benefit business strategies, including innovation, marketing choices, and customer satisfaction, with a specific emphasis on offering actionable advice for companies like Apple.
* Contribute to improving sentiment analysis methods in the ever-changing digital world, enhancing academic conversations about language processing and sentiment analysis techniques.

# **Expected Results**

Although the research project proposal outlines a thorough strategy for conducting sentiment analysis on Apple products across social media and e-commerce platforms, it's crucial to acknowledge that accurately predicting specific percentage accuracies can be difficult. The performance of sentiment analysis models relies on various elements, including the quality and diversity of the data, the harmony of the algorithms used, and the intensity in the emotions expressed within the text.

Nonetheless, the aim of the project is to attain a high degree of accuracy in predicting sentiments. The specific target for accuracy would ideally be determined during the model development and evaluation phase. It's important to keep in mind that achieving extremely high accuracy, such as 95% or 98%, may be challenging due to the inherent complexity and subjectivity of sentiment analysis. Additionally, the project may prioritize other metrics like precision, recall, or F1-score depending on the specific research goals.

The project's success would also be measured by its ability to uncover meaningful insights and trends in customer sentiments across different platforms, which can be just as valuable as high accuracy in certain contexts. Ultimately, the expected results would be a combination of accurate sentiment predictions and valuable insights for strategic decision-making in the tech industry.