**NON TECHNICAL PRESENTATION/ DATA REPORT**

**Project Overview**

The dataset comprises tweets expressing sentiments about Apple and Google products or brands, categorized as either positive or negative.

Stakeholder

* Marketing Teams: Need insights for campaign effectiveness and brand perception
* Product Management: Requires feedback on product features and customer satisfaction
* Customer Service: Needs to identify and address negative feedback quickly
* Executive Leadership: Requires high-level brand health metrics

Conclusion

Problem Statement

How can we develop a robust sentiment analysis model that accurately identifies both positive and negative customer sentiments about Apple and Google products from Twitter data, with particular emphasis on improving the detection of negative feedback

Objectives

1. Accurate Sentiment Classification: Develop a machine learning model to classify tweets into positive and negative sentiments with high accuracy, ensuring actionable insights for stakeholders.
2. Improve Negative Sentiment Detection: Focus more on recall accuracy
3. Actionable Insights for Stakeholders: Provide clear and interpretable results to marketing, customer service, and leadership teams for timely decision-making and resource allocation.
4. We will track weekly sentiment trends and market share of voice for Apple and Google, providing concise brand comparison reports to monitor consumer perception and changes over time.

Metrics of Success

Accuracy Score of X%

Precision Score of Y%

Recall

F1 score

**DATA UNDERSTANDING**

Discussing your data: Source of your data, What are the columns, What do they mean, How many rows do you have

1. tweet\_text

* Contains the actual content of tweets
* Raw text data including mentions (@), hashtags (#), and URLs
* Primary source for sentiment analysis

1. emotion\_in\_tweet\_is\_directed\_at

* Specifies the target brand/product (e.g., iPhone, iPad, Google)
* Helps identify which specific product receives the sentiment
* Useful for product-specific analysis

1. is\_there\_an\_emotion\_directed\_at\_a\_brand\_or\_product

* Sentiment classification with four categories:
  + No emotion toward brand or product: 5,389 tweets
  + Positive emotion: 2,978 tweets
  + Negative emotion: 570 tweets
  + I can't tell: 156 tweets

**DATA PREPARATION & ANALYSIS**

Discuss what to check on your data like missing, duplicate values, outliers, null values.   
Discuss how to deal with them? Percentage of the missing, null, duplicate values and is imputation, drop, forward/backfill the option to use

**Data Analysis**

Discussing EDA; Univariate, Bivariate & Multivariate Analysis

Have the visuals from the Technical work

**MODELING**

Mentioning the Models to use; have like 3 models (Baseline Model, 2nd Model, 3rd Model (hyperparameter tuned))

Explain (Justification) why you’re using the mentioned models.

Refers to your metric of success mentioned above

**EVALUATION**

Discussing each model and it performance

Mention the one that performed best and why

**CONCLUSIONS**

What are your findings from the analysis and modeling

**RECOMMENDATIONS**

Giving suggestions based on your findings; analysis and modeling

**NEXT STEPS**

Deployment for access to end users

Collecting more data points (a, b, c)