**Sentiment Analysis of Patients complain**

#### Problem:

Prior to the COVID-19 pandemic, telehealth was a growing field but faced significant barriers in terms of widespread adoption among healthcare providers and patients.

These barriers included technological challenges, regulatory issues, lack of infrastructure, and resistance due to traditional healthcare delivery preferences.

With spread of COVID-19, which has dramatically shifted the healthcare landscape, necessitating the adoption of remote healthcare practices to maintain continuity of care while adhering to social distancing guidelines.

**1. Project Overview**

This project aims to classify patients covid complain into three categories: emergency, non-emergency and follow up complains. By analyzing the sentiment expressed in the text of the complains, this model can help in understanding patients symptoms and improving recommendation systems whether to go to hospital or to stay at home and follow up the symptoms for improvement or progression. The dataset consists of around 4,000 complains, equally divided between the three sentiment classes.

**2. Data Acquisition**

The data was sourced from a comprehensive dataset containing 4,000 patients complain. Each complain is labeled with either 'emergency', ‘non-emergency’ or 'follow-up', indicating the sentiment.

**3. Data Exploration and Preprocessing**

**3.1 Data Cleaning**

- Duplicate Removal: The dataset contained no duplicate entries.

**3.2 Data Visualization & Analysis:**

* Plot a histogram that showed that complains range from 10 to 110 characters
* Then we plot histogram for emergency complains and non-emergency complains respectively.
* In general, people comment more words in the non-emergency complains than with emergency complains. However, the range of word for emergency complains are bigger than the range of non-emergency complains. It means in some cases, people give a long comments for emergency complains and people could less talk for non-emergency complains.

a wordcloud graph show the most used words in large font and the least used words in small font in emergency complains, non-emergency complains and follow-up

**3.3 Text Preprocessing**

- Preprocess the text data by removing stop-words, punctuation, removing URL links, special characters, emoji, short form and converting text to lowercase.

Then Working with the most Frequent Words and removing it.

* Tokenization and Lemmatization\*\*: complains were tokenized (breaking text into individual words) and lemmatized (reducing words to their base form).

**4. Feature Engineering**

* Convert the preprocessed text data into numerical representations suitable for machine learning models.
* Use techniques like bag-of-words (BoW) Representation: We initialize a CountVectorizer object. We fit and transform the corpus using the fit\_transform() method of CountVectorizer. This generates a sparse matrix where each row corresponds to a document in the corpus, and each column corresponds to a unique word in the corpus. The cell values represent the frequency of each word in the corresponding document.
* term frequency-inverse document frequency (TF-IDF), or word embed-dings (e.g., Word2Vec, GloVe) to represent textual features.
* Explore the use of n-grams and other text features to capture context and semantics in the complains.

**5. Model Selection and Training**

**5.1 Model Selection**

Choose appropriate machine learning algorithms for sentiment analysis, such as naive Bayes. Split the dataset into training and testing sets for model evaluation. Train the selected models on the training data and fine-tune hyperparameters if necessary.

**5.2 Model Training**

The dataset was split into an 80/20 ratio for training and testing. Models were trained using the training set, and parameters were tuned to optimize performance.

**6. Model Evaluation**

Models were evaluated based on accuracy, precision, recall, and F1-score:

- Accuracy: Measures the overall correctness of the model (94.12 %)

- Precision and Recall: Important in scenarios where false positives and false negatives have different implications.

- F1-Score: Combines precision and recall into a single metric, which is useful when seeking a balance between these two metrics.

#### Trying different n-grams: accuracy decreased with 2 (74.25%) and 3 n-grams (50.38%)

#### TF-IDF: Term Frequency-Inverse Document Frequency: accuracy increased to 94.12%

**7. Conclusion and Future Work**

The sentiment analysis model achieved satisfactory performance, effectively classifying patients complain. Future work could explore more complex models such as deep learning architectures or ensemble methods. Additionally, expanding the dataset and including complains in different languages could provide more comprehensive insights.

**8. Challenges and Learnings**

The project highlighted the importance of thorough preprocessing and the impact of feature selection on model performance. The need for computational efficiency was also a key consideration, particularly in processing large text datasets.