**LOGICALLY INTERVIEW SENTIMENTAL ANALYSIS TASK**

**DATE**:15-12-2021

**PROBLEM DEFINITION**: Given a sentence and a named entity contained within a sentence, predict the sentiment towards the named entity

**PROCEDURE :**

1. Mounting the google drive to import the dataset from the drive
2. unpacking the dataset
3. Importing the Libraries
4. Converted dataset to data frame
5. Verifying the null values in data frame
6. Checking the positive and negative sentences
7. Implementing the cleaning and text preprocessing techniques for cleaning the data frame
8. Converting the text to bag of vectors(Count Vectorizer) array
9. Convert the Labels into Numeric values using Label encoder
10. Importing the ML Models
11. Fitting the train data set
12. Predicting the test dataset
13. Plotting the confusion matrix
14. Validating the Best ML algorithm
15. Save the Result
16. Save The model

**ML Model used:**

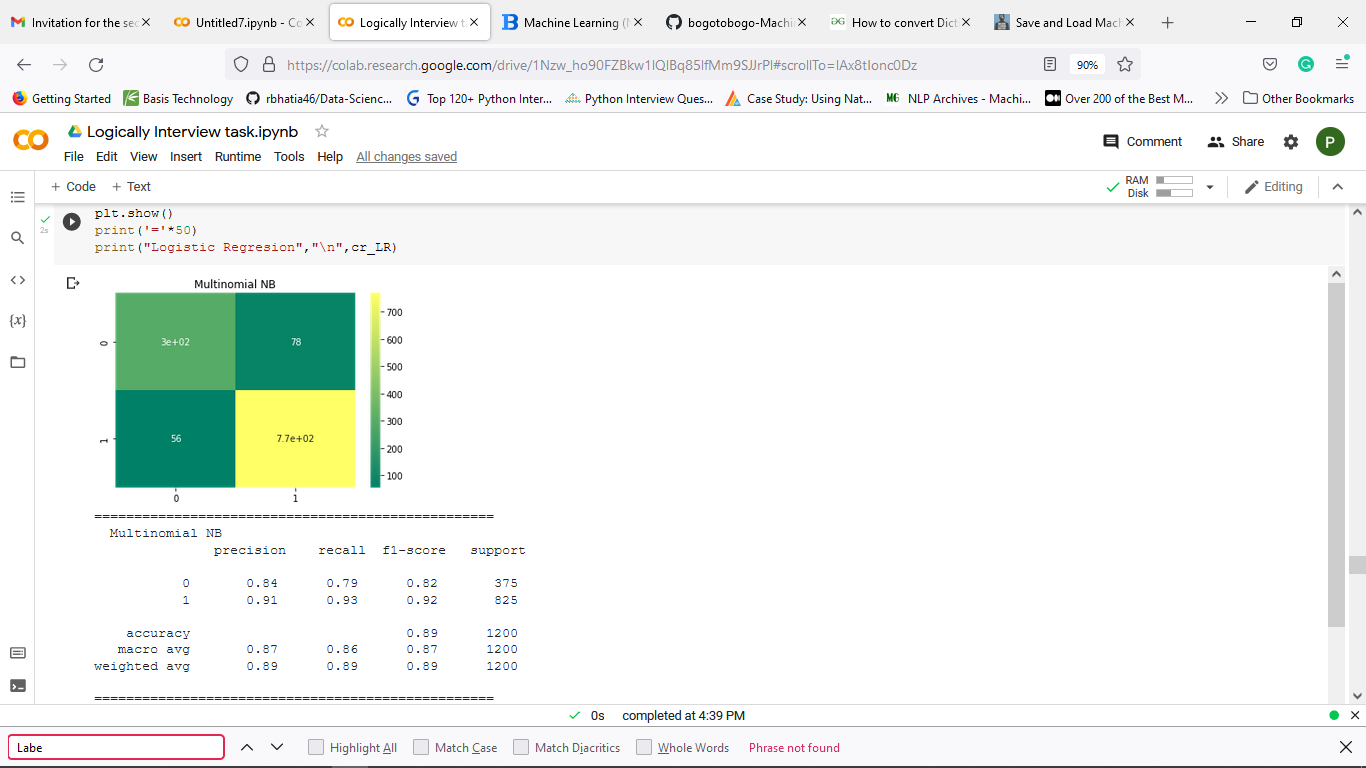
Logistic Regression, Decision Tree, Random forest, Gaussian Naive Bayes, Multinomial Naive Bayes.

**ML Model Accuracies:**

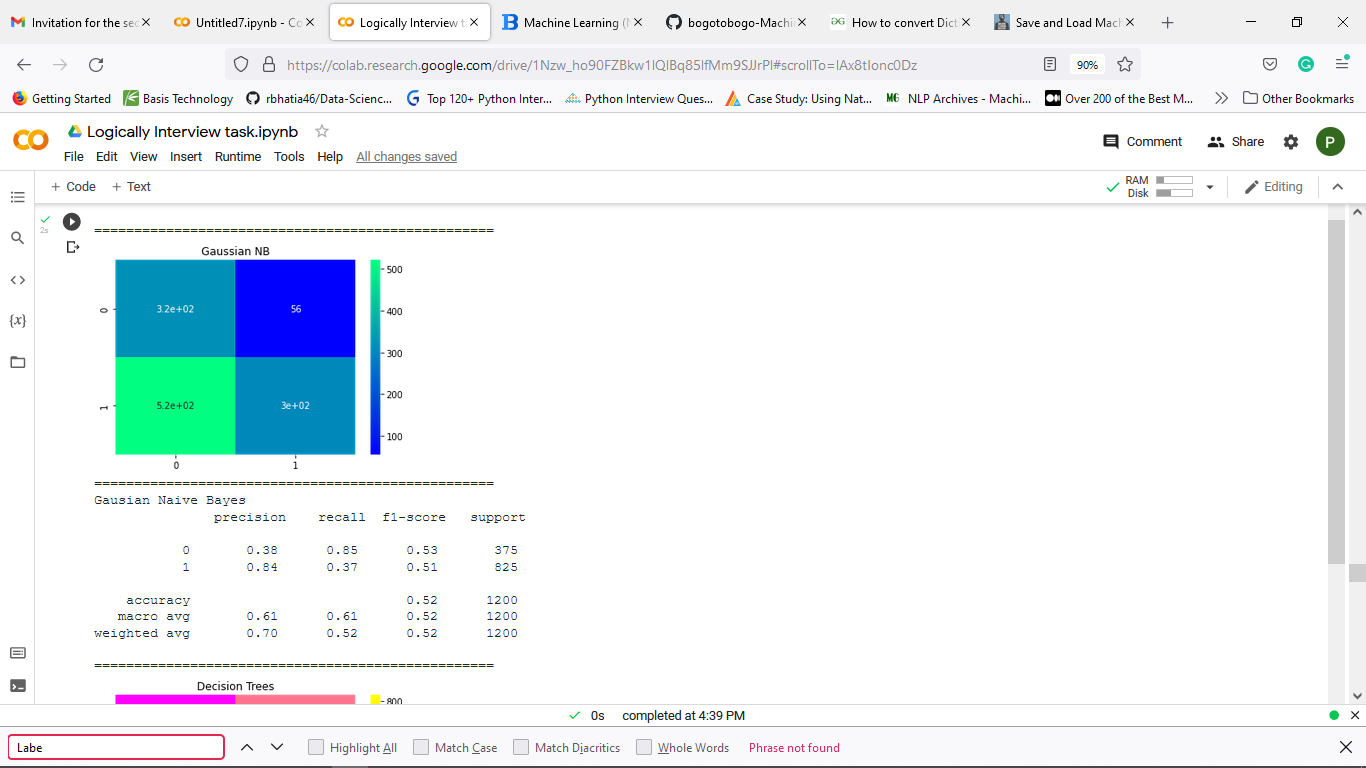
| **ML MODELS** | **Accuracy** |
| --- | --- |
| Decison Tree | 0.686144 |
| Multinomial Naive Bayes | 0.899833 |
| Gausian Naive Bayes | 0.652755 |
| Random forest | 0.979967 |
| Logistic Regresion | 0.943239 |

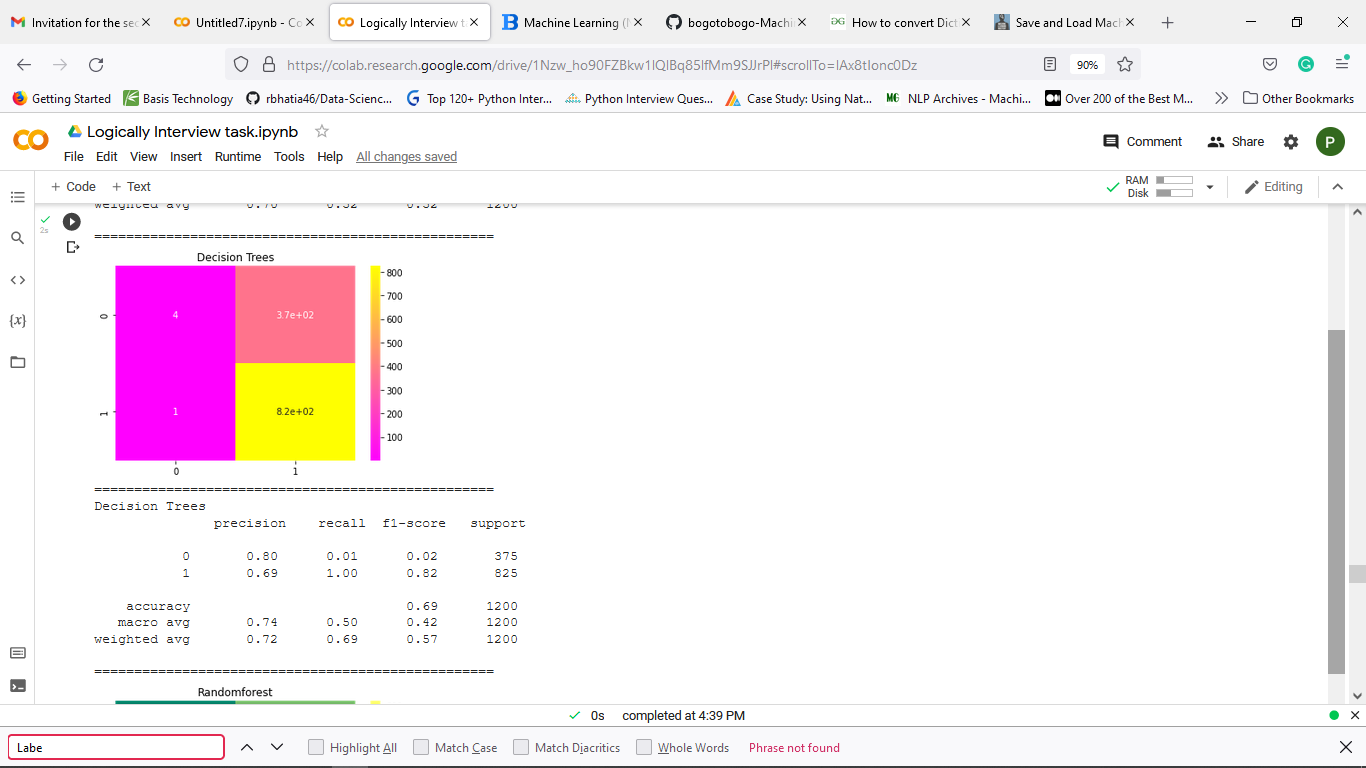
**Performance Metrices:**

**1.Multinomial Naïve Bayes**

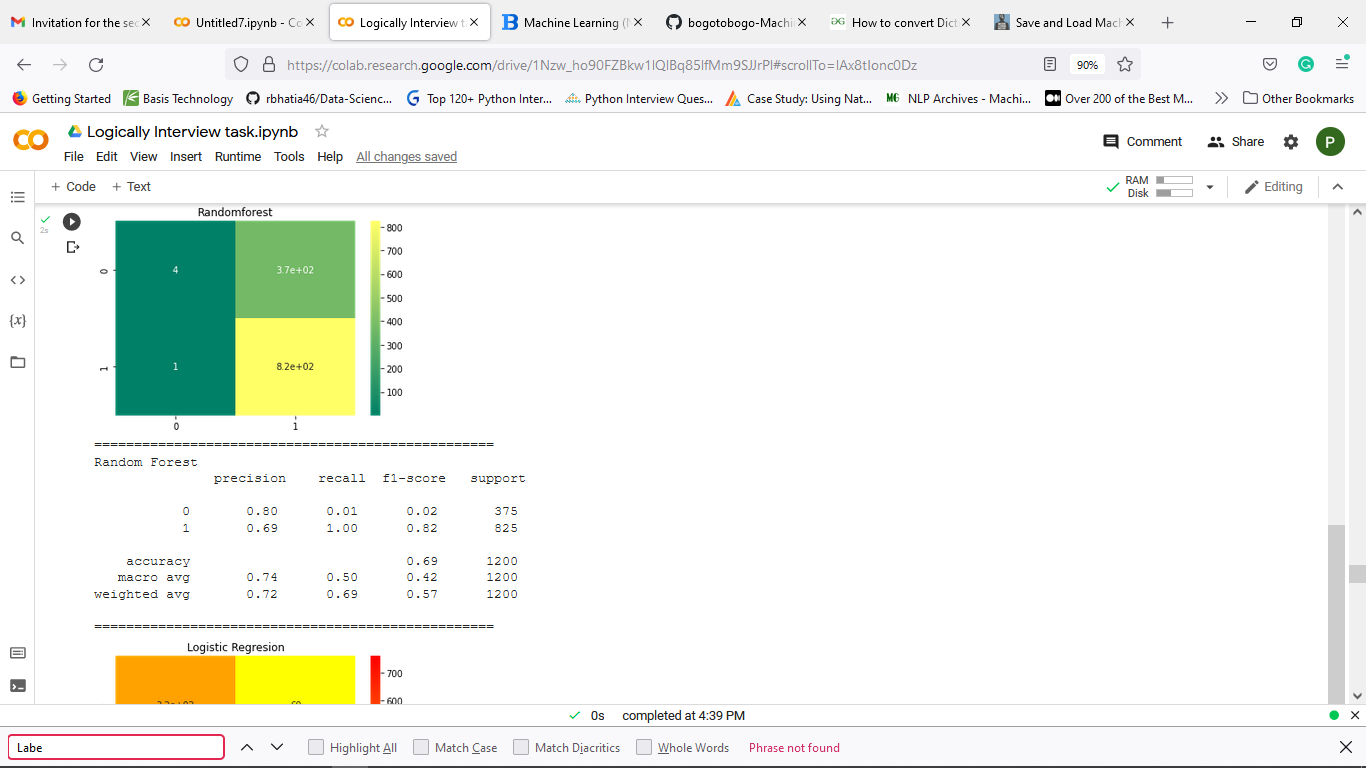


**2. Gausian Naïve Bayes**

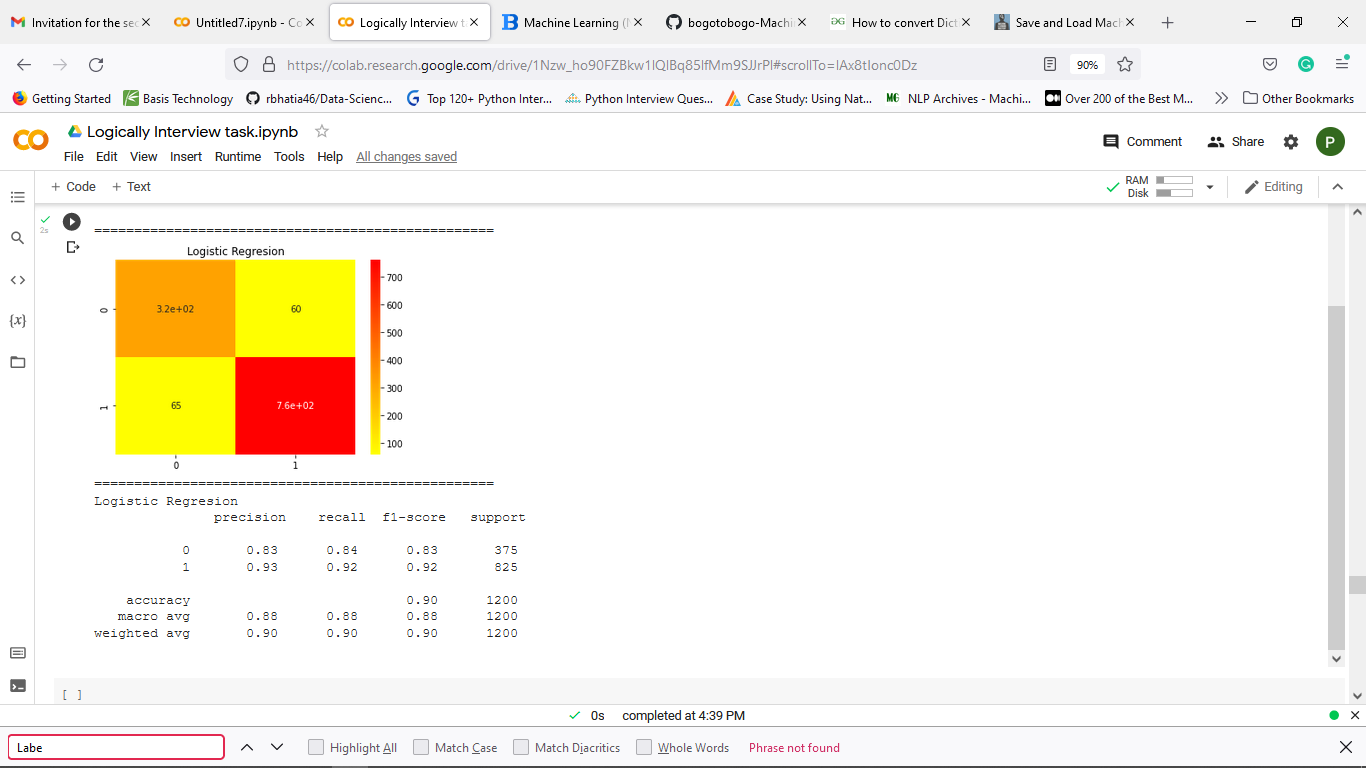


**3.Decison Tree:**

**4.Random Forest**



**5.Logistic Regresion:**



**QUESTIONS TO THINK**:

1. How does your model handle contrastive conjunction and negation?

The goal of contrastive learning is to learn such embedding space in which similar samples are close to each other while dissimilar ones are far apart. There are 3 methods for augmenting text sequences:

Back Translation, Lexical edits, cutoff

1. How do you deal with sentences that have multiple named entities and opposing sentiment (e.g. The car has great cruise control but dash controls are terrible)

we can try **Aspect-level** or **Entity-level** Sentiment Analysis. There are good efforts have been already done to find the opinions about the aspects in a sentence. You can find some of works [here](https://scholar.google.com/scholar?hl=en&q=aspect%20based%20sentiment%20analysis&btnG=&as_sdt=1%2C5&as_sdtp=). You can also go further and deeper and review those papers that are related to **feature (aspect) extraction**. What does it mean? Let me give you an example:

The car has great cruise control but dash controls are terrible."

Document-level sentiment analysis may not give us the real sense of this document here because we have one positive and one negative sentence in the document. However, by aspect-based (**aspect-level**) opinion mining, we can figure out the senses/polarities towards different entities in the document separately. By doing feature extraction, in the first step, you try to find the **features (aspects)** in different sentences (in here "quality of screen" or simply "quality" and "battery life"). Afterwards, when you have these aspects, you try to extract opinions related to these aspects ("great" for "quality" and "short" for "battery life"). In researches and academic papers, we also name features (aspects) as **target words** (those words or entities on which users comment), and the opinions as **opinion words**, the comments that have been stated about the target words.

3.How do you handle unseen vocabulary

* All missing words assigned to some unique vector (say all zeros)
* Find words similar to it and use their embedding:
* Try ngrams (prefix or suffux) of the words and check if it is in vocab
* Stem the word and check if it is in vocab
*  Simplest Solution: Use FastText. It assembles word vectors from subword n-grams which allows it to handle out of vocabulary words

4.Identify the key limitations of your approach

* Gather your data
* Clean your data
* Good Data Representation
* One-hot encoding (Bag of Words),TFIDF,Word2vec
* Visualizing the embeddings
* Classification
* Inspection

# THANKS FOR THE OPPORTNITY