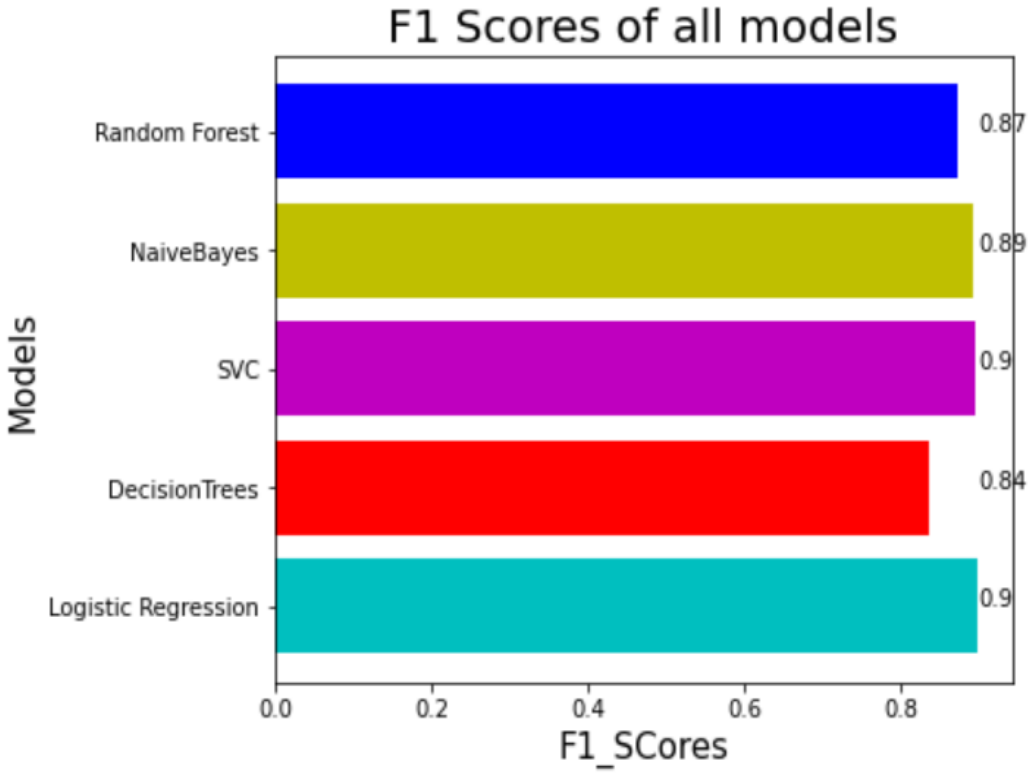
Report on Sentiment Analysis for the airline data

1. I have taken the airline data and have cleaned the dataset also checked for missing values and any duplicates then have done all the preprocessing work by making the data easy to work with and have made data analysis, modelling, tuning and predicting the same dataset.
2. The ML algorithms I used are Logistic Regression, Decision Tree, Random Forest, SVC and Naïve Bayes classifiers
3. I have done Hyperparameter tuning using RandomSearchCV method and have used best parameters for evaluation. The evaluation metrics I used are confusion matrix and f1 scores.
4. On complete analysis and evaluation, I found that logistic regression is able to predict the airline sentiment with 90% accuracy so the Logistic Regression model is the best model for this dataset.

	Model	Score
0	Logistic Regression	0.897823
1	DecisionTrees	0.835833
2	SVC	0.895817
3	NaiveBayes	0.891355
4	Random Forest	0.873040

The f1 scores of the models 1


F1 Scores of all models
Random Forest
NaiveBayes
SVC
DecisionTrees
Logistic Regression
0.87
0.89
0.9
0.84
0.9
Models
0.0
0.2
0.4
0.6
0.8
F1_SCores