Analysis Report

Problem Statement-:

An airline company collects information on all passengers who have flown in their airlines and are interested to do an analysis of customer satisfaction. They've recorded the satisfaction of the customer over several of their in-flight services.

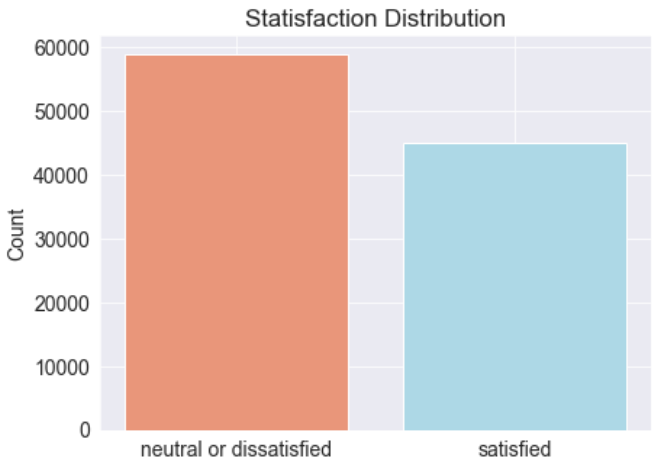
As a solution specialist the client has approached you to analyse the data and arrive at solution using machine learning to predict the customer satisfaction rate.

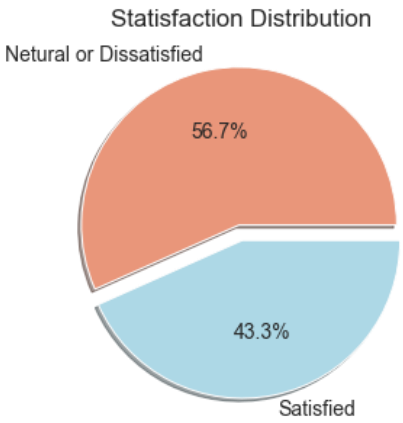
Data –

Data Contains 2 files Train and test- Train file is for training the ML model whereas the test file is to for testing Ml model

Data contains 24 Different columns

Analysis-





- The Overall train data shows that 56.7 % Passengers are dissatisfied with the services offered by the airline where as the positive rate is 43.3% which not even the 50% of the data

Gender	Dissatisfied	Satisfied
Female	29	22
Male	28	22

- The Data represents out of total passengers 29% female are dissatisfied with the services where 28% are the male passengers the positive count is 22%

❖ By Further bifurcating the data an important factor came into a light that **43%** of Loyal Customers are dissatisfied with the services where are the positive count is 39% and just 4% for Loyal and disloyal customers
Further Data represents the same data

Customer Type	Dissatisfies	Satisfied
Loyal Customer	43.7	39
disloyal Customer	13.9	4.3

- Type of travel class is also an significant factor to analyze a passengers is satisfied with the services offered to him/her , so by drilling down with the data of travel class a point is highlighted that the majority of Eco and Business class passengers are not satisfied with the services i.e 15% of business passengers and 37 % of Eco Passengers, where as the positive rate for business class is 33% which is fine but for Eco class airline services must be improved as most of the passengers opt for Eco class

Class	Dissatisfied	Satisfied
Business	14.6	33.2
Eco	36.6	8.4
Eco Plus	5.4	1.8

Feedback/Rating Analysis-

Data Contains some feedback/rating columns of passengers , different insights can be seen from the rating/feedback data of passengers

The data contains different columns as

- Inflight wifi service- Most of the passengers have rated this service as 2 and 3 star but if I talk about overall satisfaction ratio that is very less
- 2. Departure/Arrival time convenient- Majority of passengers have given 4 and 5 ratings but most of the passengers are dissatisfied with this service
- 3. Ease of Online booking- To some part passengers are satisfied with this service but the ratings given are 2 star and 3

➢ **The deep analysis and comparison in the visualization format can been seen on this link- [Feedback / Ratings Analysis]**

- **Modeling-**

  Models used-:

  * Logistic Regression * Decision Tree * Random Forest Classfier

  * Ada BoosClassfier * XGBoost Classfier

Models	Accuracy	Precision	Recall	F1
XGB Model	0.94756709	0.96097073	0.916252	0.938074
RFC Model	0.94260099	0.95980268	0.905506	0.931859
DTC Model	0.92489249	0.9289211	0.895382	0.911807

Model Which is giving the best accuracy is XGB- Xtreem Gradient Boosting

Conclusion-

We trained five different models but choose top three based on the validation scores:

- XG Boost Classification Model gave us 95% accuracy score.
- Random Forest Classification Model gave us 94% accuracy score.
- Decision Tree Classification Model gave us 92% accuracy score.

Insights taken out through Data Analysis:

- It is found out that majority of the passengers are unsatisfied/neutral with the airline services.
- Majority of the loyal passengers are dissatisfied with the airline service
- Majority of Business Travel passengers are satisfied similirly majorlity of Personal Travel passengers seems unsatisfied.
- Majority of Eco and Eco Plus Class passengers seems unsatisfied/neutral with the services.