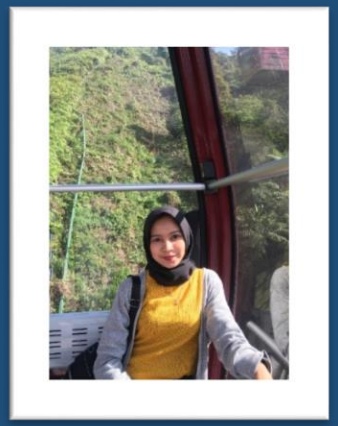
Predicting Airplane Passenger Satisfaction with Machine Learning

Self Introduction



Sarah Faradina

Bachelor of Geography
Universitas Indonesia

GIS Analyst
2018-2020

JCDS-1104-JKT

TABLE OF CONTENT

PROJECT DESCRIPTION

Dataset for this project was taken from Kaggle (US Airline Passenger). The data shows whether a passenger is satisfied with the airlines or not after travelling with them. There are several other measurement or to say feedback taken from the customers as well as their demographic data is also recorded.

Goals:
1. Predicts whether the future passengers would be satisfied or neutral/dissatisfied with service.
2. To know on which airline aspect of the services offered by them have to be emphasized more to generate more satisfied passengers.

BACKGROUND INFORMATION

U.S. Airlines are Committed to Customer Service

COMMITTED TO CUSTOMER SERVICE

Airlines are committed to delivering the highest level of customer service for the 2.4 million passengers who take to the skies every day. Historically low airfares, reliable operations and substantial investments in airports, aircraft and people are just a few ways airlines are delivering for their customers.

Carrying 2.5M **passengers** per day to/from nearly 80 countries

RELIABLE OPERATIONS

U.S. airlines

Flight completion (98.4%)
On-time arrivals (79.4%)
Baggage handling (99.7%)

IN 2018[4]

MORE AIR SERVICE

Supply of seats departing U.S. airports at all-time high

↑24%

BETWEEN 2010 – 2019[3]

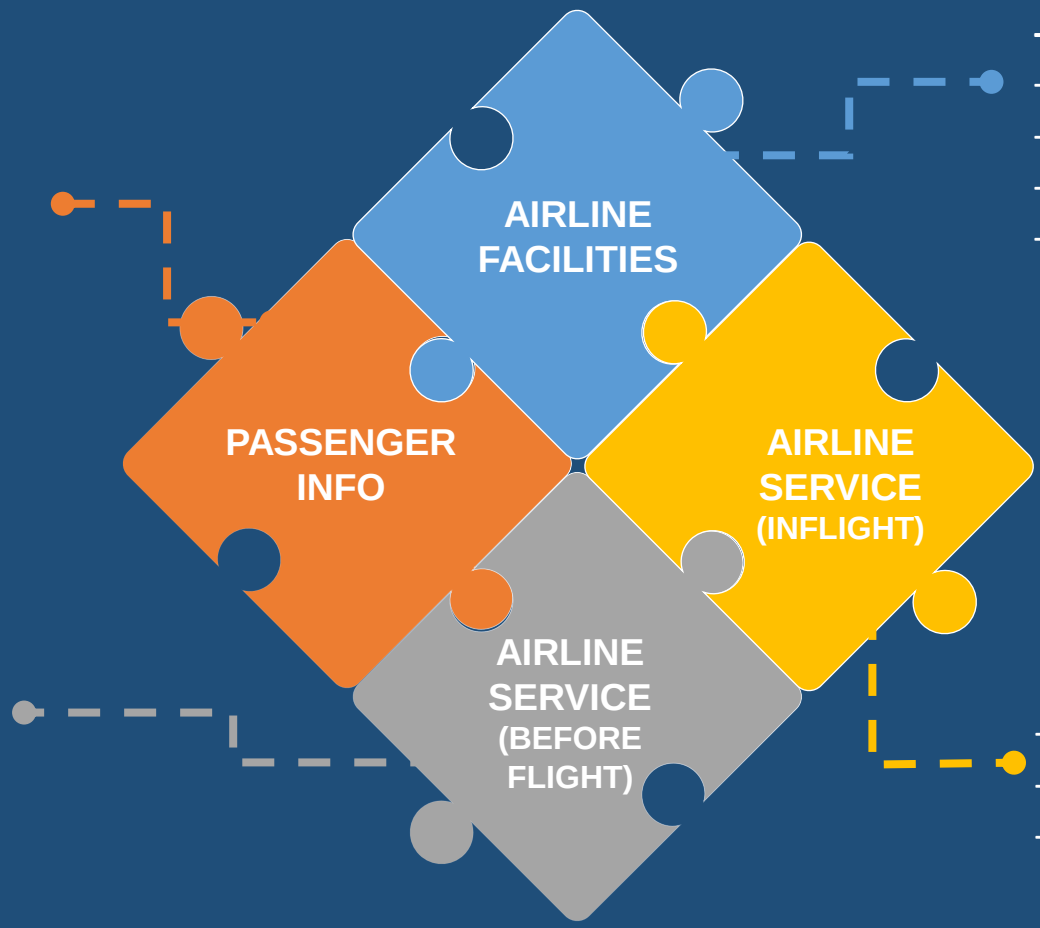
Powering 28,000 **flights** per day across the globe

Source: https://airlines.org/

Attribute Information

129980 rows | 24 columns | Kaggle

- ID
- Gender
- Age
- Customer Type
- Class
- Flight Distance
- Type of Travel
- Departure Delay
- Arrival Delay
- Satisfaction (y)



- Seat comfort
- Gate location
- Inflight entertainment
- Inflight wifi service
- Leg room service

- Departure/Arrival time convenient
- Online support
- Ease of Online booking
- Baggage handling
- Checkin service
- Online boarding

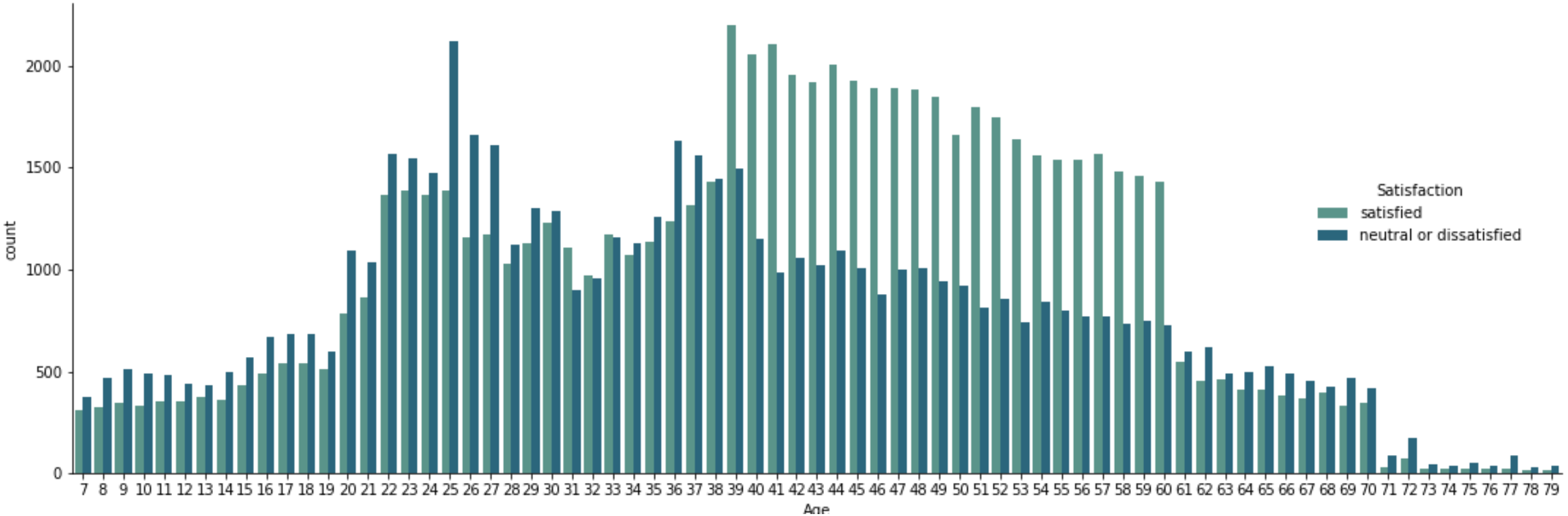
- Food and drink
- On-board service
- Cleanliness

The scale goes 0-5, where:
0 : Not Applicable
1 : Least Satisfied
5 : Most Satisfied

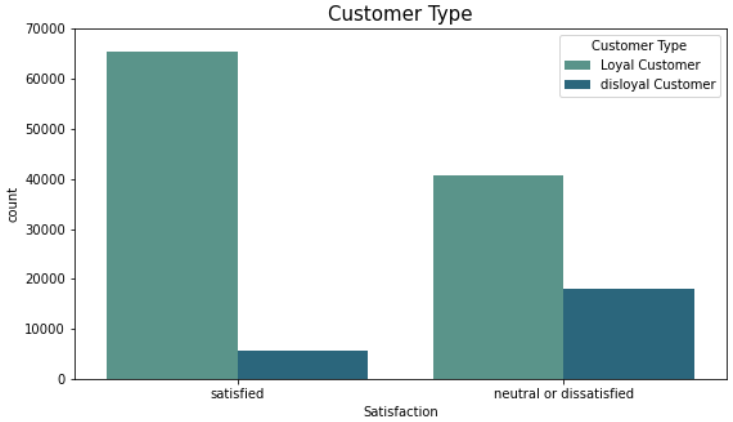
Satisfaction Insight and Analysis

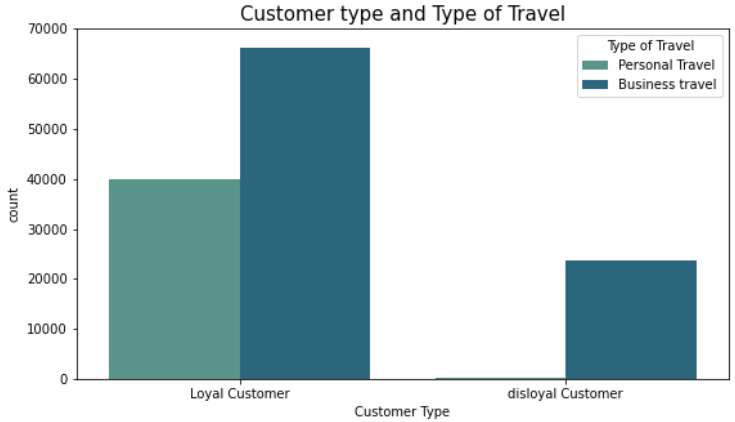
Satisfaction Report (1)

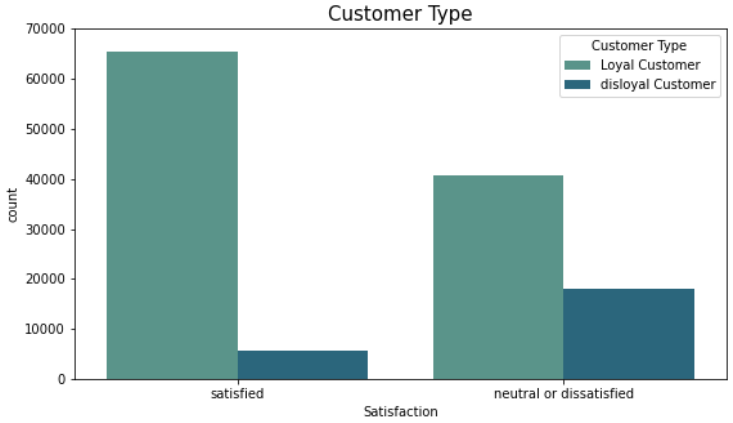
- Most of passenger who feel more satisfied after flown with the airline are passengers between the ages 39 years old and 60 years old, than younger or the older passengers.
- From the dissatisfied category, the largest number of passengers were passengers aged 25 years old.

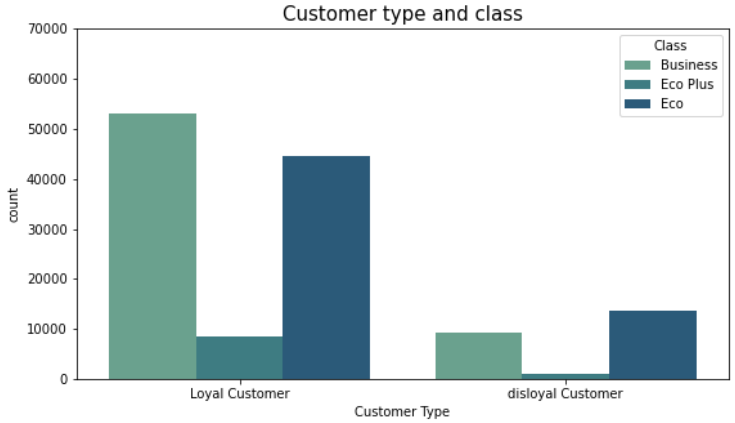


Satisfaction Report (2)



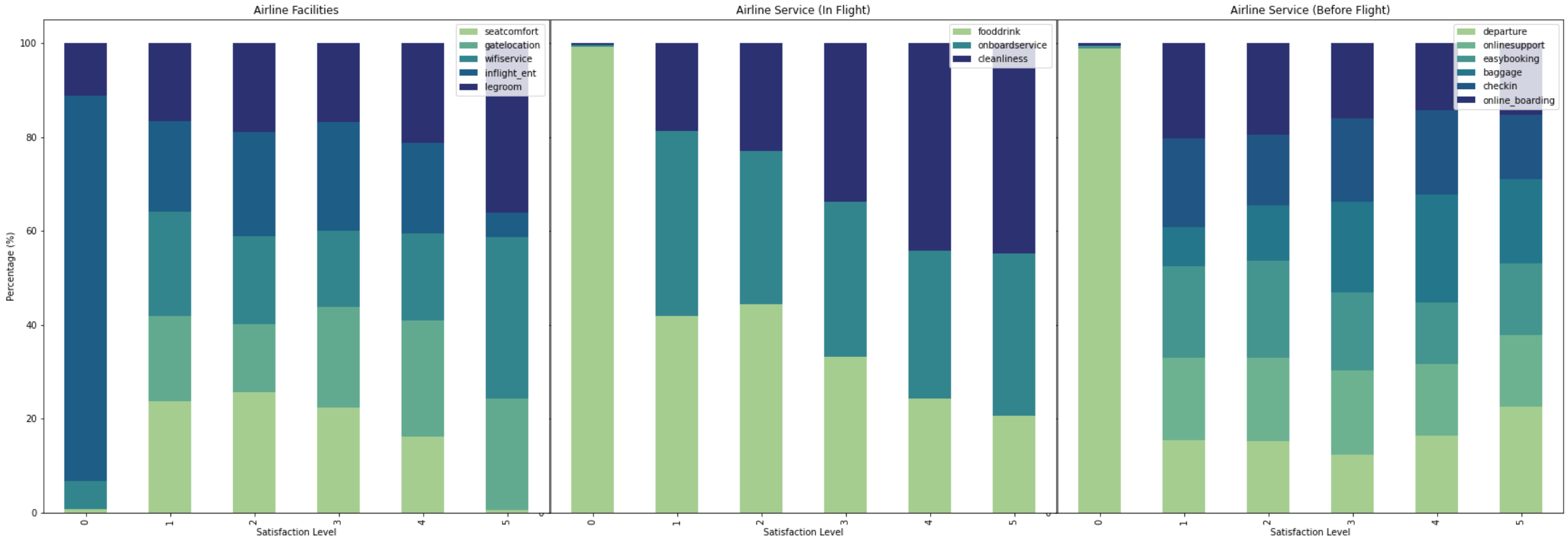






Loyal passengers are very high in satisfied and dissatisfied, likewise the loyal customer categories is the largest passengers number who travel with us both in personal travel and in business travel. Especially on business class trips where basically business ticket prices are much more expensive. And, it turns out that disloyal customers rarely travel and even then, maybe they only travel for business purposes (most likely they are traveling that paid by the office).

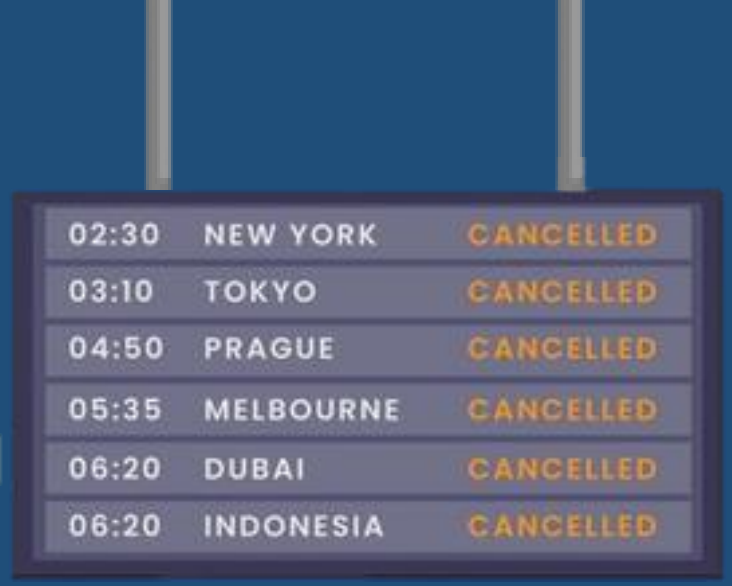
Satisfaction Report (3)



From the visualization, we can see that the worst aspect each category are:

- Service before flight : Ease of Online booking and Online boarding
- Service inflight : Food and drink
- Facilities : Seat comfort

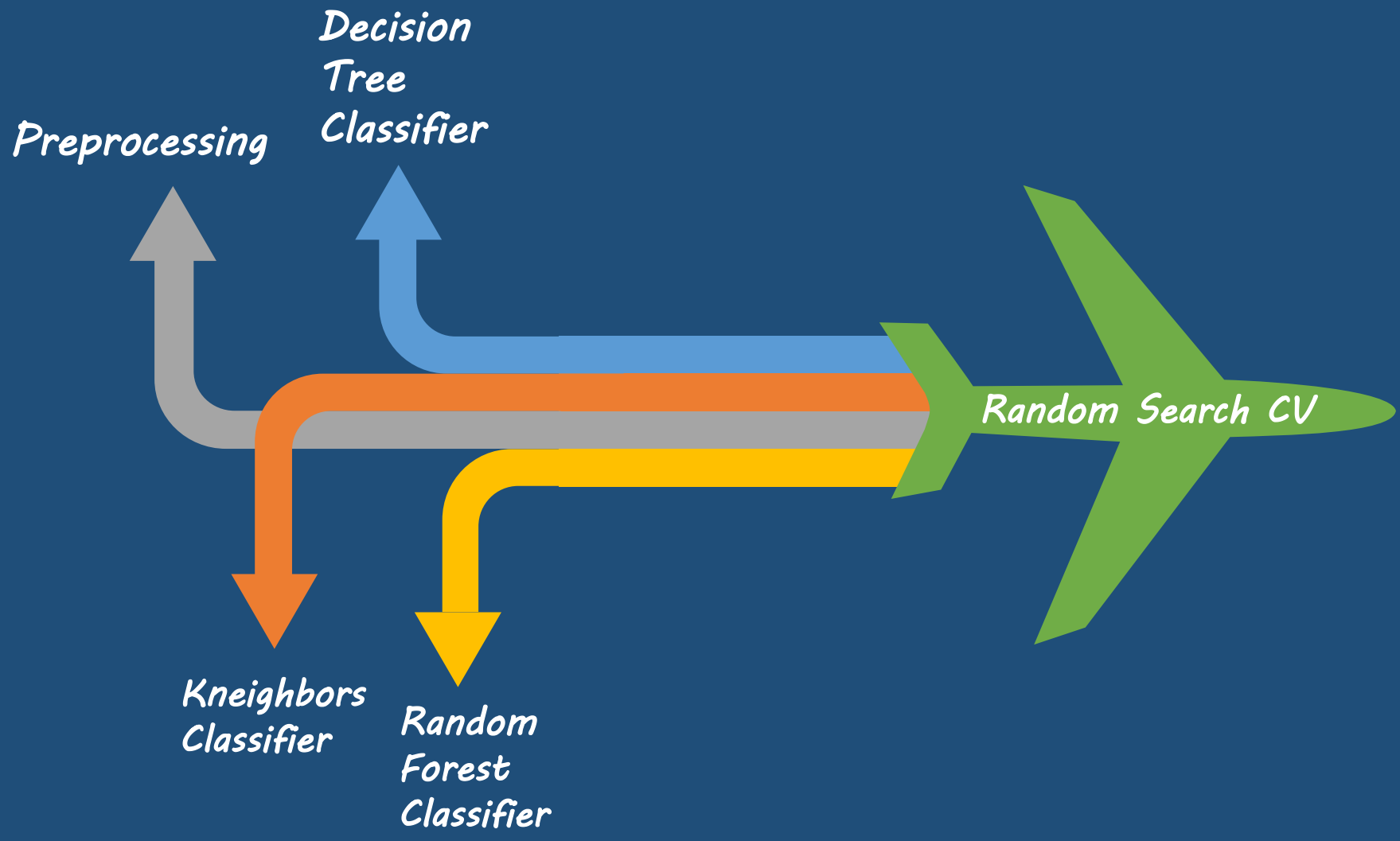
Assumption:
Aspect that need to be improved are those that have the most value in level 1 and Level 2.



MACHINE LEARNING

Supervised: Classification

Machine Learning Algoritm



Preprocessing Schema

Impute:

- SimpleImputer with median: Arrival delay in minutes

Encoding:

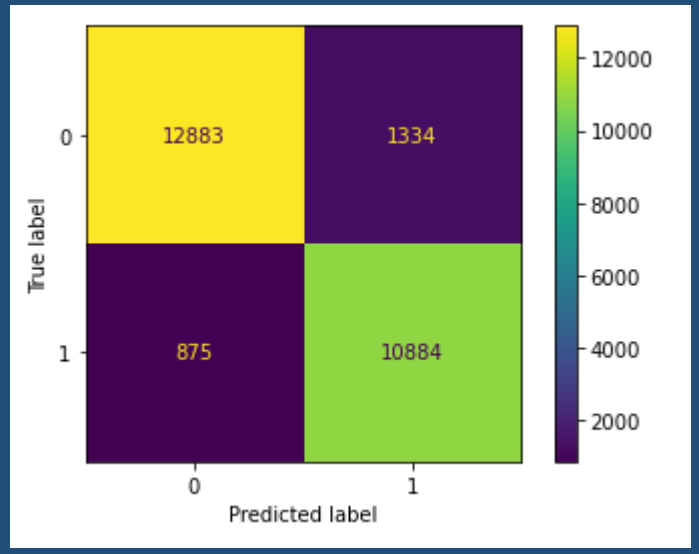
- One hot : Gender, Customer Type, Type of Travel
- Ordinal : Class

Scalling (depends on model)

- Robust : Age, Departure Delay in Minutes, Arrival Delay in Minutes

Evaluation Model

	Classifiers	Recall Before Tuned	Recall After Tuned
0	Decision Tree Classifier	0.881453	0.925589
1	K-Nearest Neighboor	0.924058	0.918360
2	Random Forest Classifier	0.838507	0.874224



Conclusion

The airline can identify the potential satisfied or dissatisfied passenger by using model (Decision Tree Classifier with hyperparameter Tuning) if the airline do satisfaction survey and the passenger fill out the survey well.

After do some analysis, the worst aspect each category are:

- Service before flight : Ease of Online booking and Online boarding services
- Service inflight : Food and drink
- Facilities : Seat comfort

There may be a problem with these services, so these should be considered and improved.

Thank You!

For full notebook of the process, please go to my github page