

## Case Study

### 1. Netflix Recommendation System:

#### Problem statement:

How can Netflix recommend personalized movies T.V shows to users based on their preference.

#### Data Science process:

##### i) Business Understanding:

Goal: The users access to view any movie, any regional movie in different language from any part of the world. Using Internet user can access any movie from anywhere at anytime.

Impact: User don't have to go to theater to watch every latest movie and don't have to pay extra money for movie because user can watch movie on Netflix with minimum cost for one month.

##### ii) Data Understanding:

We can suggest on the basis of different actor, region, language, genre, categories and the trending movies of their region. we can suggest movies web series podcast on the different users rating number of views number of reviews for more users interaction using user history we can suggest more number of content to them.

##### iii) Data Preparations:

After collecting data from users like expensive rating number of views feedback structuring data in perfect format, removing null values from data set, removing duplicate values from data set, detecting unwanted column from data set, filtering data on the categories plotting data to view the number of users interested in which genre using heat maps, different graph, box plotting.

##### iv) Modeling:

Using preprocess data set we will make model which will suggest the users most favorable movies of different genre from his history, on the basis of selected genre we can use different type of algorithm like Linear Regression, Logistic Regression, Decision Tree, Random Forest algorithm and many more to suggest favorable movies to the users.

v) Evaluation:

Training model on different data sets of different users using different algorithms like linear regression and many more and testing the generated output with different testing formula like root mean square error, mean absolute error. If the output value is closer to zero the model is being trained very well, if the output value is far from zero then the model need to get more training to get accurate output.

vi) Deployment:

Using the created model in the suggestion section where the model will generate the some joner which the user loves the more.

## 2) Predicting Customer Churn in Telecom:

### Problem Statement:

Howard Telecom Company predicts which customers are likely to stop using their services?

### Data Science Process:

#### i) Business Understanding:

Goal: In this we have to understand that or predict that which customers are likely to stop their services.

Impact: If we will successfully able to predict and reduce customer chure then we can see a large positive impact on telecom company like increased revenue, reduced cost and improved customer value etc.

#### ii) Data Understanding:

To prepare this predicting model we will collect raw data through interview, survey etc... and filter it based on the requirement. We will also analyze the contract type Internet services. And also monitor the monthly charges tenured which help us to spot any universal outlier. Here we can add some features like monthly charges and tenure in form of scale which will help in strong predictive. Power we can also show all these in the form of chart with different color in it.

#### iii) Data Preparations:

After collecting data from user experience feedback structuring data in perfect chart format. We remove all null values from data set duplicate values etc... Then plotting data to view to number of user interested in which network using heat maps, different graphs, box plotting.

#### iv) Modeling:

Using pre-processes dataset, we will make model which will suggest the user most proper Internet connection of different companies for different area. We use different type of algorithms like logistic regression use random forest to predict network.

v) Evaluation:

Making an algorithm which automatically conducts surveys in the rural area when the network fluctuates by the company which are not used by the users. Algorithm which will automatically define the problem and will send the solution with the problem to the responsible person. And also make an automatic call for conducting survey. And the solution person use matrix like accuracy precision recall to solve the problem.

vi) Deployment:

Automated call, Conduct survey provide alert for high risk customer to customer services team. Implement proactive offer and discount to retain customers.

### 3) Predictive Diabetes for Health Care:

#### Problem Statement:

Can we predict the likelihood of a patient developing diabetes based on their health indicators?

#### Data Science Process:

##### i) Business Understanding:

Goal: We can able for early detection of diabetes based on their health. Predict the diabetes so Dr will tell you as early as possible which type of care, treatment is needed.

Impact: We can improve patient outcomes by providing patients with different sources like in hospital clinic providing facilities to understand avoid high risk and low risk diabetes.

##### ii) Data Understanding:

Explore different data set to check description features like BP, Glucose, Kidney Failure and analyze data that how many are diabetic and Non-diabetic patients. Identify relation between indicators diabetes outcomes like high glucose on BMI.

##### iii) Data Platform:

After collecting data from user experience, rating, number of views, feedback removing null values from data set, duplicate values, filtering data and plotting graph to view the number of users interested in software using head map, different graph, box plotting.

##### iv) Modeling:

Using pre-process dataset, we will make model which will suggest the user health to make his body better day by day. We use different type of algorithm like Random forest or xBoost for high accuracy complex relationship.

##### v) Evaluation:

Access Model performance using matrix like accuracy, precision, recall, F1 score and confusion matrix. Focus on recall to ensure high detection of diabetes patient use cross verification to check consistency and avoid overfitting.

vi) Deployment:

Integrate the model into hospital/clinic system or model app for real time prediction. Provide early warning to Doctor for high risk patient.