***Intelligent Customer Retention: Using Machine Learning for  
Enhanced Prediction of Telecom Customer Churn***

***Abstract:***

*Customers are the base for any business success and that is why firms become aware of the significance of acquiring satisfaction of customers. Customer churn is an essential lissue and it is regarded as one of the most essential concerns among firms because of increasing rivalry among firms, increased significance of marketing strategies and customers conscious behaviour in present years. Organizations must develop different strategies to resolve the churn issues relying on the services they offer. Customer churn practice is essential in competitive and rapidly developing in telecom sector. The process of migrating from one service provider to another telecom service provider occurs due to good services or rates or due to various advantages which rivalry firm provides customers when signing up. Due to the greater cost related with acquiring new customers the prediction of customer churn has developed as an indispensable part of planning process and strategic decision making in telecom sector. The main aim of the study is to explore the customer churn prediction in telecom using in big machine learning data platform. Machine learning techniques have been used for estimating the customer probability to churn. This study makes use of logistic regression and KNN with big data for predicting consumer churn in the telecom sector. Logistic regression has been used widely to estimate the probability of churn as a function of variables set or features of customers. Similarly, for churn K-Nearest Neighbour is used to examine if a customer churns or not based on their feature’s proximity to customers in every class. This study uses Kaggle website for dataset in predicting and analysing churn. The results of the studies how that the accuracy rate of prediction in consumer churn is found to be 0.80 percent and area under curve is found to be 0.71 percent.*

***1.Introduction****:*

***1.1 Overview: Customer churn (or customer attrition)****is a tendency of customers to abandon a brand and stop being a paying client of a particular business. The percentage of customers that discontinue using a company’s products or services during a particular time period is called a customer churn (attrition) rate*. *One of the ways to calculate a churn rate is to divide the number of customers lost during a given time interval by the number of acquired customers, and then multiply that number by 100 percent. For example, if you got 150 customers and lost three last month, then your monthly churn rate is 2 percent.*

*Churn rate is a health indicator for businesses whose customers are subscribers and paying for services on a recurring basis, notes head of data analytics department at Science Soft*[*Alex Bekker*](https://twitter.com/alexlbekker)*, “Customers [of subscription-driven businesses] opt for a product or a service for a particular period, which can be rather short – say, a month. Thus, a customer stays open for more interesting or advantageous offers. Plus, each time their current commitment ends, customers have a chance to reconsider and choose not to continue with the company. Of course, some natural churn is inevitable, and the figure differs from industry to industry. But having a higher churn figure than that is a definite sign that a business is doing something wrong.”*

*There are many thing brands may do wrong, from complicated onboarding when customers aren’t given easy-to-understand information about product usage and its capabilities to poor communication, e.g. the lack of feedback or delayed answers to queries. Another situation: Long time clients may feel unappreciated because they don’t get as many bonuses as the new ones.*

*In general, it’s the overall customer experience that defines brand perception and influences how customers recognize value for money of products or services they use.*

*The reality is that even loyal customers won’t tolerate a brand if they’ve had one or several issues with it. For instance, 59 percent of US respondents to the*[*survey by PricewaterhouseCoopers*](https://www.pwc.com/us/en/advisory-services/publications/consumer-intelligence-series/pwc-consumer-intelligence-series-customer-experience.pdf#page=8)*(PwC) noted that they will say goodbye to a brand after several bad experiences, and 17 percent of them after just one bad experience.*

*1.2 Purpose: Customer churn is a common problem across businesses in many sectors. If you want to grow as a company, you have to invest in acquiring new clients. Every time a client leaves, it represents a significant investment lost. Both time and effort need to be channelled into replacing them. Being able to predict when a client is likely to leave, and offer them incentives to stay, can offer huge savings to a business.*

*As a result, understanding what keeps customers engaged is extremely valuable knowledge, as it can help you to develop your retention strategies, and to roll out operational practices aimed at keeping customers from walking out the door.*

*Predicting churn is a fact of life for any subscription business, and even slight fluctuations in churn can have a significant impact on your bottom line. We need to know: “Is this customer going to leave us within X months?” Yes or No? It is a binary classification task.*

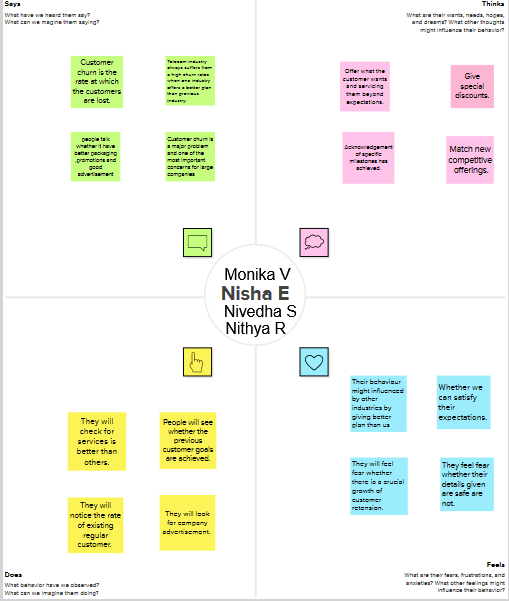
*In the beginning, only telecom companies could afford to use churn prediction models and pay big data scientists to analyze their customer data. Now, customer churn rates and customer churn risk can be analyzed using churn data more easily.*

*One essential aspect of Churn prediction is in its use for preventing customer churn.*

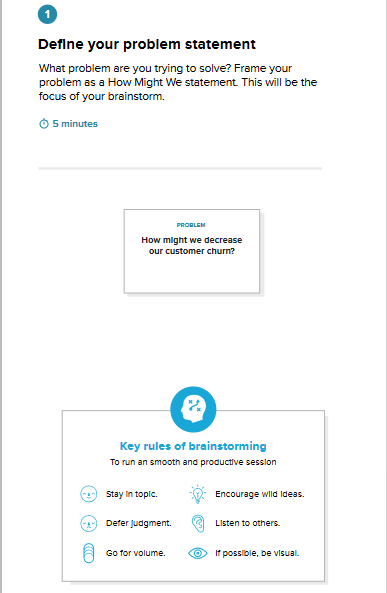
*In this blog post, the basis of customer churn will be discussed, and major aspects of churn prediction such as churn prediction models and some of its use cases.*

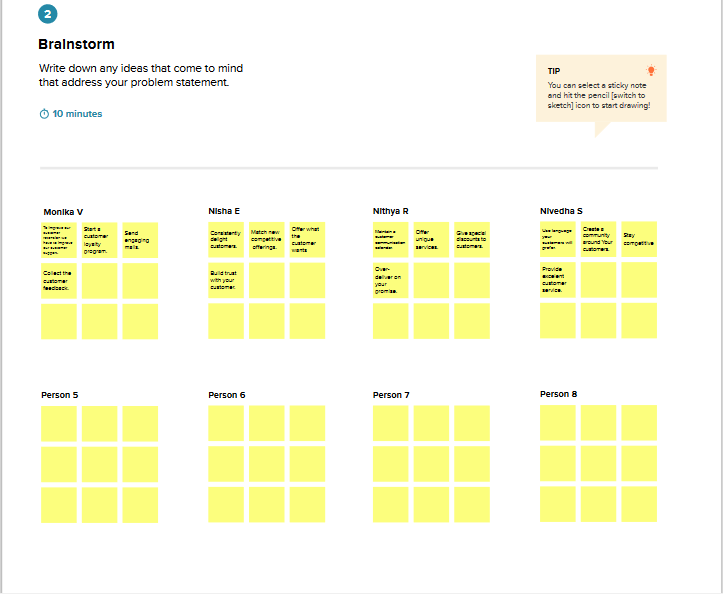
***2.Problem Definition & Design Thinking:***

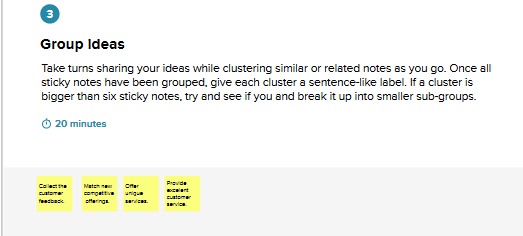
*2.1.Empathy map:*

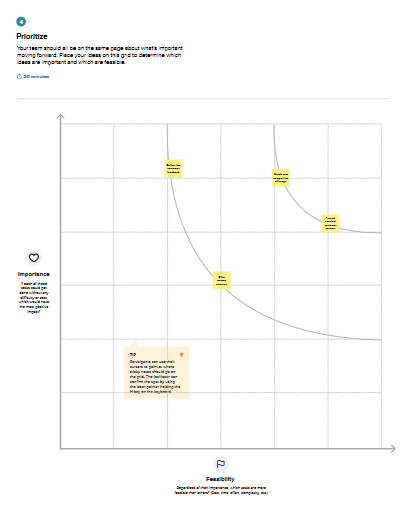
******

*2.2.Ideation & Brainstorming map:*

******

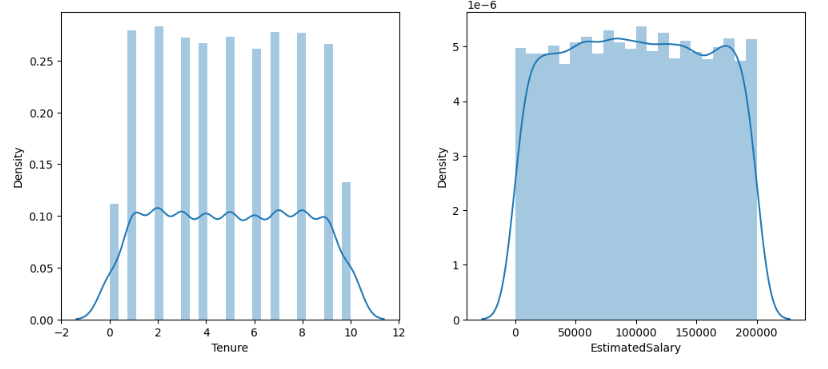
******

******

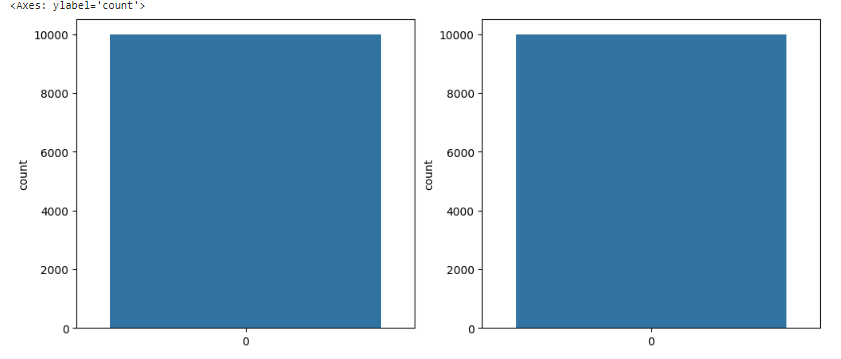
******

***3.Result:***

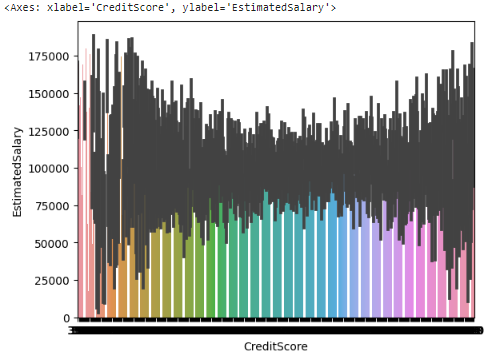
*Univariate analysis:*

******

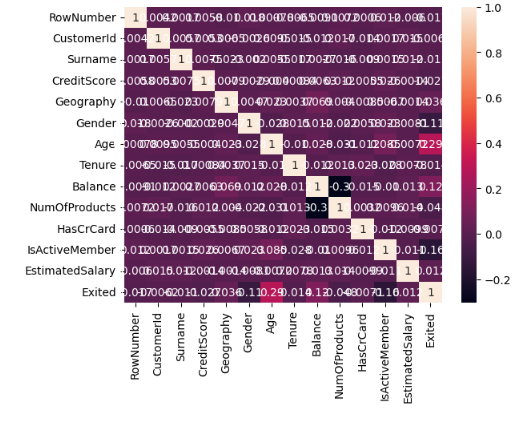
*Count plot:*

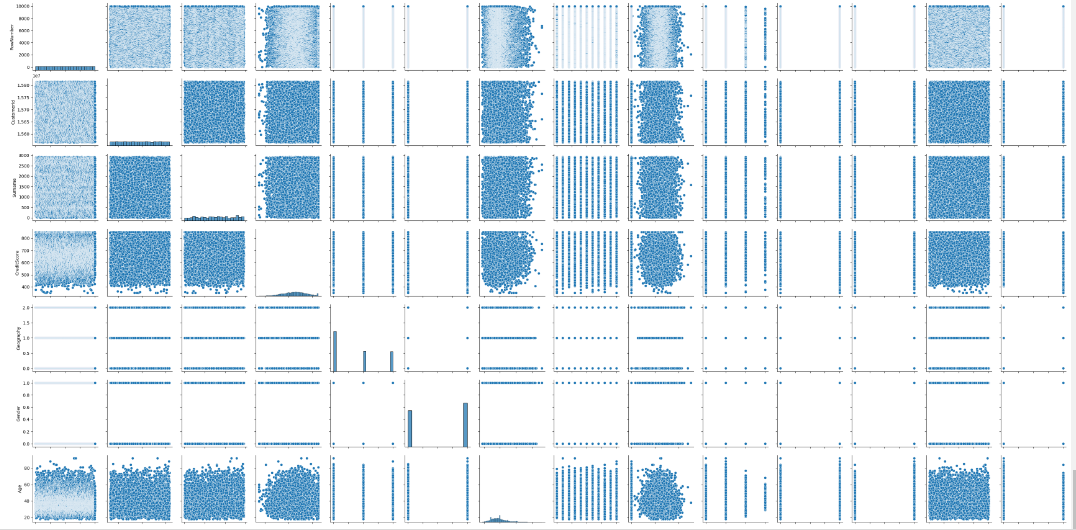
******

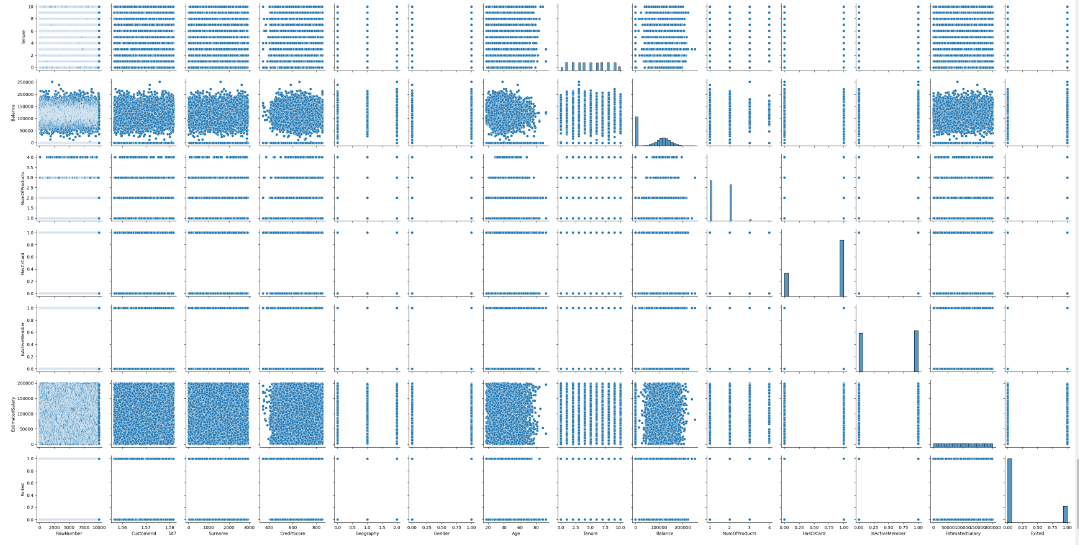
*Bivariate analysis:*

******

*Multivariate analysis:*

******

******

******

***4. Advantages:***

*Because of its efficient and straightforward nature, doesn't require high computation power, easy to implement, easily interpretable, used widely by data analyst and scientist. Also, it doesn't require scaling of features. Logistic regression provides a probability score for observations.*

*The ability to predict churn before it happens allows businesses to take proactive actions to keep existing customers from churning .this could look like :customer success teams reaching out to those high risk customers to provide support or to gauge what needs may not be being met.*

*More predictable and stable revenue streams. Shifting from a pay-once model to subscription-based billing allows for better insight into your revenue forecast.*

***Disadvantages:***

*Logistic regression is not able to handle a large number of categorical features/variables. It is vulnerable to overfitting. Also, can't solve the non-linear problem with the logistic regression that is why it requires a transformation of non-linear features. Logistic regression will not perform well with independent variables that are not correlated to the target variable and are very similar or correlated to each other.*

*There is a direct relationship between customer lifetime value and the ability to grow your business . As such ,the higher your customer churn rate ,the lower your chances of growing your business .*

***5. Application:***

*Churn prediction is about identifying customers who are likely to churn. Perfecting this prediction process allows businesses to leverage reliable information about their current customers, giving them information to build effective customer retention and marketing strategies.*

***6.Conclusion:***

*In the competitive telecom sector standardization and public policies of mobile communication permits customers to switch over from one carrier to another carrier easily resulting in a competitive market. The prediction of churn or the task of recognizing customers who are probable to discontinue service use is a lucrative and essential issue of telecom sector. Customer churn is often a critical problem for the telecom sector as customers do not delay to leave if they do not predict what they are viewing for. Customers mainly need value for money, competitive cost and greater service quality. Customer churning is associated directly to satisfaction of customer. It is a known fact that the customer acquisition cost is larger than customer retention cost that makes the retention a difficult prototype of business. There is no standard approach which resolves the churning problems of worldwide service providers of telecom industry accurately. Big data analytics with machine learning technique is used for customer churn which sets warning bells for customers before any damage could occur, providing telecom firms the chance to take precautionary steps. These techniques are used to find the churn in customers by constructing models and studying from historical information. Conducting trials with perspective of end users, collecting their views on network, normalization of data, data set pre-processing, using feature selection, removing missing values and class imbalance and changing existing variables with derived variables develops the churn prediction accuracy which supports the telecom sector to retain their customers much efficiently. It can be concluded that Big data analytics with machine learning were predicted to be an effective way for recognizing churn in customers.*

*In the telecom sector it is essential to identify and manage customers who are probable to churn which is always characterized by volatile markets and strong rivalry. Proper customer management who are probable to churn can reduce the churn probability while extend the profit of the telecom sector. Telecom firms are realizing the importance of customer churn prediction as a way of producing huge number of profits. Constructing a model of churn prediction will simplify the retention process of customer and the telecom sector will gain success and increase in competitive market. The churn prediction model is relied strongly on the process of data mining and data mining techniques due to a developed performance produced by machine learning algorithms. Customer churn is used to make a model that encompasses the customer hazard and customer survival functions accurately. The machine learning algorithms have been used for accurate prediction in telecom sector. The telecom churn prediction has been identified to be of varied domain of application to churn prediction in comparison to other telecom-based sector due to its volume, biases and variety of dataset. It can be concluded that the customer churn models are one of the significant solutions to remain competitive in telecom sector.*

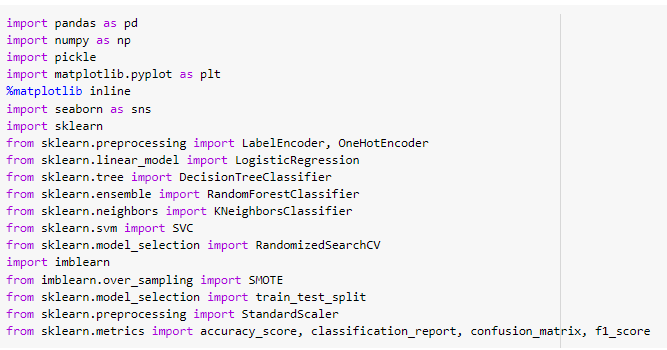
*In nowadays highly competition between firms and in digital world the factors of customer churn are an essential undertaking for every service provider to make profitable and long-term relation with particular customers. The factors influencing customers switching behaviour would be the quality of call, satisfaction level, level of tariff, image of brand, handsets, tenure and income. Some customers are much cost sensitive and move to other telecom service provider if they get better cost and also, they chose the service provider chosen by their friends and family. Another factors which impact a customer to select a service provider is cost and communication followed by responsiveness to their service complaints. Quality is also one of the factors which impact customers to move from one service provide to another service provider. The rate of call also plays an essential part for making a decision on moving to another service provider followed by coverage of network, customer care and value added service. The impact of family is also another essential factor of churn rate in telecom sector. The factors that influence customers probability defecting to rivalry involves insufficient or slow response to billing errors and complaints. Some other factors namely packaging costs, insufficient characteristics and older techniques may also impact customers churn to affect the rivalry.*

*It can be concluded that telecom service providers must provide much attraction to the above stated factors for their customers to reduce the churn rate flexibly and effectively. The cost of obtaining new customers can be greater than that of customer retention. One of the best way for customer retention is to reduce customers churn rate where churn refers to migration of customer from one service provider to another service provider or terminating particular services over particular periods for several reasons that can be predicted previously if the firm examines its records of data and uses machine learning technique which enhances the firms to find customers who are probable to churn. Several algorithms are available to reduce the churn rate in telecom companies. The telecommunication service providers use advanced analytics algorithms to mine through huge number of data of customers. This algorithm is smart enough to recognize hidden characteristics to find which customers are much probable to churn. Data mining plays an essential role in telecom firms and their effort to reduce overall churn develops good marketing strategies, recognize fraudulent activities and consumers and manage their network better. A proper algorithm is chosen relying on the problem nature and that of feasible data. It can be concluded that machine learning algorithms is regarded as one of the best solutions for telecom sector to reduce the churn rate.*

***7.Future scope:***

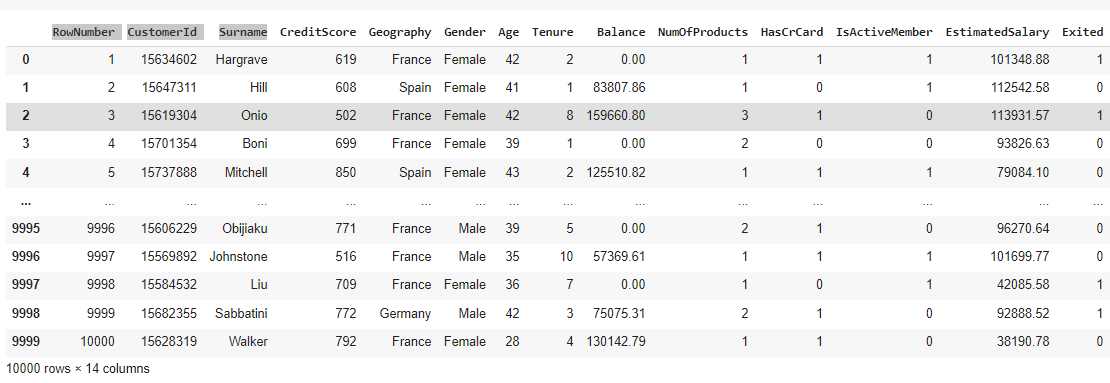
*Our study investigates how machine learning can be used to predict customer churning in the B2B context. As mentioned above, chum prediction is one of two parts in customer chum management, and for future work, it would be interesting to investigate what features to use and how they impact chum prediction in the context of B2B. It would also be interesting to investigate, based on the variables, what measures could be taken related to retention strategies and how organizations should actively act towards customers that are predicted to churn. Another aspect, which would be interesting to further investigate, is what other methods can be used for feature selection and sampling and how they would impact the result. In addition, what other algorithms that could be used and how they could be modified in order to achieve the best result.*

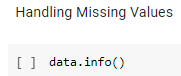
***8.Appendix:***

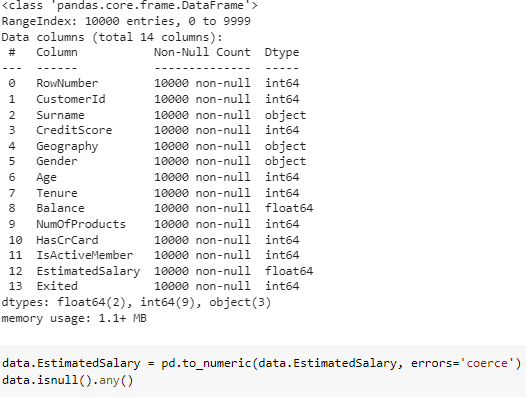
*Importing the Libraries *

*Read the Dataset*

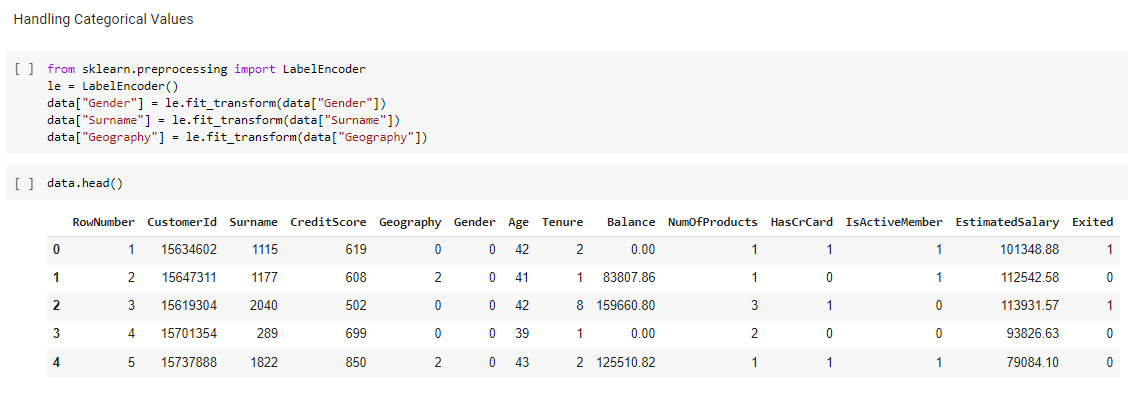
**

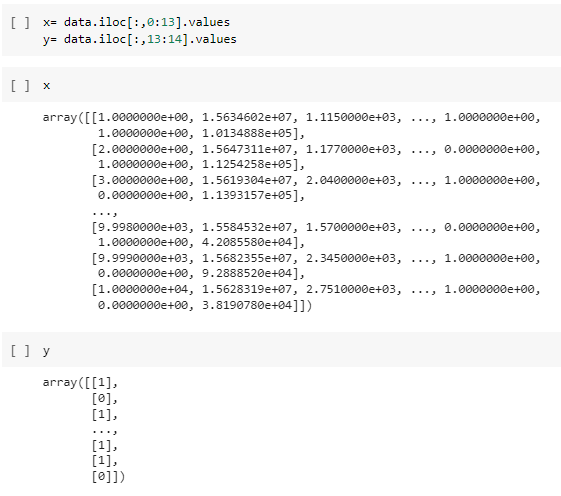
**

**

**

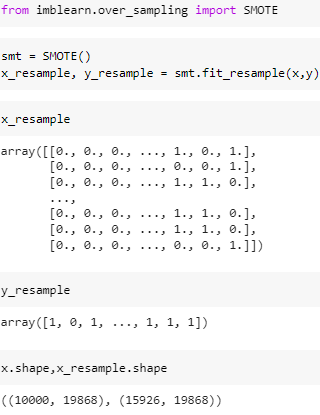
**

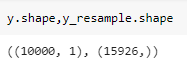
**

**

**

*Handling Imbalance Data*

**

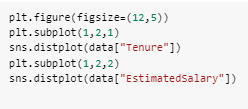
*  
TASK 3*

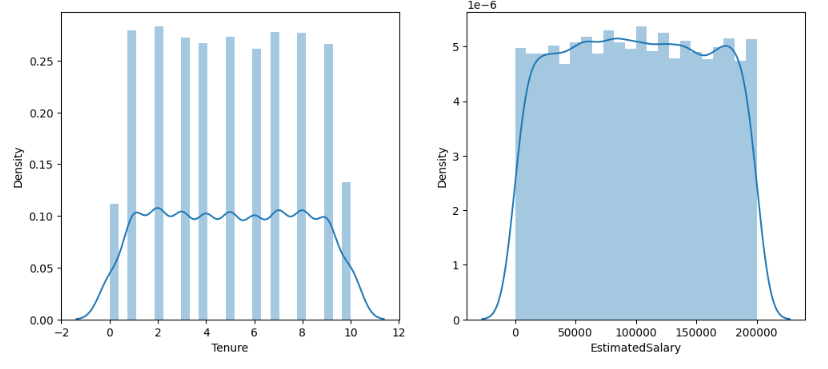
*Descriptive Statistical*

**

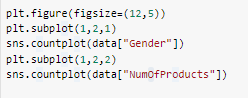
**

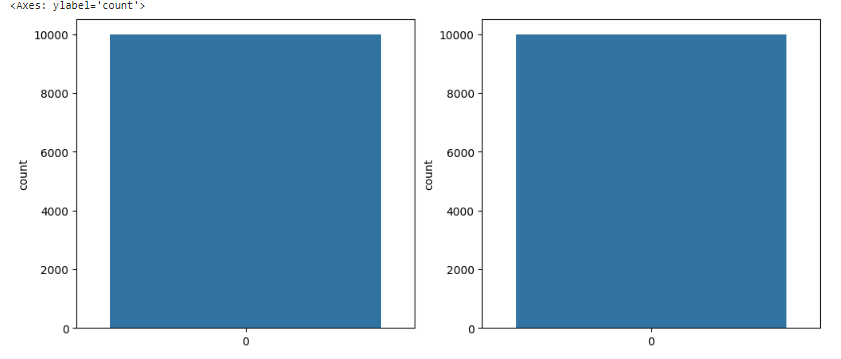
*Univariate Analysis*

**

******

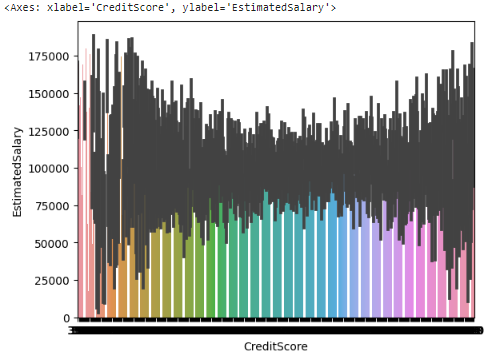
*CountPlot*

**

******

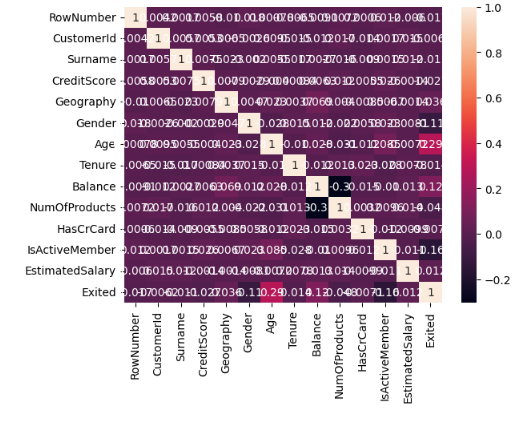
*Bivariate Analysis*

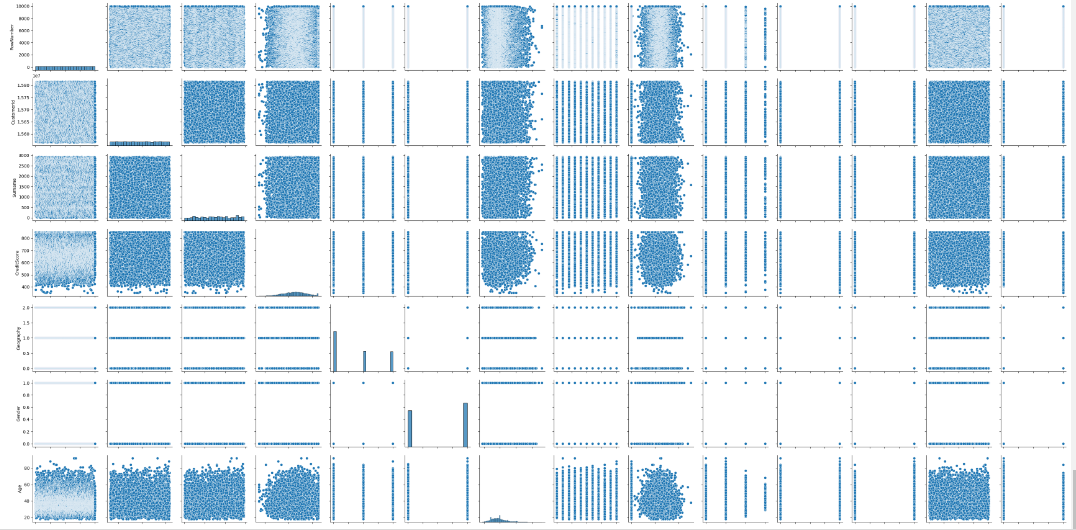
**

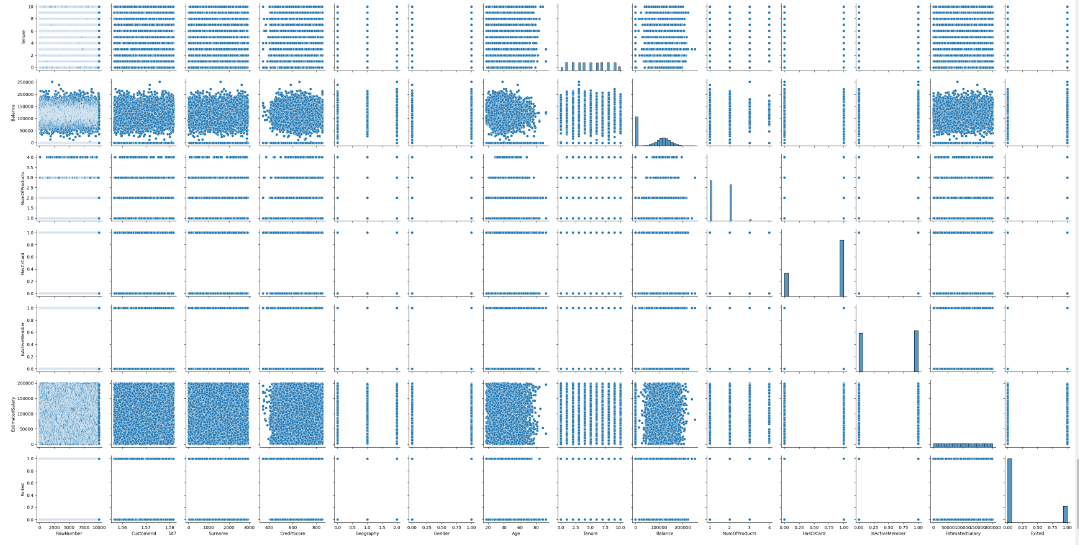
******

*Multivariate Analysis*

**

******

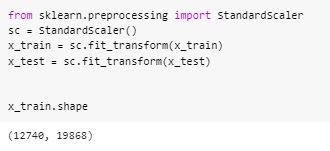
******

******

**

**

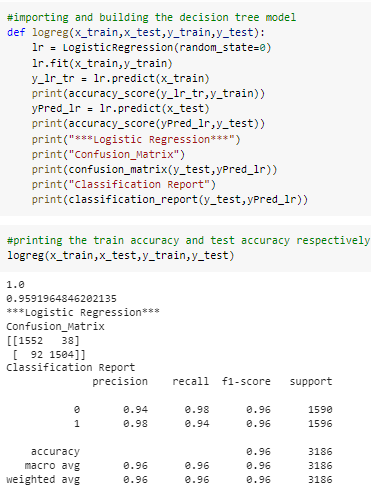
*Scalling the Data*

**

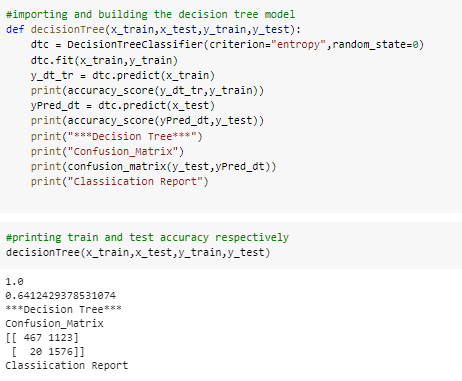
*TASK 4*

*Modeling Building*

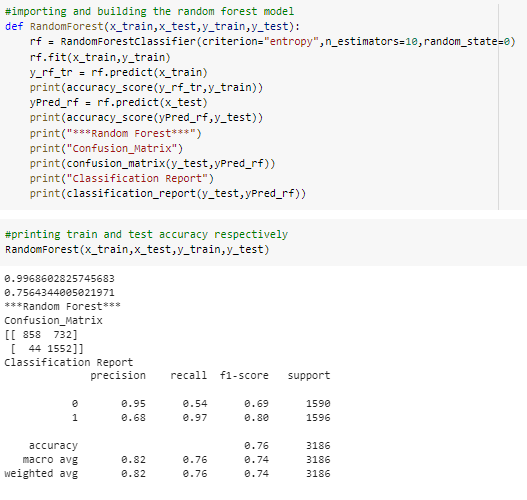
*Logistic Regression Model*

**

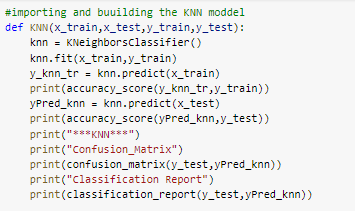
*Decision Tree Model*

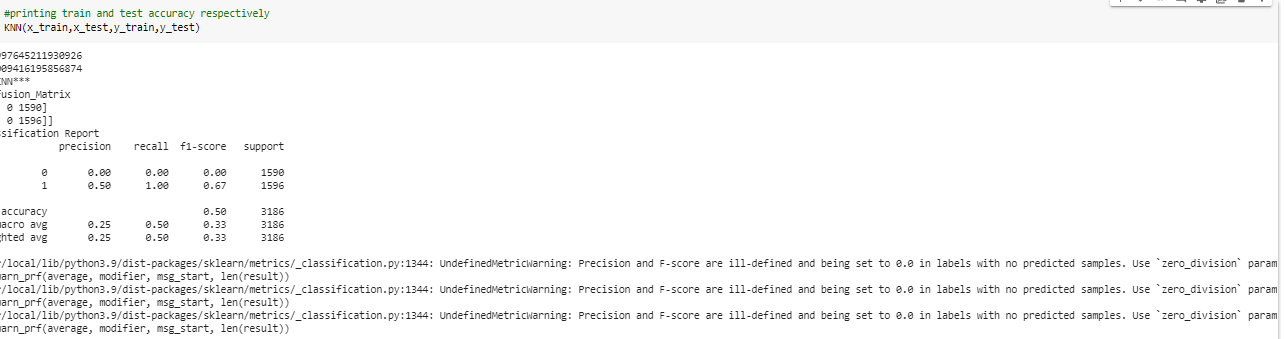
**

*Random Forest Model*

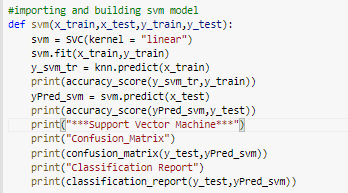
**

*KNN Model*

**

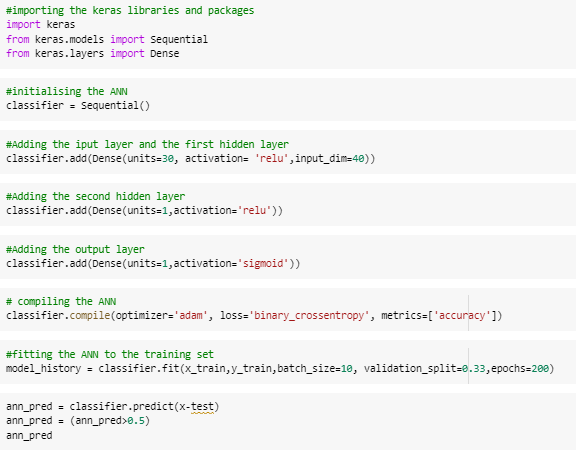
**

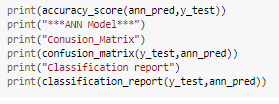
*SVM Model*

**

**

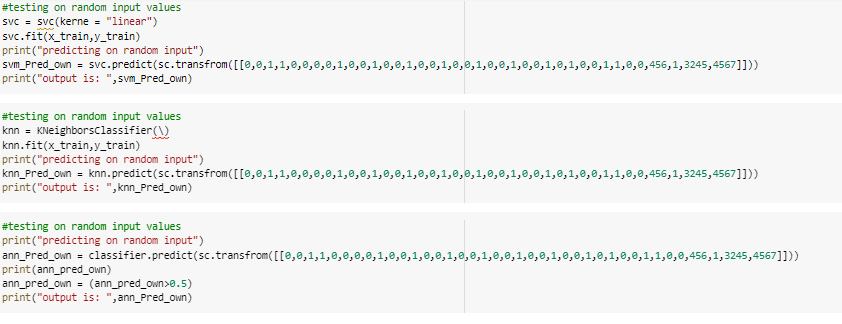
*ANN Model*

**

**

*Testing The Model*

**

**

*TASK 5*

*Performance Testing & Hyperparameter Tuning*

*Compare The Model*

**