MODEL TO PREDICT CUSTOMERS LEAVING TELECOMMUNICATION COMPANIES

Team members:

1. Rajagopal Naidu Kodavati 11596822
2. Jaya Naga Satya Pavan Ganesh Kotipalli 11594936
3. Jaya Sindhu Edara 11708783
4. Sri Mounish Seeni 11549237

Motivation:

1. The motivation for developing churn prediction models in a telecommunications company is to improve customer satisfaction, reduce costs, and gain a competitive edge by making data-driven decisions and proactively addressing customer churn.
2. This project has the potential to lead to significant financial and operational benefits for the company.
3. These models are a key tool for increasing revenue and driving business growth.
4. It's a strategic initiative that combines data analysis, modeling, and business intelligence to drive tangible results and long-term success.

Significance:

1. It leverages data-driven insights to inform decision-making and improve the company's competitiveness in a dynamic industry.
2. It directly impacts revenue, customer satisfaction, and overall business success.

3.Customer churn is a significant concern for telecommunications companies.

4. Losing customers can result in a loss of revenue, and it can be more costly to acquire new customers than to retain existing ones. Predicting and preventing churn can significantly impact a company's bottom line.

Objectives:

1.Initially we are planning to categorize our data into numerical and categorical values and perform exploratory data analysis like (univariate, bivariate, etc.) to get a complete idea on what the dataset can do.

2.After finding out the capabilities of the dataset we are planning to perform feature engineering on the given data and develop some features which are useful for better finding out the customer churn.

3.Identifying the factors which contribute customer churn and reduce them.

4.The main objective is to improve the performance of the machine learning models by training them with more relatable features.

Features:

The features which we are using are categorical and numerical

1. Categorical features: Customer_ID, Gender, Partner, Dependents, Phone_Service, Multiple_lines, Internet_Service, Online Security, Online Backup, Device Protection, Tech Support, Stream Movies, Paperless Billing, Payment Method.

2. Numerical features: Tenure, Monthly_Charges, Totla_Charges.

3. Also Planning on perfoming feature engineering with these features and create useful features which contribute to customer churn

Increment – 1:

Related Work:

Our project was on prediction of employees who leaves company or not, so for this we take data form online sources which has all the required fields to categorize the situation of employee and we make necessary changes to our model to make good features for extraction. With the required features we build a model to predict the outcome.

Dataset: https://www.kaggle.com/datasets/blastchar/telco-customer-churn

We take the dataset from Kaggle platform which has required columns and categories to our model and we make changes according to best fit of our model.

Implementation:

- Initially we imported our dataset from Kaggle and make necessary changes to make good fit for the model
- The changes we made in dataset is to convert some categorial data to numerical data which helps to us to make easy prediction
- The conversion is on type of 1 = True and 0 = False in many cases
- After the data was ready we briefly summarize the data into two categories which are categorial data and numerical data for

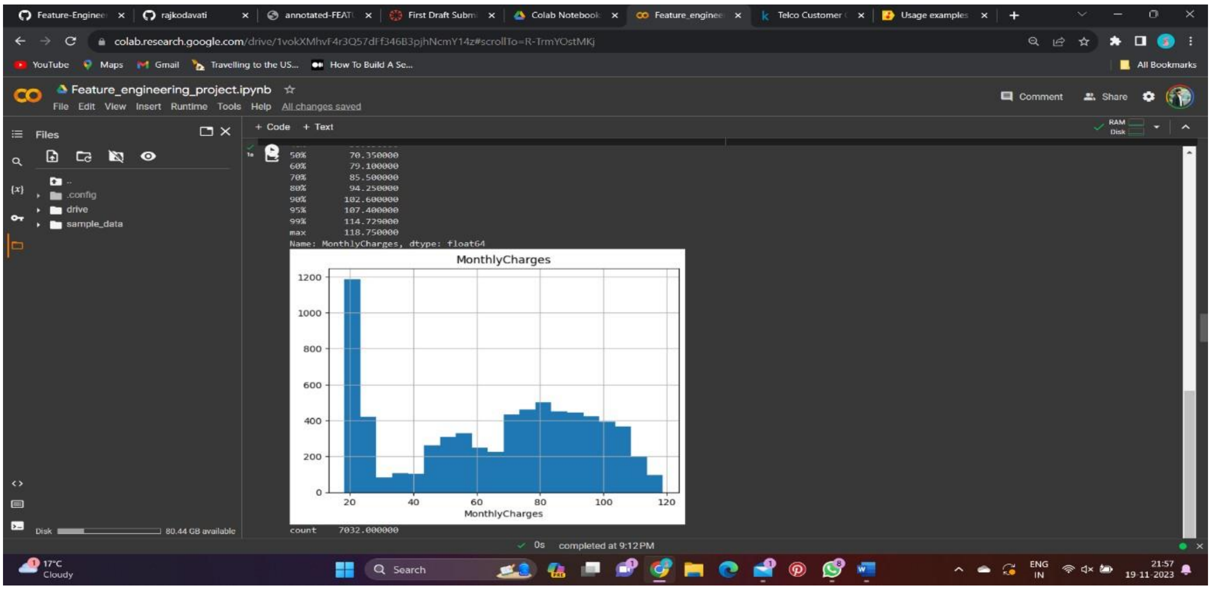
better understanding to build the model and make prediction easy

- After summarization we print corelation matrix for clear visualization to user
- Performs one-hot encoding technique which is one of the popular encoding techniques to categorial data and we split the data to test and train to find accuracy, recall, precision of the data. For this we also use catboost library.

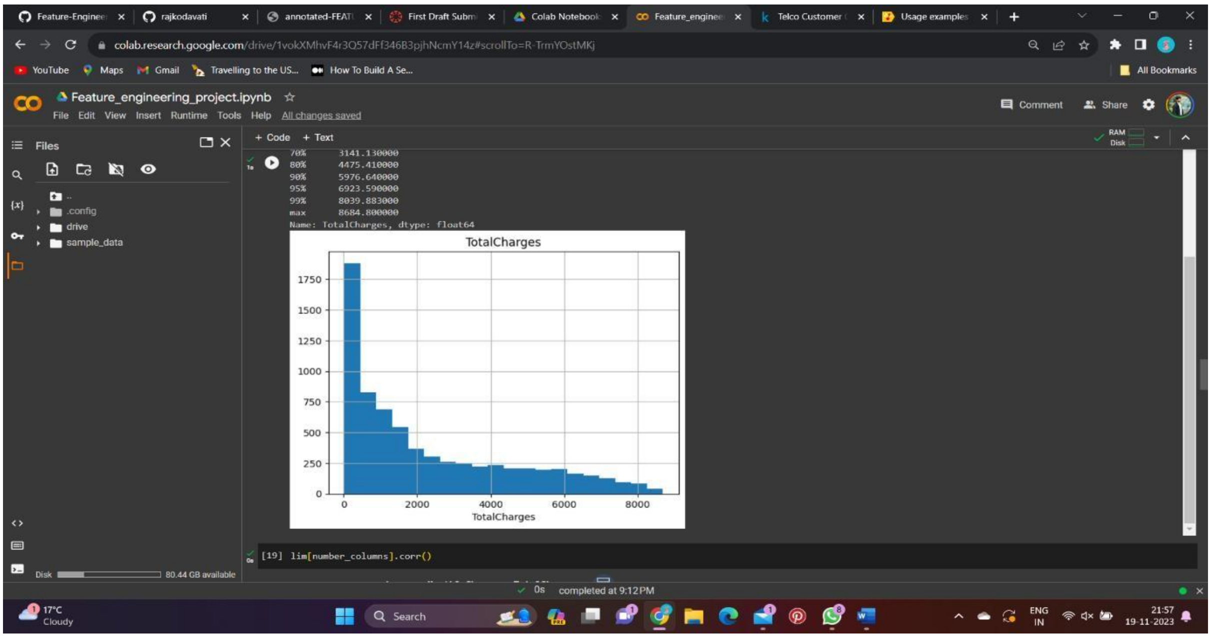
Analysis:

As the process mention above was done successful we analyse the data and display the data in graph formats and text formats using matplotlib libraries and displays how accurate the data which we are using.

Analysis we made on numerical data:



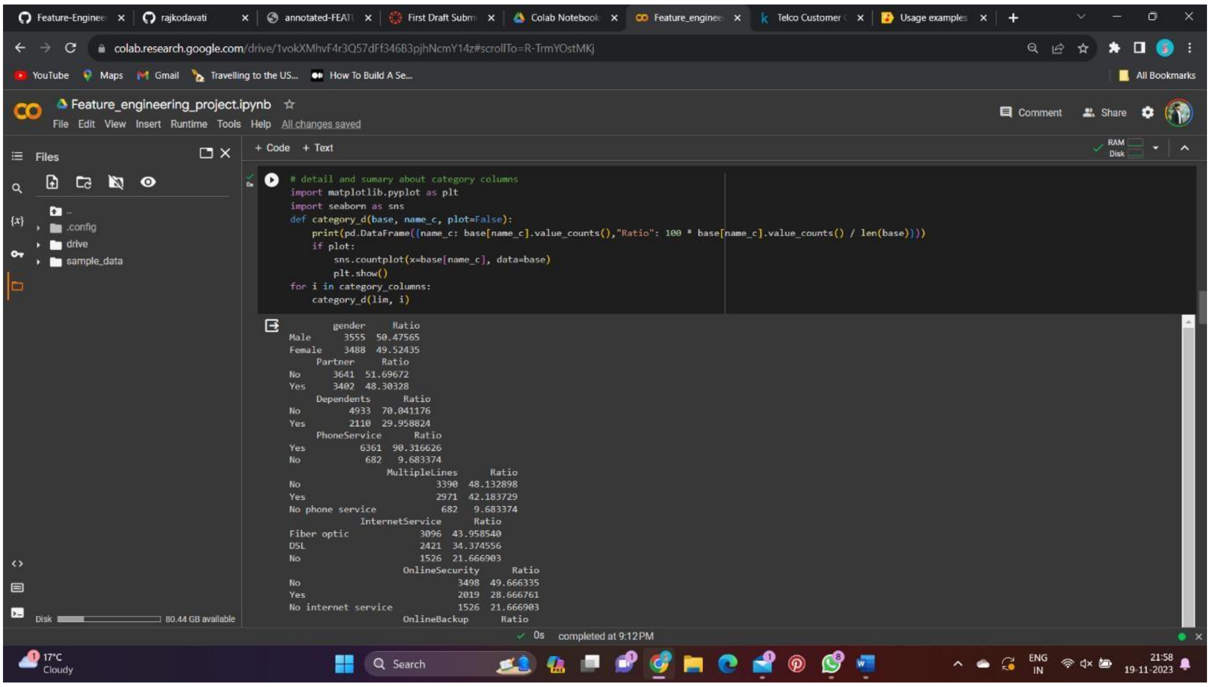
```
50%        70.350000
60%        79.100000
70%        85.500000
80%        94.250000
90%       102.600000
95%       107.400000
99%       114.729000
max       118.750000
Name: MonthlyCharges, dtype: float64
```



```
70%       3141.130000
80%       4475.410000
90%       5976.640000
95%       6923.590000
99%       8039.883000
max       8684.800000
Name: TotalCharges, dtype: float64
```

```
[19] lim[number_columns].corr()
```

Analysis we made on categorial data:

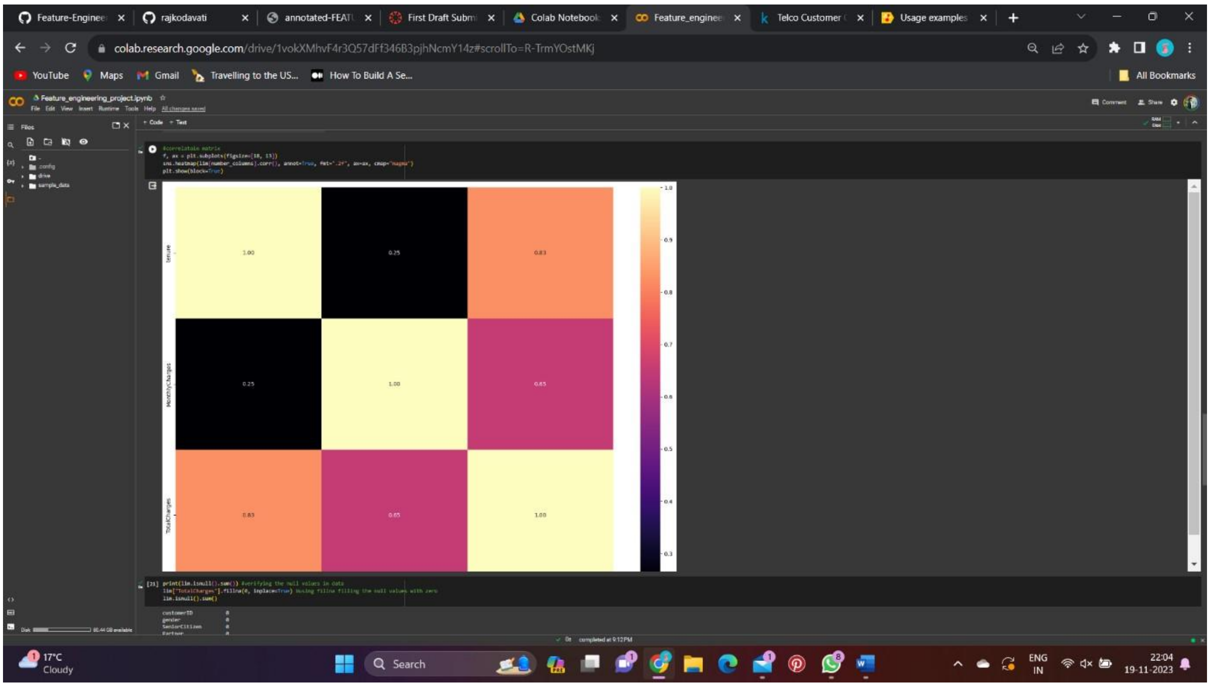


```
# detail and sumary about category columns
import matplotlib.pyplot as plt
import seaborn as sns
def category_d(base, name_c, plot=False):
    print(pd.DataFrame({name_c: base[name_c].value_counts(),"Ratio": 100 * base[name_c].value_counts() / len(base)}))
    if plot:
        sns.countplot(x=base[name_c], data=base)
        plt.show()
for i in category_columns:
    category_d(lim, i)
```

```
        gender      Ratio
Male      3555  50.47565
Female    3488  49.52435
     Partner      Ratio
No      3641  51.69672
Yes     3402  48.30328
     Dependents       Ratio
No         4933  70.041176
Yes        2110  29.958824
     PhoneService       Ratio
Yes          6361  90.316626
No            682   9.683374
                  MultipleLines       Ratio
No                         3390  48.132898
Yes                        2971  42.183729
No phone service            682   9.683374
             InternetService       Ratio
Fiber optic             3096  43.958540
DSL                     2421  34.374556
No                      1526  21.666903
                    OnlineSecurity       Ratio
No                            3498  49.666335
Yes                           2019  28.666761
No internet service           1526  21.666903
                    OnlineBackup       Ratio
```

Preliminary results:

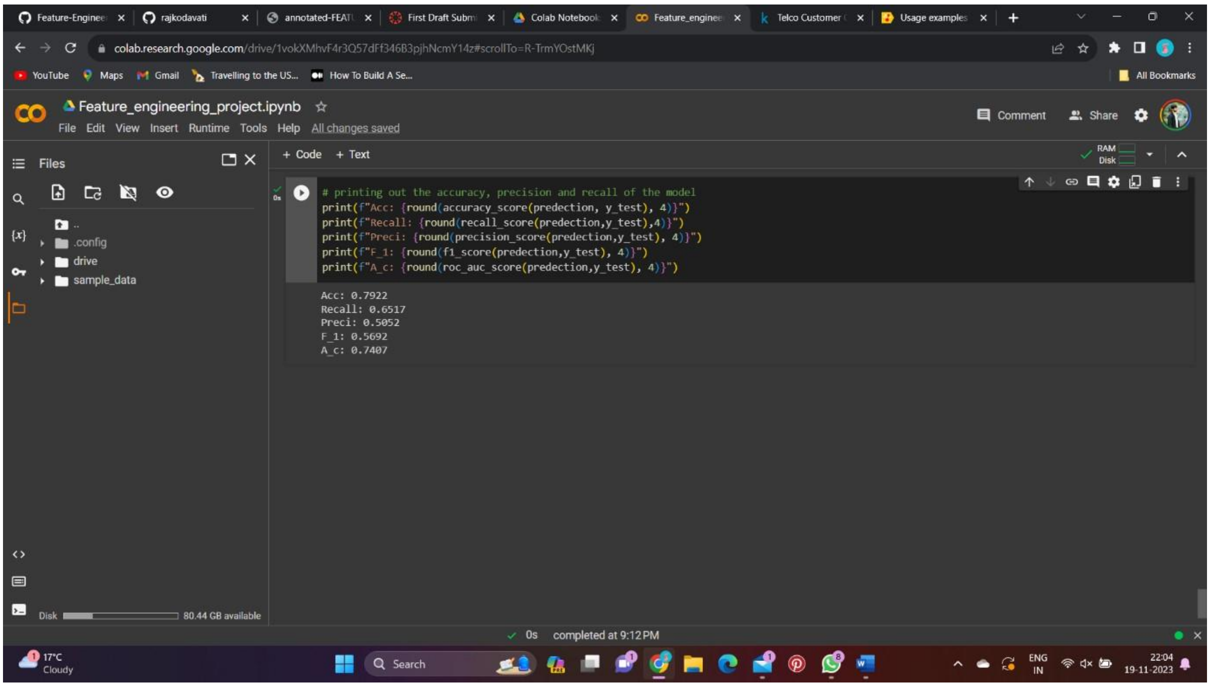
We print the corelation matrix of the model and we split our model to test and train dataset to find prediction and accuracy of our dataset.



```
#correlatoin matrix
f, ax = plt.subplots(figsize=(18, 13))
sns.heatmap(lim[number_columns].corr(), annot=True, fmt=".2f", ax=ax, cmap="magma")
plt.show(block=True)
```

```
[21] print(lim.isnull().sum()) #verifying the null values in data
     lim["TotalCharges"].fillna(0, inplace=True) #using fillna filling the null values with zero
     lim.isnull().sum()
```

```
customerID        0
gender            0
SeniorCitizen     0
Partner           0
```



```
# printing out the accuracy, precision and recall of the model
print(f"Acc: {round(accuracy_score(predection, y_test), 4)}")
print(f"Recall: {round(recall_score(predection,y_test),4)}")
print(f"Preci: {round(precision_score(predection,y_test), 4)}")
print(f"F_1: {round(f1_score(predection,y_test), 4)}")
print(f"A_c: {round(roc_auc_score(predection,y_test), 4)}")
```

```
Acc: 0.7922
Recall: 0.6517
Preci: 0.5052
F_1: 0.5692
A_c: 0.7407
```

Project Management:

Work completed:

We make our dataset perfect to our model and we perform prediction and deep analysis of our data, and we train and test the data to find how accurate is our dataset and we make features ready to extract to build our model, for predicting the final out as employees will leave the company or not.

Responsibility:

Rajagopal Naidu Kodavati make all the necessary requirements to prepare the data and perform the tasks to get best accuracy and brief analysis of data.

Jaya naga satya pavan ganesh kotipalli and Jaya sindhu Edara helped me in documentation for first draft and making proposal

Sri Mounish Seeni helps me in making proposal **Contribution:**

Rajagopal Naidu Kodavati – 50%

Jaya naga satya pavan ganesh – 25%

Jaya Sindhu Edara – 20%

Sri Mounish Seeni – 5%

Work to be completed:

After we make our dataset and analysis perfect to our model and make features, next step we need to perform good feature selection and extraction of selected features. After the successful extraction of features, we build our model by using encoding techniques and predict the results as final output. Also displays how accurate our model.

Responsibility:

Rajagopal Naidu Kodavati – Feature Extraction

Jaya naga satya pavan ganesh kotipalli – Encoding and build model

Jaya sindhu - Making in final project documentation

Sri Mounish Seeni – Displays final output and split the predicted output to perform operation to show accuracy of our model
Contibution:

Rajagopal Naidu Kodavati – 25%

Jaya naga satya pavan ganesh kotipalli – 25%

Jaya sindhu – 25% Sri

Mounish Seeni – 25%

References:

1. Makhtar M, Nafis S, Mohamed M, Awang M, Rahman M, Deris M. Churn classification model for local telecommunication company based on rough set theory. J Fundam Appl Sci. 2017;9(6):854–68 2. Suchánek P, Králová M: Customer satisfaction, loyalty, knowledge and competitiveness in the food industry. Economic Research-Ekonomska Istraživanja. 2019;32(1):1237–1255. 10.1080/1331677x.2019.1627893 [CrossRef] [Google Scholar]

3. Pope L: How to Prevent Customer Churn with Retention Marketing. G2. 2020, 27th August. Reference Source [Google Scholar]

INCRIMENT – 2

Introduction:

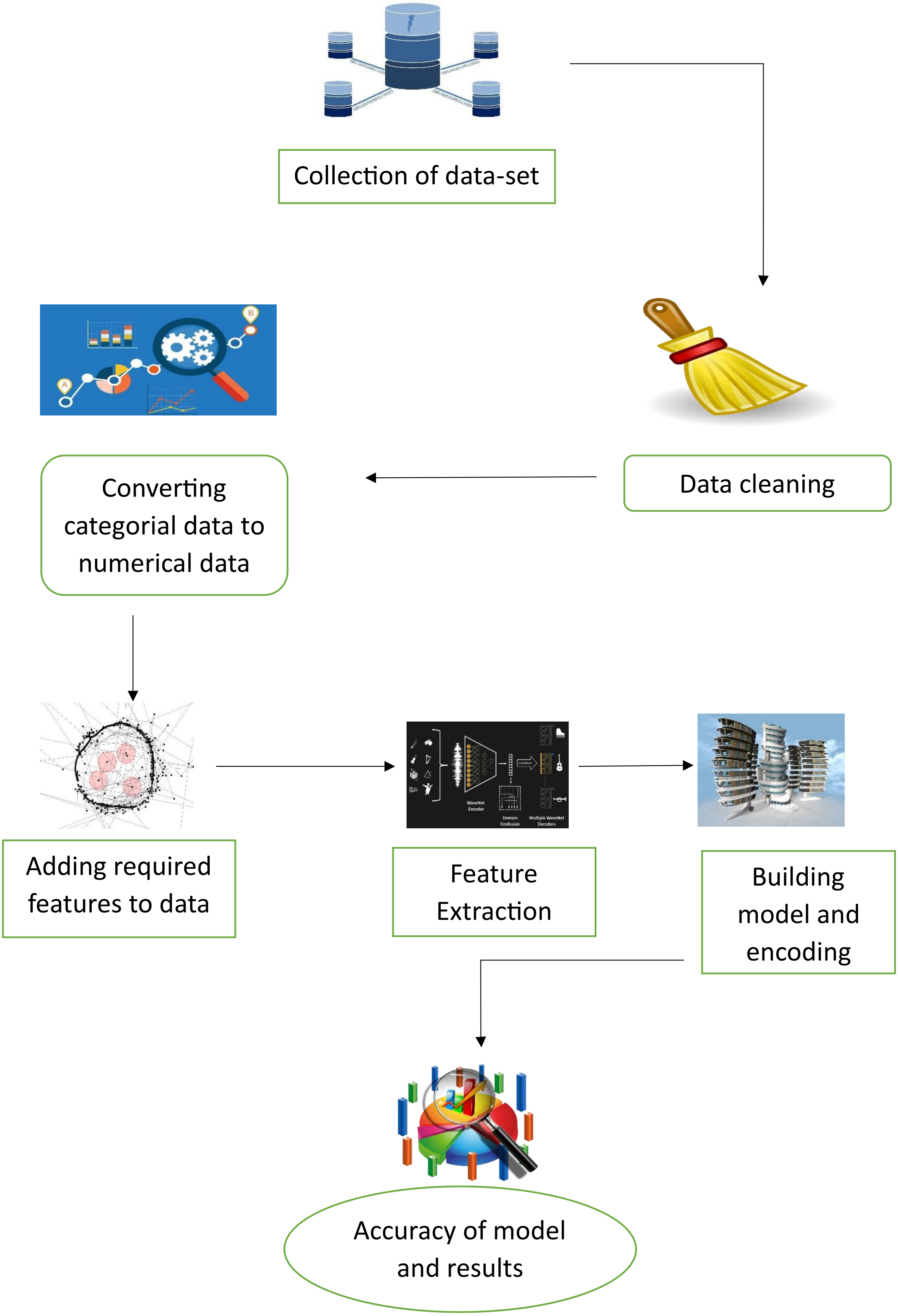
This project is on to develop a model that can predict customers leaving telecommunication companies, this is basically to improve the satisfaction of the customers. The main plot of this model is to bring back churned customers, in general companies will identify those customers and offer them with different deals to attract them back due to lot of clients it will eventually remines a loss to the companies. This model helps to identify the churned clients and reason behind it so the companies can predict the cause that is responsible for dissatisfaction of the users and then resolve it in the earlier stage. So that the user will not leave the services. The telecommunication companies must spend money on these churned customers to make them to remain in their services without any estimation they might spent more money than needed on these churns or maybe less this will affect the economic status of the telecommunication company. By using feature extraction, data preprocessing and modeling we built a model that can predict a churned client in advance. There are different types of processes to identify different types of churned users. So that the companies can plan different safekeeping techniques.

Background:

(1)(Makhtar M, Nafis S, Mohamed M, Awang M, Rahman M, Deris M. Churn classification model for local telecommunication company based on rough set theory.) there research has basically focused on identifying churners and non-churners in the telecommunication company by using data splitting, feature selection, data discretization, attribute reduction, and filtering process. It helped us with selection of most important feature to our model and data pre-processing.(2)(J Fundam Appl Sci. 2017;9(6):854–68 2. Suchánek P, Králová M: Customer satisfaction, loyalty, knowledge and competitiveness in the food industry. Economic Research-Ekonomska Istraživanja. 2019;32(1):1237–1255. 10.1080/1331677x.2019.1627893 [CrossRef] [Google Scholar]) there research is about direct effect of product knowledge on users satisfaction and competition between the companies this can also be influenced by the loyalty from their work we have used encoding techniques those are one hot encoding technique and labelled encoding technique. (3)(Pope L: How to Prevent Customer Churn with Retention Marketing. G2. 2020, 27th August.) there paper is about the issue of user churn identification by telecommunication companies at which rate if frequencies it has been raised. From their research splitting the data and training the data to the model are the techniques that are used for this project.

Model:

Architecture diagram:



Collection of data-set

Data cleaning

Converting categorial data to numerical data

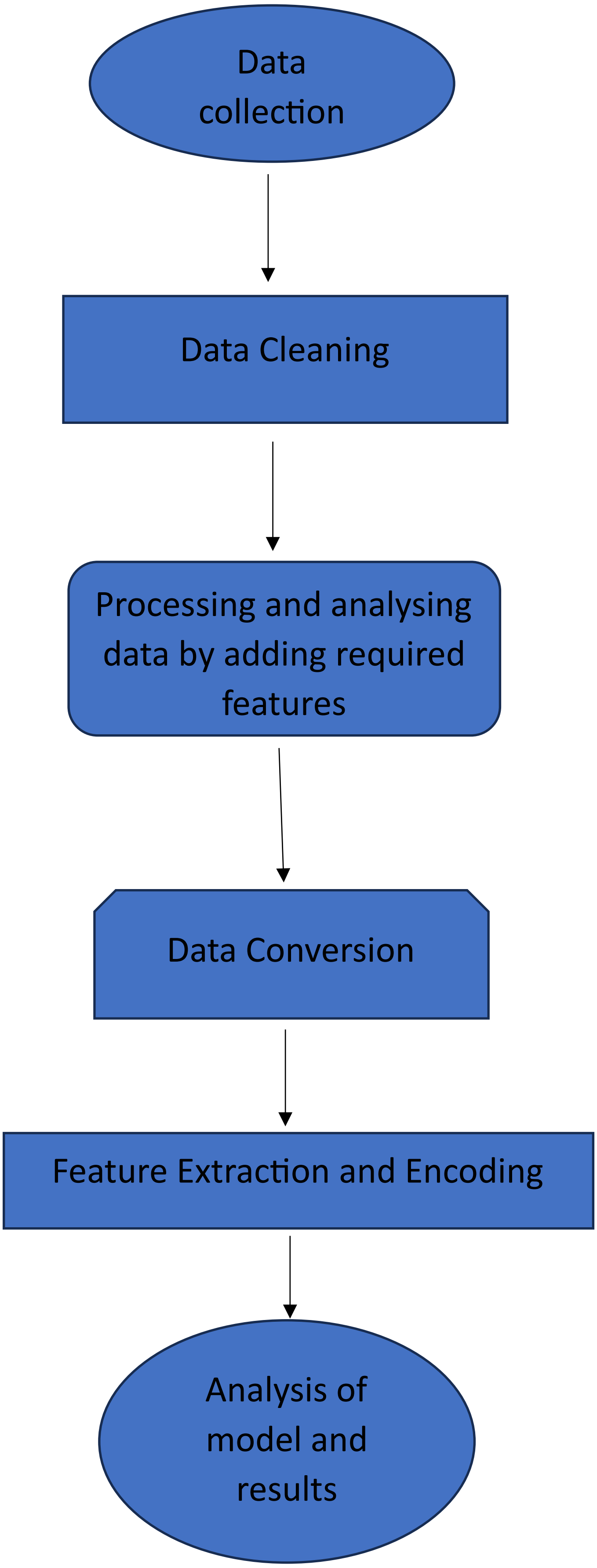
Adding required features to data

Feature Extraction

Building model and encoding

Accuracy of model and results

Workflow diagram:


Data collection
Data Cleaning
Processing and analysing data by adding required features
Data Conversion
Feature Extraction and Encoding
Analysis of model and results

Dataset:

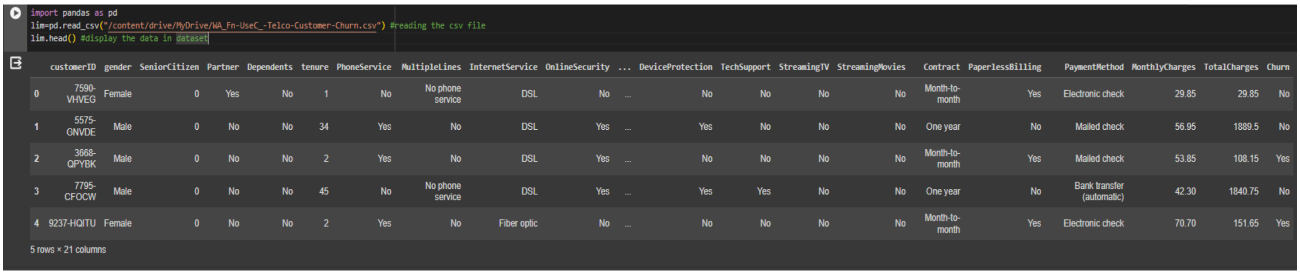
Dataset link:

https://www.kaggle.com/datasets/blastchar/telco-customer-churn

We take this dataset from kagle we choose this dataset because the features which are present in the dataset are best fit for model, as we are working on telecommunication companies to predict whether the customer leaves company or not, so here our dataset was belongs to one of the telecommunication company. For our model we required to have complete employee’s details like age, gender, experience, hobbies, martial status and more. In this data set we have most of the features and we also add more features to dataset, to make our model more accurate.

Design of features in dataset:

Here we list all the columns (features) of our dataset and first 5 rows of sample data.



```
import pandas as pd
lim=pd.read_csv("/content/drive/MyDrive/WA_Fn-UseC_-Telco-Customer-Churn.csv") #reading the csv file
lim.head() #display the data in dataset
```

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No	No	No	No	Month-to-month	Yes	Electronic check	29.85	29.85	No
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	Yes	No	No	No	One year	No	Mailed check	56.95	1889.5	No
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No	No	No	No	Month-to-month	Yes	Mailed check	53.85	108.15	Yes
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes	Yes	No	No	One year	No	Bank transfer (automatic)	42.30	1840.75	No
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	No	No	No	No	Month-to-month	Yes	Electronic check	70.70	151.65	Yes

5 rows × 21 columns

Column’s: We labelled some important features in our dataset

Customer id

Gender

SeniorCitizen

Partner

Dependents

tenure

PhoneService

Analysis of data:

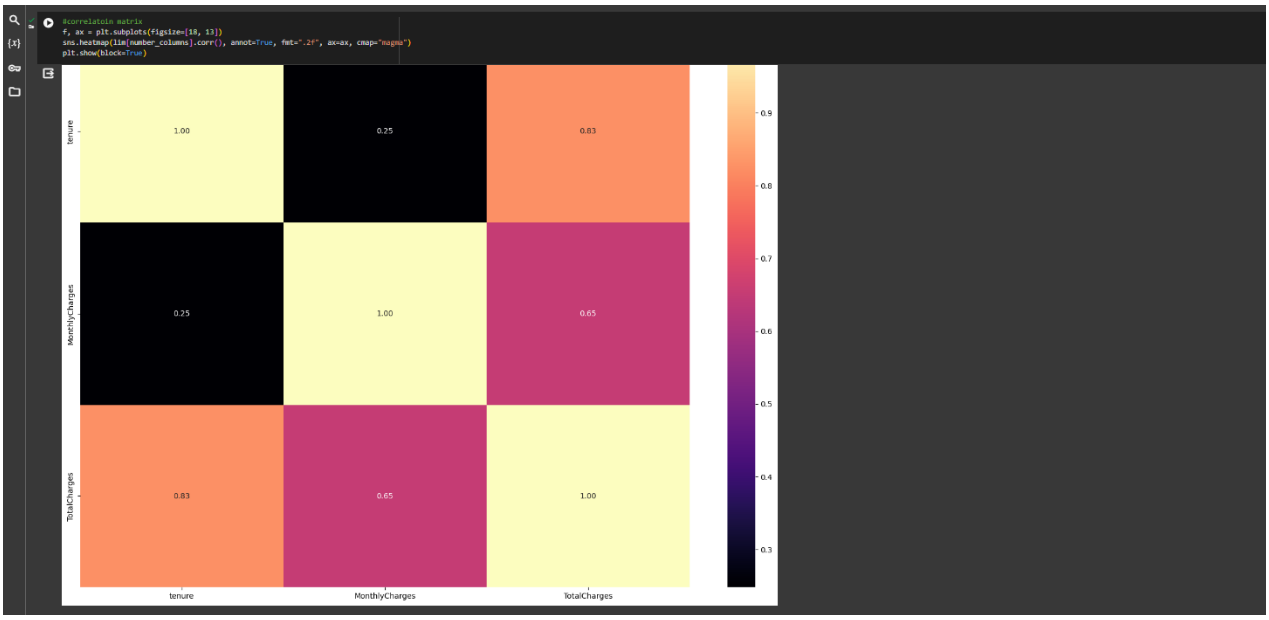
Initially we check for the null values on rows and columns. Then we look into the dimensions of data. We now check the info. Of each variable like which data type each variable is. Then we converted the "TotalCharges" column to numerical values and transform the "Churn" column to binary values 1to ‘yes’, 0 to ‘No’. This is useful cuz it replaces any unconverted values to numeric in turn avoiding errors. We then use ‘col’ function it classifies the data frame in three lists. Category columns for object-type columns, Number_b_columns for numeric columns with unique values below x, and Category_b_columns for object columns with unique values exceeding y. These help us provide distribution and characteristics of categorical and numerical columns in a dataset. We then define a function to get a detailed summaries for categorical columns. Here we fing the percentage ratios of each category. We have done the similar analysis to the numerical column data.

We then performed univariate analysis for each variable Starting with tenure, We have calculated the count , SD, mean which gives distribution of customer tenure in months. Followed by monthly charges and total charges these analyses helped us in understanding characteristics, identifying the outliners and help us getting better modelling decisions. We then performed the correlation matrix, it helped us in providing insights about the linear relationship between variables and shows how strong two variables are related with the help of heatmap. This helped us in better selecting the features and model building.

Before building and training the model we made sure that there are no null values in the dataset.

Graph model:

We finally analysed our rawdataset with some changes in confusion matrix



```
#correlatoin matrix
f, ax = plt.subplots(figsize=[18, 13])
sns.heatmap(lim[number_columns].corr(), annot=True, fmt=".2f", ax=ax, cmap="magma")
plt.show(block=True)
```

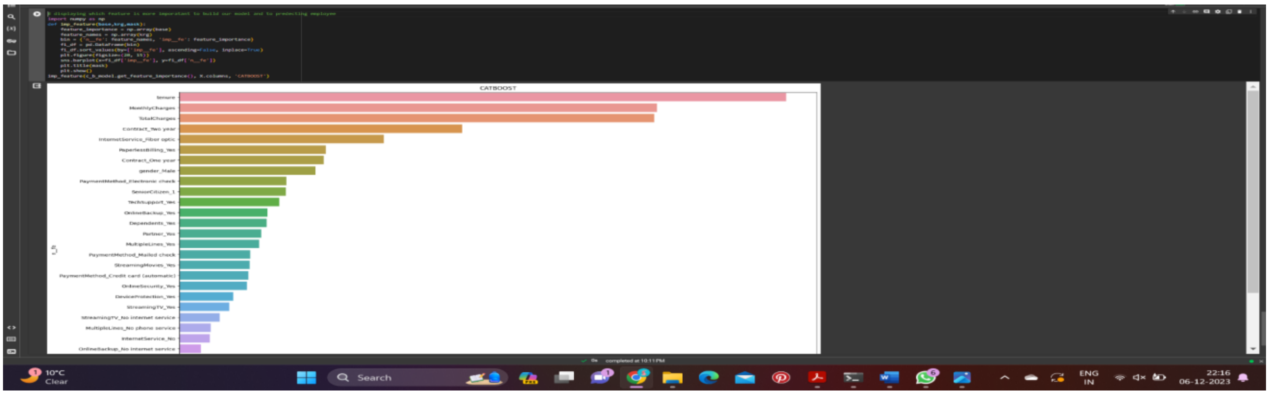
Implementation:

Algorithms:

For this model we use simple algorithms like training the data, performing encoding techniques to the extracted data for making predictions and we mainly work on feature extraction by making necessary features and alter the previous features which are in raw dataset.

Finally, we implement the importance of features to prediction and also display the features which are mostly considered to decide whether the employee leaving company or not.

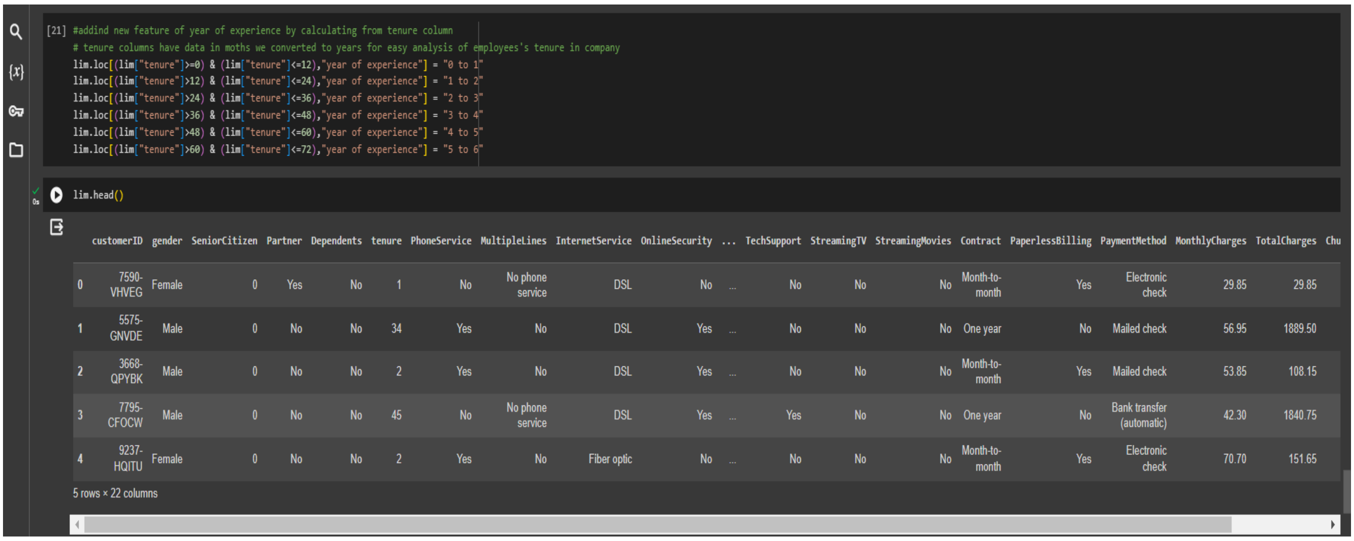
We also implement the accuracy, precision,recall of our prodel to know how accurate our model was.



```
# displaying which feature is more imporatant to build our model and to predecting employee
import numpy as np
def imp_feature(base,krg,mask):
    feature_importance = np.array(base)
    feature_names = np.array(krg)
    bin = {'n__fe': feature_names, 'imp__fe': feature_importance}
    fi_df = pd.DataFrame(bin)
    fi_df.sort_values(by=['imp__fe'], ascending=False, inplace=True)
    plt.figure(figsize=(20, 15))
    sns.barplot(x=fi_df['imp__fe'], y=fi_df['n__fe'])
    plt.title(mask)
    plt.show()
imp_feature(c_b_model.get_feature_importance(), X.columns, 'CATBOOST')
```

Results:

After the completion of first draft submission we worked on adding some extra features to our dataset which helps to make prediction of employees, we merge some old features to make easy processing.



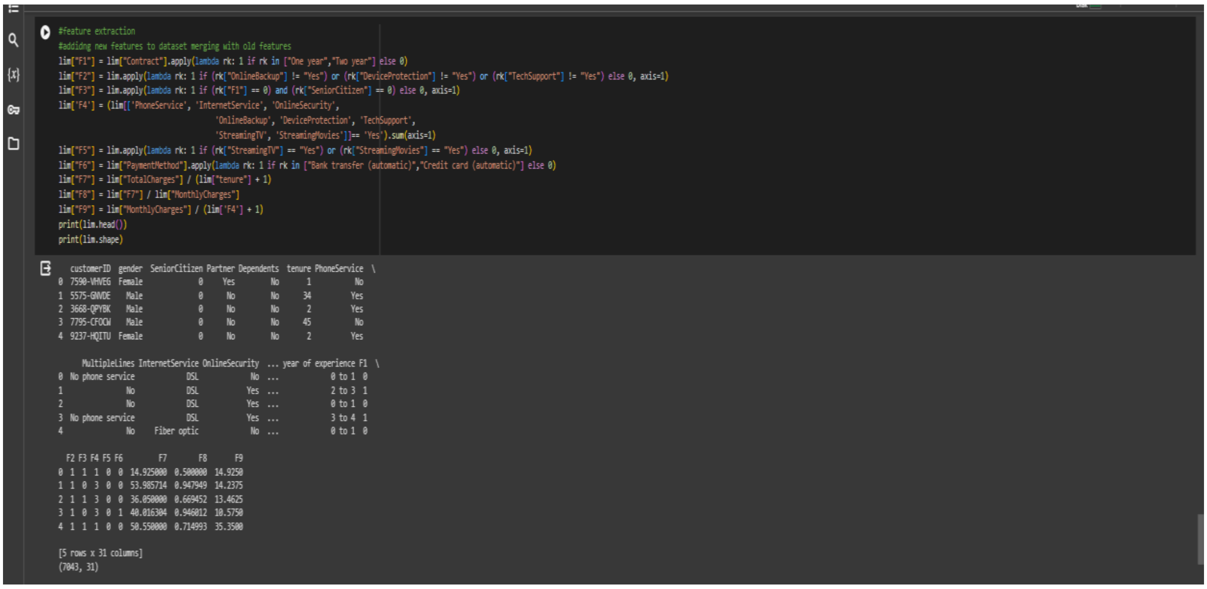
```
[21] #addind new feature of year of experience by calculating from tenure column
     # tenure columns have data in moths we converted to years for easy analysis of employees's tenure in company
     lim.loc[(lim["tenure"]>=0) & (lim["tenure"]<=12),"year of experience"] = "0 to 1"
     lim.loc[(lim["tenure"]>12) & (lim["tenure"]<=24),"year of experience"] = "1 to 2"
     lim.loc[(lim["tenure"]>24) & (lim["tenure"]<=36),"year of experience"] = "2 to 3"
     lim.loc[(lim["tenure"]>36) & (lim["tenure"]<=48),"year of experience"] = "3 to 4"
     lim.loc[(lim["tenure"]>48) & (lim["tenure"]<=60),"year of experience"] = "4 to 5"
     lim.loc[(lim["tenure"]>60) & (lim["tenure"]<=72),"year of experience"] = "5 to 6"
```

```
lim.head()
```

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Chu
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No	No	No	Month-to-month	Yes	Electronic check	29.85	29.85	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	No	No	No	One year	No	Mailed check	56.95	1889.50	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No	No	No	Month-to-month	Yes	Mailed check	53.85	108.15	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes	No	No	One year	No	Bank transfer (automatic)	42.30	1840.75	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	No	No	No	Month-to-month	Yes	Electronic check	70.70	151.65	

5 rows × 22 columns

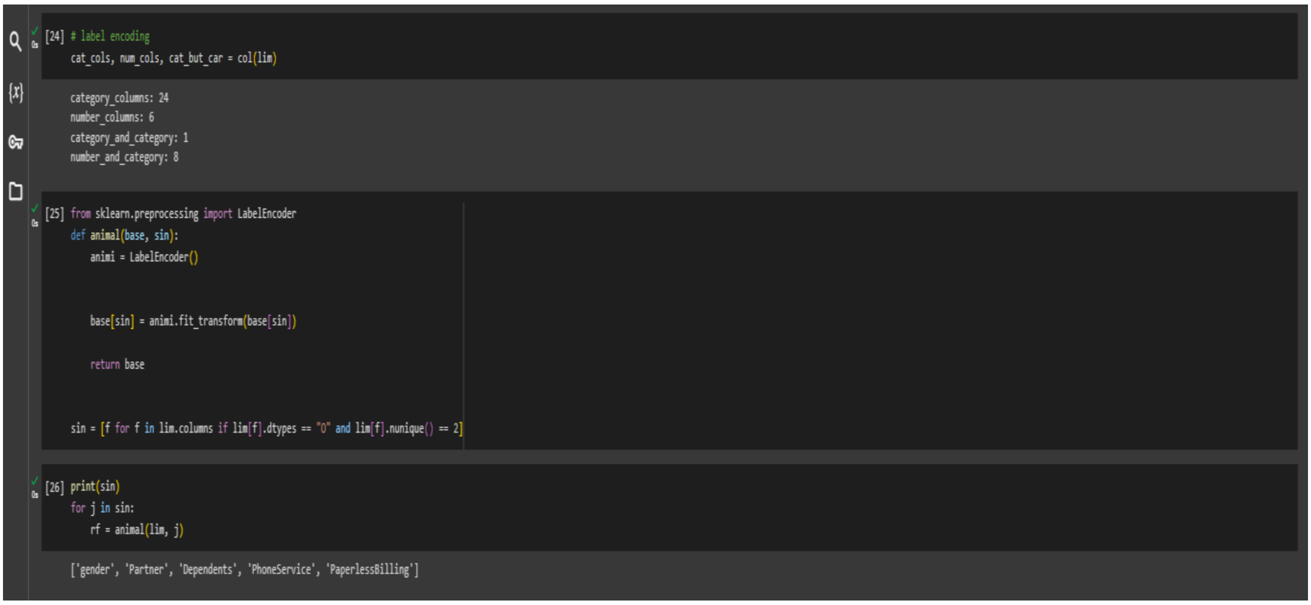
After adding the required feature we enter into the main step which is feature extraction here we extracted all the required feature which helps to our model and make new features by merging the old features also taking easy columns which are altered above to fit in our model.



```
#feature extraction
#addidng new features to dataset merging with old features
lim["F1"] = lim["Contract"].apply(lambda rk: 1 if rk in ["One year","Two year"] else 0)
lim["F2"] = lim.apply(lambda rk: 1 if (rk["OnlineBackup"] != "Yes") or (rk["DeviceProtection"] != "Yes") or (rk["TechSupport"] != "Yes") else 0, axis=1)
lim["F3"] = lim.apply(lambda rk: 1 if (rk["F1"] == 0) and (rk["SeniorCitizen"] == 0) else 0, axis=1)
lim['F4'] = (lim[['PhoneService', 'InternetService', 'OnlineSecurity',
                  'OnlineBackup', 'DeviceProtection', 'TechSupport',
                  'StreamingTV', 'StreamingMovies']]== 'Yes').sum(axis=1)
lim["F5"] = lim.apply(lambda rk: 1 if (rk["StreamingTV"] == "Yes") or (rk["StreamingMovies"] == "Yes") else 0, axis=1)
lim["F6"] = lim["PaymentMethod"].apply(lambda rk: 1 if rk in ["Bank transfer (automatic)","Credit card (automatic)"] else 0)
lim["F7"] = lim["TotalCharges"] / (lim["tenure"] + 1)
lim["F8"] = lim["F7"] / lim["MonthlyCharges"]
lim["F9"] = lim["MonthlyCharges"] / (lim['F4'] + 1)
print(lim.head())
print(lim.shape)
```

```
   customerID  gender  SeniorCitizen Partner Dependents  tenure PhoneService  \
0  7590-VHVEG  Female              0     Yes         No       1           No
1  5575-GNVDE    Male              0      No         No      34          Yes
2  3668-QPYBK    Male              0      No         No       2          Yes
3  7795-CFOCW    Male              0      No         No      45           No
4  9237-HQITU  Female              0      No         No       2          Yes

      MultipleLines InternetService OnlineSecurity  ... year of experience F1  \
0  No phone service             DSL             No  ...             0 to 1  0
1                No             DSL            Yes  ...             2 to 3  1
2                No             DSL            Yes  ...             0 to 1  0
3  No phone service             DSL            Yes  ...             3 to 4  1
4                No     Fiber optic             No  ...             0 to 1  0

  F2 F3 F4 F5 F6         F7        F8       F9
0  1  1  1  0  0  14.925000  0.500000  14.9250
1  1  0  3  0  0  53.985714  0.947949  14.2375
2  1  1  3  0  0  36.050000  0.669452  13.4625
3  1  0  3  0  1  40.016304  0.946012  10.5750
4  1  1  1  0  0  50.550000  0.714993  35.3500

[5 rows x 31 columns]
(7043, 31)
```

After feature extraction we perform encoding process to our data for building the model for this we use to main encoding techniques the first is labelled encoding and one hot encoding.

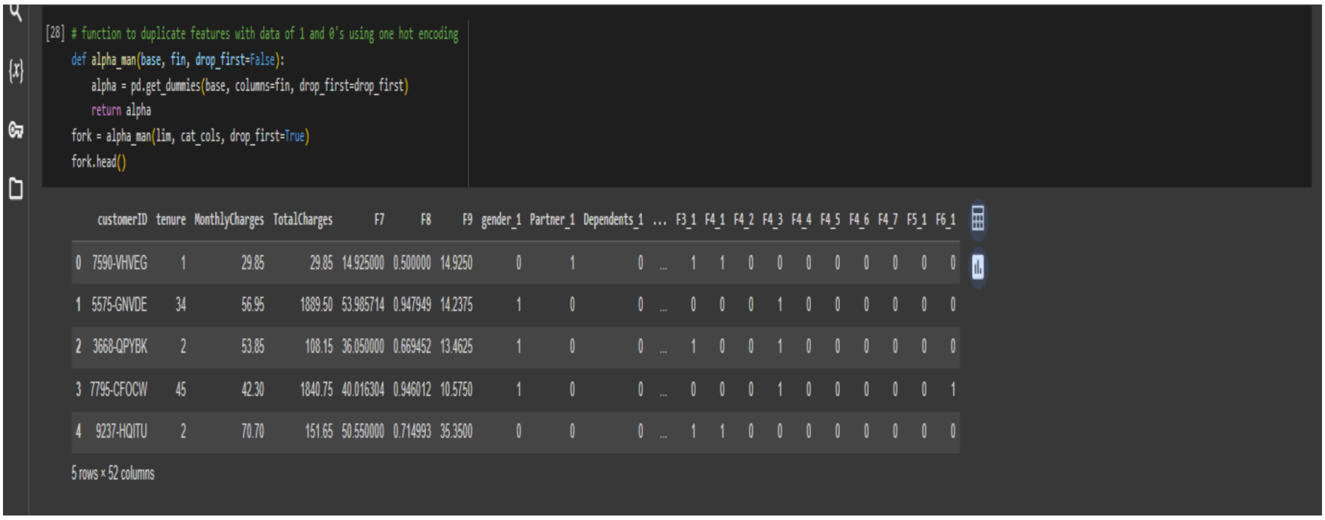


```
[24] # label encoding
     cat_cols, num_cols, cat_but_car = col(lim)

     category_columns: 24
     number_columns: 6
     category_and_category: 1
     number_and_category: 8
```

```
[25] from sklearn.preprocessing import LabelEncoder
     def animal(base, sin):
         animi = LabelEncoder()


         base[sin] = animi.fit_transform(base[sin])

         return base


     sin = [f for f in lim.columns if lim[f].dtypes == "O" and lim[f].nunique() == 2]
```

```
[26] print(sin)
     for j in sin:
         rf = animal(lim, j)

     ['gender', 'Partner', 'Dependents', 'PhoneService', 'PaperlessBilling']
```

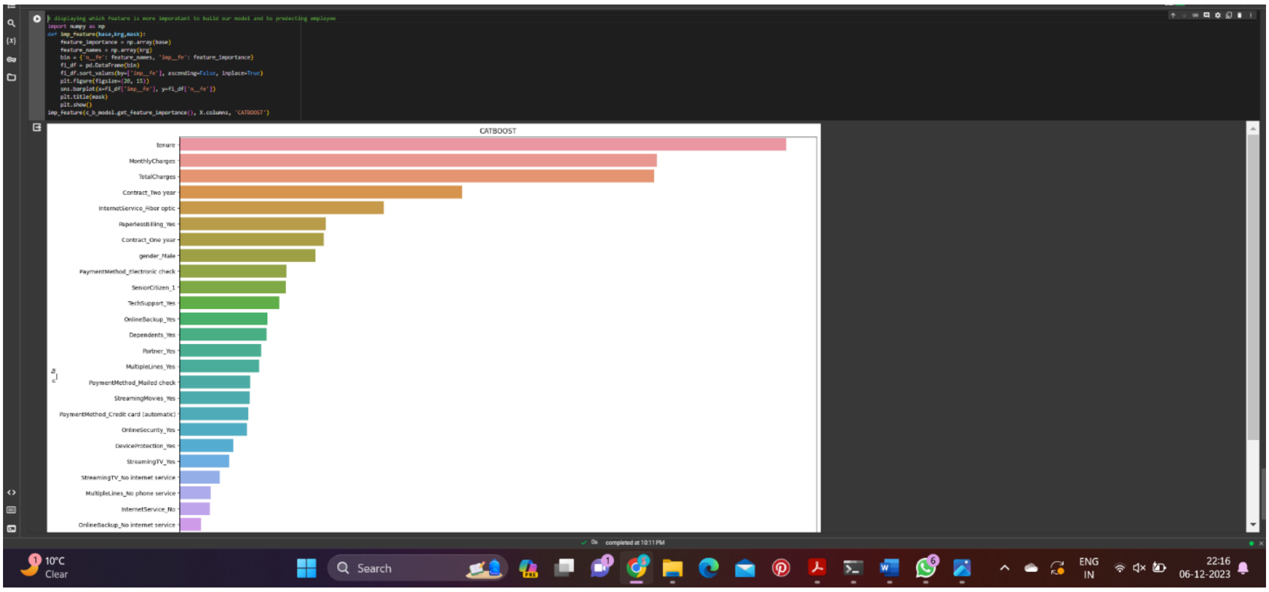
One hot encoding



```
[28] # function to duplicate features with data of 1 and 0's using one hot encoding
     def alpha_man(base, fin, drop_first=False):
         alpha = pd.get_dummies(base, columns=fin, drop_first=drop_first)
         return alpha
     fork = alpha_man(lim, cat_cols, drop_first=True)
     fork.head()
```

```
     customerID  tenure  MonthlyCharges  TotalCharges         F7        F8       F9  gender_1  Partner_1  Dependents_1  ...  F3_1  F4_1  F4_2  F4_3  F4_4  F4_5  F4_6  F4_7  F5_1  F6_1
  0  7590-VHVEG       1           29.85         29.85  14.925000  0.500000  14.9250         0          1             0  ...     1     1     0     0     0     0     0     0     0     0
  1  5575-GNVDE      34           56.95       1889.50  53.985714  0.947949  14.2375         1          0             0  ...     0     0     0     1     0     0     0     0     0     0
  2  3668-QPYBK       2           53.85        108.15  36.050000  0.669452  13.4625         1          0             0  ...     1     0     0     1     0     0     0     0     0     0
  3  7795-CFOCW      45           42.30       1840.75  40.016304  0.946012  10.5750         1          0             0  ...     0     0     0     1     0     0     0     0     0     1
  4  9237-HQITU       2           70.70        151.65  50.550000  0.714993  35.3500         0          0             0  ...     1     1     0     0     0     0     0     0     0     0

  5 rows × 52 columns
```

Later on we display how important features are to build our model printing each feature and its importance to predict the customers leaving company or not.



```
# displaying which feature is more imporatant to build our model and to predecting employee
import numpy as np
def imp_feature(base,krg,mask):
    feature_importance = np.array(base)
    feature_names = np.array(krg)
    bin = {'n__fe': feature_names, 'imp__fe': feature_importance}
    fi_df = pd.DataFrame(bin)
    fi_df.sort_values(by=['imp__fe'], ascending=False, inplace=True)
    plt.figure(figsize=(20, 15))
    sns.barplot(x=fi_df['imp__fe'], y=fi_df['n__fe'])
    plt.title(mask)
    plt.show()
imp_feature(c_b_model.get_feature_importance(), X.columns, 'CATBOOST')
```

Project Management:

Work completed:

Description:

We have completed all of our project model now it is ready to work and for making this project successful we split the work among our team, as we are team of four members we split the work like handling data, handling source code, handling the documentation part. Our project works on predicting the employee leaving company or not for this we analyse the data of the employee by making data perfect with all the required feature we can assume the chance of employee leaving or not. We use skitlearn to train the data and fit into model also we use encoding techniques in our model.

Responsibilities:

Rajagopal Naidu Kodavati – Works most of the coding part and team lead of the project

Jaya naga satya pavan ganesh – Handel the data set and working in analysing and cleaning of data

Jaya sindhu – Works with proposal and first draft

Mounish – Works with documentation part in final submission

Contribution:

The overall contribution of our project by us is labelled in percentage

Rajagopal Naidu Kodavati – 40%

Jaya naga satya pavan ganesh – 25%

Mounish – 25%

Jaya sindhu – 25%

Issues/concerns:

We face little issues on handling of data in initial stage by choosing perfect dataset to fit our model and later on we discussed on selecting required features to best fit our model so it helps to sort out things and in some cases we use internet for references.

References:

We have use the same references to entire project as I submitted in first draft and in final submission

1. Makhtar M, Nafis S, Mohamed M, Awang M, Rahman M, Deris M. Churn classification model for local telecommunication company based on rough set theory. J Fundam Appl Sci. 2017;9(6):854–68 2. Suchánek P, Králová M: Customer satisfaction, loyalty, knowledge and competitiveness in the food industry. Economic Research-Ekonomska Istraživanja. 2019;32(1):1237–1255. 10.1080/1331677x.2019.1627893 [CrossRef] [Google Scholar]

3. Pope L: How to Prevent Customer Churn with Retention Marketing. G2. 2020, 27th August. Reference Source [Google Scholar]

Github link:

https://github.com/rajkodavati/Model-to-predict-customers-leavingcomanies-using-feature-engineering