Software Requirements Specification
for

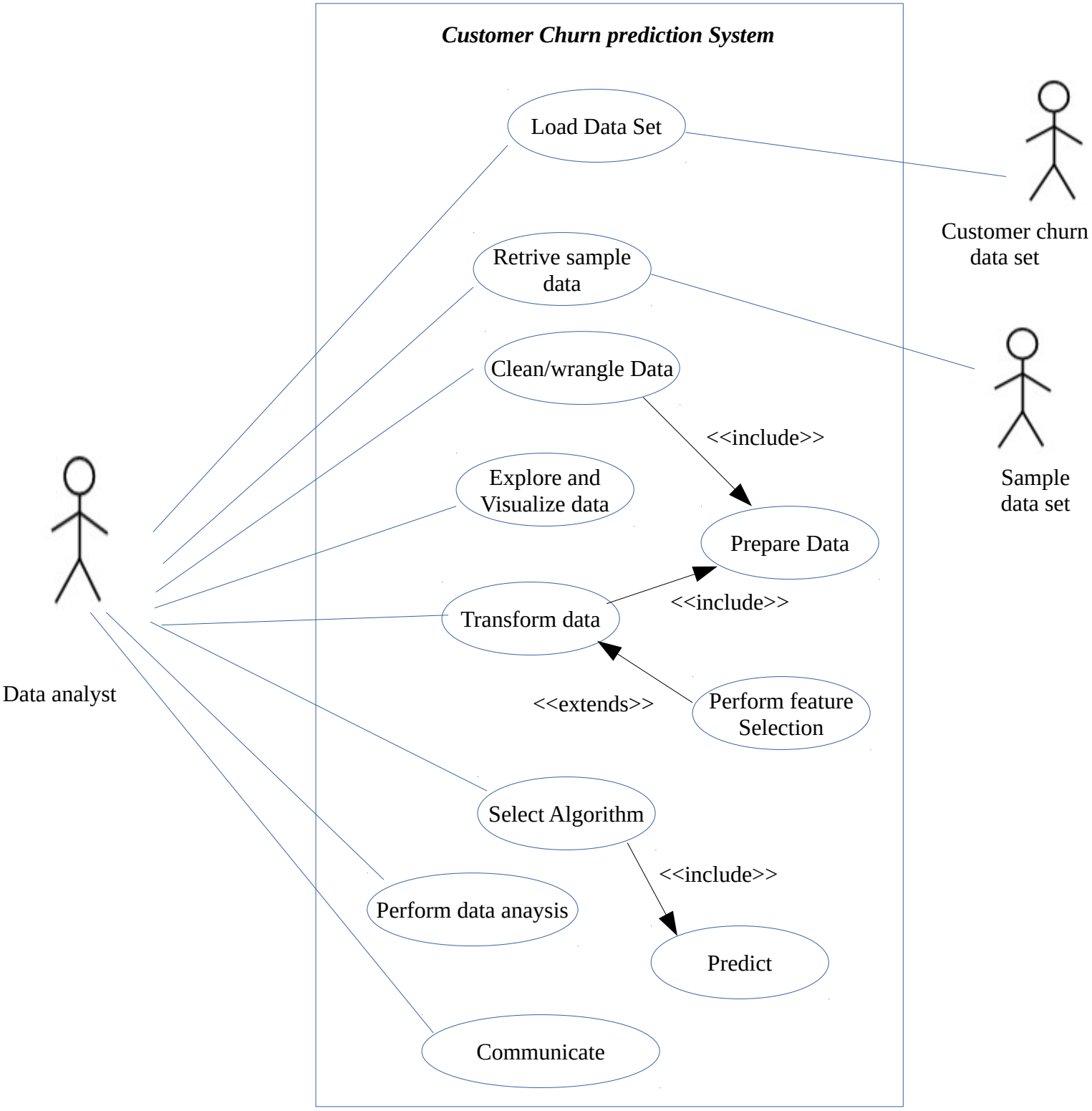
Telecom Customers Churn Prediction System

Prepared by

	Names	Registration Number	Signature
1	KYANZI HASSAN MUSISI	17/U/19979/EVE	
2	WAMALA EDGAR WATSON	17/U/20633/EVE	
3	KASOZI JAFFAR	17/U/4741/EVE	
4	BWAYO JOEL	17/U/3777/EVE	

Makerere University

The diagram below shows the Use case diagram and its description customer



The use case has one primary actor and two secondary actors actors are Data analyst , customer churn data set and sample data set .

Data analyst
performs the tasks like loading data set , creating sample data from the original data set to be used in performing analysis, cleaning/wrangling of data, selecting algorithm for prediction, exploring data and performing data analysis.

customer churn data set CSV file
This is the file containing the data to be loaded in the system. Sample data will be imported from this file using pandas library that contains the read_csv() function.

Sample data set CSV
This the file that will store the sample data to be used in the system for analysis.

Clean/ wrangle data
This involves filling of missing values, label encoding, one hot encoding and errors in the data set Libraries like numpy, pandas and sk-learn shall be used to perform the above task to make sure that our data ready to used for the next step . The include arrow shows that as data is being cleaned is being prepared for the next step in the pipeline.

Explore and Visualize data
Here graphs are plotted using different features against the target feature. For our case the target feature is churn and draw an insight or more understanding of the features from the visual graphs .The following Libraries shall be used.
matplotlib and seaborn libraries.

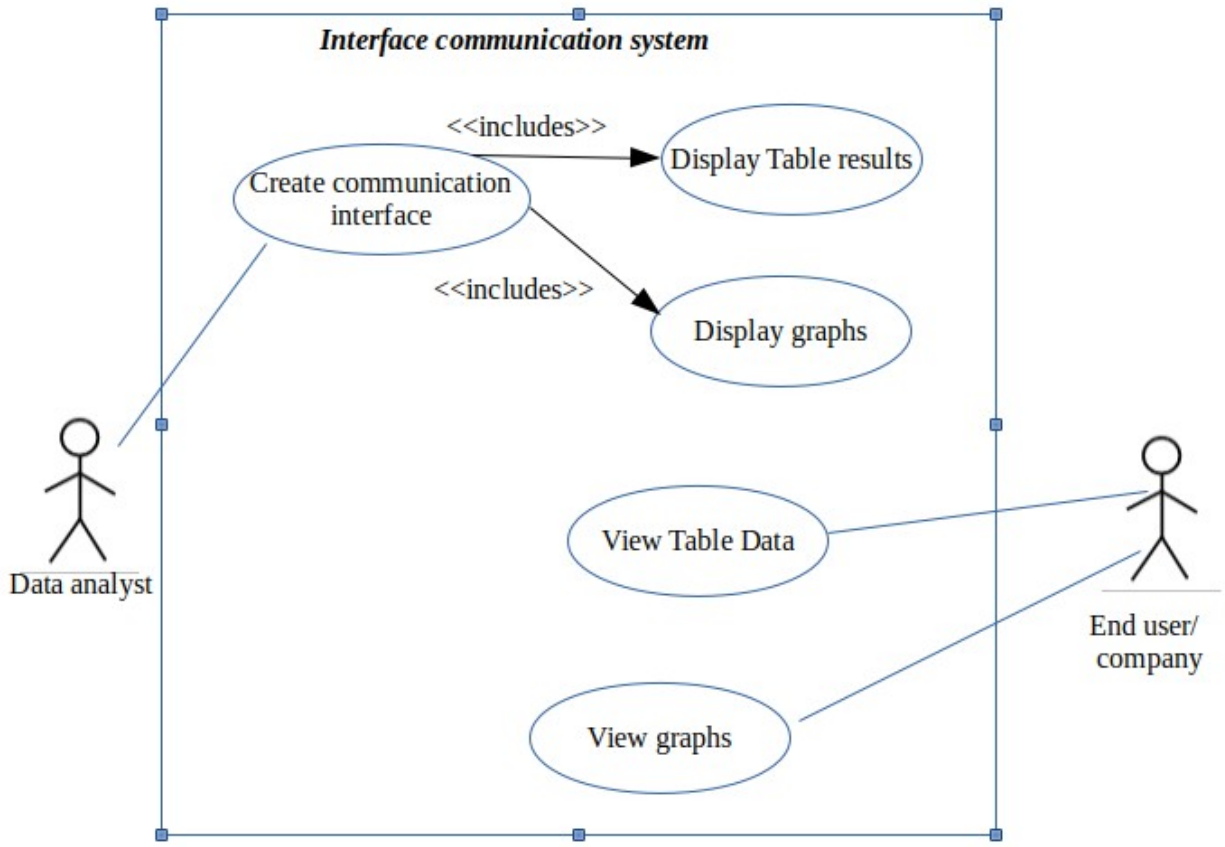
Transform data
This involves transformation into machine learning or modeling. This will help us to go into deeper understanding of the features and if possible specify or select the features to use in our model. This helps to remove irrelevant features.

Select algorithm
for our case we shall use logistic regression to train data and help us in making predictions.

Perform data analysis
From the graphs plotted analysis will be performed and insights drawn basing on visualization.

Communicate
This means how the end results are going to be communicated or deployed to the intended audience For our case we shall use dashboard to communicate our insight draw from visualization and also the accurate score . For our case we shall use dash python to communicate.

Use case diagram showing communication system of customer churn



Here we have two actors that's the Data analyst and End user or Company
Data analyst performs only one task that's creating communication interface.

End user/company
He can view table data and graphs which helps him to draw an insight.

Create communication interface
This involves creating dashboard which is used for displaying graphs and tables.
For our case we shall use dash or plotly to create a dashboard.