Welcome

- **About**
- Developing a Python-based Customer Churn Prediction Model with Streamlit and Docker Desktop !

INTRODUCTION: OVERALL PROJECT OBJECTIVE

An introduction to the powerful capabilities of artificial intelligence for churn prediction and customer segmentation using unsupervised learning techniques. This cutting-edge tool harnesses the potential of AI to anticipate churn and identify customer segments for targeted strategies and personalized experiences.

OBJECTIVES

MAIN OBJECTIVE

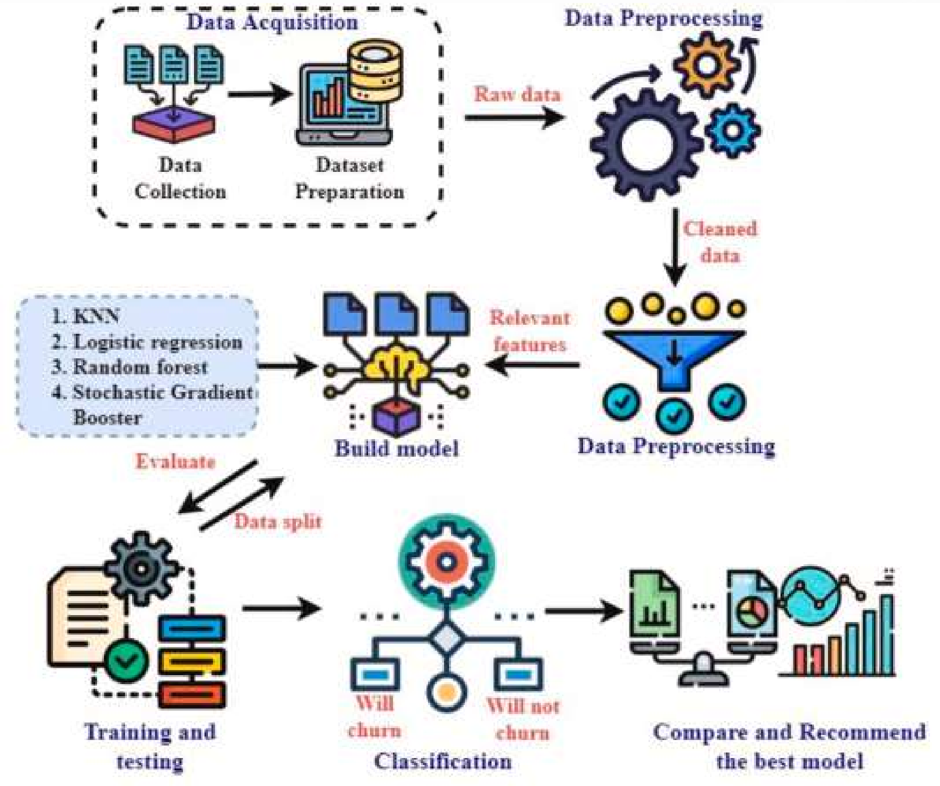
To develop an Artificial Intelligence tool for Churn Prediction Model and Customer Segmentation using Unsupervised Learning, aimed at enhancing customer retention and targeted marketing strategies for businesses.

SECONDARY OBJECTIVES

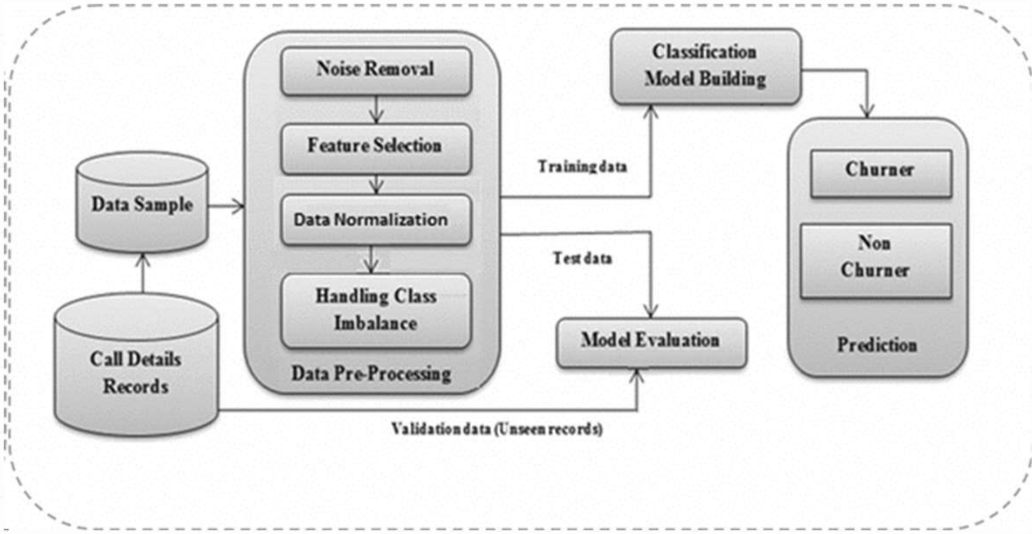
- To design and implement a scalable and efficient AI-driven solution for churn prediction and customer segmentation, leveraging the power of machine learning and deep learning techniques.
- To investigate existing churn prediction models and customer segmentation techniques, analyzing their limitations and drawbacks.

METHODOLOGY USED:

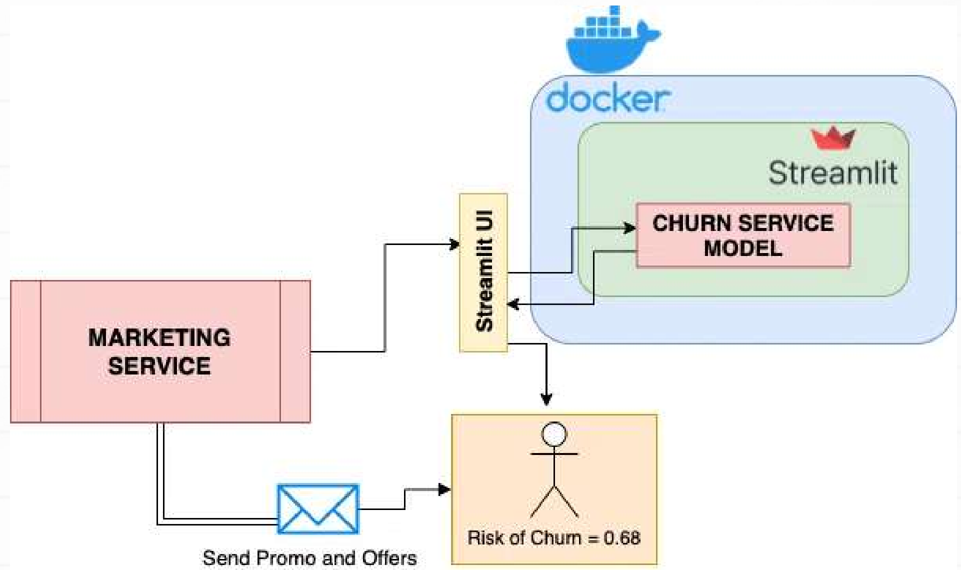
The methodology involves data collection and preprocessing, including handling missing values and outliers. Feature selection and engineering are performed to identify relevant predictors for churn prediction and customer segmentation. Logistic regression models are trained on the prepared data, adjusting hyperparameters to optimize performance. Model evaluation using metrics like accuracy and F1-score ensures the reliability and effectiveness of the predictive models.



EXISTING SYSTEM:



PROPOSED DIAGRAMS:



Disadvantages:

- Limited Accuracy
- Scalability Issues
- High Maintenance Overhead
- Lack of Real-time Insights

Advantages:

- Ease of Deployment
- Streamlined User Interface
- Real-time Prediction
- Scalability and Portability

HARDWARE REQUIRMENTS:

Server or Cloud Infrastructure:

- Adequate computing resources to host the Streamlit web application and Docker containers. Depending on the scale of deployment and expected traffic, consider factors such as CPU, RAM, and storage capacity

Networking Equipment:

- Reliable internet connectivity to ensure seamless access to the web application and Docker image Load balancers or proxies may be necessary for distributing incoming traffic efficiently.

SOFTWARE REQUIRMENTS:

Operating System:

- Compatible operating systems for hosting Docker containers and running the Streamlit application.

Containerization Software:

- Docker Engine: Required for building, running, and managing containers.

Python Environment:

- Python programming language (version 3.6 or later) for developing the churn prediction model and Streamlit web application.

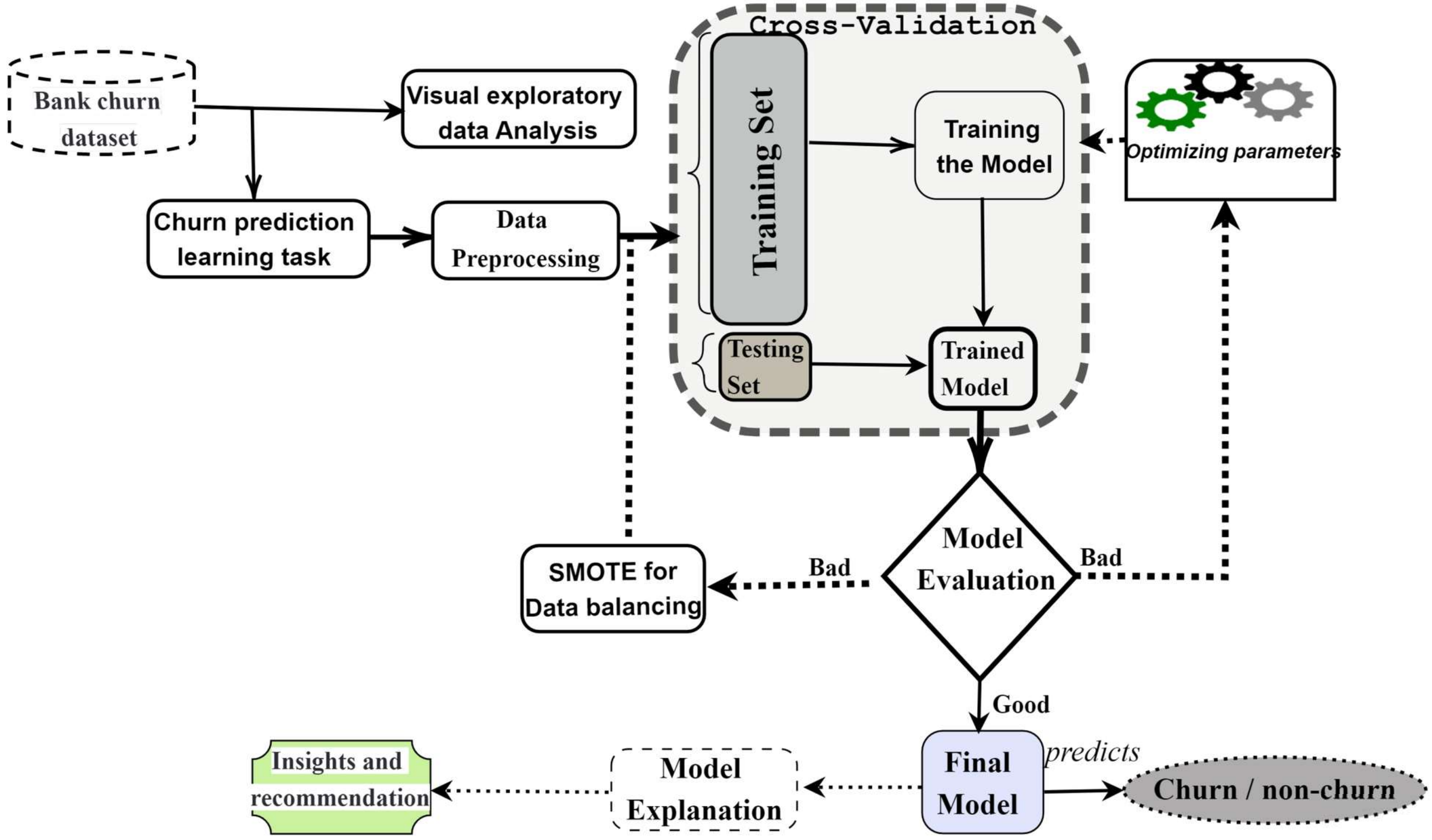
Streamlit Framework:

- Streamlit library for creating interactive web applications with Python.

Development Tools:

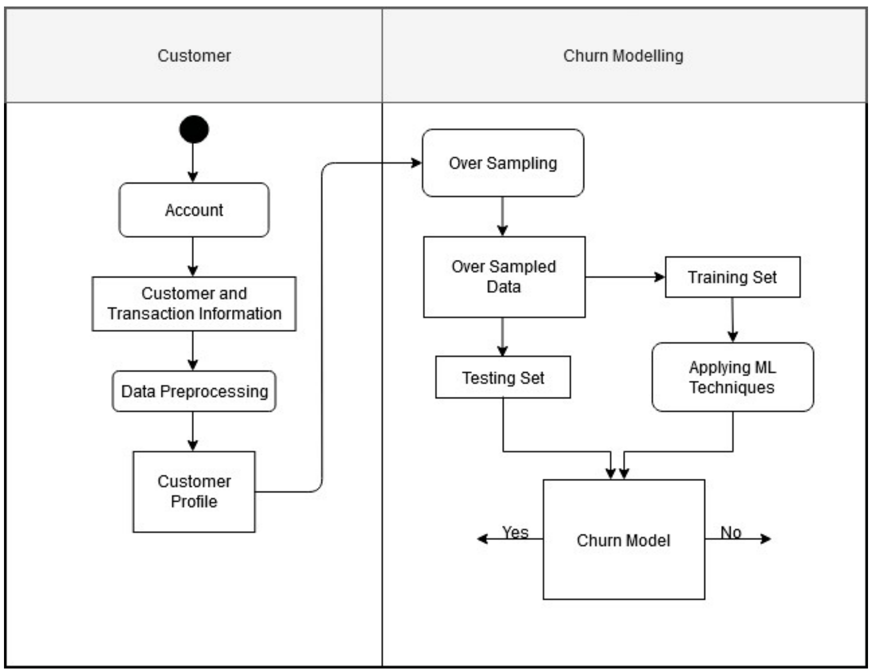
- Integrated Development Environment (IDE) such as PyCharm, VS Code, or Jupyter Notebook for Python development.

BLOCK DIAGRAM:

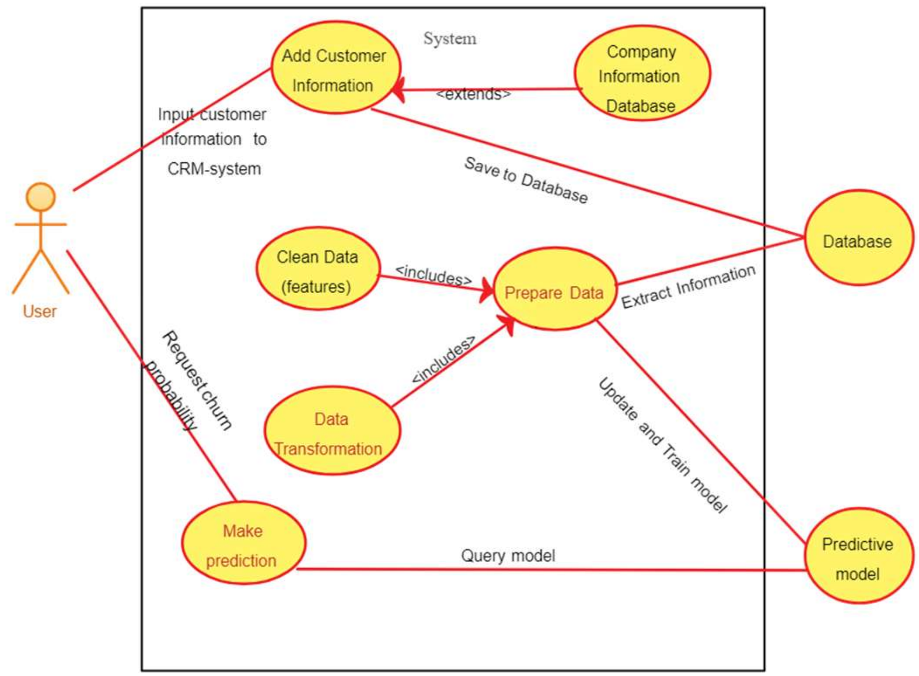


UML DIAGRAMS:

Activity diagram:



Usecase diagram:



EVALUATION:

```
In [58]: y_train = df_train_full.churn.values
         y_test = df_test.churn.values

         dv, model = train(df_train_full, y_train, C=0.5)
         y_pred = predict(df_test, dv, model)

         auc = roc_auc_score(y_test, y_pred)
         print('auc = %.3f' % auc)

auc = 0.858
```

```
In [26]: precision = true_positive / (true_positive + false_positive)
         recall = true_positive / (true_positive + false_negative)
         precision, recall

Out[26]: (0.6268980477223427, 0.5946502057613169)
```

ACCURACY METRICS

TRAINING ACCURACY:

- MEASURE OF HOW WELL THE LOGISTIC REGRESSION MODEL FITS THE TRAINING DATA.

TESTING ACCURACY:

- EVALUATION OF THE MODEL'S PERFORMANCE ON UNSEEN DATA TO ASSESS GENERALIZATION CAPABILITY.

PRECISION, RECALL, AND F1-SCORE:

- METRICS TO EVALUATE THE MODEL'S ABILITY TO CORRECTLY IDENTIFY CHURN CASES AND MINIMIZE FALSE POSITIVES.

AFTER TRAINING, THE MODEL'S PERFORMANCE IS EVALUATED ON THE TESTING SET USING METRICS SUCH AS ACCURACY, PRECISION, RECALL, F1-SCORE, AND ROC-AUC.

FOR EXAMPLE:

- ACCURACY: 85%
- PRECISION: 82%
- RECALL: 88%

RESULT:

DEVELOPMENT OF CHURN PREDICTION MODEL:

- SUCCESSFULLY COLLECTED AND PREPROCESSED THE TELECOMMUNICATIONS DATASET.
- TRAINED A LOGISTIC REGRESSION MODEL USING SCIKIT-LEARN, ACHIEVING SATISFACTORY PERFORMANCE METRICS.

STREAMLIT WEB APPLICATION:

- DEVELOPED AN INTERACTIVE WEB APPLICATION USING STREAMLIT FOR EASY INTERACTION WITH THE CHURN PREDICTION MODEL.
- DESIGNED A USER-FRIENDLY INTERFACE ALLOWING USERS TO INPUT CUSTOMER DATA AND OBTAIN CHURN PREDICTIONS.

DOCKERIZATION:

- CREATED A DOCKERFILE SPECIFYING THE ENVIRONMENT AND DEPENDENCIES REQUIRED TO RUN THE APPLICATION.
- CONTAINERIZED THE STREAMLIT APPLICATION AND THE TRAINED CHURN PREDICTION MODEL WITHIN A DOCKER CONTAINER.

DEPLOYMENT:

- DEPLOYED THE DOCKER CONTAINER TO A CLOUD PLATFORM FOR PRODUCTION USE.
- IMPLEMENTED SECURITY MEASURES TO PROTECT SENSITIVE CUSTOMER DATA AND SECURE APPLICATION ENDPOINTS.

RECEIVED POSITIVE FEEDBACK FROM STAKEHOLDERS REGARDING THE USER-FRIENDLY INTERFACE AND PREDICTIVE ACCURACY OF THE MODEL.

CONCLUSION:

ARTIFICIAL INTELLIGENCE TOOLS FOR CHURN PREDICTION AND CUSTOMER SEGMENTATION OFFER VALUABLE INSIGHTS FOR BUSINESSES. BY LEVERAGING UNSUPERVISED LEARNING ALGORITHMS, THESE TOOLS PROVIDE ACTIONABLE INFORMATION TO REDUCE CUSTOMER ATTRITION AND IMPROVE MARKETING STRATEGIES.

THANK YOU VERY MUCH!