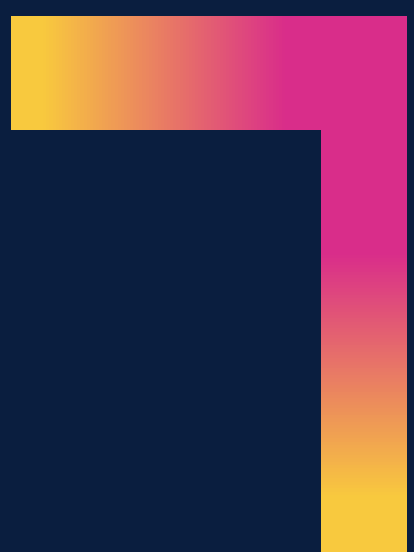




Churn Detection Modeling

Students: Jairo Adarmes (Capstone II)

Mentor: Lina Sokol, Product Owner (Kaseya)

Instructor/Faculty: *Masoud Sadjadi*, Florida International University

PROBLEM

Our sponsor (Kaseya) wants to identify clients that suspend or are most likely going to suspend their services and why are they stopping the service, whether if it is due to Kaseya solving their problem, or client did not find out how to use the service at full potential or any reason that could be presented. Identifying and classifying factors and possible reasons why a client came to churn the service would give Kaseya an overview of how to interfere, improve, and prevent future clients to end their service by providing them their needs.

VERIFICATION

Our model went into testing by our own members of the team.

We tested accuracy with:

- Logistic Regression (very low accuracy since unbalanced data provided)
- SMOTE to oversample our data (Accuracy was improved with this technique).
- Added K-Fold validation (Accuracy became more consistent in the model).

Predictive accuracy achieved was ~68%

(with Duration, Won Opportunities, Renewed, and One-time purchase – being the biggest factor in to play when customer churn)

SUMMARY

Our Churn Detection Model Project is a machine learning predictive model that classifies whether a client of Kaseya would churn or renew, by identifying possible factors that can contribute into the client's decision with the membership of the service, which were Kaseya identifying certain factors on customers can help the business improve/offer better service to clients in order to continue business with customers. Model has lots of room for accuracy improvement and expansions with more better-balanced data set.

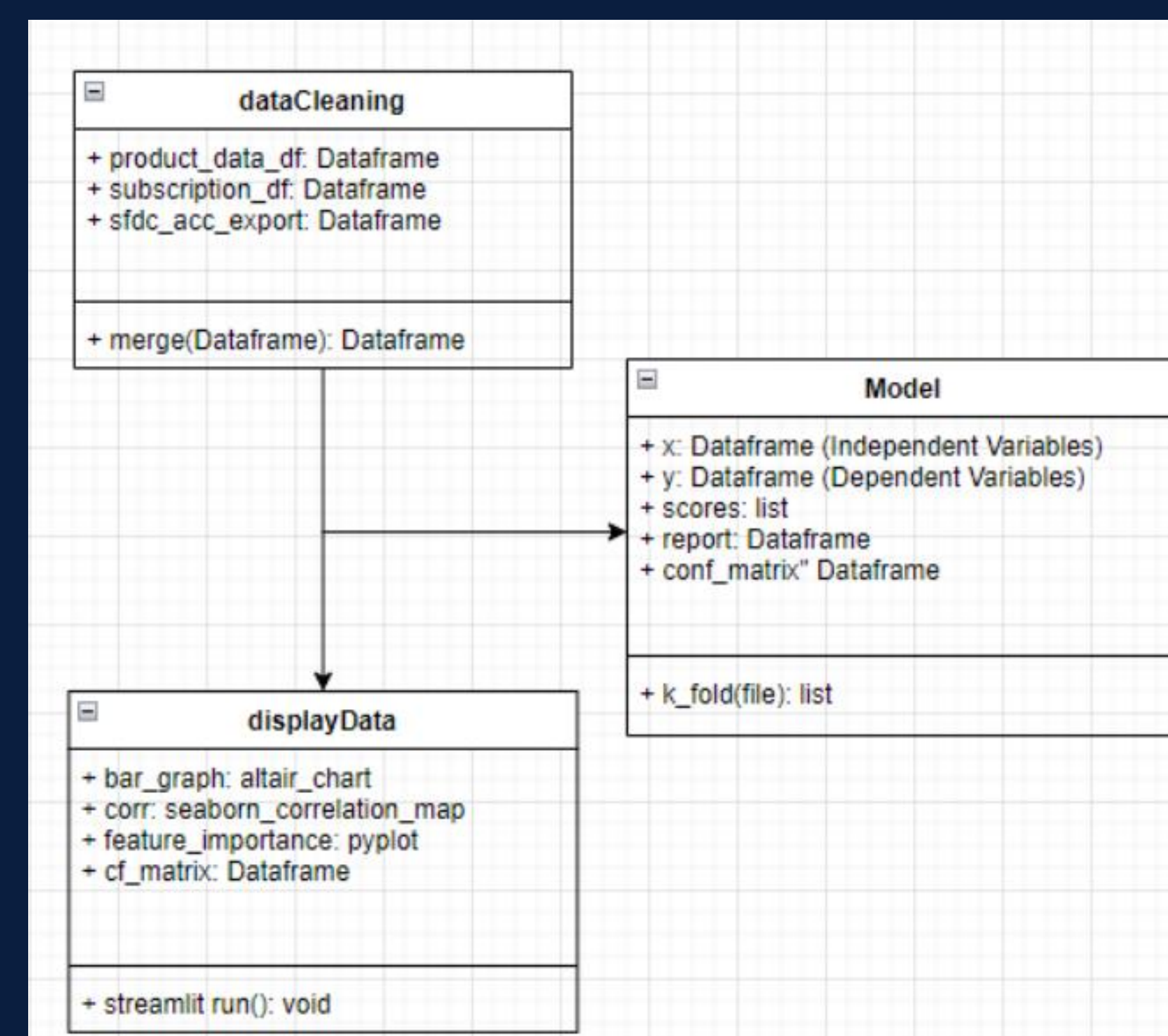
ACKNOWLEDGEMENT

The material presented in this poster is based upon the work done by (Capston II & I team members). Special thanks to our product owner and mentor, Lina Sokol, and our instructor Masoud Sadjadi.

CURRENT SYSTEM

A machine learning predictive model is needed to decrease the churn rate of the Kaseya clients and improve business. There is no current system to identify factors as in why client tend to end their services with the company, therefore this project's goal is to create an accurate model to predict and identify these factor mentioned before.

SYSTEM DESIGN



Screenshots



REQUIREMENTS

For the elaboration of the project was needed:

- Python 3.8.10
- Visual Studio Code
- Datasets from Kaseya
- Streamlit
- GitHub
- Trello

IMPLEMENTATION

- The data set given by Kaseya was cleaned and binarized.
- Data frame was cleaned, and features were added.
- Model was programmed in python
- Oversampling technique SMOTE was applied.
- K-Fold validation with Logistic regression was added for more balanced data.
- Data frame tables were added to stream lit
- Correlation heat maps were created for better visual display.
- Confusion Matrix was added to display predictability.

