Customer Revenue Prediction (CRP)

Sujan Barama

Thejas Kiran

Introduction & Motivation

- Identify high-value and loyal customers
- Optimize marketing strategy to allocate budget more accurately for customer outreach
- Guide product development to what the user requires
- Create customer segments for personalized campaigns with target sales

Dataset

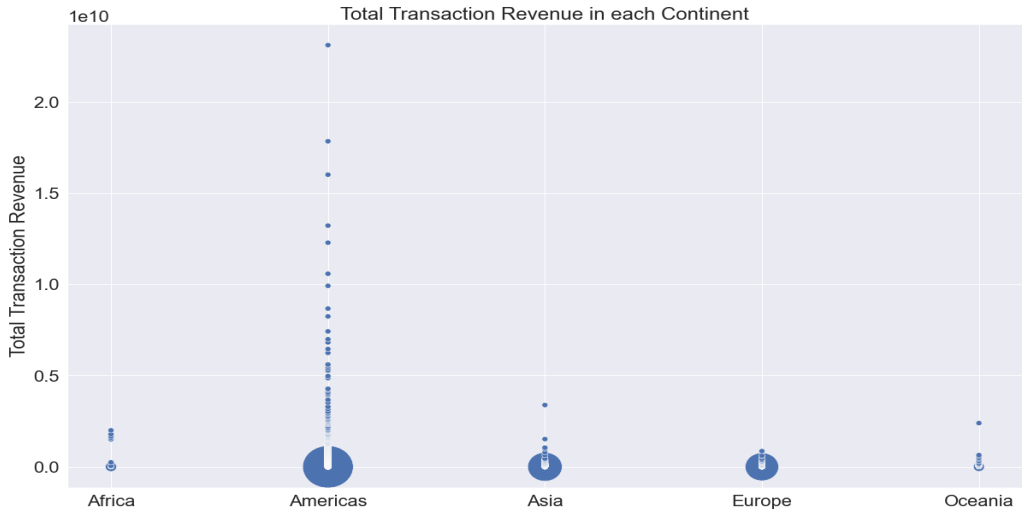
Google Analytics Customer Revenue Prediction (Kaggle)

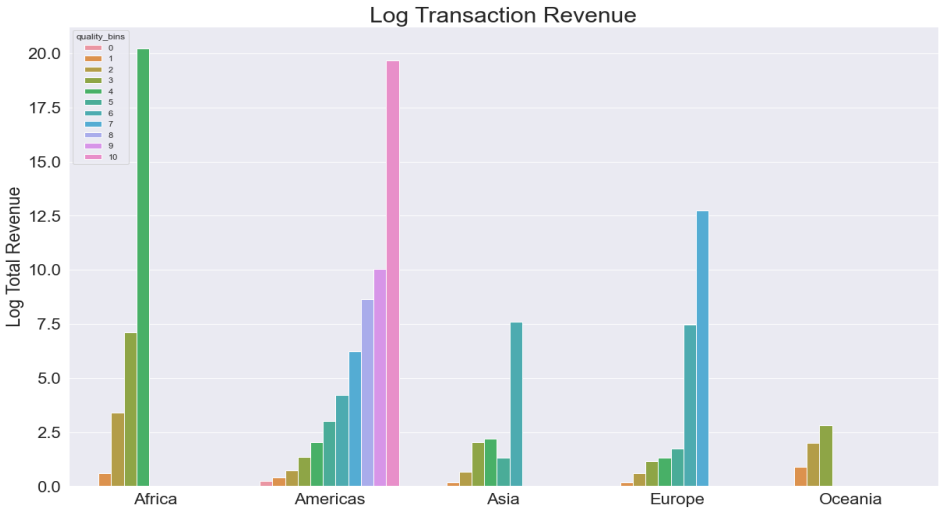
13 columns with a lot of them as JSON objects

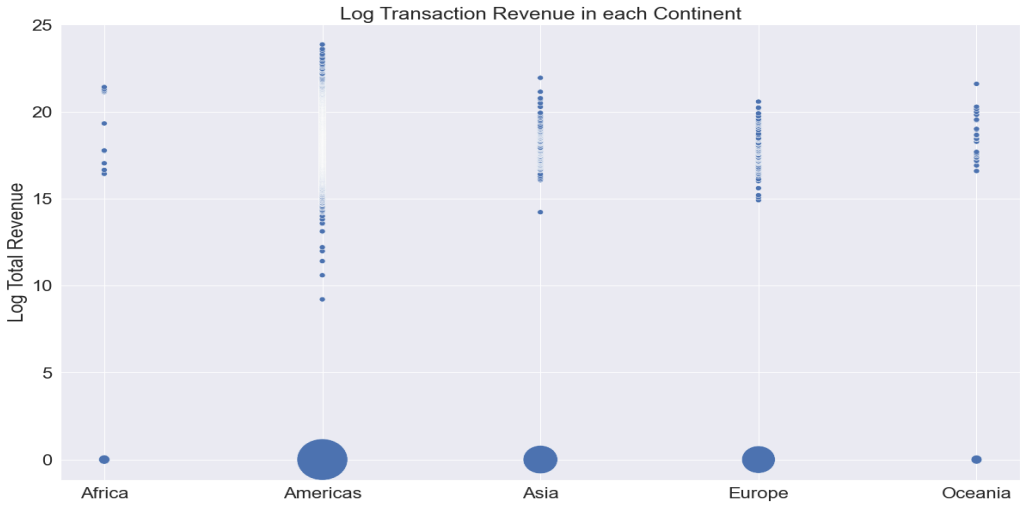
1.71 Million data points

Target is revenue

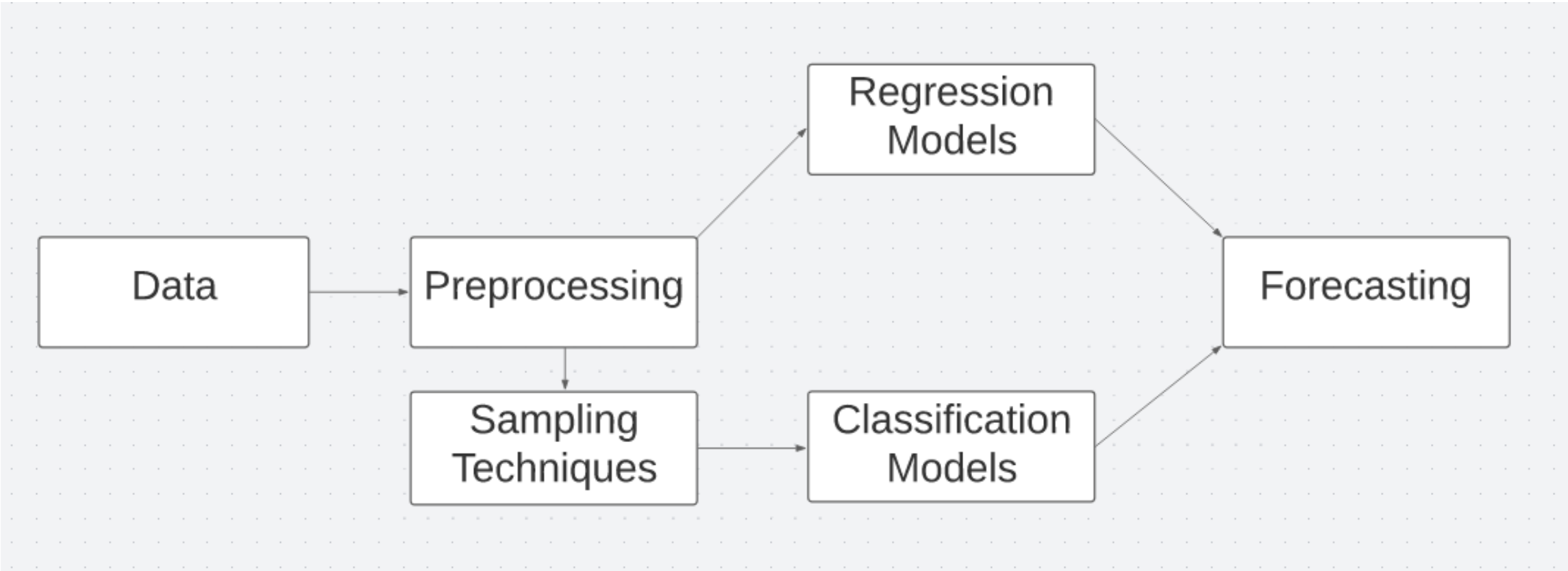
Data Understanding







Proposed Architecture



Modeling

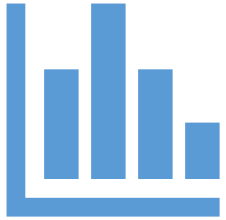
Classification Algorithms		
MODEL NAME	**SPECIFICITY**	**ACCURACY**
Logistic Regression	53.82	98.21
Decision Tree Classifier	81.15	96.78
XGBoost Classifier	**84.19**	97.15

Regression Algorithms	
MODEL NAME	**RMSE**
Linear Regression	1.708
Decision Tree Regressor	1.669
XGBoost Regressor	**1.623**

Final Results

	Linear Regression	**Decision Tree Regressor**	**XGBoost Regressor**
Logistic Regression	1.738	1.707	*1.621*
Decion Tree Classifier	1.696	1.666	**1.612**
XGBoost Classifier	1.690	1.659	**1.599**

Conclusion



Integrating the results of classification and regression was slightly beneficial

XGBoost regression gave the best results for all the classifiers

A custom accuracy metrics should have been used

Future Work

Build a custom accuracy metrics

Scale the project from customer revenue prediction to customer lifetime revenue prediction

Dealing with attenuation error

References

- Google Analytics Customer Revenue Prediction | Kaggle
- Default channel definitions - Analytics Help (google.com)
- Google Analytics Customer Revenue Prediction | Jessica Guo

Thank You
Questions?