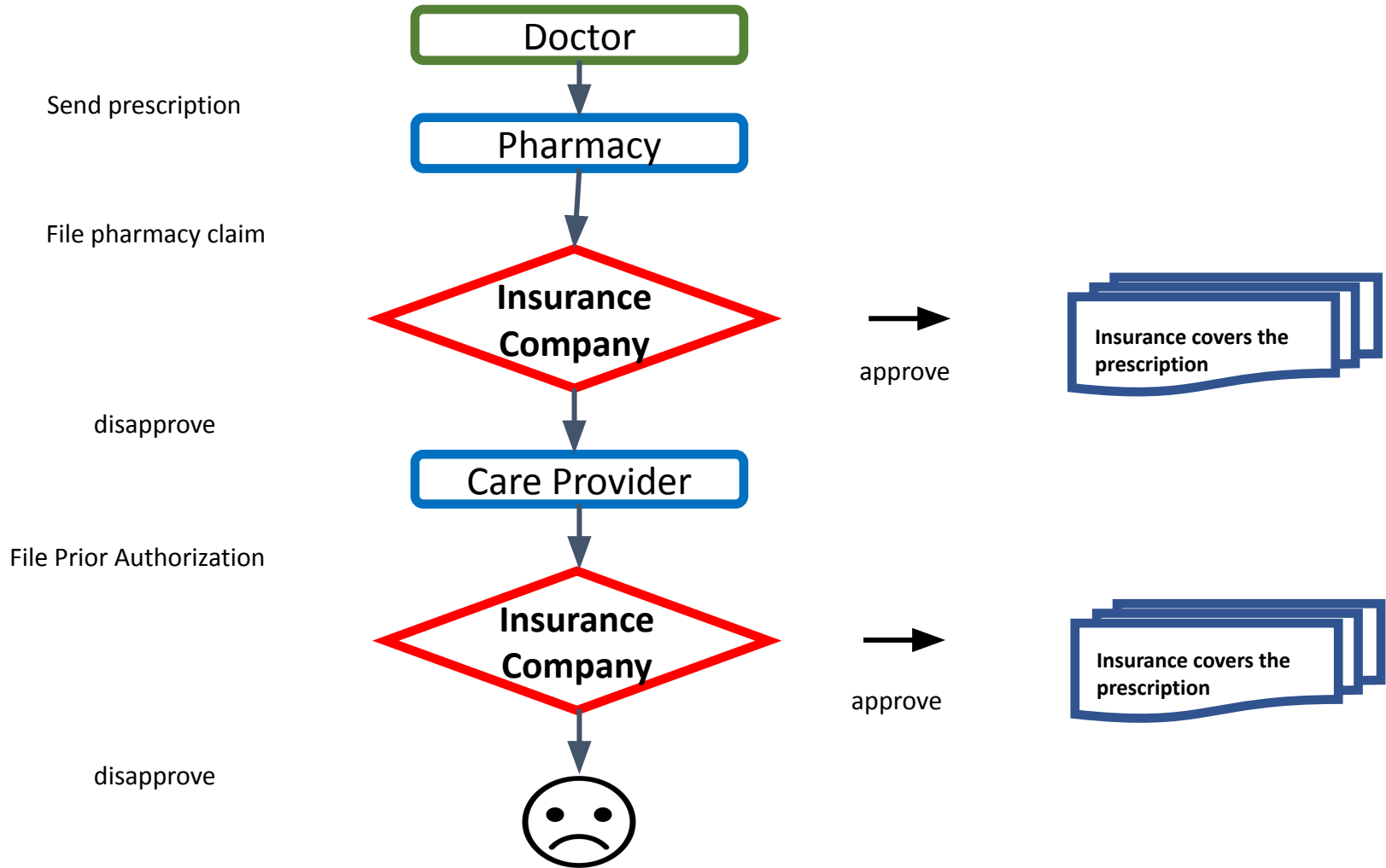


covermymeds®

Team Spruce : Jeeuhn Kim, Xichen Kong

Repo : [https://github.com/ XKong97/CoverMyMeds2022](https://github.com/XKong97/CoverMyMeds2022)

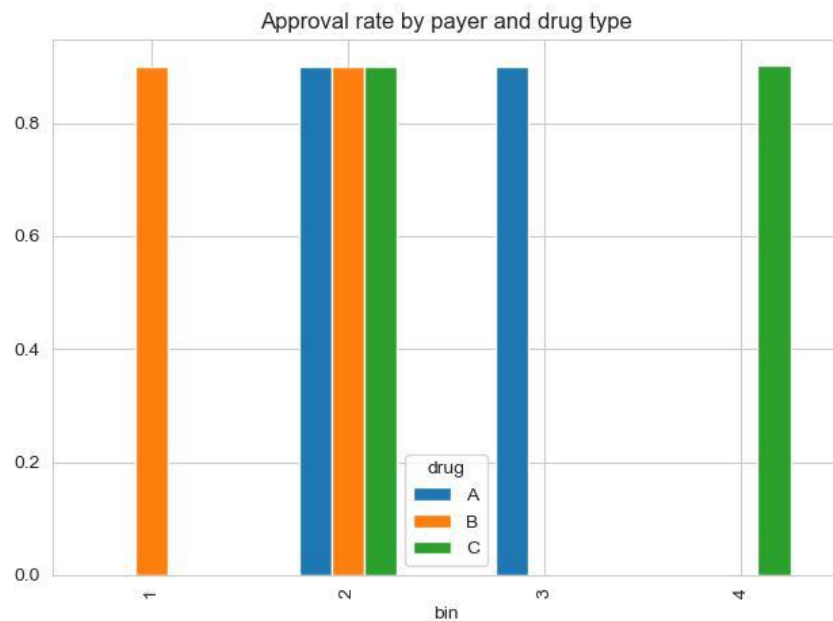
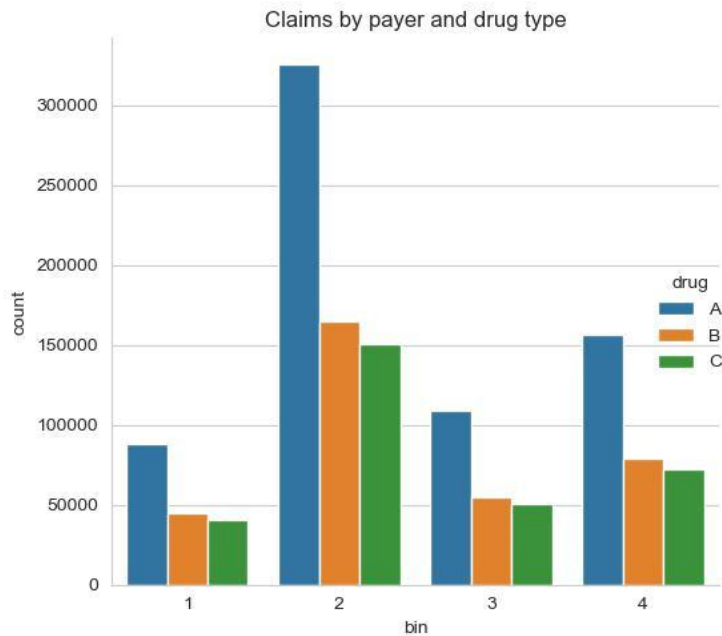


Summary of the Data set

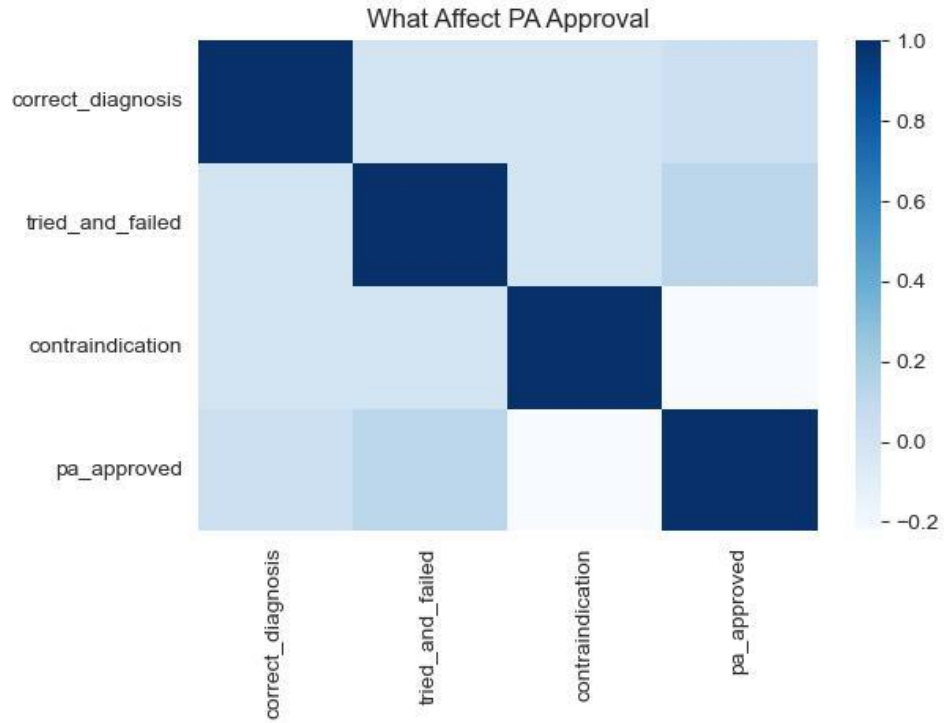
By **covermymeds**[®]

	The number of pharmacy claims
Approved	779625
Not Approved	555951
Total	1335576

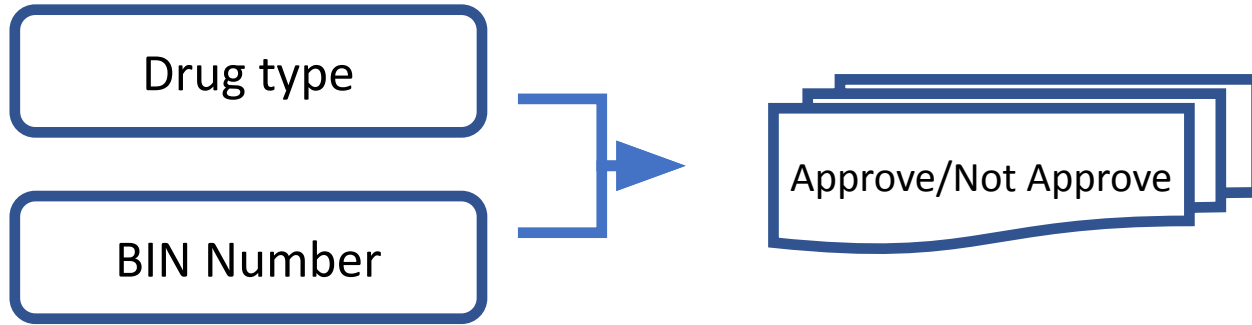
A look into the data



A look into the data - continued



Pharmacy claim approval prediction



	The number of pharmacy claims
Approved	779625
Not Approved	555951
Total	1335576

Key Performance Index

True Positive Rate

+

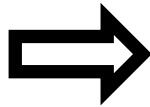
False Positive Rate

=

ROC-AUC

Results

	Accuracy Score	Precision	TPR	FPR	F1 Score	ROC-AUC
Baseline	0.584	0.584	1.000	0.100	0.737	0.500
Logistic	0.935	0.900	1.000	0.155	0.948	0.922
LDA	0.762	0.835	0.739	0.206	0.784	0.852
SVM	0.624	0.627	0.858	0.704	0.717	0.633
Decision Tree	0.935	0.900	1.000	0.155	0.948	0.922
Random Forest	0.935	0.900	1.000	0.155	0.948	0.922
Extra Trees	0.935	0.900	1.000	0.155	0.948	0.922
AdaBoost	0.935	0.900	1.000	0.155	0.948	0.922
XGBoost	0.935	0.900	1.000	0.155	0.948	0.922



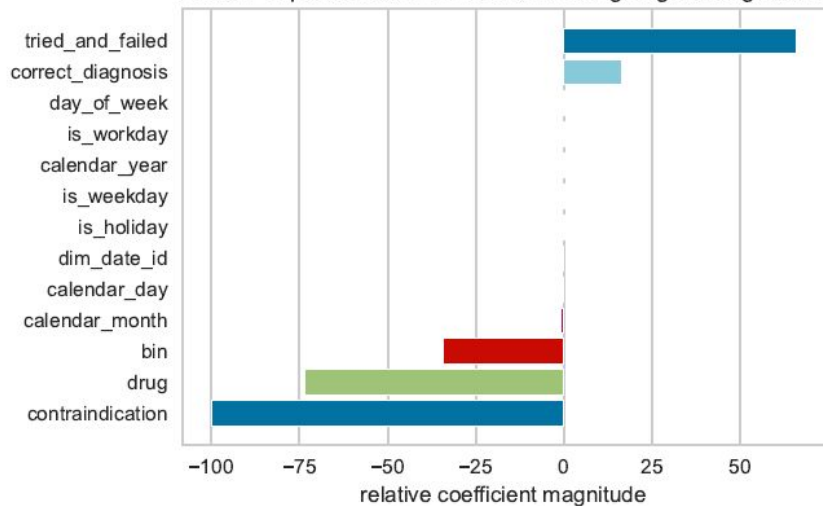
Logistic Regression or Decision Tree

Modelling: PA Approval

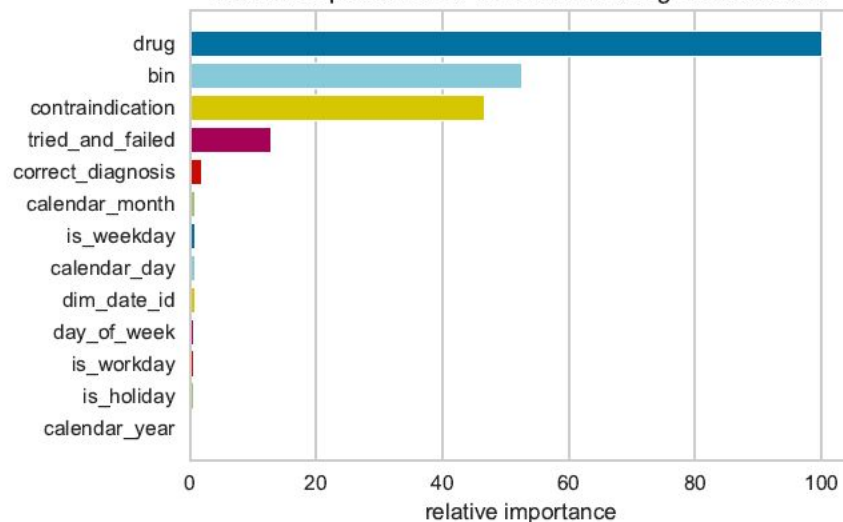
1. Read merged claim data and define X and y
2. Select features using all data
3. Split train and test data and obtain a random subset of train data for tuning
4. Separately tune hyperparameters and select best Logistic Regression and XGBoost model based on F1 score
5. Choose best threshold for each model using train data
6. Generate predictions using test data with selected models and thresholds
7. Compare performance based on metrics

Modelling: Feature Selection

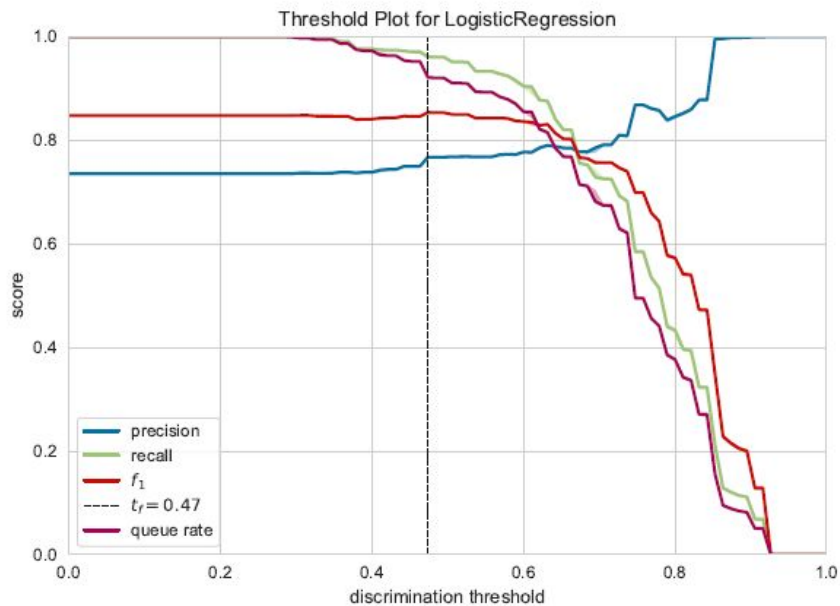
Feature Importances of 13 Features using LogisticRegression



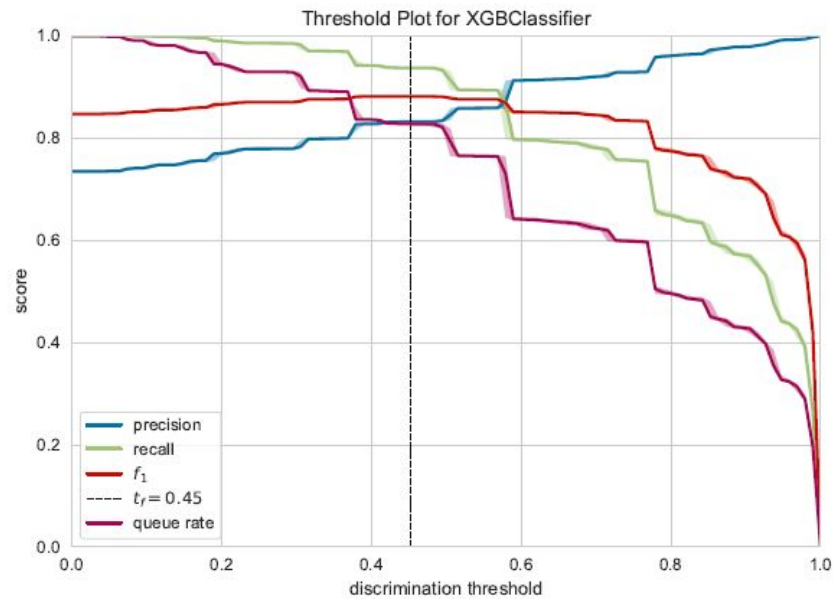
Feature Importances of 13 Features using XGBClassifier



Modelling: Choosing Threshold



Cutoff = 0.7



Cutoff = 0.75

Modelling: Compare Performance

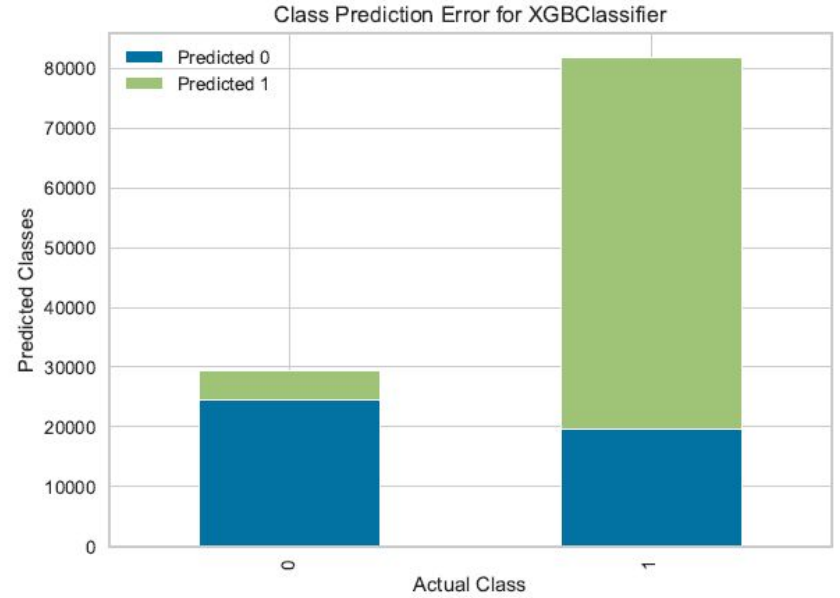
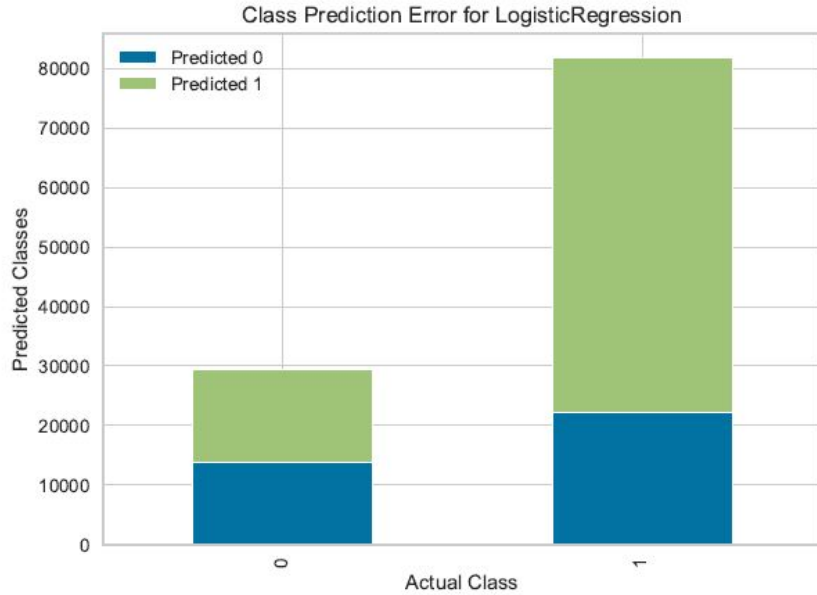
Confusion Matrix for Logistic Regression

True Negative 13759	False Positive 15768
False Negative 22286	True Positive 59378

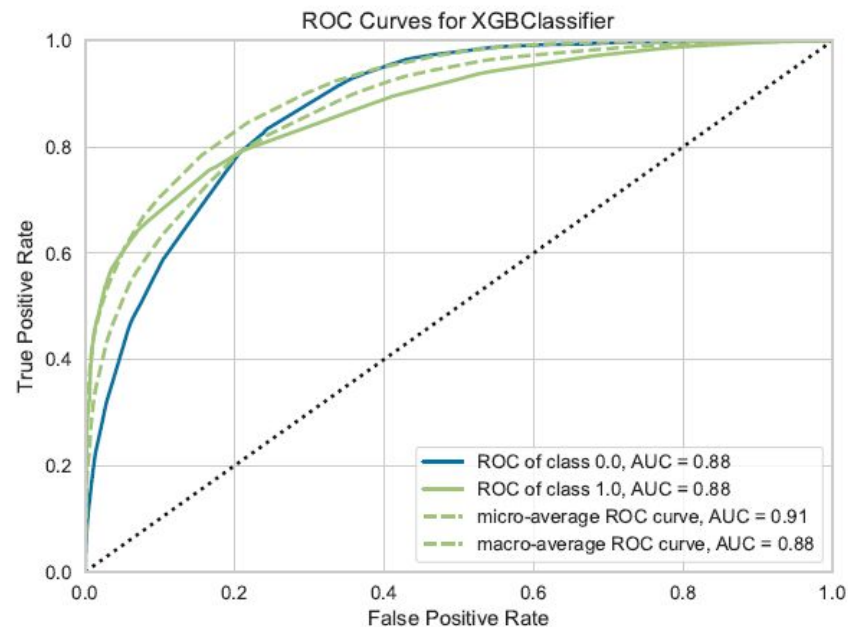
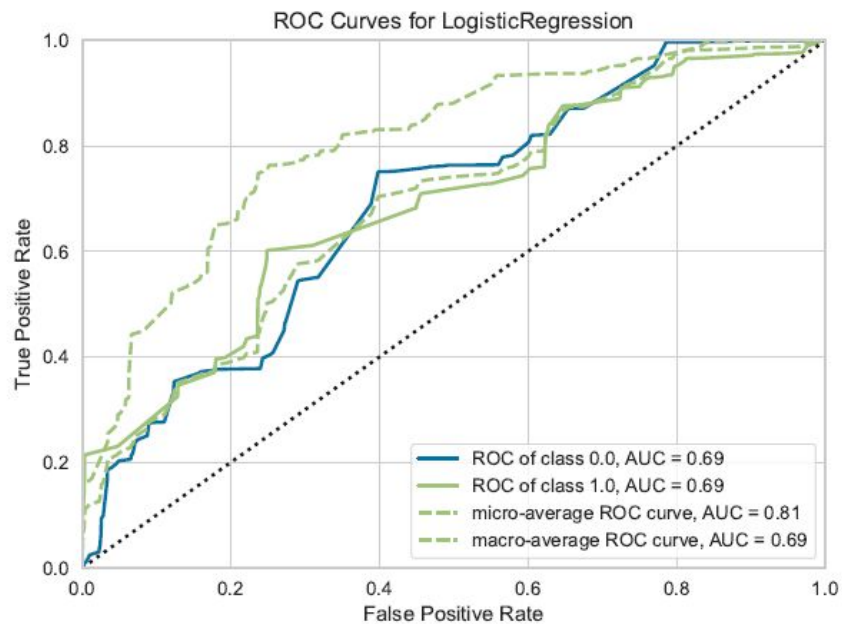
Confusion Matrix for XGBoost

True Negative 24570	False Positive 4957
False Negative 19760	True Positive 61904

Modelling: Compare Performance



Modelling: Compare Performance



Recommendations

- Encourage providers to provide information indicating that the patient has an associated contraindication for the medication requested wherever applicable
- Encourage providers to provide information indicating that the patient has tried and failed the relevant generic alternatives wherever applicable
- Understand the formulary for each payer

Thank you!