Google Merchandise Store Sales Prediction System

Objective

To build a prediction model to determine if new customers will transact from site or not

Assumptions

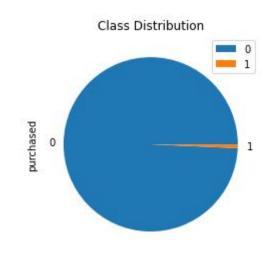
- New customer is anyone with 'Visit Number' as 1
- Anyone with 'purchased' as not null, will be considered having made transaction.
- Any bounced session will not have purchase made

Final Model

Following parameters in Decision Tree gave best results: random_state=42,criterion='entropy',max_feat ures='auto',splitter='best',class_weight={0:

2474, 1: 350782}

Data was highly imbalance, with positive class (people who made purchase) consisting of only 0.03 %.



Hence we used Synthetic Minority Over sampling technique, for improving performance.