



# **SPRINT 2 REPORT**

**BY MITCHELL MCCALLUM**



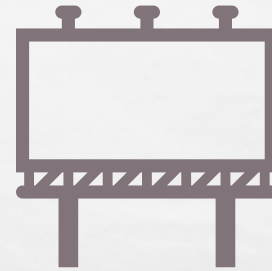


# OVERVIEW



Problem: US Telecoms company losing **4.16%** of customers over roughly the span of a year

Lost Customers on average are **14.39%** more profitable



Solution: Multiclass and multilevel Classification for telecom companies

Designed to enhance customer satisfaction and develop targeted retention strategies.

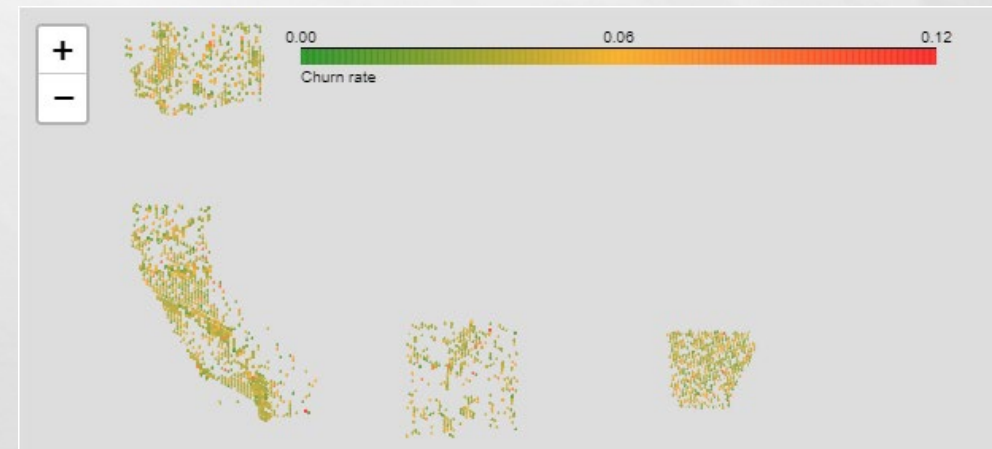
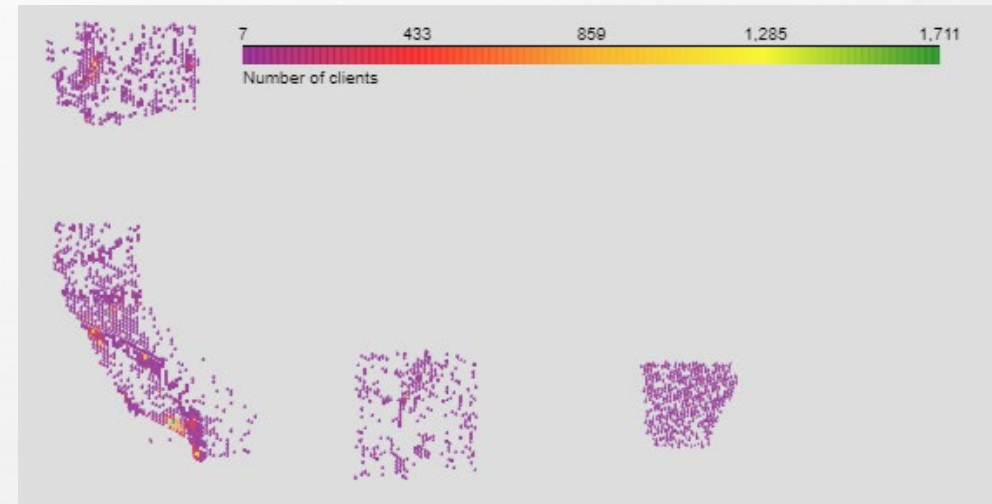
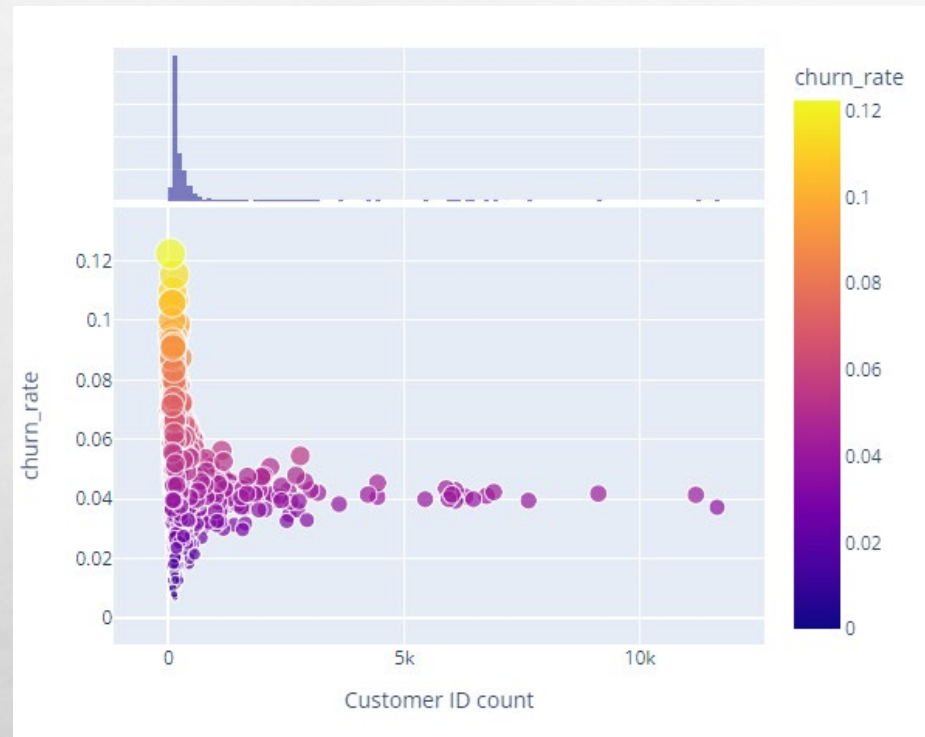
# DATA HANDLING

- IRRELEVANT FEATURES/IDS
- HANDLING BAD DATA IN TARGET DATA SET (CHURN REASON)
- PREPROCESSING
  - ENCODING TEXT VARIABLES
  - BALANCING CHURN CATEGORIES
    - CLASS UNDER SAMPLING
    - SYNTHETIC MINORITY OVER SAMPLING TECHNIQUE (SMOTE)



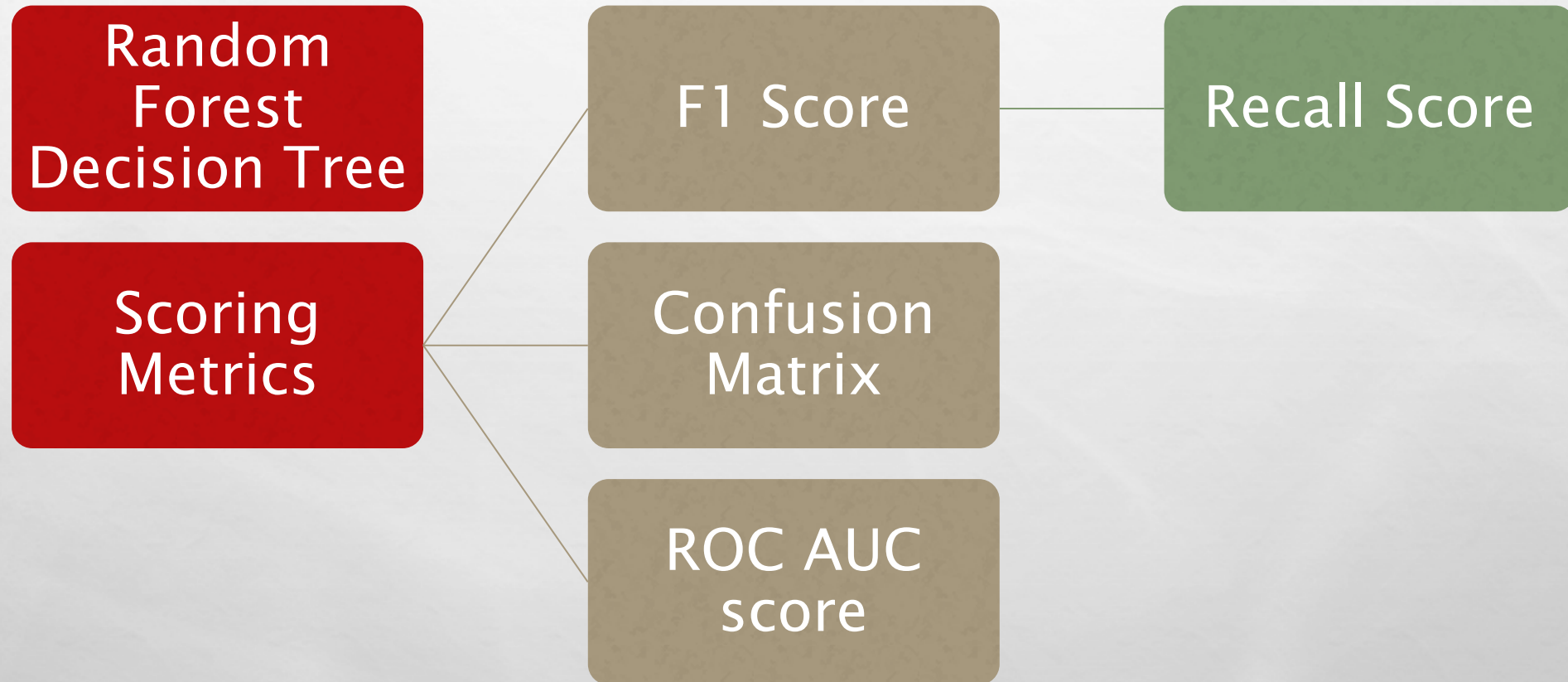


# CUSTOMER DISTRIBUTION AND CHURN RATE





# MODELS AND METRICS



# FUTURE WORK

- GRADIENT BOOSTING => XGBOOST
- HYPERPARAMETER TUNING
- DEEP NEURAL NETWORK
  - MULTILEVEL, MULTILABEL CLASSIFICATION
  - MULTILABEL BINARIZER
  - TRAIN A DEEPER MULTI-LAYER PERCEPTRON
- REFINEMENT OF DATA PREPROCESSING TO ENSURE NO LEAKAGE AND DRIFT

