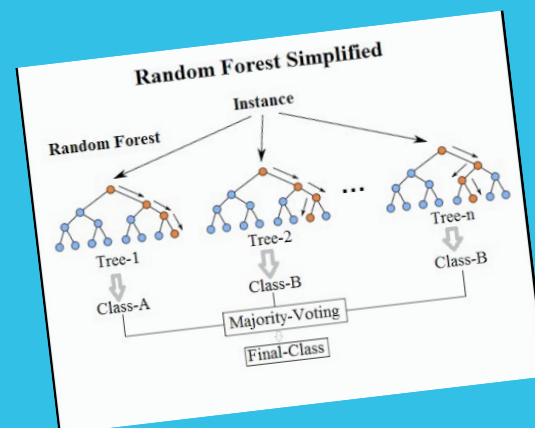
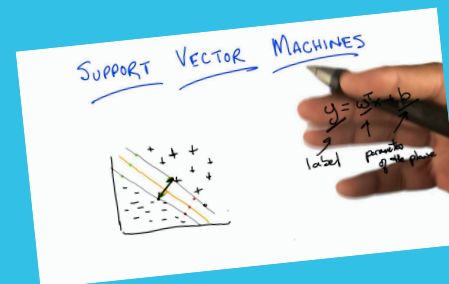
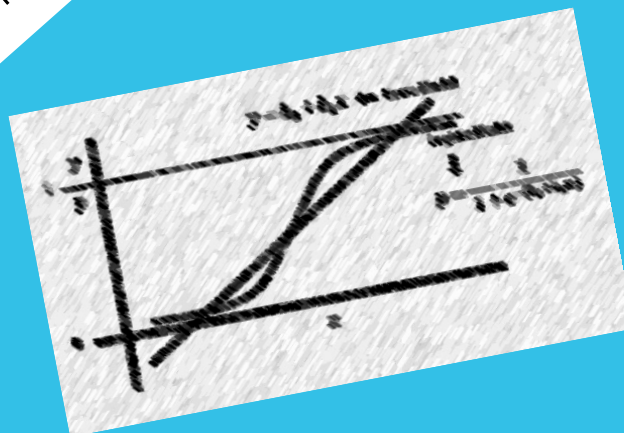


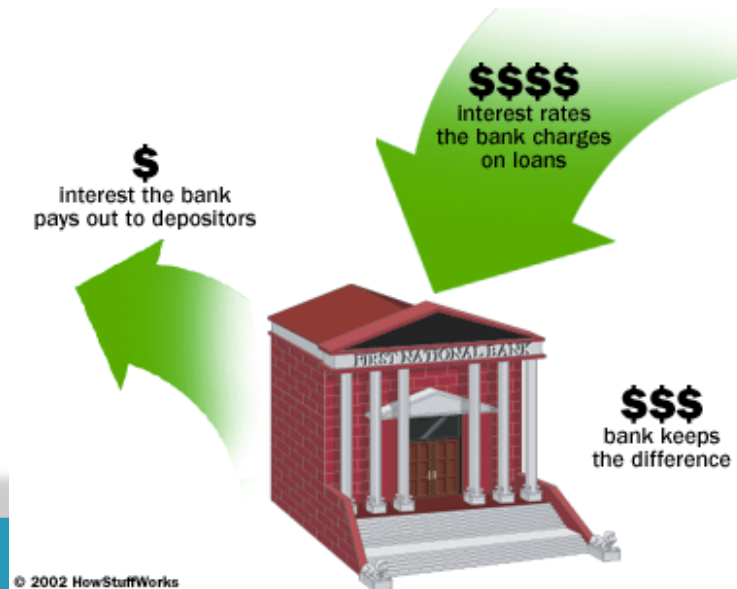
MAXIMIZING BANK'S MARKETING CAMPAIGN PROFITABILITY

THROUGH MACHINE LEARNING



PROBLEM BACKGROUND

- Portuguese retail bank marketing long-term deposit offer to existing customers
- Long-term deposits
 - Fixed investment term, usually 1 to 5 years
 - Safe investments
 - Very appealing to conservative, low-risk investors
- Marketing channel used – Telemarketing



DATA AT A GLANCE

- **Data Source**
 - Publicly available on UCI website
 - CSV format
 - Data from external sources has been used during cost-benefit analysis
- **Data**
 - Data collected is from May 2008 to Nov 2010
 - 21 attributes and 41188 observations
 - 20 independent variables – client data, call data, socio-economic factors and campaign data



GOAL

- Prediction about customers that are most likely to accept term deposit offer
- Metric – Campaign profitability
- End goal is to Maximize the metric
- Test multiple machine learning algorithms and shortlist the one with highest Profitability




EXPLORATORY DATA ANALYSIS

- Chi-square
- Data Normalization
- Multi-collinearity among macroeconomic factors
 - Principal Component Analysis
 - Eigen values and scree plot



MACHINE LEARNING MODELS

- **Logistic Regression**
 - Performance on original data
 - Data rebalancing
 - Regularization
 - Model Evaluation
 - **Random Forest Classifier**
 - Regularization
 - Model Evaluation
 - **Support Vector Machine**
 - Regularization
 - Model Evaluation
- 

EVALUATION METRICS

- Profitability
 - Cost-Benefit Analysis
 - Profitability= $\$43 * (\text{True Positives}) - \$11 * (\text{True Positives} + \text{False Positives})$
- ROC AUC

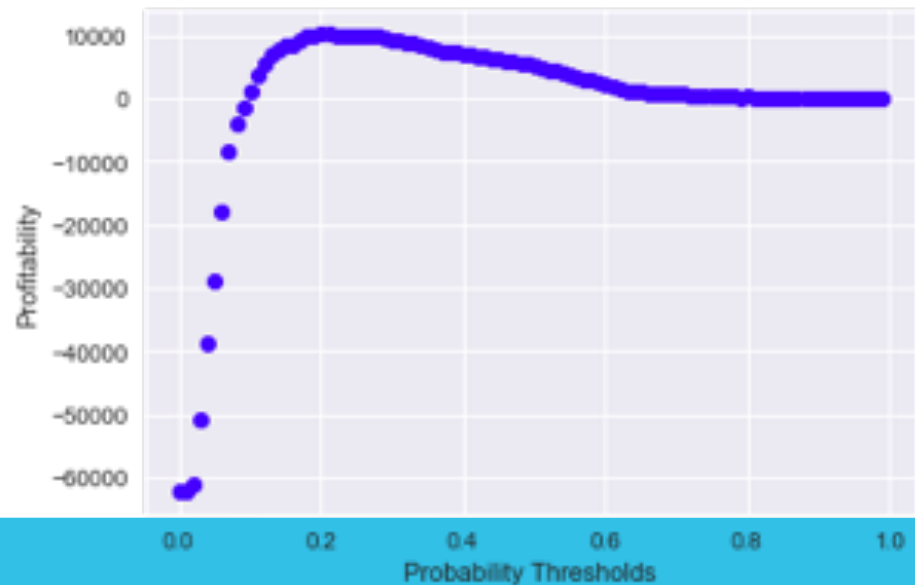
BASELINE PERFORMANCE

- Conversion rate for this campaign - 11.2%
- Out of the total 41188 customers contacted,
 - 4613 accepted offer
 - 36575 rejected offer
- Profitability = $\$43(\text{Total converts}) - \$11(\text{Total customers contacted})$
= $\$43(4613) - \$11(41188)$
= $\$1,98,359 - \$4,53,068$
= $(-\$2,54,709)$
- Overall, this campaign made a loss of \$2,54,709
- Goal is not only to obtain profits but also to develop a model that can maximize it

LOGISTIC REGRESSION

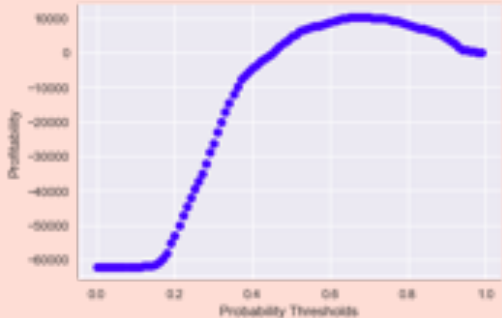
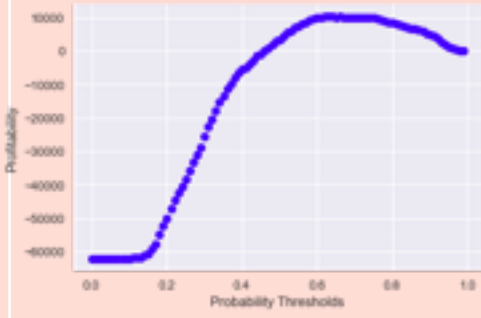
ORIGINAL IMBALANCED DATA

- Threshold at which profitability is the highest is: 0.21
- Regularization Parameter C : 100
- Maximum achievable profitability with the model is: \$ 10,159
- ROC AUC: 0.7746



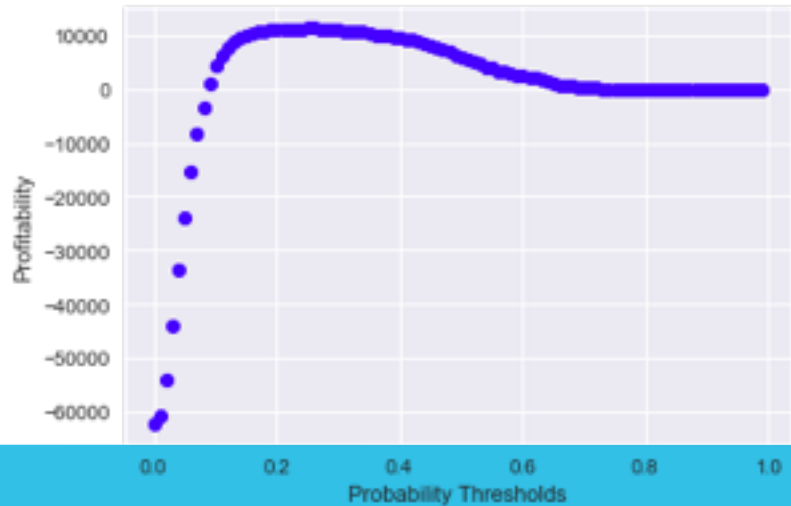
LOGISTIC REGRESSION

RESAMPLED DATA

	Upsampled Data	Downsampled Data
Final Regularization parameters	$C = 100$	$C = 100$
Threshold	0.67	0.63
ROC AUC	0.7783	0.7792
Profitability	\$ 10,157	\$ 10,114
Profitability at varying thresholds		

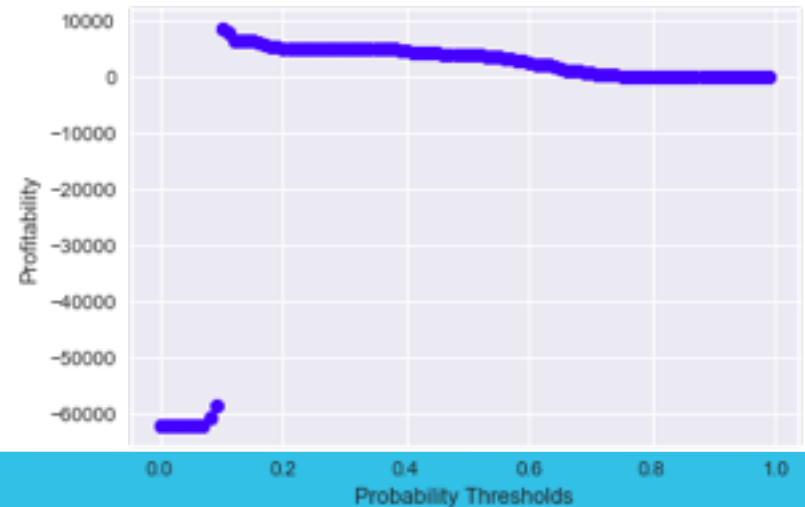
RANDOM FOREST CLASSIFIER

- Threshold at which profitability is the highest is: 0.26
- Regularization Parameters
 - min_samples_leaf = 5
 - min_samples_split = 10
 - n_estimators = 50
- Maximum achievable profitability with the model is: \$ 11,359
- ROC AUC: 0.7955



SUPPORT VECTOR MACHINE

- Threshold at which profitability is the highest is:
- Regularization Parameters
 - Kernel= RBF
 - $C = 0.1$
 - $\gamma = \text{auto}$
- Maximum achievable profitability with the model is: \$ 8,536
- ROC AUC: 0.7095



CONCLUSION

	Profitability	ROC AUC
Logistic Regression	\$10,159	0.7746
Random Forest Classifier	\$11,359	0.7955
Support Vector Machines	\$8,536	0.7095

- Random Forest Classifier
- Provided an increase in profitability by 104.5 % over baseline model
- Campaign profits of \$ 11,359 with this model

FUTURE STUDY

- Extend study to address
 - Resource allocation
 - Manpower planning