Project: User Churn Prediction

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions and Recommendations

Business Problem

- A telecommunications company provides telephone service to its customers.
- They would like to investigate the business insights about customer churn and the factors leading to these actions.
- We work on this project to provide insights to this telecommunications company

Business Problem

Business insights to investigate:

- Predict if a customer would soon churn and stop the service
- What factors or any predictable patterns are affecting the customer churn actions ?

Data

Source:

- Dataset of SyriaTel Telecommunications Company through Kaggle Competition

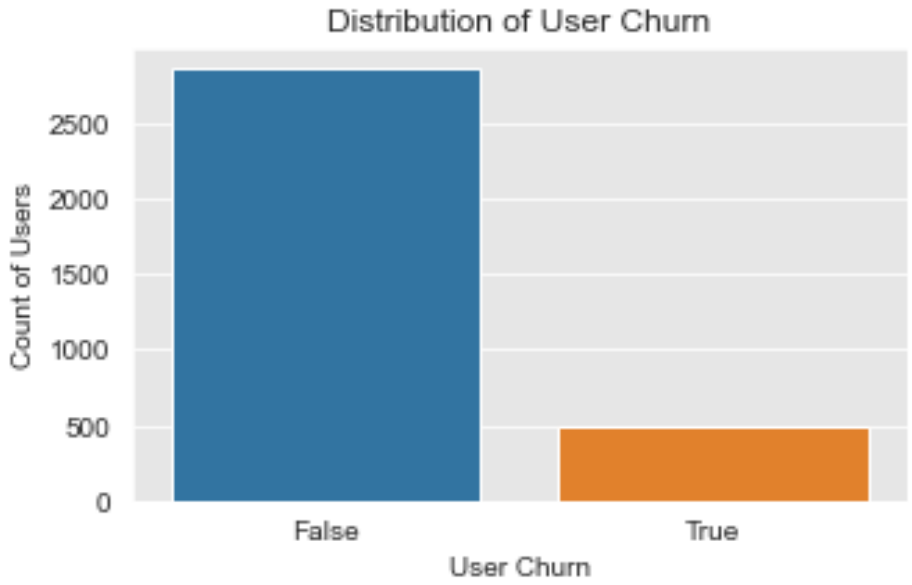
Information:

- Classification Labels: Churn (Yes, No)
- Voice Mail Plan, International Plan
- Total Day Minutes, Total Day Calls
- Total Eve Minutes, Total Eve Calls
- Total Intl Minutes, Total Intl Charges
- Customer Service Calls
- Number Vmail Messages

Methods

- Classification Approach
- Train Classification Model for customer churn prediction
- Investigate Feature importance of the predictive model

Results



User Churn Labels:
False: 2850
True : 483

Key Features:

- ‘Customer Service Calls’, ‘Total Day Minutes’, ‘Total Day Calls’
- ‘International Plan’, ‘ Total Eve Minutes’

Results – Modeling

Model : Classification Models

- Base model :
  - Linear Model, logistic Regression
- Second Model :
  - Non-Linear Model, Random Forest
- Final Model:
  - Random Forest, with tuned parameters

- Model improvement with accuracy

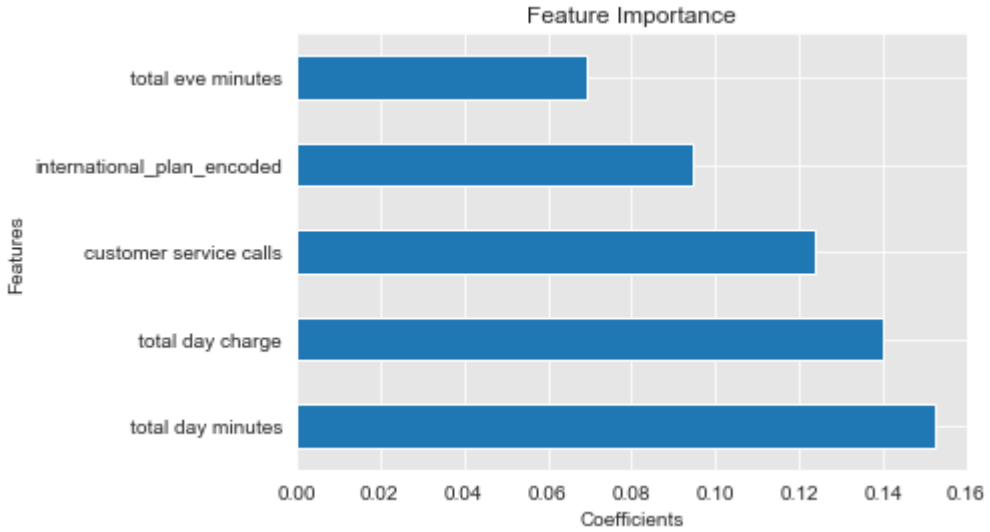
Results

	User Churn	Prediction
1994	0	0
3078	0	0
1551	1	0
52	0	0
307	1	1
...	...	...
2458	0	0
3264	0	0
1521	0	0
2595	1	1
1931	0	0

Model Performance

- Accuracy : >0.9
- Precision/Recall : 0.95, 0.86
- ROC Aera under Curve : 0.91

Results




```
Feature importance ranking:
1. feature 2 total day minutes: 0.152388
2. feature 4 total day charge: 0.140177
3. feature 14 customer service calls: 0.123639
4. feature 15 international_plan_encoded: 0.094776
5. feature 5 total eve minutes: 0.069128
6. feature 7 total eve charge: 0.061867
7. feature 12 total intl calls: 0.047132
8. feature 13 total intl charge: 0.043996
9. feature 11 total intl minutes: 0.039468
10. feature 8 total night minutes: 0.038498
11. feature 10 total night charge: 0.034284
12. feature 9 total night calls: 0.029350
13. feature 1 number vmail messages: 0.027603
14. feature 3 total day calls: 0.026368
15. feature 0 account length: 0.026201
16. feature 6 total eve calls: 0.025031
17. feature 16 voice_mail_plan_encoded: 0.020095
```

Conclusions

Classification models for User Churn Prediction

- Model prediction for binary classification outcome
- Relevant features affecting user decisions

Recommendations

- Telecommunications company : Use model to identify the users of high probability of churning
- Improve customer service in the areas of important features

Next Steps

- Generality of Model Performance
  - Limitation : small dataset
- Feature Engineering for Model Improvement
- More data

Thank You!

Email: aliang@gmail.com
GitHub: @username