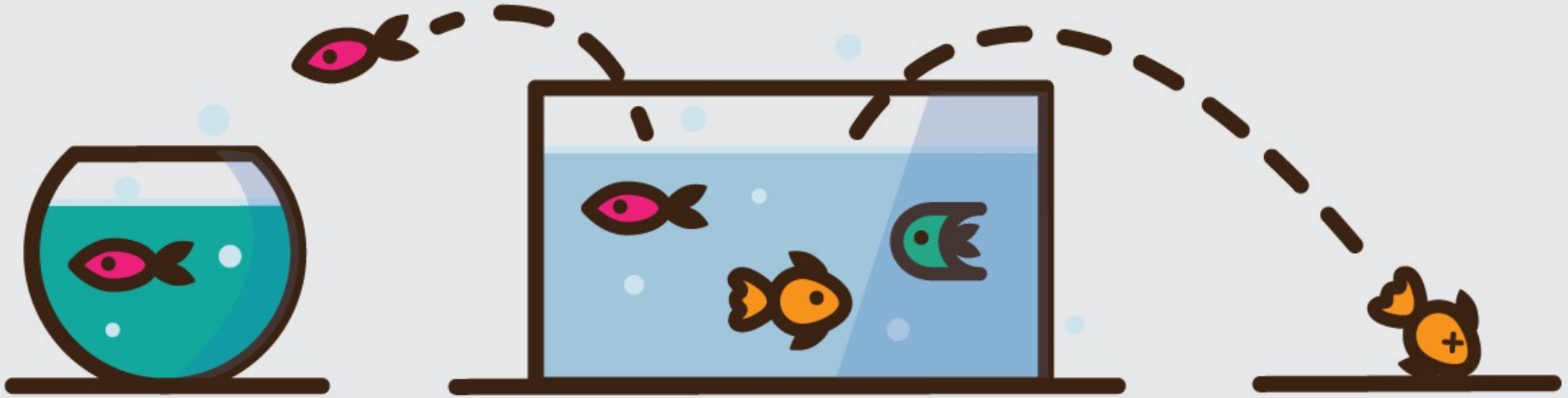


Telcom Customer Churn Analysis

Colin Shu – 7/29/2019

How bad is churn?

- Churn analysis is **vital** to a wide range of companies, such as communication companies, newspaper companies, and dotcom companies.
- The major reason is that for these companies highly rely on subscription services, **the extra cost to gain new customers** are often **higher** than **the maintenance costs to retain existing ones.**
- Companies make prediction of the customers churn activity and **refine their business plan** to retain their customers before they leave.

How bad is churn?

Source: https://thekinigroup.com/customer-churn-rate/

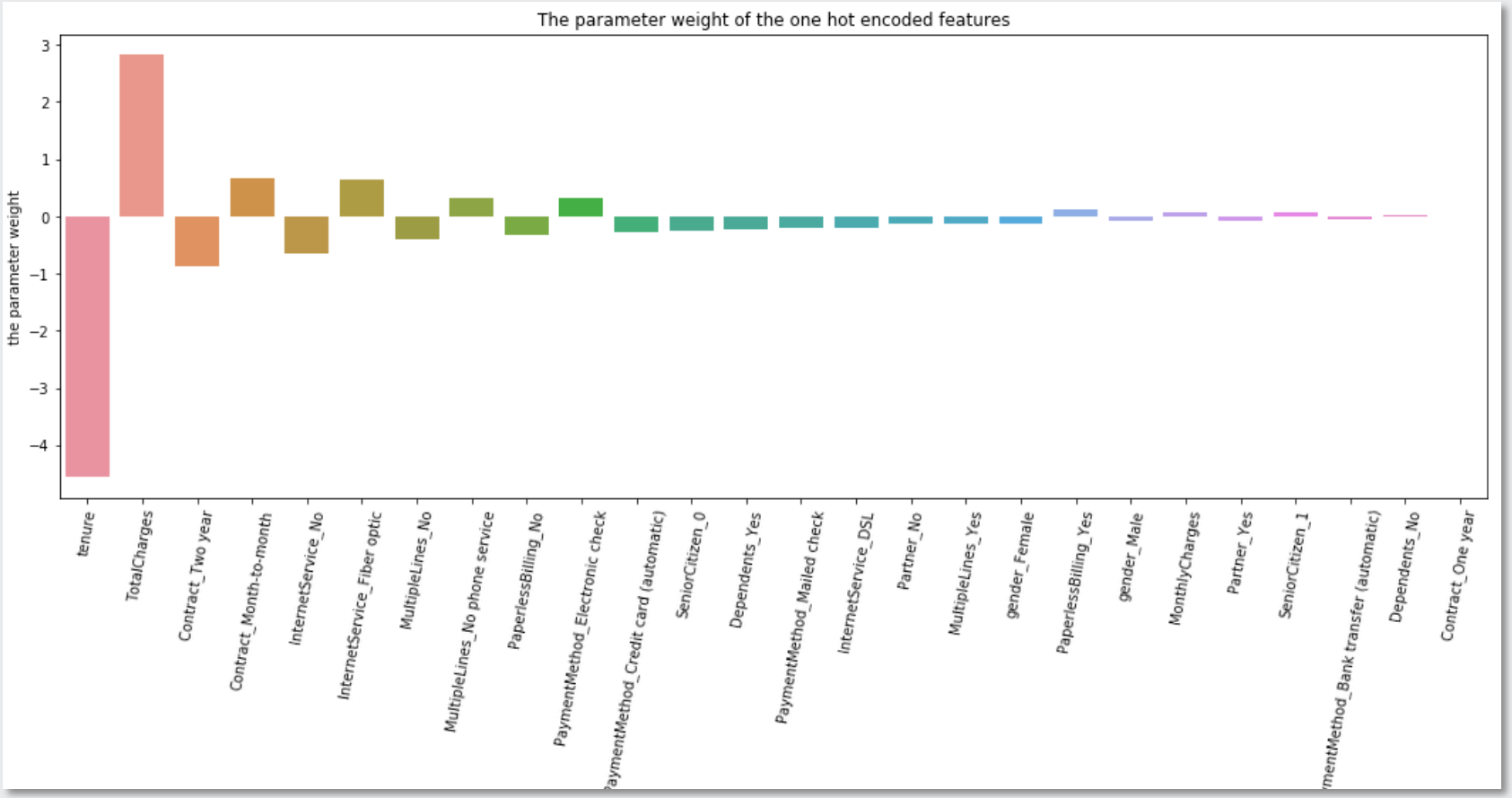
How to predict a churn?

- Since the number of customers and their data has been growing rapidly, churn analysis is also becoming one of the big data use case in business.
- Churn analysis is usually defined as a **classification** problem, requiring multiples features to predict the churn activity (**churn or not**) of customers.
- Multiple **machine learning model** can be used to classify these binary results, such as Logistic regression, Random Forest model, XGBoost, Support Vector Machine, K-Nearest Neighbors, etc.

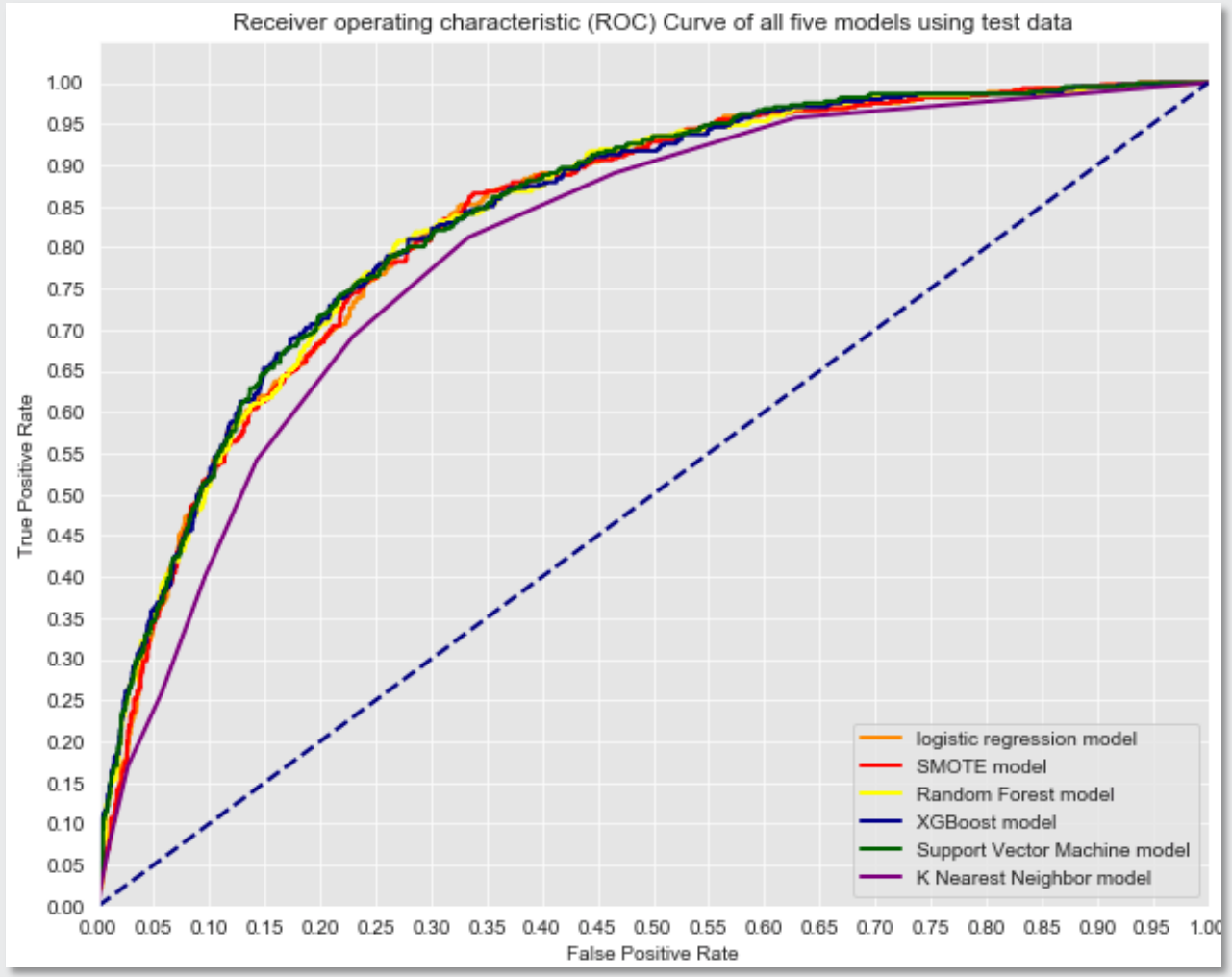
A brief roadmap of the model prediction

- Data Source: data set from **Kaggle** website (https://www.kaggle.com/blastchar/telco-customer-churn)
- Target: Customers who left **within the last month** – the column is called Churn
- Features: 20 features in total, after feature selection, **12** features were chose.
- Set baseline: After split the training and testing data set, by calculating the ratio of churn to the whole testing data size, the baseline is set as **0.254**.
- Models: The **classification models** that were used include: Logistic regression, Random Forest model, XGBoost, Support Vector Machine, K-Nearest Neighbors, etc.
- Model evaluation: Multiple **metric scoring methods** were used: precision, recall, specificity, accuracy, F1 score, the AUC value.

The results of the models



The results of the models

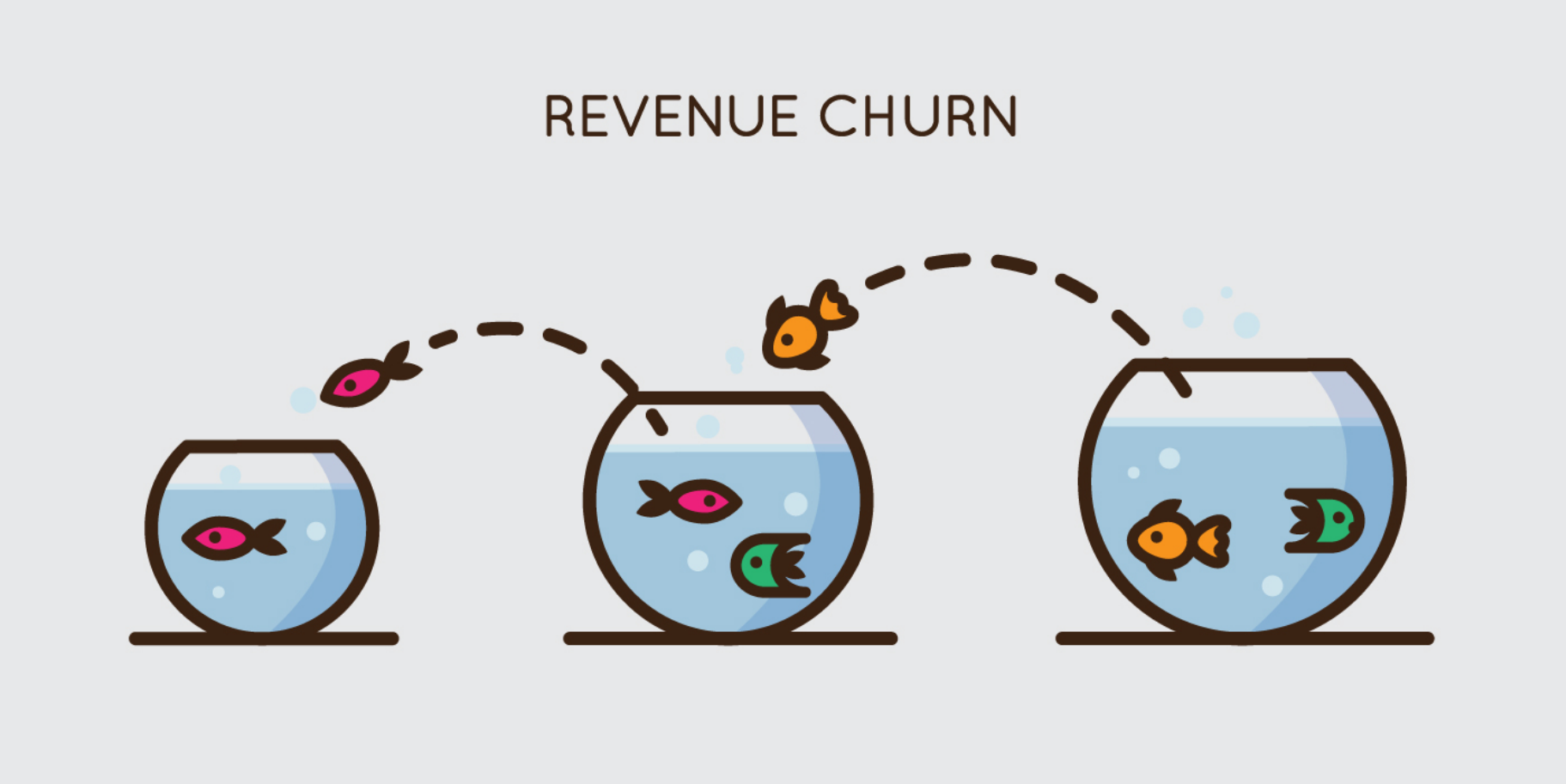


The results of the models

Future work

- This analysis demonstrates a good churn detection results. For future work, we could look at the churn customers deeper and perform the clustering on the churn customers to find out what kind of customer might be churn lately. Furthermore, we could look at the false positive and false negative prediction and perform the **clustering** on them to look at what kind of customers might be hard to predict the churn activity.
- After talking to marketing team, we can discuss how to retain the customers. For example, if we would like to send emails or give coupons. We can includes the cost and predicted profit to calculate the **expected values**.
- Before any marketing practice, our team can perform the **AB test** to see whether **the new solution is worthy**.

Thank you for your listening!



Pic source: https://www.livechatinc.com/blog/churn-rate/