MKT 869

Thanks Emmerson, we already have first model, it is not perfect, I am going to change some **feature engineering steps** to improve the result. There are three sections, first is improvement direction, second is compare the different versions to find the best one. Third is breakdown the key relationship.

Improvement direction

The **predicted positive** churn group is our target customer. Since the company's budget of reducing customer churn is limited, we will focus on improving the precision rate. **It makes the cost-effective.**

Improve model

Version 1, this version we change the outliers’ step. we use **Winsorization to handle** the outliers instead of removing them from dataset. But the result is not better than original model. So, we give up this.

Version 2, we directly remove the outliers and make a feature scaling. **This movement is significantly improving 10% precision rate.** So, we keep this change

Version 3: this version we change the missing value step. we used regression **imputation** to fill in the missing values instead of mean imputation. The result slightly better than version 2, **version 3 of the model performs the best.**

Key relationship.

Let’s breaking down the key relationship in our model.

Age has highest **Exponentiated Beta Coefficients**, it means that **older** customers are **more likely** to churn. The second highest is female, it means being female leads to 70% higher odds of churning.

There are some strong factors that lead to reduced churn. For example, a customer in France or Spain is **reducing the churn rate by** more than 50%. When a customer is an active member, it reduces the churn rate by more than 65%. Other features have a slight impact on the churn rate.