Telecom Machine Learning

Nirmal Patel

November 13, 2018

Machine Learning Questions

How do you frame your main question as a machine learning problem?

My telecom churn analysis question can be frames as a machine learning question such as which factors are most influencial in customer churning. This would allow me to run machine learning models such as logistic regression as a binomial to figure out if a customer churns or not.

Is it a supervised or unsupervised problem?

The dataset I have is more of a supervised problem because the data is given in a well structured with varous input and output variables. in this case the output variable we are testing it the churning of a customer.

If it is supervised, is it a regression or a classification?

My problem is a regression problem due to churning being a binary variable with two factors such as yes or no. Though if we wanted to figure out which contract type a person would be more likely to choose then it would be classification.

What are the main features (also called independent variables or predictors) that you'll use?

Some of the independent variables include gender, SeniorCitizen, tenure, MultipleLines, InternetService, Contract, PaperlessBilling, PaymentMethod, and MonthlyCharges. The Dependent Variable is Churn of the customer.

Which machine learning technique will you use?

I will use logistic learning for figuring out churning of the customers.

How will you evaluate the success of your machine learning technique? What metric will you use?

The success of my model will be evaluated by the accuracy of logistic regression.

Machine Learning Models

```
churnmodel1 <-
glm(Churn~gender+SeniorCitizen+tenure+MonthlyCharges+MultipleLines+InternetSe
rvice+Contract+PaperlessBilling+PaymentMethod,data=telecom,
family="binomial")
plot(allEffects(churnmodel1))
```