



Multivariate Logistic Regression - Telecom Churn Example

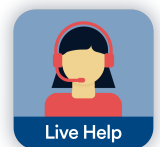
Let's now look at the process of building a logistic regression model in Python.

You will be looking at the **telecom churn prediction** example. You will use 21 variables related to customer behaviour (such as the monthly bill, internet usage etc.) to predict whether a particular customer will switch to another telecom provider or not (i.e. churn or not).

Problem Statment

You have a telecom firm which has collected data of all its customers. The main types of attributes are:

- Demographics (age, gender etc.)
- Services availed (internet packs purchased, special offers taken etc.)
- Expenses (amount of recharge done per month etc.)





means that the customer has churned and 0 means the customer has not churned.

Please find the churn dataset [here](#).

Please find the internet_data dataset [here](#).

Please find the customer_data dataset [here](#).

Please find the data dictionary [here](#)

Please find the Logistic Regression code file [here](#)

So, here's what the data frame churn_data looks like:

	customerID	tenure	PhoneService	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
0	7590-VHVEG	1	No	Month-to-month	Yes	Electronic check	29.85	29.85	No
1	5575-GNVDE	34	Yes	One year	No	Mailed check	56.95	1889.5	No
2	3668-QPYBK	2	Yes	Month-to-month	Yes	Mailed check	53.85	108.15	Yes
3	7795-CFOCW	45	No	One year	No	Bank transfer (automatic)	42.30	1840.75	No
4	9237-HQITU	2	Yes	Month-to-month	Yes	Electronic check	70.70	151.65	Yes

Data Frame 1: churn_data

Also, here's the data frame customer_data:

	customerID	gender	SeniorCitizen	Partner	Dependents
0	7590-VHVEG	Female	0	Yes	No
1	5575-GNVDE	Male	0	No	No
2	3668-QPYBK	Male	0	No	No
3	7795-CFOCW	Male	0	No	No
4	9237-HQITU	Female	0	No	No

Data Frame 2: customer_data

Lastly, here's the data frame internet_data:



Now, as you can clearly see, the first 5 customer IDs are exactly the same for each of these data frames. Hence, using the column customer ID, you can collate or merge the data into a single data frame. We'll start with that in the next segment.

Coming Up

In the next segment, you will start with reading and inspecting the dataframes and then move on to preparing that data for model building.

[Report an error](#)

PREVIOUS
Introduction

NEXT

Data Cleaning and
Preparation - I

