**Dataset:** *Telecom.csv*

**Notebook:** *sagemaker-spark.ipynb*

**Description:**  It consists of cleaned customer activity data (features), along with a churn label specifying whether a customer canceled the subscription, will be used to develop predictive models.

**Attributes:**

1. *CustomerID:* Customer ID
2. *Gender:* male/female
3. *SeniorCitizen :* customer is a senior citizen or not (1, 0)
4. *Partner:*  customer has a partner or not (Yes, No).
5. *Dependents:* customer has dependents or not (Yes, No)
6. *Tenure:* Number of months the customer has stayed with the company
7. *PhoneService:* customer has a phone service or not (Yes, No)
8. *MultipleLines:* Whether the customer has multiple lines or not (Yes, No, No phone service)
9. *InternetService:* Customer’s internet service provider (DSL, Fiber optic, No).
10. *OnlineSecurity:* customer has online security or not (Yes, No, No internet service)
11. *OnlineBackup:* customer has online backup or not (Yes, No, No internet service).
12. *DeviceProtection:* customer has device protection or not (Yes, No, No internet service).
13. *TechSupport:* tech support or not (Yes, No, No internet service)
14. *StreamingTV:* customer has streaming TV or not (Yes, No, No internet service)
15. *StreamingMovies:* customer has streaming movies or not (Yes, No, No internet service)
16. *Contract:* The contract term of the customer (Month-to-month, One year, Two year).
17. *PaperlessBilling:* Whether the customer has paperless billing or not (Yes, No)
18. *PaymentMethod:* The customer’s payment method (Electronic check, Mailed check, Bank transfer (automatic), Credit card (automatic))
19. *MonthlyCharges:* The amount charged to the customer monthly
20. *TotalCharges:* The total amount charged to the customer
21. *Churn:* churned or not.

**Source:** uci.edu