Telecom Customer Churn Prediction Using Watson Auto AI

Sahiti Nallamolu

Internship Title : RSIP Career Basic ML 122

Project ID : SPS_PRO_285

INTRODUCTION

OVERVIEW

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves. It involves collection of data followed by data cleaning which is analyzed and represented in various visualization techniques. The cleaned data along with the output is fed to build a model which recognizes the pattern in data further this model can be used to test with unknown data to predict certain outcome.

Prediction Analysis in Machine Learning trains the model to find patterns hidden in data fed to it, using this the model will predict the outcome for given data.
These models are used to predict anything from customer churn and sports to corporate earnings.
Predictive modeling is useful because it gives accurate insight into any question and allows users to create forecasts. To maintain a competitive advantage, it is critical to have insight into future events and outcomes that challenge key assumptions.

This improves a company's chance of gaining and retaining of customers/employees by identifying the underlying patterns.

PURPOSE

Customer retention is one of the primary growth pillars for products with a subscription-based business model. Several bad experiences – or even one – would lead a customer to quit. And if droves of unsatisfied customers churn at a clip, both material losses and damage to reputation would be enormous. Churn rate is a health indicator for businesses whose customers are subscribers and paying for services. Hence understanding the churn rate of a company is important as it will determine the upcoming business expansion strategies. This will help in strategising the improvement in service so as to increase the customer traffic.

LITERATURE SURVEY

PROBLEM

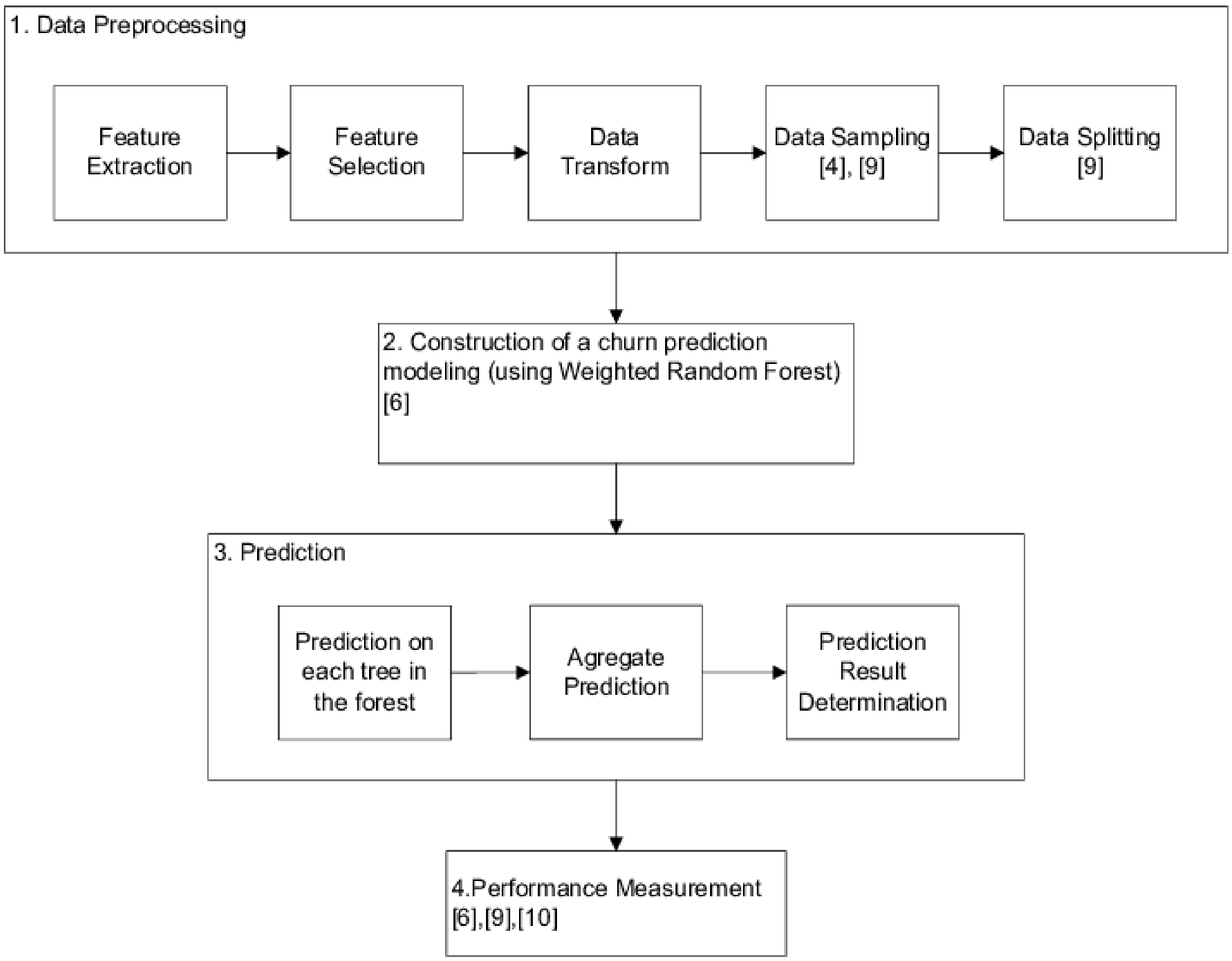
Customer churn is a major problem and one of the most important concerns for large companies. Due to the direct effect on the revenues of the companies, especially in the telecom field, companies are seeking to develop means to predict potential customer to churn.
Telecommunication industry always suffers from a very high churn rates when one industry offers a better plan than the previous there is a high possibility of the customer churning from the present due to a better plan in such a scenario it is very difficult to avoid losses but through prediction we can keep it to a minimal level. Therefore, finding factors that increase customer churn is important to take necessary actions to reduce this churn. Churn prediction helps in identifying those customers who are likely to leave a company.

PROPOSED SOLUTION

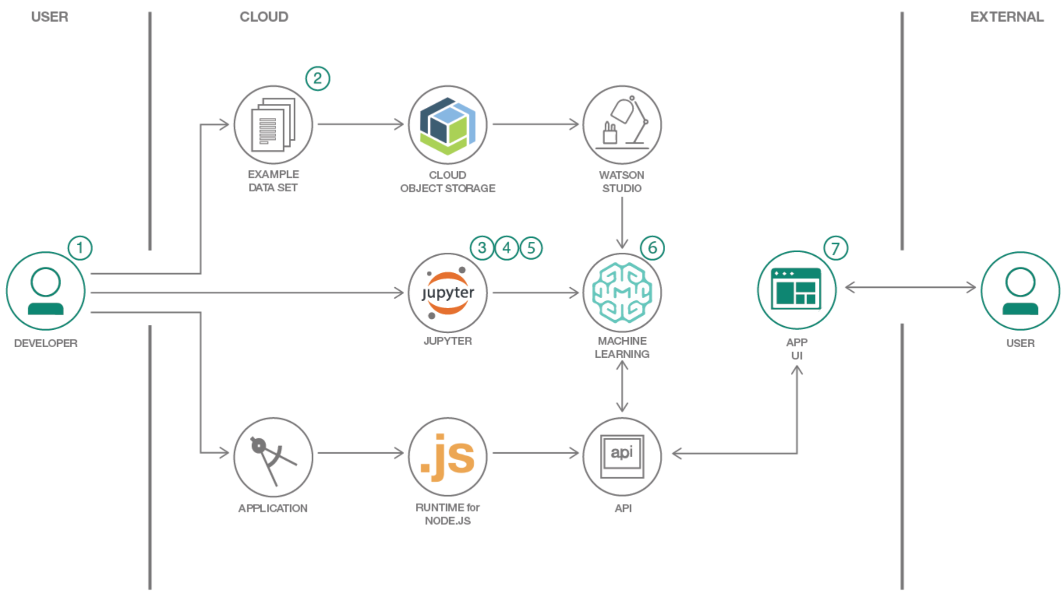
We are building a Machine Learning model to predict the customer churn using IBM Watson AutoAI Machine Learning Service. The main contribution of our work is to develop a churn prediction model which assists telecom operators to predict customers who are most likely subject to churn. The model developed in this work uses machine learning techniques on IBM platform and builds a new way of features' engineering and selection.
The model is deployed on IBM cloud to get scoring end point which can be used as API in mobile app or web app building. We are developing a web application which is built using node red service. We make use of the scoring end point to give user input values to the deployed model. The model prediction is then showcased on User Interface built using IBM Node Red.

THEORITICAL ANALYSIS

BLOCK DIAGRAM



HARDWARE/SOFTWARE DESIGNING

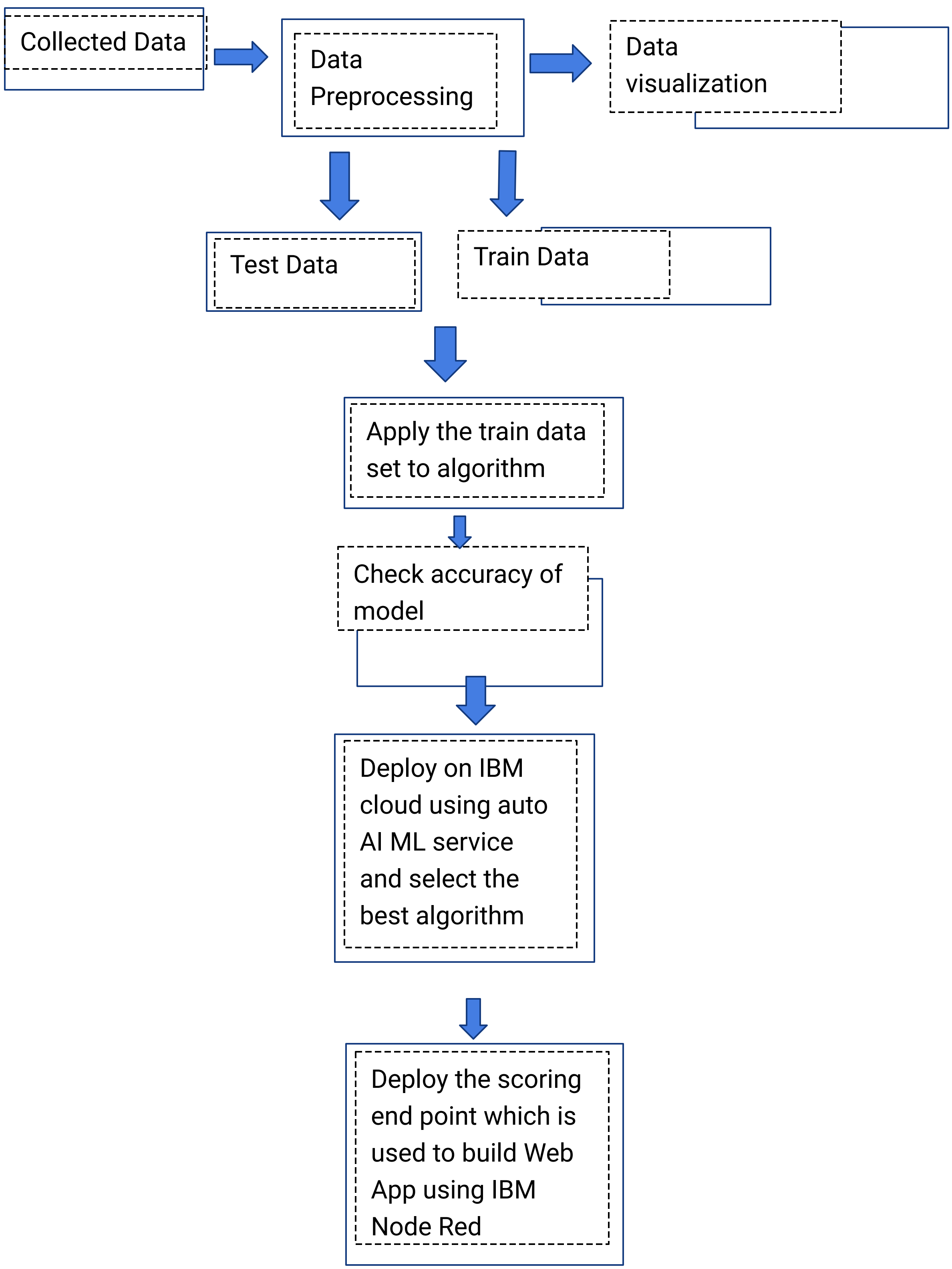


EXPERIMENTAL INVESTIGATIONS

From the given dataset, the following investigations have been derived:

1. According to the geography, Germany has more customer churning.
2. The customers with less than 2 products are more likely to exit the company.
3. There are more female customers who exit from the company than male customers.

FLOWCHART


Collected Data
Data Preprocessing
Data visualization
Test Data
Train Data
Apply the train data set to algorithm
Check accuracy of model
Deploy on IBM cloud using auto AI ML service and select the best algorithm
Deploy the scoring end point which is used to build Web App using IBM Node Red

RESULT

The model deployed using IBM Auto AI ML service gives the prediction and probability rate of whether a customer will exit the company or not.
The model takes the several parameter details of the customer as an input and gives an output of 2 values that is the prediction value and the probability value.
The deployed model is implemented using the scoring end point, which is used as API in web service for which we built the UI using IBM Node Red.
When a user opens the URL (scoring end point), a web page will open containing a form, where the user has to enter the details of a customer and submit. Then the details are applied to the model deployed with the reference of ML instance id, and the model calculates and displays the prediction and probability values on the screen.

ADVANTAGES & DISADVANTAGES

Advantages

- The model gives an accurate prediction and probability values.
- This model can be of use to provide special schemes to that particular customer whose churn prediction value is high so as to retain the customer.
- Calculation speed is very high. It takes only a few miliseconds to get the output, hence time saving.
- There is no requirement of hardware, as the model is deployed on IBM cloud. Hence, cost saving.
- Auto AI makes it easy to apply the most accurate algorithm to the data.
- Employees with no coding knowledge can operate the model.
- Easy to make changes in the Web service or model.

Disadvantages

- The model predicts only for 1 customer, which means user must enter the details of each customer. Not feasible if we have large data to be checked.
- There may be a few errors in prediction due to inefficient data.

APPLICATIONS

1. Telecom Business
2. E-commerce
3. Internet Service
4. Online streaming platforms (with subscription)
5. Corporate Services

CONCLUSION

Customer Churn Prediction (CCP) is a challenging activity for decision makers and machine learning community because most of the time, churn and non-churn customers have resembling features. From different experiments on customer churn and related data, it can be seen that a classifier shows different accuracy levels for different zones of a dataset. In such situations, a correlation can easily be observed in the level of classifier's accuracy and certainty of its prediction.
Finding an accurate classifier is a challenging job for employees, Hence we use Auto AI ML service , which automates this process of finding the best fit algorithm to give accurate results.
Customer Churn Predcition model is highly used to retain existing customers and generate new customers, expanding business which has direct effects on the company's revenue leading to business expansion.

FUTURE SCOPE

The future scope of this paper will use hybrid classification techniques to point out existing association between churn prediction and customer lifetime value. The passive and the dynamic nature of the industry ensure that data mining has become increasingly significant aspect in the telecommunication industry prospect.

BIBILOGRAPHY

https://www.wikipedia.org/
https://ieeexplore.ieee.org/
https://medium.com/

APPENDIX

SOURCE CODE

[{"id":"5e5e8b8b.c37bc4","type":"tab","label":"Flow
2","disabled":false,"info":""},{"id":"7ad5be0a.4829f","type":"function","z":"5e5e8b8b.c37bc4","name":"PreToken
","func":"global.set(\"id\",msg.payload.id)\nglobal.set(\"sn\",msg.payload.sn)\nglobal.set(\"cs\",msg.paylo
ad.cs)\nglobal.set(\"lo\",msg.payload.lo)\nglobal.set(\"ge\",msg.payload.ge)\nglobal.set(\"ag\",msg.paylo
ad.ag)\nglobal.set(\"te\",msg.payload.te)\nglobal.set(\"ba\",msg.payload.ba)\nglobal.set(\"np\",msg.payl
oad.np)\nglobal.set(\"cc\",msg.payload.cc)\nglobal.set(\"am\",msg.payload.am)\nglobal.set(\"sa\",msg.p
ayload.sa)\nvar apikey=
\"9I6JC3Y30z7aBLkMMXG6YXHU5tFGFjvjGqnaxFy2c2fL\";\nmsg.headers={\"content-type\":\"application
/x-www-form-urlencoded\"}\nmsg.payload={\"grant_type\":\"urn:ibm:params:oauth:grant-type:apikey\",\"a
pikey\":apikey}\nreturn
msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","x":300,"y":140,"wires":[["b49b42f7.eefac"]]},{"id":"b49b42f
7.eefac","type":"http
request","z":"5e5e8b8b.c37bc4","name":"","method":"POST","ret":"obj","paytoqs":"ignore","url":"https://iam.clo
ud.ibm.com/identity/token","tls":"","persist":false,"proxy":"","authType":"","x":310,"y":240,"wires":[["69763c57.1
755a4"]]},{"id":"f6ad3a71.d095c8","type":"debug","z":"5e5e8b8b.c37bc4","name":"","active":true,"tosidebar":tr
ue,"console":false,"tostatus":false,"complete":"payload","targetType":"msg","statusVal":"","statusType":"auto","
x":790,"y":100,"wires":[]},{"id":"69763c57.1755a4","type":"function","z":"5e5e8b8b.c37bc4","name":"Pre
Prediction","func":"var rn = global.get('rn')\nvar id = global.get('id')\nvar sn = global.get('sn')\nvar cs =
global.get('cs')\nvar lo = global.get('lo')\nvar ge = global.get('ge')\nvar ag = global.get('ag')\nvar te =
global.get('te')\nvar ba = global.get('ba')\nvar np = global.get('np')\nvar cc = global.get('cc')\nvar am =
global.get('am')\nvar sa = global.get('sa')\nvar token=msg.payload.access_token\nvar
instance_id=\"d5fa10bd-940b-4d43-bc26-7cd64f3adcc2\"\nmsg.headers={'Content-Type':
'application/json',\"Authorization\":\"Bearer
\"+token,\"ML-Instance-ID\":instance_id}\nmsg.payload={\"input_data\": [{\"fields\":
[\"RowNumber\",\"CustomerID\", \"Surname\", \"CreditScore\", \"Location\", \"Gender\", \"Age\",
\"Tenure\",\"Balalnce\",\"NumberofProducts\",\"CreditCard\", \"ActiveMember\",\"Salary\"], \"values\":
[[rn,id,sn,cs,lo,ge,ag,te,ba,np,cc,am,sa]]}]}\nreturn
msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","x":440,"y":340,"wires":[["9ce6be06.df08e"]]},{"id":"9ce6be0
6.df08e","type":"http
request","z":"5e5e8b8b.c37bc4","name":"","method":"POST","ret":"obj","paytoqs":"ignore","url":"https://us-sout
h.ml.cloud.ibm.com/v4/deployments/93a9b88c-68da-4798-8f58-c266bed990bb/predictions","tls":"","persi
st":false,"proxy":"","authType":"","x":530,"y":160,"wires":[["30d25c03.f97e14","f6ad3a71.d095c8","b952accf.16
ce4"]]},{"id":"ada1d2b0.b4b24","type":"ui_form","z":"5e5e8b8b.c37bc4","name":"","label":"","group":"cbabaea3.
3112c","order":1,"width":0,"height":0,"options":[{"label":"CustomerID","value":"id","type":"number","required":tru
e,"rows":null},{"label":"Surname","value":"sn","type":"text","required":true,"rows":null},{"label":"Creditscore","val
ue":"cs","type":"number","required":false,"rows":null},{"label":"Location","value":"lo","type":"text","required":true,
"rows":null},{"label":"Gender","value":"ge","type":"text","required":true,"rows":null},{"label":"Age","value":"ag","typ
e":"number","required":true,"rows":null},{"label":"Tenure","value":"te","type":"number","required":true,"rows":null

},{"label":"Balance","value":"ba","type":"number","required":true,"rows":null},{"label":"Numofproducts","value":"np","type":"number","required":false,"rows":null},{"label":"CreditCard","value":"cc","type":"number","required":false,"rows":null},{"label":"ActiveMember","value":"am","type":"number","required":true,"rows":null},{"label":"Salary","value":"sa","type":"number","required":true,"rows":null}],"formValue":{"id":"","sn":"","cs":"","lo":"","ge":"","ag":"","te":"","ba":"","np":"","cc":"","am":"","sa":""},"payload":"","submit":"submit","cancel":"cancel","topic":"","x":90,"y":240,"wires":[["7ad5be0a.4829f"]]},{"id":"30d25c03.f97e14","type":"function","z":"5e5e8b8b.c37bc4","name":"","func":"msg.payload=msg.payload.predictions[0].values[0][1][0]\nreturn msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","x":760,"y":180,"wires":[["94838814.4c0c98"]]},{"id":"94838814.4c0c98","type":"ui_text","z":"5e5e8b8b.c37bc4","group":"cbabaea3.3112c","order":2,"width":0,"height":0,"name":"","label":"Prediction :","format":"{{msg.payload}}","layout":"row-spread","x":970,"y":200,"wires":[]},{"id":"f422a00e.75d1d","type":"ui_text","z":"5e5e8b8b.c37bc4","group":"cbabaea3.3112c","order":3,"width":0,"height":0,"name":"","label":"Probability :","format":"{{msg.payload}}","layout":"row-spread","x":950,"y":320,"wires":[]},{"id":"b952accf.16ce4","type":"function","z":"5e5e8b8b.c37bc4","name":"","func":"msg.payload=msg.payload.predictions[0].values[0][1][1]\nreturn msg;","outputs":1,"noerr":0,"initialize":"","finalize":"","x":740,"y":300,"wires":[["f422a00e.75d1d"]]},{"id":"cbabaea3.3112c","type":"ui_group","z":"","name":"Customer Churn ","tab":"b59f5041.1edff","order":1,"disp":true,"width":"6","collapse":false},{"id":"b59f5041.1edff","type":"ui_tab","z":"","name":"Home","icon":"dashboard","disabled":false,"hidden":false}]

UI OUTPUT SCREENSHOT