Report on Churn Analysis for Small Businesses

Salvader Ron Nathaniel

Abstract

In this report I have discussed about churn analysis and why churn analysis is important for a company and the benefits churn analysis provides to the company in understanding the customer and market. I have also created a prototype using an openly available dataset from kaggle. The term 'Churn' means leaving the company and it is very important for a business to known why their customers are switching to other companies for good and services. Having a robust and accurate churn prediction model helps businesses to take actions to prevent customers from leaving the company. Churn analysis can be applied from a very small business to large scale industry as well. By finding out the churn rate and keeping it low ensures company that it is highly unlikely to have customer retention. Churn analysis also shows the number of re-current user for the company.

Deloitte research says, "Competition is stronger, the landscape is more dynamic, and consumers have more options to discover content and find value"

So churn analysis should be mandatory and helpful for company growth and development.

1. Problem Statement

Every business or company sells their good and services either to customers or companies and one aim of the provider is to know their consumer, it is majorly done through churn analysis which aims in finding how many customers annually have increased or decreased and at which rate, according to that the provider company must improve the quality of the products based on the analysis.

2. Market/Customer/Business Assessment

2.1 **Market Assessment: -** Churn analysis helps to understand why customers don't return for repeat business, it helps to understand that needs of the market and changes that are happening due to which the repetitive customer's rate is decreasing; by analysing it the company must make changes in their required product to survive in the market.

2.2 **Customer Assessment: -** Customer churn analysis is done to calculate the rate at which the customers are opting or shifting to other brand rather, the deeper analysis of customer churn also tells us about the ratio at which people are not returning for purchase to that of customers that are constantly buying from your brand. It tells us about the customers need that are changing with course of time and what updates a company products are required to fulfil the needs of customers.

2.3 **Business Assessment: -** A business assessment ensures your business has what it needs to deliver on your goals and it helps to grow the business in the way that is smart and strategic. It provides a solid foundation for making sound opinions. Most importantly, the real power of a business assessment is the process itself.

3. Target Specifications and Characterization

Churn analysis can be applied to businesses such as credit card companies, cable service providers, OTT service providers, and telecommunication companies worldwide. It helps to improve customer retention for business. It can also be implemented to small scale businesses or newly opened start-ups for expanding business and making strategy for it. Churn analysis is like a health indicator for businesses. Churn analysis is carried out by SAAS (Software as a Service) companies to understand the cause behind

the customer churn and implement the required strategy and ensure customer retention.

4. External Search

The sources that I have used for understanding churn analysis are:-

1. https://whatagraph.com/blog/articles/churn-analysis
2. www.google.com
3. https://www.ventureradar.com/keyword/Churn%20Prediction
4. https://www.softwaresuggest.com/churn-prediction-software

5. Bench marking alternative products

Below mention are the companies that perform churn analysis for various businesses:-

i. **Churnly –** Churnly's AI is powerful and impressive enough to help you deal with customer churn at every stage. Churnly provides end-to-end solutions designed for Windows. This online Customer Success system offers Account Alerts, Customer Engagement, Health Score, on boarding, Usage Tracking / Analytics at one place.

ii. **Qymatix -** Qymatix is a boon for B2B organizations. By using high-grade machine learning, this churn prediction software can track the customer journey, do instant analysis, and create behaviour-based modelling.

iii. **Trifacta -** Trifacta gathers data stored in different locations and do predictive analysis to reduce customer churn. This churn prediction software enables you to work directly with the data available and remove IT dependencies. Its churn predictions are data-driven and realistic.

iv. **Data Science Studio (DSS) -** Data Science Studio (DSS) is what you should get now if you want to do real-time churn prediction from available data. With its great data exploration ability, this churn prediction software ensures that not a single data goes waste.

6. Applicable Patents

i. **Churn Prediction and management System -** A system and method for managing churn among the customers of a business is provided. The system and method provide for an analysis of the causes of customer churn and identifies customers who are most likely to churn in the future.

Link - https://patents.google.com/patent/US20140278779A1/en

ii. **Predicting customer churn in a telecommunications network environment -** This churn analysis can be used by telecommunications companies and analysing how many customers place repetitive orders and how many customers are shifting to other telecommunications brand.

Link - https://patents.google.com/patent/US20150310336A1/en

7. Applicable Constraints

The constraint that needs to be applied while performing a churn analysis is: -

i. Understanding problem and final goal
ii. Data Collection
iii. Data preparation and processing
iv. Modelling and testing
v. Model Deployment and monitoring

8. Business Opportunity

Churn analysis help business to understand why customers don't return for repeat business, understanding churn analysis helps to improve the business products and it also helps to understand the current market needs that the customer requires. Churn analysis must be used in every business to analyse customer retention and make a solid foundation for business strategies.

9. Concept Generation

For any business to be successful one should understand the customer and it needs and how the needs are changing day-to-day

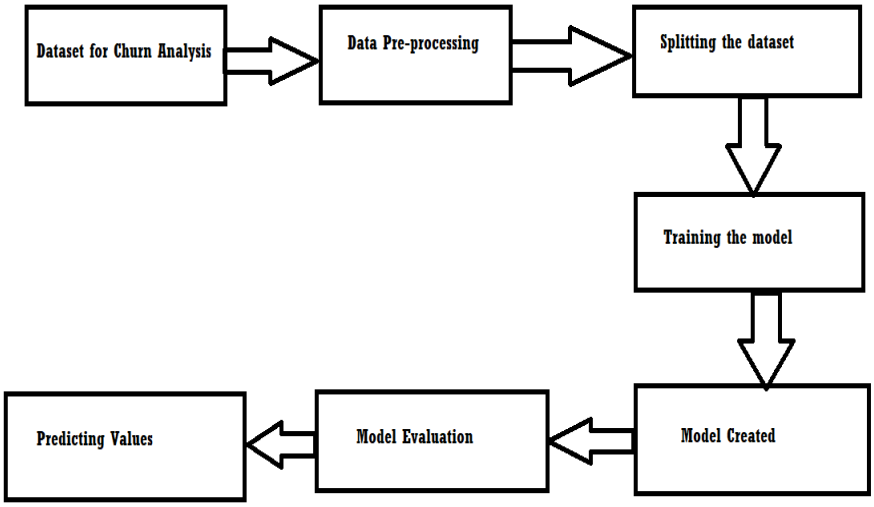
and how to ensure the customer trust on the company product. Company and stakeholders are investing more time and money finding out the reason and predict the type of customers that can switch to other brand and minimize it. The sudden downfall of user interaction with the company product also gives rise to perform churn analysis and finding the cause of it.

10. Concept Development

In order to build this project I have used a sample churn analysis dataset of why employees leave a particular company from kaggle and a machine learning model is being created with the help of the columns like age, time period for which he/she is in the company, if he is promoted to manager or not and like this and output is being predicted whether an employee will leave the company after a particular time period or not for which I have used random forest algorithm, we can also use other machine learning algorithm like logistic regression, decision tree, Naïve Bayes based or XGBoost on the dataset.

11. Final Product Prototype with Schematic Diagram

The following is the model prototype diagram for how churn analysis can be done a company's dataset:-



12. Product Details

12.1 How does it work?

A web portal can be created with textboxes and labels which take the input for the model for the independent values and then those values at backend will be fed to the model and output will be shown to the user easily.

12.2 Data Sources

A dataset from kaggle is being used for creating a prototype model of when the employees of a particular company quit their jobs and why they do so

Link -

https://drive.google.com/file/d/1lxBAYa_AAiyYPGqDM2ue006ONmiII45h/view?usp=sharing

12.3 Algorithms and framework used

For creating the model I have used random forest algorithm but we can also use other supervised machine learning algorithms as well. The most powerful and effective algorithm used is XGBoost.

For deployment we can either use ngrok or flask-ngrok for creating a frontend.

12.4 Team required to develop

Team required for creating churn analysis software are:-

i. Data analyst
ii. Machine Learning engineer
iii. Front end Developer
iv. Backend Developer
v. Database Administrator

13. Code Implementation

Below are some of the snapshots of the code implementation of the prototype along with github repo link:-

Importing required libraries and dataset

```
#importing all the necessary libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
from google.colab import drive
drive.mount('/content/drive')
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
```

```
#import dataset from
data=pd.read_csv('customer_churn.csv')
```

```
data
```

	Names	Age	Total_Purchase	Account_Manager	Years	Num_Sites	Onboard_date	Location	Company	Churn
0	Cameron Williams	42.0	11066.80	0	7.22	8.0	2013-08-30 07:00:40	10265 Elizabeth Mission Barkerburgh, AK 89518	Harvey LLC	1
1	Kevin Mueller	41.0	11916.22	0	6.50	11.0	2013-08-13 00:38:46	6157 Frank Gardens Suite 019 Carloshaven, RI 1...	Wilson PLC	1
2	Eric Lozano	38.0	12884.75	0	6.67	12.0	2016-06-29 06:20:07	1331 Keith Court Alyssahaven, DE 90114	Miller, Johnson and Wallace	1
3	Phillip White	42.0	8010.76	0	6.71	10.0	2014-04-22 12:43:12	13120 Daniel Mount Angelabury, WY 30645-4695	Smith Inc	1
4	Cynthia Norton	37.0	9191.58	0	5.56	9.0	2016-01-19 15:31:15	765 Tricia Row Karenshire, MH 71730	Love-Jones	1
...	...	...	...	...	...	...	...	...	...	...
895	Paul Miller	42.0	12800.82	1	3.62	8.0	2007-12-01 13:29:34	9316 Julian Fort Suite 328 North Leslie, ME 43961	Evans-Lucero	0

Applying Data pre-processing

Churn_Analysis_Feyn.ipynb

File Edit View Insert Runtime Tools Help Last edited on July 3

Comment Share

+ Code + Text

Connect Editing

```
data.info()
```

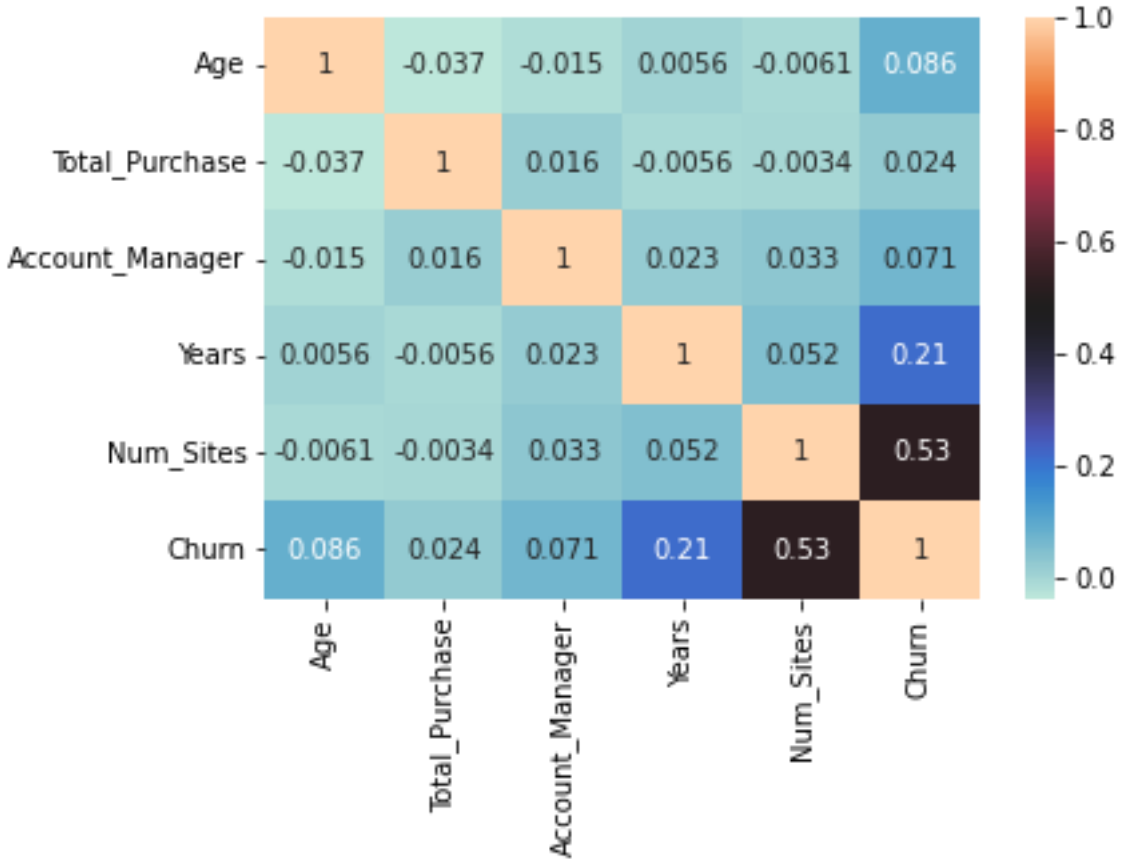
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 900 entries, 0 to 899
Data columns (total 10 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   Names            900 non-null    object
 1   Age              900 non-null    float64
 2   Total_Purchase   900 non-null    float64
 3   Account_Manager  900 non-null    int64
 4   Years            900 non-null    float64
 5   Num_Sites        900 non-null    float64
 6   Onboard_date     900 non-null    object
 7   Location         900 non-null    object
 8   Company          900 non-null    object
 9   Churn            900 non-null    int64
dtypes: float64(4), int64(2), object(4)
memory usage: 70.4+ KB
```

```
data.dtypes
```

```
Names              object
Age               float64
Total_Purchase    float64
Account_Manager     int64
Years             float64
Num_Sites         float64
Onboard_date       object
Location           object
Company            object
Churn               int64
dtype: object
```

```
data.isnull().sum()
```

Visualizing Dataset



Model Creation

+ Code + Text — Connect — Editing

```
# Splitting the dataset into the Training set and Test set
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(X, Y, test_size = 0.25, random_state =40)
#in train_test_split test size is 25% split and random state is starting of split data
```

```
# Feature Scaling
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
X_train = sc.fit_transform(x_train)
X_test = sc.transform(x_test)
```

```
from sklearn.ensemble import RandomForestClassifier
classifier = RandomForestClassifier(n_estimators=300,criterion = 'gini')
classifier.fit(X_train, y_train)
```

```
RandomForestClassifier(n_estimators=300)
```

```
y_pred = classifier.predict(X_test)
y_pred
```

```
array([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0,
       0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
       0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1,
       0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0,
       0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0,
       0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0])
```

Model Evaluation

Link - https://github.com/salvaderron/Feynn_Labs_Internship.git

14. Conclusion

From the above report it can be concluded how important is churn analysis is for a company and how it helps to understand the customer and market latest need and to take a smart and innovative decision for business expanding