

PRESENTATION TOPIC:

Customer Churn Prediction


MUHAMMAD ZUBAIR 695-2019

MUHAMMAD HUNAIN IQBAL 699-2019

Course: Artificial Intelligence lab

SUBMITTED TO : Sir Abid Ali

Objectives

- Introduction
 - Benefits of churn prediction
 - Data set of churn prediction
 - Data model And Workflow
 - conclusion
- 
- A series of three parallel white diagonal lines on the right side of the slide, extending from the middle towards the bottom right corner.

1ST PART

INTRODUCTION

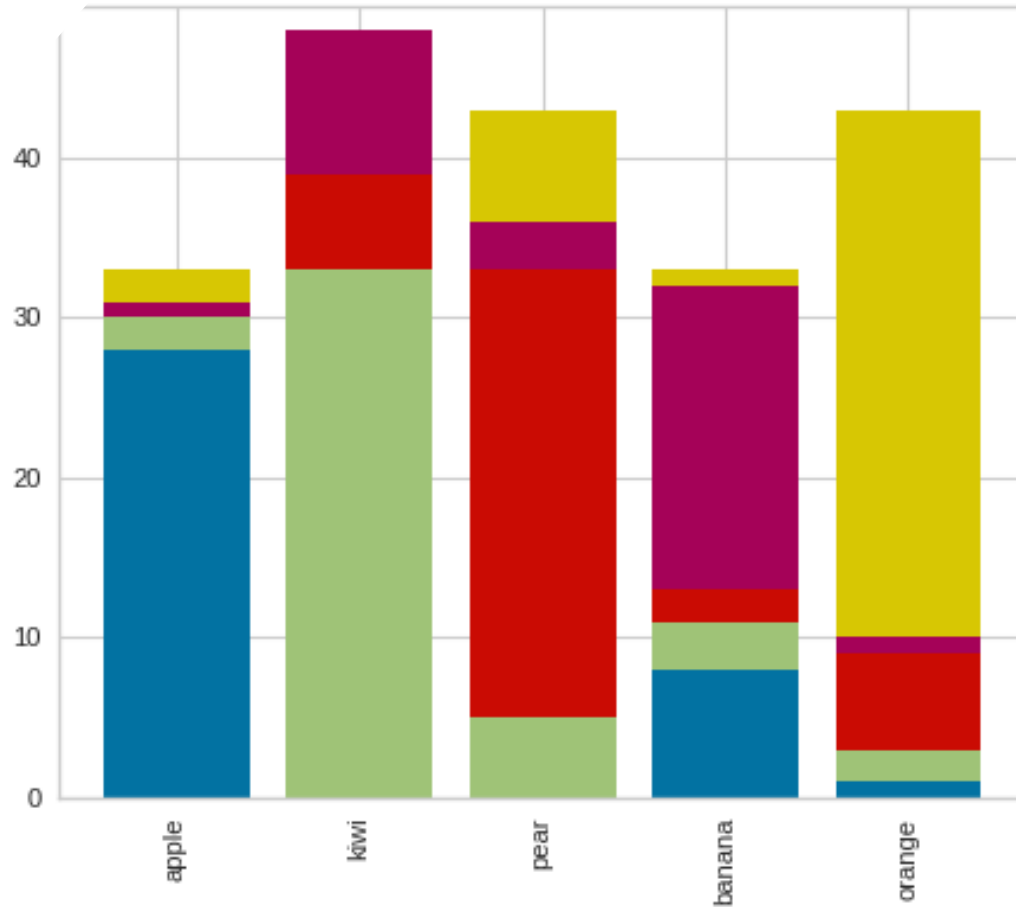


INTRODUCTION OF CUSTOMER CHURN


- ▶ Customer attrition, also known as customer churn, customer turnover, or customer defection, is the loss of clients or customers.

CUSTOMER CHURN PREDICTION

- ▶ Churn Prediction is essentially predicting which clients are most likely to cancel a subscription i.e. 'leave a company' based on their usage of the service



BENEFITS OF CUSTOMER CHURN

- Identify at-risk customers
 - Identify pain points
 - Identify methods to implement
- 
- Three white lines of varying lengths and slopes are positioned in the bottom right corner of the slide, extending from the right edge towards the center.

Deep Learning (ANN)

- Why we used ANN for churn prediction
- Is there any alternative of deep learning.

2nd PART


Customer Churn Prediction

Dataset And Model Training

TELCOM CUSTOMER CHURN DATASET


customerID	gender	SeniorCitize	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	Paperless	PaymentMethod	MonthlyCh	TotalCharges	Churn
7590-VHVEG	Female	0	Yes	No	1	No	No phone	DSL	No	Yes	No	No	No	No	Month-to-month	Yes	Electronic check	29.85	29.85	No
5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	No	Yes	No	No	No	One year	No	Mailed check	56.95	1889.5	No
3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	No	No	No	No	Month-to-month	Yes	Mailed check	53.85	108.15	Yes
7795-CFOCW	Male	0	No	No	45	No	No phone	DSL	Yes	No	Yes	Yes	No	No	One year	No	Bank transfer (auto)	42.3	1840.75	No
9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	No	No	No	No	No	Month-to-month	Yes	Electronic check	70.7	151.65	Yes
9305-CDSKC	Female	0	No	No	8	Yes	Yes	Fiber optic	No	No	Yes	No	Yes	Yes	Month-to-month	Yes	Electronic check	99.65	820.5	Yes
1452-KIOVK	Male	0	No	Yes	22	Yes	Yes	Fiber optic	No	Yes	No	No	Yes	No	Month-to-month	Yes	Credit card (auto)	89.1	1949.4	No
6713-OKOMC	Female	0	No	No	10	No	No phone	DSL	Yes	No	No	No	No	No	Month-to-month	No	Mailed check	29.75	301.9	No
7892-POOKP	Female	0	Yes	No	28	Yes	Yes	Fiber optic	No	No	Yes	Yes	Yes	Yes	Month-to-month	Yes	Electronic check	104.8	3046.05	Yes
6388-TABGU	Male	0	No	Yes	62	Yes	No	DSL	Yes	Yes	No	No	No	No	One year	No	Bank transfer (auto)	56.15	3487.95	No
9763-GRSKD	Male	0	Yes	Yes	13	Yes	No	DSL	Yes	No	No	No	No	No	Month-to-month	Yes	Mailed check	49.95	587.45	No
7469-LKBCI	Male	0	No	No	16	Yes	No	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two year	No	Credit card (auto)	18.95	326.8	No
8091-TTVAX	Male	0	Yes	No	58	Yes	Yes	Fiber optic	No	No	Yes	No	Yes	Yes	One year	No	Credit card (auto)	100.35	5681.1	No
0280-XJGEX	Male	0	No	No	49	Yes	Yes	Fiber optic	No	Yes	Yes	No	Yes	Yes	Month-to-month	Yes	Bank transfer (auto)	103.7	5036.3	Yes
5129-JLPIS	Male	0	No	No	25	Yes	No	Fiber optic	Yes	No	Yes	Yes	Yes	Yes	Month-to-month	Yes	Electronic check	105.5	2686.05	No
3655-SNQYZ	Female	0	Yes	Yes	69	Yes	Yes	Fiber optic	Yes	Yes	Yes	Yes	Yes	Yes	Two year	No	Credit card (auto)	113.25	7895.15	No
8191-XWSZG	Female	0	No	No	52	Yes	No	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	One year	No	Mailed check	20.65	1022.95	No
9959-WOFKT	Male	0	No	Yes	71	Yes	Yes	Fiber optic	Yes	No	Yes	No	Yes	Yes	Two year	No	Bank transfer (auto)	106.7	7382.25	No
4190-MFLUW	Female	0	Yes	Yes	10	Yes	No	DSL	No	No	Yes	Yes	No	No	Month-to-month	No	Credit card (auto)	55.2	528.35	Yes
4183-MYFRB	Female	0	No	No	21	Yes	No	Fiber optic	No	Yes	Yes	No	No	Yes	Month-to-month	Yes	Electronic check	90.05	1862.9	No
8779-QRDMV	Male	1	No	No	1	No	No phone	DSL	No	No	Yes	No	No	Yes	Month-to-month	Yes	Electronic check	39.65	39.65	Yes
1680-VDCWW	Male	0	Yes	No	12	Yes	No	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	One year	No	Bank transfer (auto)	19.8	202.25	No
1066-JKSGK	Male	0	No	No	1	Yes	No	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Month-to-month	No	Mailed check	20.15	20.15	Yes
3638-WEABW	Female	0	Yes	No	58	Yes	Yes	DSL	No	Yes	No	Yes	No	No	Two year	Yes	Credit card (auto)	59.9	3505.1	No
6322-HRPFA	Male	0	Yes	Yes	49	Yes	No	DSL	Yes	Yes	No	Yes	No	No	Month-to-month	No	Credit card (auto)	59.6	2970.3	No
6865-JZNKO	Female	0	No	No	30	Yes	No	DSL	Yes	Yes	No	No	No	No	Month-to-month	Yes	Bank transfer (auto)	55.3	1530.6	No
6467-CHFZW	Male	0	Yes	Yes	47	Yes	Yes	Fiber optic	No	Yes	No	No	Yes	Yes	Month-to-month	Yes	Electronic check	99.35	4749.15	Yes
8665-UTDHZ	Male	0	Yes	Yes	1	No	No phone	DSL	No	Yes	No	No	No	No	Month-to-month	No	Electronic check	30.2	30.2	Yes
5248-YGIJN	Male	0	Yes	No	72	Yes	Yes	DSL	Yes	Yes	Yes	Yes	Yes	Yes	Two year	Yes	Credit card (auto)	90.25	6369.45	No
8773-HHUOZ	Female	0	No	Yes	17	Yes	No	DSL	No	No	No	No	Yes	Yes	Month-to-month	Yes	Mailed check	64.7	1093.1	Yes
3841-NFECX	Female	1	Yes	No	71	Yes	Yes	Fiber optic	Yes	Yes	Yes	Yes	No	No	Two year	Yes	Credit card (auto)	96.35	6766.95	No
4929-XIHWV	Male	1	Yes	No	2	Yes	No	Fiber optic	No	No	Yes	No	Yes	Yes	Month-to-month	Yes	Credit card (auto)	95.5	181.65	No
6827-IEAUQ	Female	0	Yes	Yes	27	Yes	No	DSL	Yes	Yes	Yes	Yes	No	No	One year	No	Mailed check	66.15	1874.45	No
7310-EGVHZ	Male	0	No	No	1	Yes	No	No	No internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Month-to-month	No	Bank transfer (auto)	20.2	20.2	No
3413-BMNZE	Male	1	No	No	1	Yes	No	DSL	No	No	No	No	No	No	Month-to-month	No	Bank transfer (auto)	45.25	45.25	No
6234-RAAPL	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	Yes	Yes	No	Yes	Yes	No	Two year	No	Bank transfer (auto)	99.9	7251.7	No
6047-YHPVI	Male	0	No	No	5	Yes	No	Fiber optic	No	No	No	No	No	No	Month-to-month	Yes	Electronic check	69.7	316.9	Yes

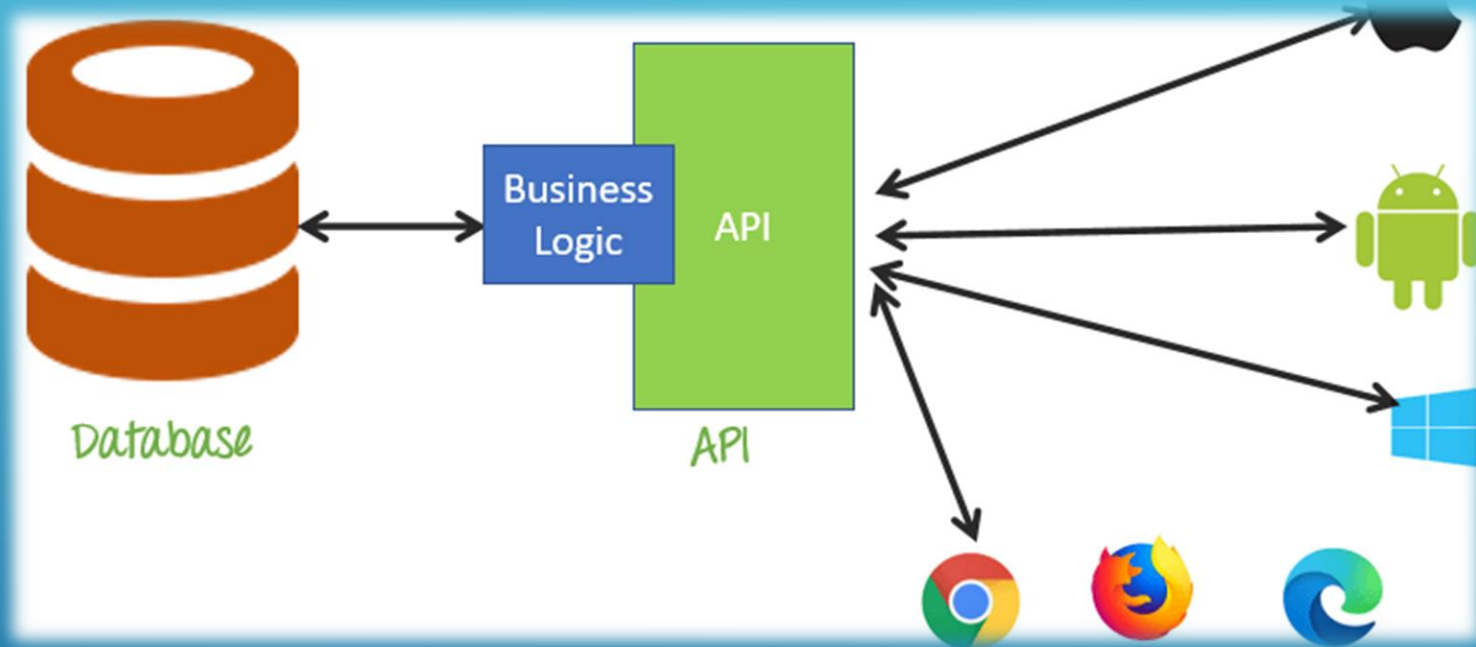
DATA EXPLORATION AND DATA PREPROCESSING

- ▶ What is Data Exploration
 - ▶ What is Data Preprocessing
 - ▶ Data Preprocessing Techniques
 - Data Cleaning
 - Data integration
 - Data Reduction
- 
- A series of three parallel white diagonal lines in the bottom right corner of the slide, extending from the middle of the right edge towards the bottom left.

What Is An Api

When you use an application on your mobile phone, the application connects to the internet and sends data to a server.

Several white lines of varying lengths and angles are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.



SCREENSHOTS

PROJECT



```
from kaggle.api.kaggle_api_extended import KaggleApi
```

```
api = KaggleApi()  
api.authenticate()
```

```
api.dataset_download_file('blastchar/telco-customer-churn',  
                           file_name = 'WA_Fn-UseC_-Telco-Customer-Churn.csv')
```



```
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
```

```
data_frame = pd.read_csv("Downloads/WA_Fn-UseC_-Telco-Customer-Churn.csv")
data_frame.head()
```

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProtection	TechSupp
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	Yes	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	No	

5 rows × 21 columns

```
data_frame.shape
```

(7043, 21)

```
In [6]: data_frame.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7043 entries, 0 to 7042
Data columns (total 21 columns):
 #   Column                Non-Null Count  Dtype  
---  -
 0   customerID            7043 non-null   object 
 1   gender                7043 non-null   object 
 2   SeniorCitizen         7043 non-null   int64  
 3   Partner               7043 non-null   object 
 4   Dependents            7043 non-null   object 
 5   tenure                7043 non-null   int64  
 6   PhoneService          7043 non-null   object 
 7   MultipleLines         7043 non-null   object 
 8   InternetService       7043 non-null   object 
 9   OnlineSecurity        7043 non-null   object 
10  OnlineBackup          7043 non-null   object 
11  DeviceProtection      7043 non-null   object 
12  TechSupport           7043 non-null   object 
13  StreamingTV           7043 non-null   object 
14  StreamingMovies       7043 non-null   object 
15  Contract              7043 non-null   object 
16  PaperlessBilling      7043 non-null   object 
17  PaymentMethod         7043 non-null   object 
18  MonthlyCharges        7043 non-null   float64 
19  TotalCharges          7043 non-null   object 
20  Churn                 7043 non-null   object 
dtypes: float64(1), int64(2), object(18)
memory usage: 1.1+ MB
```

```
In [7]: data_frame.drop("customerID", axis='columns', inplace=True)
data_frame.dtypes
```

```
Out[7]: gender                object
```


DTYPE CONVERT

```
data_frame[pd.to_numeric(data_frame.TotalCharges, errors='coerce').isnull()]
```

security	OnlineBackup	DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract	PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges
Yes	No	Yes	Yes	Yes	No	Two year	Yes	Bank transfer (automatic)	52.55	
internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two year	No	Mailed check	20.25	
Yes	Yes	Yes	No	Yes	Yes	Two year	No	Mailed check	80.85	
internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two year	No	Mailed check	25.75	
Yes	Yes	Yes	Yes	Yes	No	Two year	No	Credit card (automatic)	56.05	
internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two year	No	Mailed check	19.85	
internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two year	No	Mailed check	25.35	
internet service	No internet service	No internet service	No internet service	No internet service	No internet service	Two year	No	Mailed check	20.00	
internet service	No internet service	No internet service	No internet service	No internet service	No internet service	One year	Yes	Mailed check	19.70	
No	Yes	Yes	Yes	Yes	No	Two year	No	Mailed check	73.35	
Yes	Yes	No	Yes	No	No	Two year	Yes	Bank transfer (automatic)	61.90	

```
data_frame.iloc[488]["TotalCharges"]
```

```
new_data_frame = data_frame[data_frame.TotalCharges != ' ']  
new_data_frame.shape
```

```
(7032, 20)
```

```
new_data_frame.TotalCharges = pd.to_numeric(new_data_frame.TotalCharges)
```

D:\jupy\lib\site-packages\pandas\core\generic.py:5168: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-vs-copying

ONE HOT ENCODING & DUMMY VARIABLES

```
In [49]: count_uniqu_value(new_data_frame)
```

```
gender : ['Female' 'Male']
Partner : ['Yes' 'No']
Dependents : ['No' 'Yes']
PhoneService : ['No' 'Yes']
MultipleLines : ['No' 'Yes']
InternetService : ['DSL' 'Fiber optic' 'No']
OnlineSecurity : ['No' 'Yes']
OnlineBackup : ['Yes' 'No']
DeviceProtection : ['No' 'Yes']
TechSupport : ['No' 'Yes']
StreamingTV : ['No' 'Yes']
StreamingMovies : ['No' 'Yes']
Contract : ['Month-to-month' 'One year' 'Two year']
PaperlessBilling : ['Yes' 'No']
PaymentMethod : ['Electronic check' 'Mailed check' 'Bank transfer (automatic)'
                 'Credit card (automatic)']
Churn : ['No' 'Yes']
```

```
In [50]: yes_no_columns = ["Partner", "Dependents", "PhoneService", "MultipleLines", "OnlineSecurity", "OnlineBackup",
                          "DeviceProtection", "TechSupport", "StreamingTV", "StreamingMovies", "PaperlessBilling", "Churn"]
        ]
```

```
In [63]: for col in yes_no_columns:
        new_data_frame[col].replace({'Yes':1, "No":0}, inplace=True)
```

D:\jupyter\lib\site-packages\pandas\core\series.py:4563: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
return super().replace(

```
for column in new_data_frame:
    print(f'{column} : {new_data_frame[column].unique()}')
```

```
gender : ['Female' 'Male']
SeniorCitizen : [0 1]
Partner : [1 0]
Dependents : [0 1]
tenure : [ 1 34  2 45  8 22 10 28 62 13 16 58 49 25 69 52 71 21 12 30 47 72 17 27
          5 46 11 70 63 43 15 60 18 66  9  3 31 50 64 56  7 42 35 48 29 65 38 68
          32 55 37 36 41  6  4 33 67 23 57 61 14 20 53 40 59 24 44 19 54 51 26 39]
PhoneService : [0 1]
MultipleLines : [0 1]
InternetService : ['DSL' 'Fiber optic' 'No']
OnlineSecurity : [0 1]
OnlineBackup : [1 0]
DeviceProtection : [0 1]
TechSupport : [0 1]
StreamingTV : [0 1]
StreamingMovies : [0 1]
Contract : ['Month-to-month' 'One year' 'Two year']
PaperlessBilling : [1 0]
PaymentMethod : ['Electronic check' 'Mailed check' 'Bank transfer (automatic)'
                 'Credit card (automatic)']
MonthlyCharges : [29.85 56.95 53.85 ... 63.1  44.2  78.7 ]
TotalCharges : [ 29.85 1889.5  108.15 ... 346.45 306.6 6844.5 ]
Churn : [0 1]
```

```
new_data_frame['gender'].replace({'Female':1, "Male":0}, inplace=True)
new_data_frame.gender.unique()
```

```
new_data_frame['gender'].replace({'Female':1,"Male":0},inplace=True)
```

```
new_data_frame.gender.unique()
```

```
array([1, 0], dtype=int64)
```

```
data_frame2 = pd.get_dummies(data=new_data_frame,columns=["InternetService","Contract","PaymentMethod"])
```

```
data_frame2.shape
```

```
(7032, 27)
```

```
data_frame2.sample(10)
```

SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	OnlineSecurity	OnlineBackup	DeviceProtection	...	InternetService_DSL	InternetService_
0	0	0	18	1	0	0	1	1	...	0	
0	1	1	12	1	1	0	0	0	...	0	
0	0	0	8	1	0	0	1	0	...	1	
1	0	0	2	1	0	0	1	1	...	0	
0	0	0	62	1	1	1	1	0	...	0	
0	1	1	8	1	0	0	0	1	...	1	
0	1	1	43	1	0	0	0	0	...	0	
0	0	0	5	1	0	0	0	0	...	0	
0	1	1	46	1	0	0	0	0	...	0	
1	0	0	1	1	0	1	0	1	...	1	

umns

FEATURE SCALING AND TRAIN_TEST_SPLIT

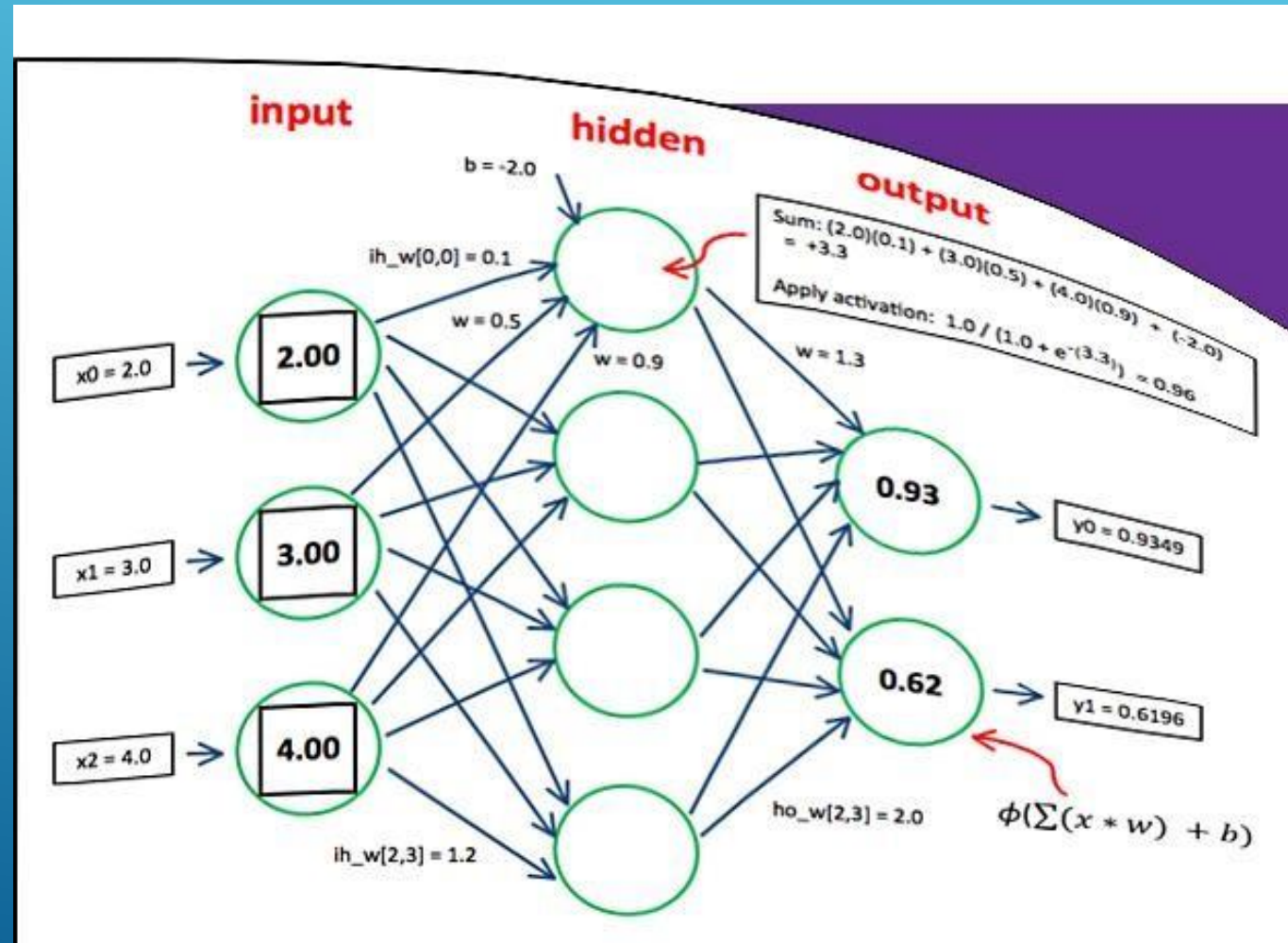
```
: col_to_scale = ["TotalCharges","tenure","MonthlyCharges"]
from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
data_frame2[col_to_scale] = scaler.fit_transform(data_frame2[col_to_scale])

: for column in data_frame2:
    print(f'{column} : {data_frame2[column].unique()}')

gender : [1 0]
SeniorCitizen : [0 1]
Partner : [1 0]
Dependents : [0 1]
tenure : [0.         0.46478873 0.01408451 0.61971831 0.09859155 0.29577465
 0.12676056 0.38028169 0.85915493 0.16901408 0.21126761 0.8028169
 0.67605634 0.33802817 0.95774648 0.71830986 0.98591549 0.28169014
 0.15492958 0.4084507 0.64788732 1.         0.22535211 0.36619718
 0.05633803 0.63380282 0.14084507 0.97183099 0.87323944 0.5915493
 0.1971831 0.83098592 0.23943662 0.91549296 0.11267606 0.02816901
 0.42253521 0.69014085 0.88732394 0.77464789 0.08450704 0.57746479
 0.47887324 0.66197183 0.3943662 0.90140845 0.52112676 0.94366197
 0.43661972 0.76056338 0.50704225 0.49295775 0.56338028 0.07042254
 0.04225352 0.45070423 0.92957746 0.30985915 0.78873239 0.84507042
 0.18309859 0.26760563 0.73239437 0.54929577 0.81690141 0.32394366
 0.6056338 0.25352113 0.74647887 0.70422535 0.35211268 0.53521127]
PhoneService : [0 1]
MultipleLines : [0 1]
OnlineSecurity : [0 1]
OnlineBackup : [1 0]
DeviceProtection : [0 1]
TechSupport : [0 1]
StreamingTV : [0 1]
StreamingMovies : [0 1]
PaperlessBilling : [1 0]
MonthlyCharges : [0.11542289 0.38507463 0.35422886 ... 0.44626866 0.25820896 0.60149254]
TotalCharges : [0.0012751 0.21586661 0.01031041 ... 0.03780868 0.03321025 0.78764136]
Churn : [0. 1.]
InternetService_DSL : [1 0]
InternetService_Fiber optic : [0 1]
InternetService_No : [0 1]
Contract_Month-to-month : [1 0]
Contract_One year : [0 1]
Contract_Two year : [0 1]
PaymentMethod_Bank transfer (automatic) : [0 1]
PaymentMethod_Credit card (automatic) : [0 1]
PaymentMethod_Electronic check : [1 0]
PaymentMethod_Mailed check : [0 1]
```

```
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test = train_test_split(X,y,test_size=0.2,random_state=5)
X_train.shape
y_train.shape
```

- ▶ `import tensorflow as tf`
- ▶ `from tensorflow import keras`
- ▶ `model = keras.Sequential([`
- ▶ `keras.layers.Dense(20,input_shape=(26,),activation='relu'),`
`keras.layers.Dense(15,activation='relu'),`
`keras.layers.Dense(1,activation='sigmoid')`
- ▶ `])`
- ▶ `model.compile(optimizer='adam',loss = "binary_crossentropy",`
`metrics=['accuracy'])`
- ▶ `model.fit(X_train,y_train,epochs =100)`



PROJECT OUTPUT

```
import seaborn as sn
cm = tf.math.confusion_matrix(labels=y_test,predictions=y_pred)

plt.figure(figsize = (10,7))
sn.heatmap(cm, annot=True, fmt='d')
plt.xlabel('Predicted')
plt.ylabel('Truth')
```

Text(69.0, 0.5, 'Truth')



C
O
N
F
U
S
I
O
N
M
A
T
R
I
X

THE END

Thank You