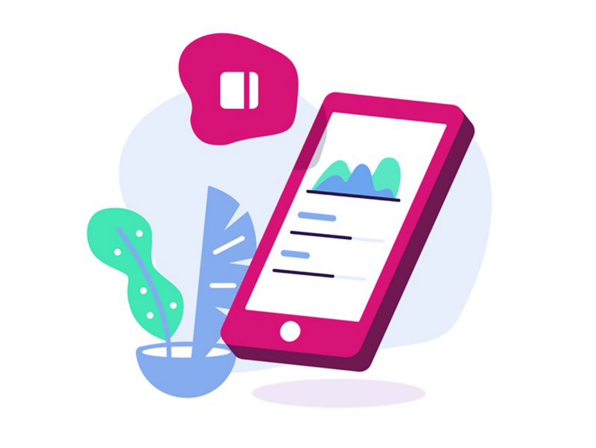
STOCK MARKET PREDICTION

TABLE OF CONTENTS



01

INTRODUCTION

What is Stock Market Prediction?

- It is an act trying to determine the future stock value using its historical data.
- It can be for single organization such as Morgan Stanley.
- It can be for the particular stock exchange. For example, New York Stock Exchange (NYSE).
- Based on the **Efficient-Market Hypothesis (EMH).**

Why Stock Market Prediction ?

- Primary source to raise funds for many listed organizations.
- It's also plays important role in the country's economic state.
- Based on the concept of demand and supply.
- It's a dynamic entity in order to predict certain results.
- It posses many theoretical and experimental challenges.

02

DATA UNDERSTANDING

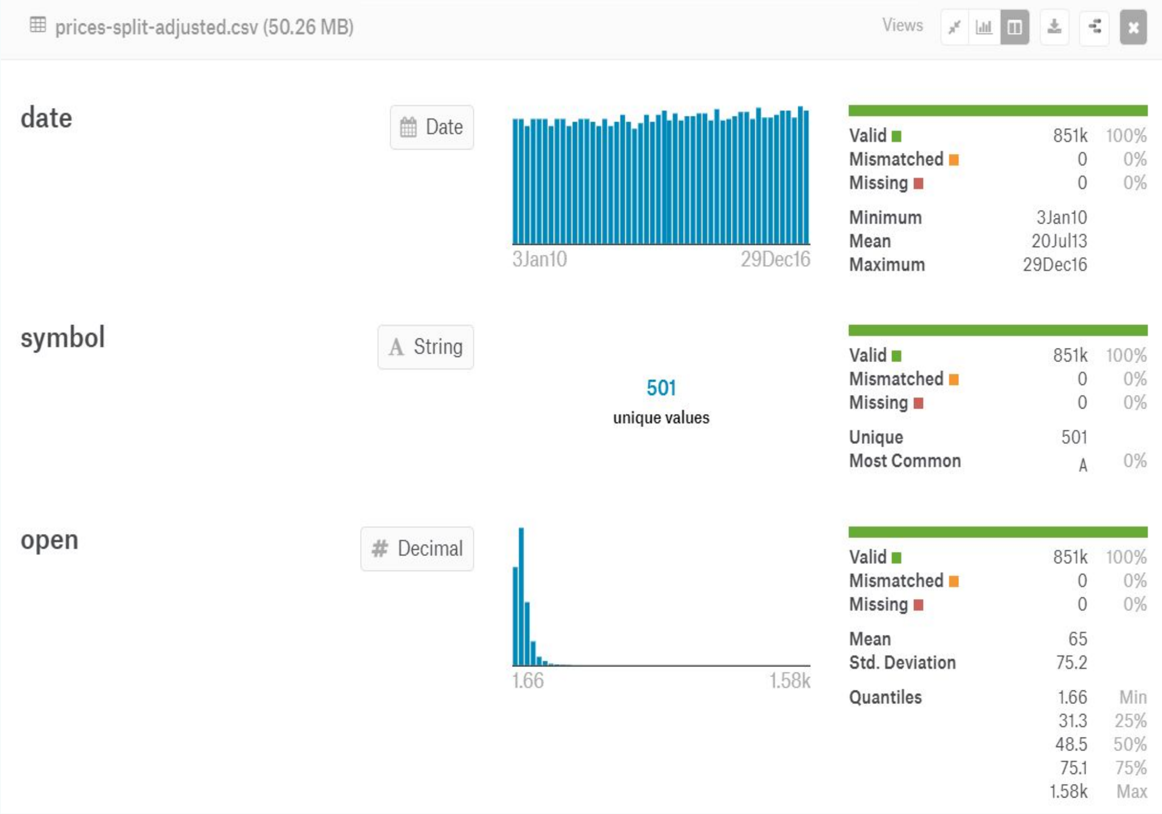
Dataset Source

- **Data Source:** Kaggle
- **Market:** New York Stock Exchange (NYSE)
- **Number of Rows:** 851264
- **Number of Columns:** 7
- **Date Range:** 5/Jan/2016 - 30/Dec/2016

Dataset Snapshot

	A	B	C	D	E	F	G
1	date	symbol	open	close	low	high	volume
2	1/5/2016	WLTW	123.43	125.84	122.31	126.25	2163600
3	1/6/2016	WLTW	125.24	119.98	119.94	125.54	2386400
4	1/7/2016	WLTW	116.38	114.95	114.93	119.74	2489500
5	1/8/2016	WLTW	115.48	116.62	113.5	117.44	2006300
6	1/11/2016	WLTW	117.01	114.97	114.09	117.33	1408600
7	1/12/2016	WLTW	115.51	115.55	114.5	116.06	1098000
8	1/13/2016	WLTW	116.46	112.85	112.59	117.07	949600
9	1/14/2016	WLTW	113.51	114.38	110.05	115.03	785300
10	1/15/2016	WLTW	113.33	112.53	111.92	114.88	1093700
11	1/19/2016	WLTW	113.66	110.38	109.87	115.87	1523500
12	1/20/2016	WLTW	109.06	109.3	108.32	111.6	1653900
13	1/21/2016	WLTW	109.73	110	108.32	110.58	944300
14	1/22/2016	WLTW	111.88	111.95	110.19	112.95	744900
15	1/25/2016	WLTW	111.32	110.12	110	114.63	703800
16	1/26/2016	WLTW	110.42	111	107.3	111.4	563100
17	1/27/2016	WLTW	110.77	110.71	109.02	112.57	896100
18	1/28/2016	WLTW	110.9	112.58	109.9	112.97	680400
19	1/29/2016	WLTW	113.35	114.47	111.67	114.59	749900
20	2/1/2016	WLTW	114	114.5	112.9	114.85	574200
21	2/2/2016	WLTW	113.25	110.56	109.75	113.86	694800
22	2/3/2016	WLTW	113.38	114.05	109.64	114.64	896300

Descriptive Statistics Measures



Calculated following measures for the dataset:

- **Missing Values**
- **Unique Values**
- **Most Common Value (Categorical Features)**
- **Box Plot measures:**
  - Mean
  - Standard Deviation
  - Minimum
  - 25% (Q1)
  - 50% (Median/ Q2)
  - 75% (Q3)
  - Maximum

.

1. Missing Values

In statistics, missing data, or missing values, occur when no data value is stored for the variable in an observation

Some common missing value representation are :

1) NA
2) NaN
3) Blank (no value)

```
In [36]:  1  df.isnull().sum()

Out[36]: date      0
         symbol    0
         open      0
         close     0
         low       0
         high      0
         volume    0
         dtype: int64
```

close # Decimal

Valid	851k	100%
Mismatched	0	0%
Missing	0	0%
Mean	65	
Std. Deviation	75.2	
Quantiles	1.59	Min
	31.3	25%
	48.5	50%
	75.1	75%
	1.58k	Max

low # Decimal

Valid	851k	100%
Mismatched	0	0%
Missing	0	0%
Mean	64.3	
Std. Deviation	74.5	
Quantiles	1.5	Min
	30.9	25%
	48	50%
	74.4	75%
	1.55k	Max

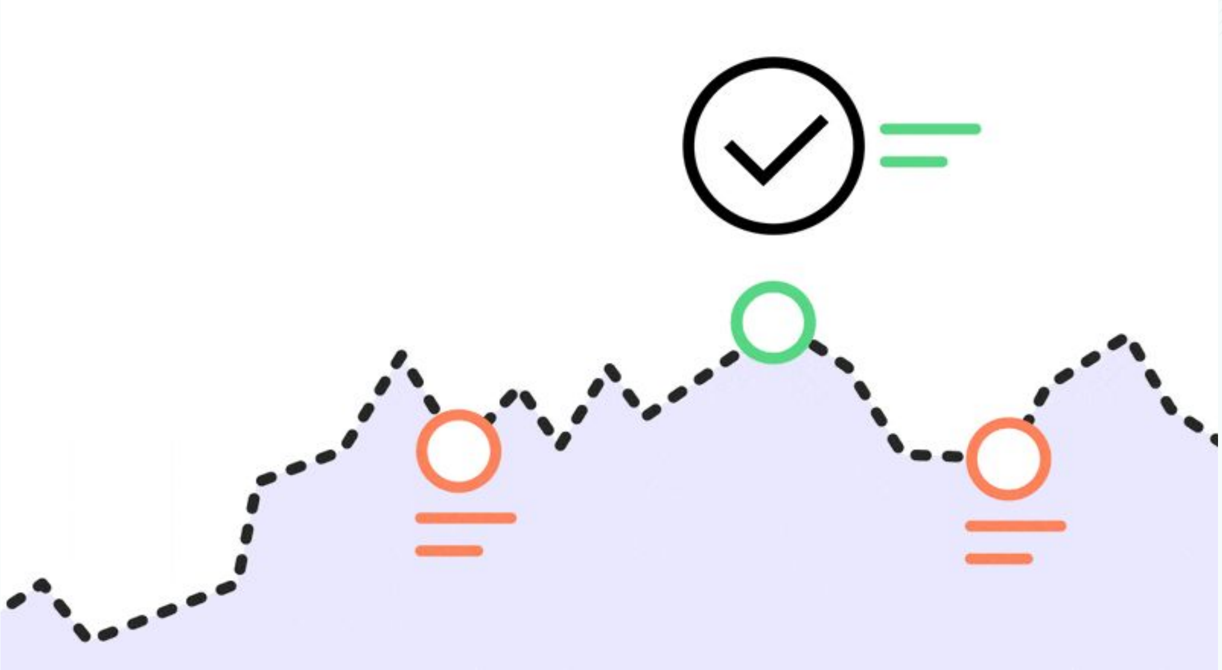
high # Decimal

Valid	851k	100%
Mismatched	0	0%
Missing	0	0%
Mean	65.6	
Std. Deviation	75.9	
Quantiles	1.81	Min

volume # Decimal

Valid	851k	100%
Mismatched	0	0%
Missing	0	0%
Mean	5.42m	
Std. Deviation	12.5m	
Quantiles	0	Min
	1.22m	25%
	2.48m	50%
	5.22m	75%
	860m	Max

Data Problems



2. Outliers Detection

Outliers are extreme values that deviate from other observations on data , they may indicate a variability in a measurement, experimental errors or a novelty.

Box Plot

A boxplot is a standardized way of displaying the distribution of data based on a five number summary (“**Minimum**”, **First Quartile (Q1)**, **Median (Q2)**, **Third Quartile (Q3)**, and “**Maximum**”) whereas, **IQR = Q3-Q1**

Range: **Q1- 1.5* IQR** to **Q3 + 1.5*IQR**

3. Standardization - Feature Scaling

- **Feature scaling** is a method used to normalize the range of independent variables or features of data

- We have normalize the dataset using standardization mechanism

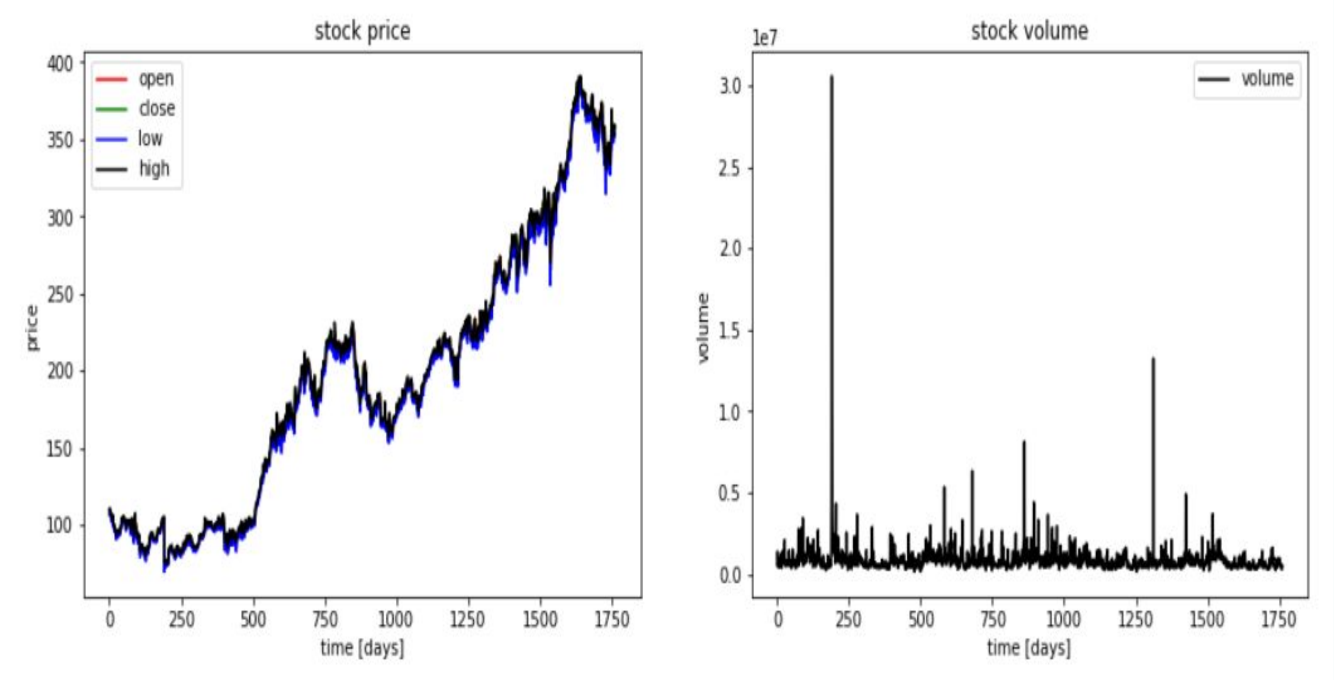
- Feature **standardization** makes the values of each feature in the data have zero-mean (when subtracting the mean in the numerator) and unit-variance.

Feature scaling the vector for better model performance.

In [45]:

```
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler(feature_range=(0, 1))
stocks = scaler.fit_transform(stocks)
```

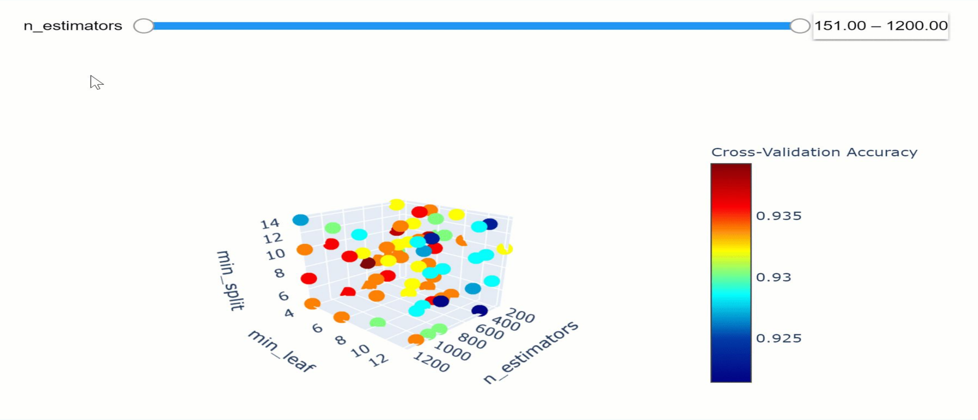
Data Features Relationship Visualization



- Snapshot of **Walmart Stock**
- Showing **relationship** between **Price** and **Time** features
- Showing relationship between **Volume** and **Time** features

Data Splitting

```
In [48]:  1  train = int(len(stocks) * 0.80)
          2  test = len(stocks) - train
```



Dataset is divided into two groups
Training dataset and Test dataset

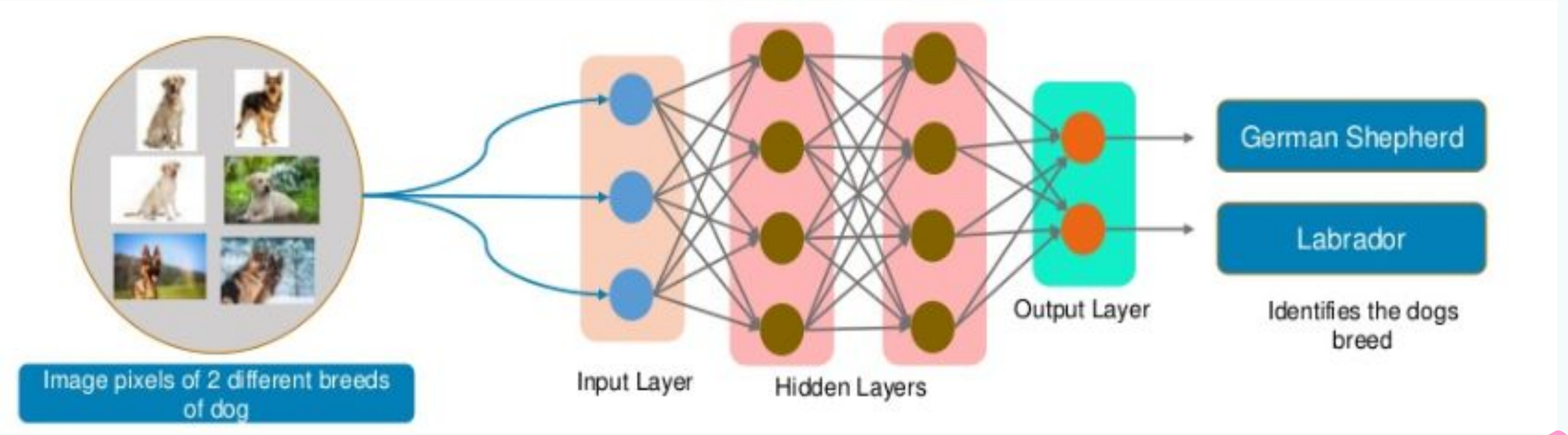
- **Total Samples=** 851,264
- **Training Samples=** 80%
- **Testing Samples=** 20%

MODELING

03

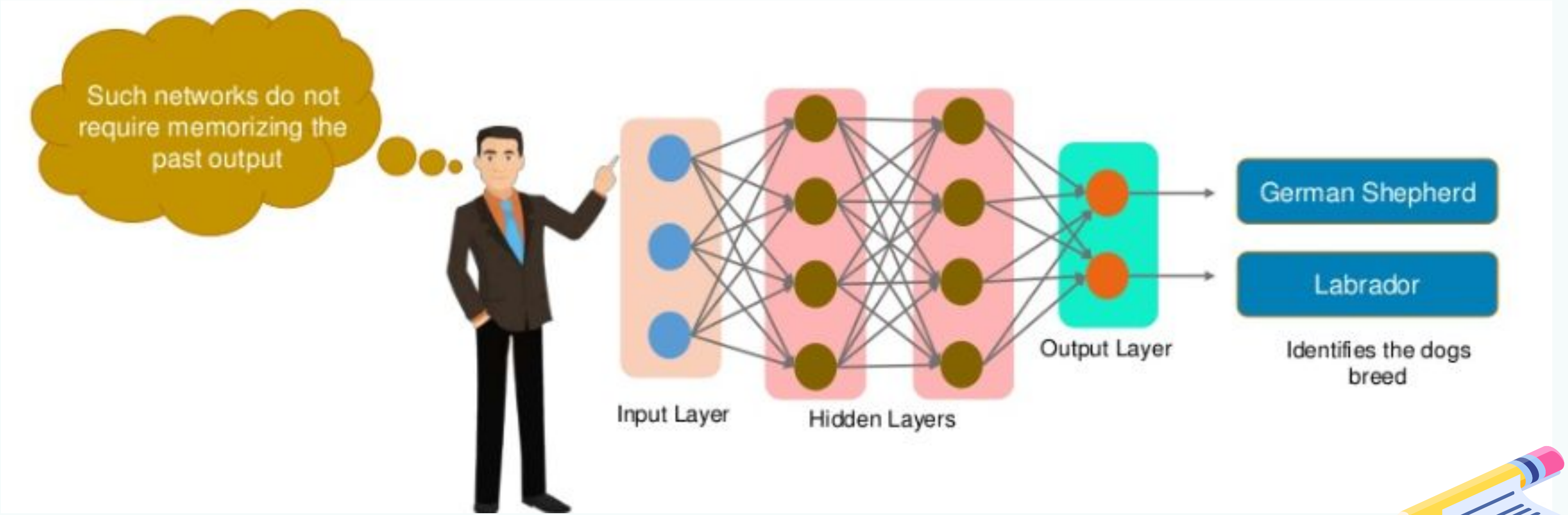
What is a Neural Network?

Neural Networks used in Deep Learning, consists of Different layers connected to each other and work on the structure and functions of a human brain.

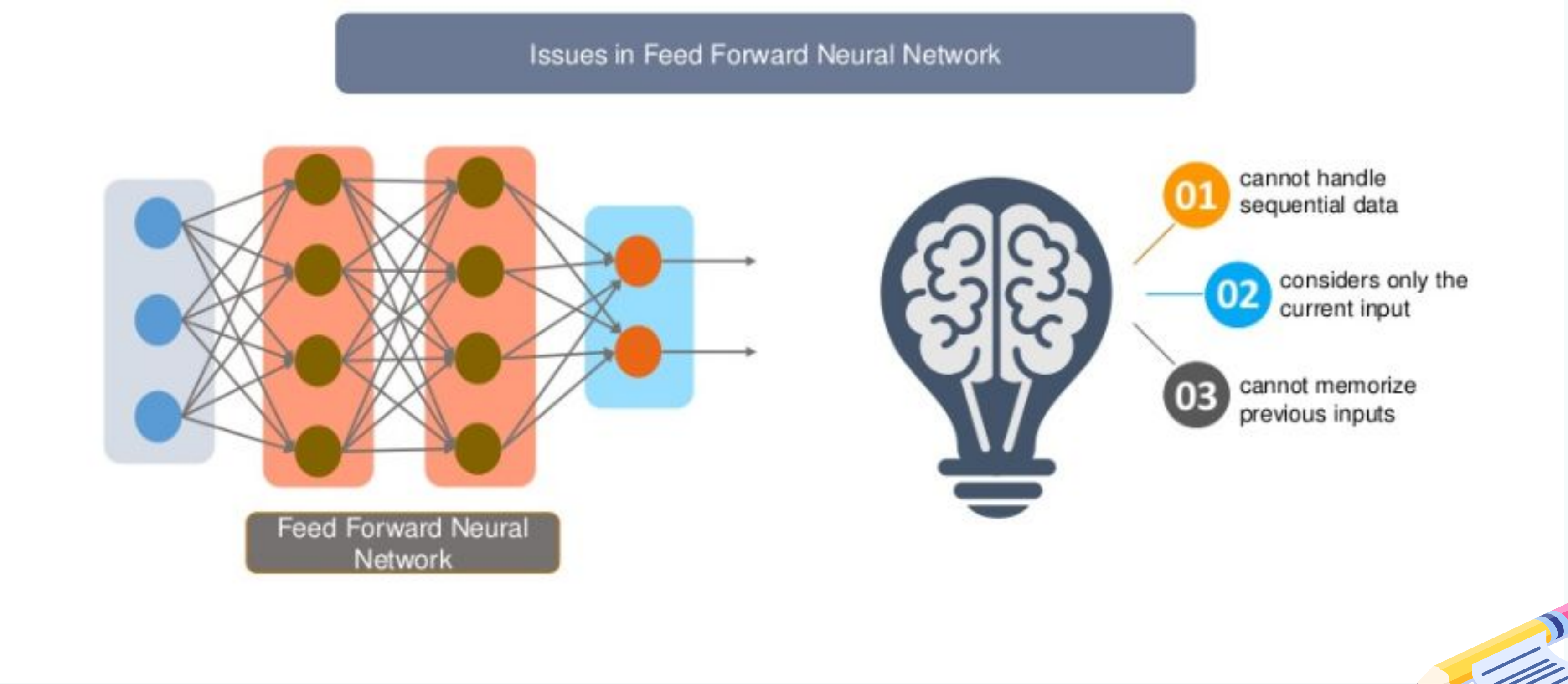


What is a Neural Network?

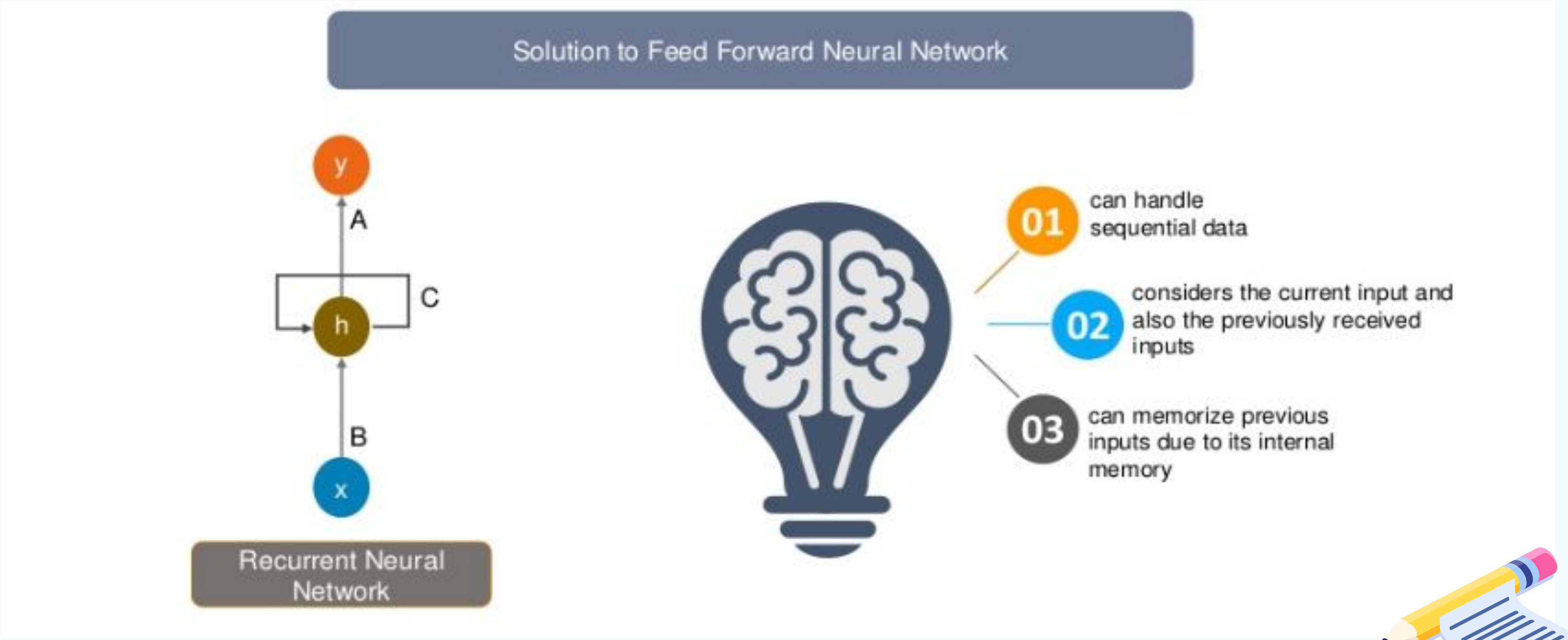
Neural Networks used in Deep Learning, consists of Different layers connected to each other and work on the structure and functions of a human brain.



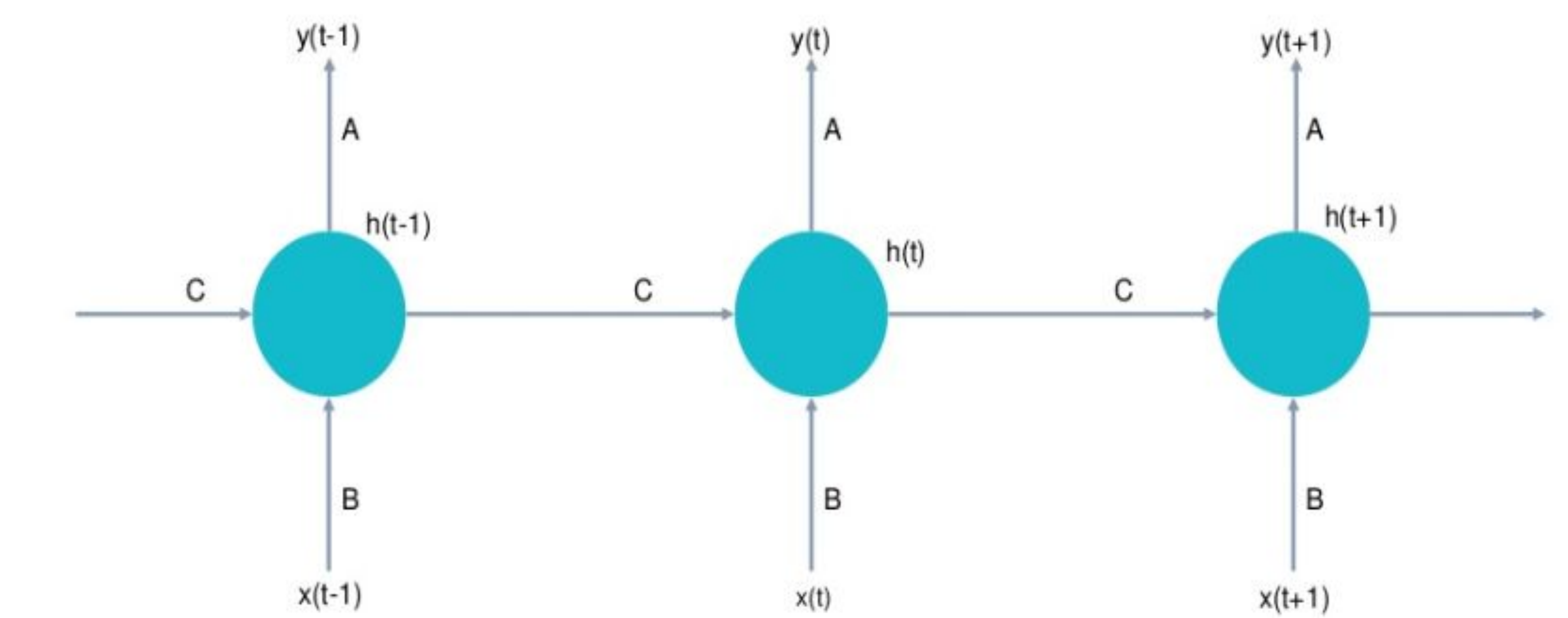
Why Recurrent Neural Network?



Why Recurrent Neural Network?



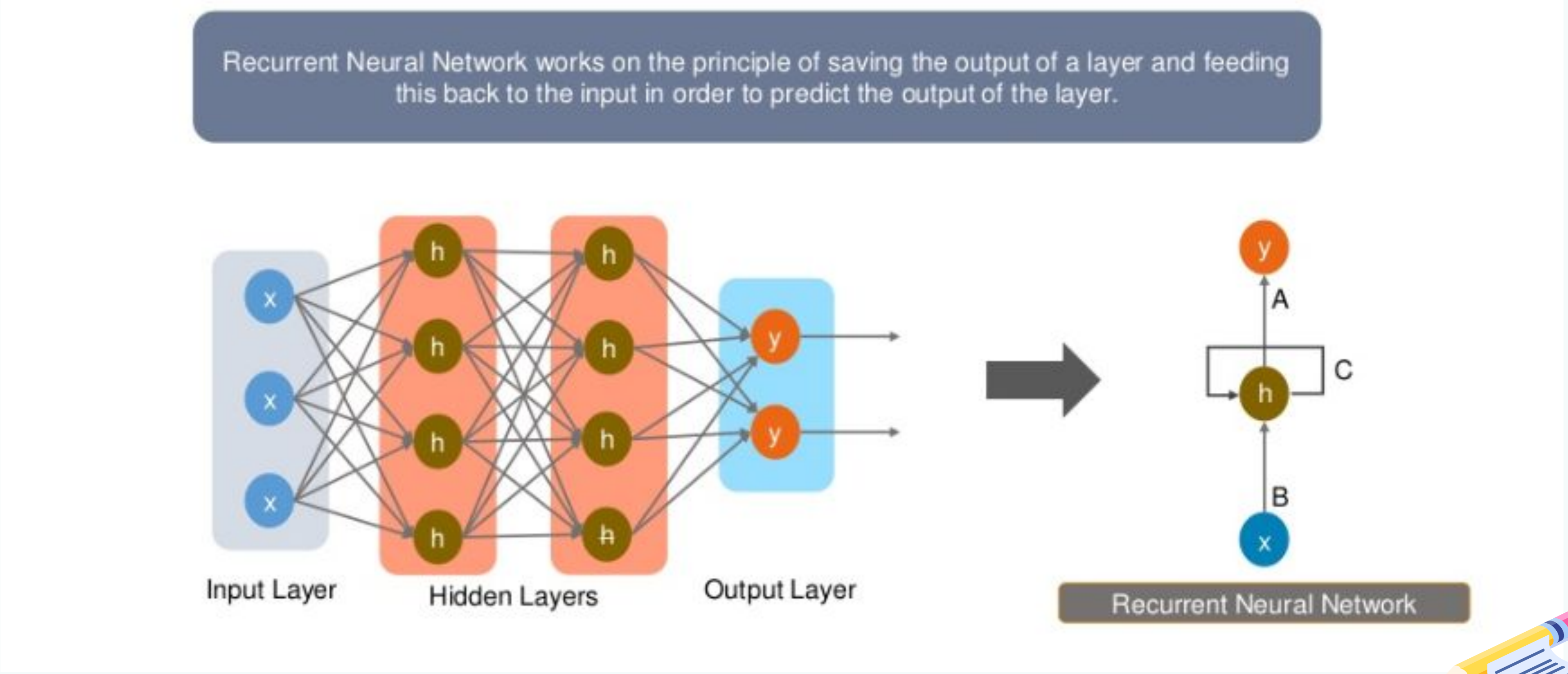
How does a RNN work?



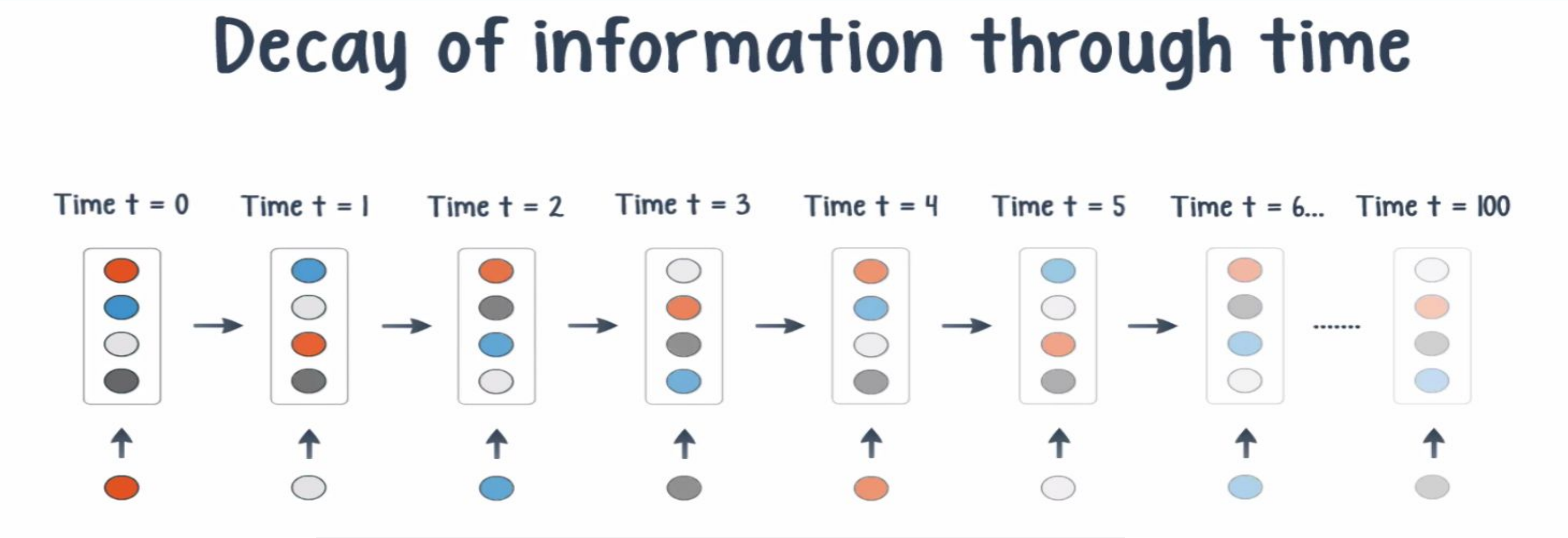
$$h(t) = f_c\,(h(t-1)\ ,\ x(t))$$

$h(t)$ = new state
f_c = function with parameter c
$h(t-1)$ = old state
$x(t)$ = input vector at time step t

What is a Recurrent Neural Network?



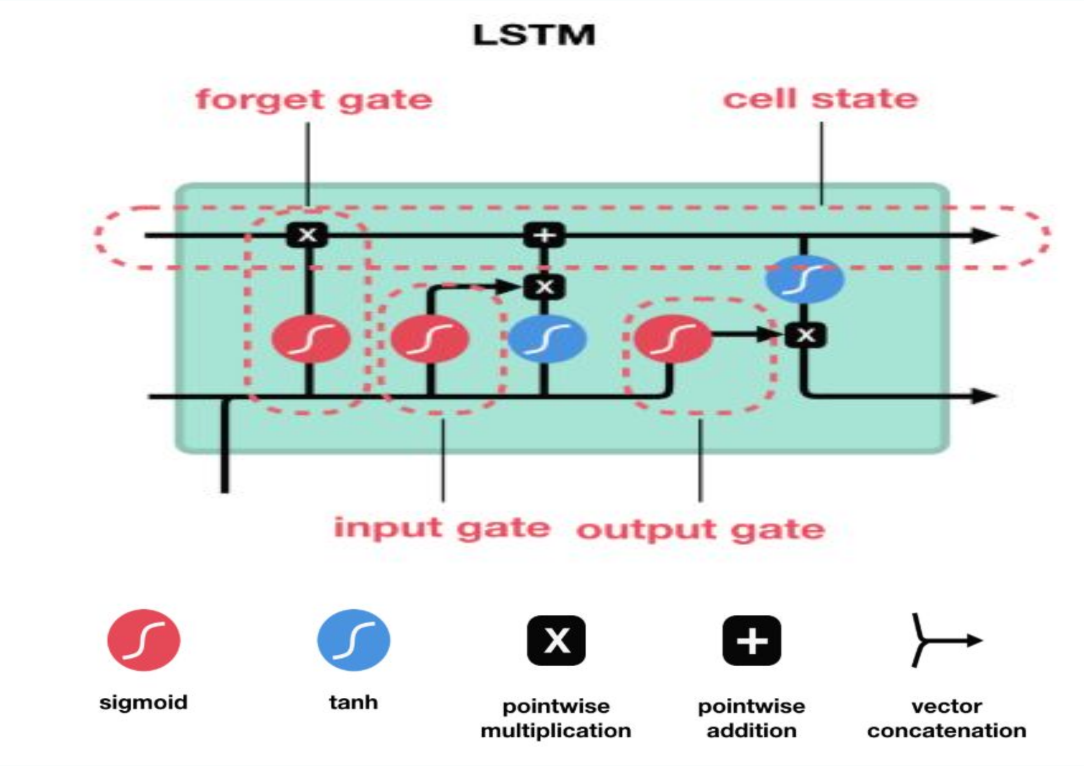
Problems with Traditional RNN



Gradient Update Rule

new weight = weight - learning rate*gradient

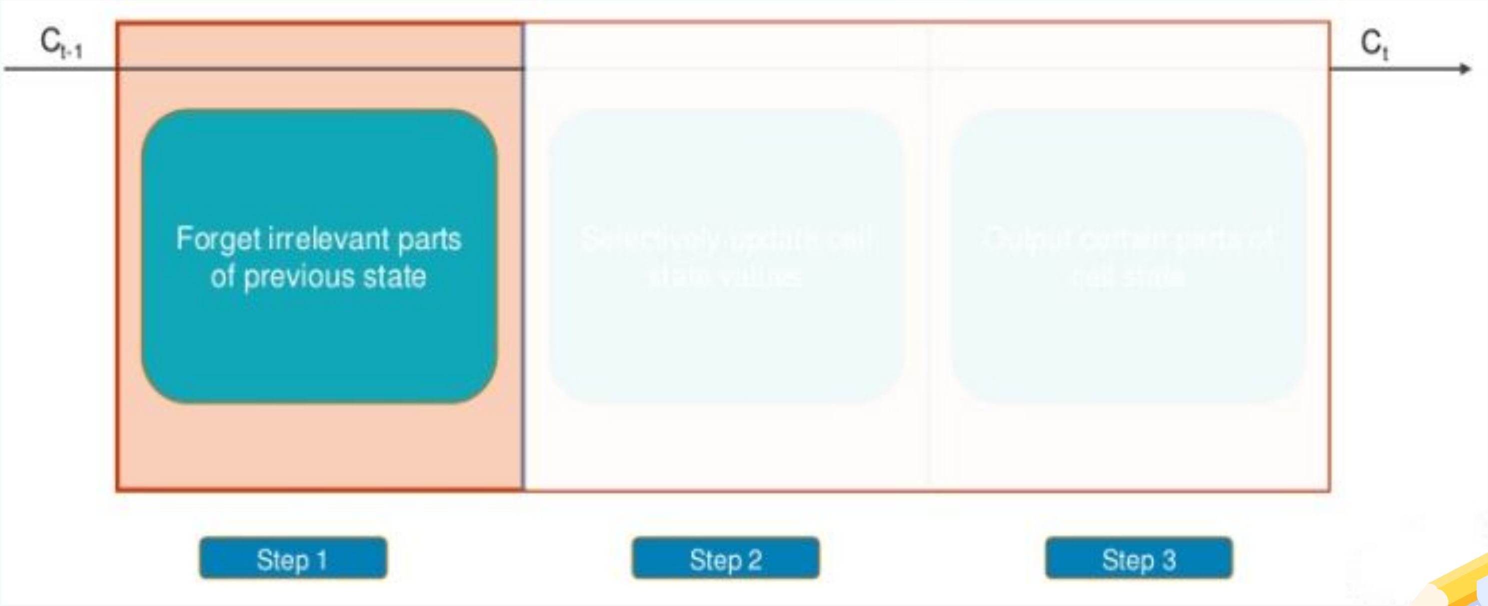
Long Short-Term Memory Network (LSTM)



- Similar to RNN in structure except the gates integration in LSTM.
- It contains four gates irrespective to the LSTM i.e.,
  - Input gate
  - Forget gate
  - Cell State gate
  - Output gate

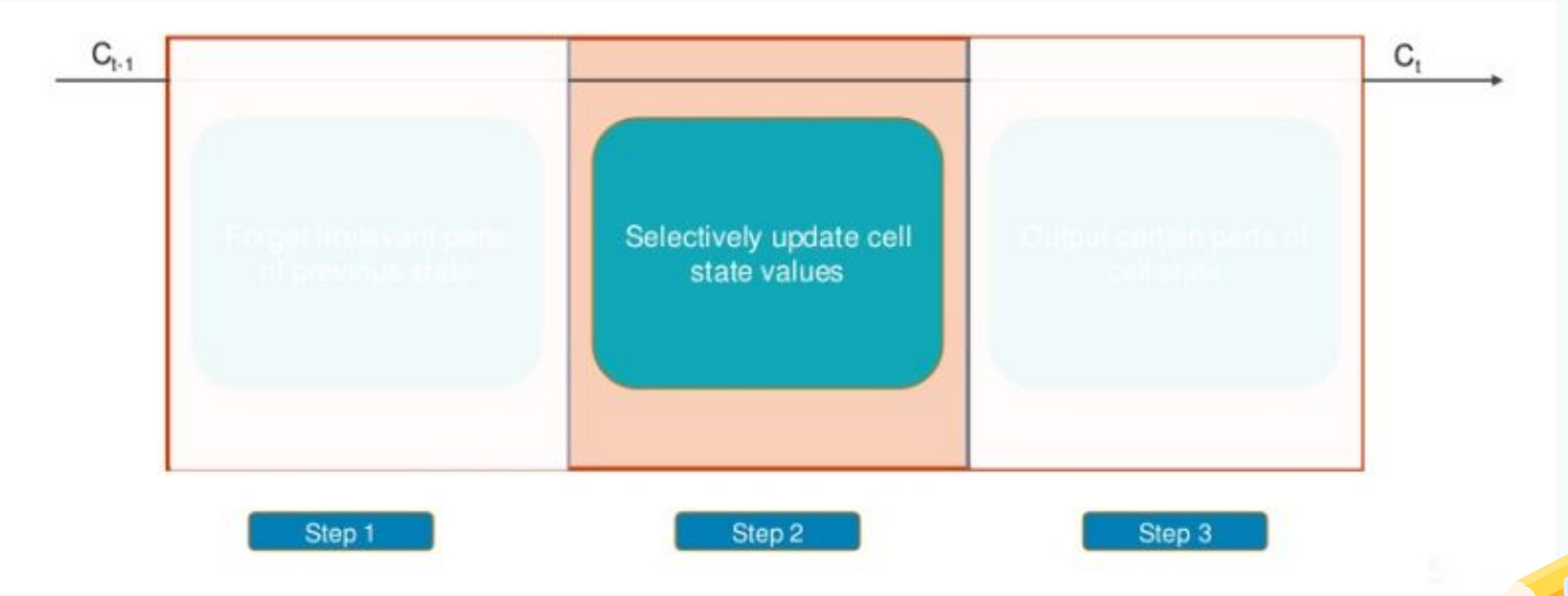
Long Short-Term Memory Network (LSTM)

3 step process of LSTMs



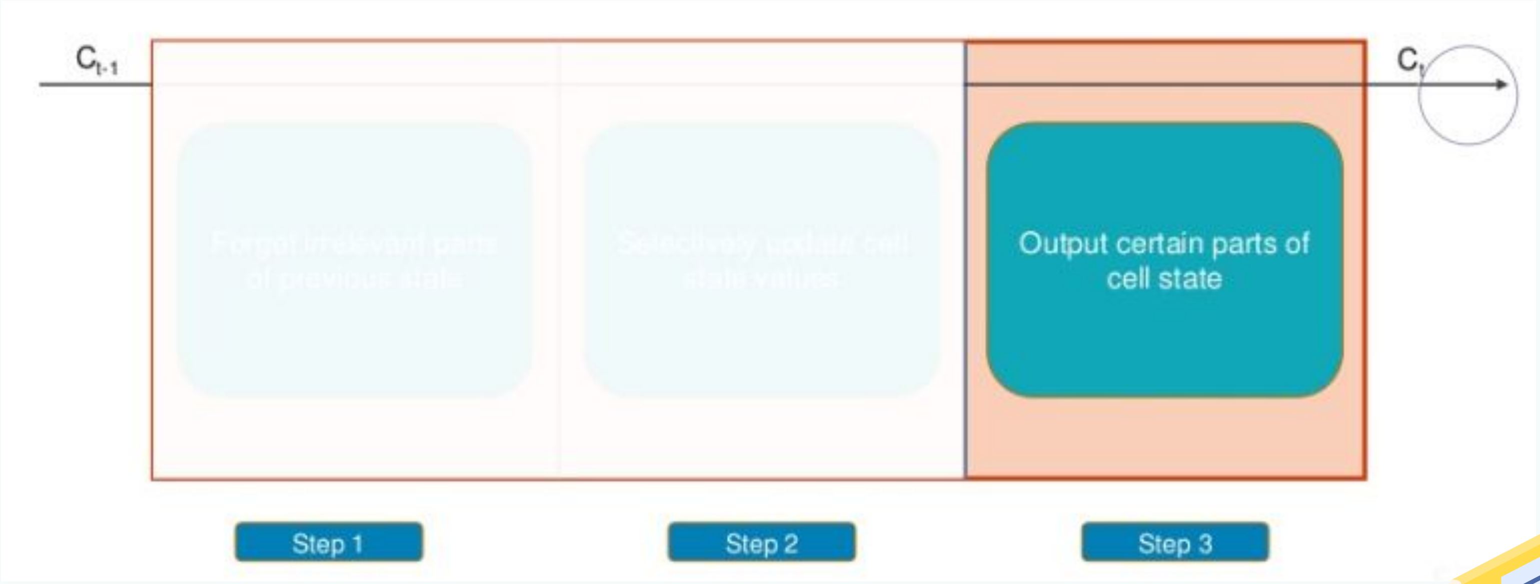
Long Short-Term Memory Network (LSTM)

3 step process of LSTMs

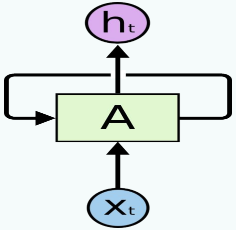


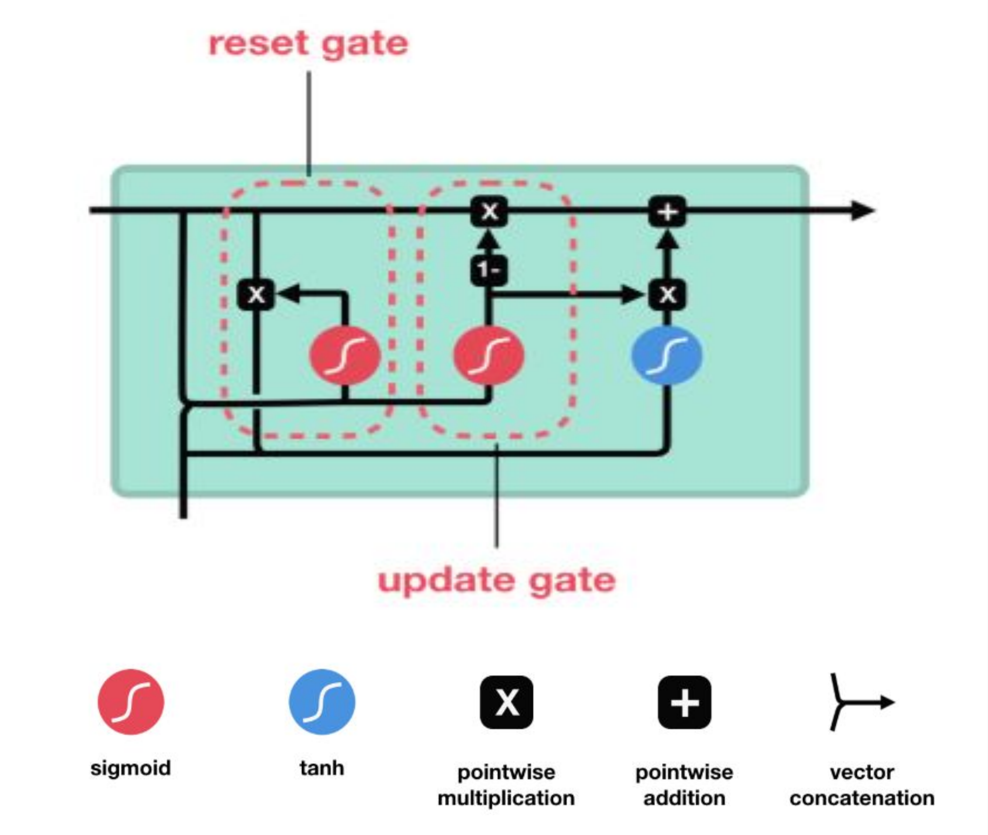
Long Short-Term Memory Network (LSTM)

3 step process of LSTMs



Gated Recurrent Unit (GRU)





- Similar to LSTM in structure.
- It contains only two gates irrespective to the LSTM i.e.,
  - a reset gate
  - a update gate.
- Majorly used in time-series analysis problems.

Model 1 (LSTM)

SPECIFICATION

- **Inputs** = 4
- **Hidden Layers (LSTM based)**= 2 (256, 128 Neurons)
- **N_outputs** = 4

```
model = Sequential()
model.add(LSTM(256))
model.add(Dense(128 ,  activation = 'relu'))
model.add(Dense(4))
print(model.summary())
```

Model 2 (GRU)

SPECIFICATION

- **Inputs** = 4
- **Hidden Layers (GRU based)** = 2 (256, 128 Neurons)
- **Outputs** = 4

```
model = Sequential()
model.add(GRU(256))
model.add(Dense(128 ,  activation = 'relu'))
model.add(Dense(4))
print(model.summary())
```

Model Tuning Parameters

MODEL 1 (LSTM)

- **Activation Function:** Rectified Linear Units (ReLU)
- **Weight Optimizer:** Adam Optimizer
- **Kernel Initializer:** Normal Distribution
- **Epochs:** 100
- **Batch_Size:** 128

MODEL 2 (GRU)

- **Activation Function:** Rectified Linear Units (ReLU)
- **Weight Optimizer:** Adam Optimizer
- **Kernel Initializer:** Normal Distribution
- **Epochs:** 100
- **Batch_Size:** 128

```
model.compile(loss='mean_squared_error', optimizer=Adam(lr = 0.0005) , metrics = ['mean_squared_error'])
```

MODEL EVALUATION

04

Model Evaluation Measures

1. **Mean Square Error (MSE)**

   MSE measures the average magnitude of the errors in a set of predictions, without considering their direction.

$$MSE = \frac{1}{n}\Sigma\underbrace{\left(y - \widehat{y}\right)^2}_{\text{The square of the difference between actual and predicted}}$$

2. **Root Mean Square Error (RMSE)**

   RMSE is a quadratic scoring rule that also measures the average magnitude of the error.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}$$

Model Evaluation Measures

Model 1 (LSTM)

- **Mean Square Error (MSE)**

  0.0045 MSE

- **Root Mean Square Error (RMSE)**

  0.067 RMSE

Model 2 (GRU)

- **Mean Square Error (MSE)**

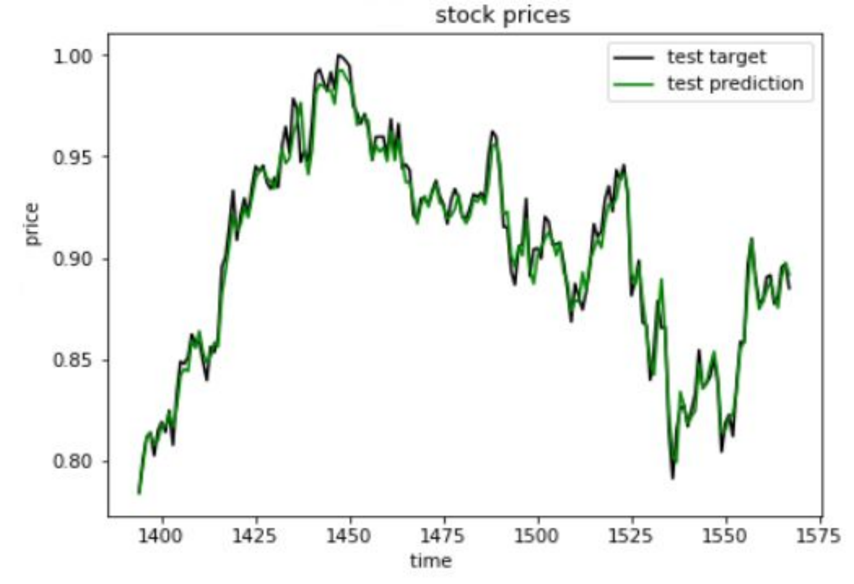
  0.00033 MSE

- **Root Mean Square Error (RMSE)**

  0.02 RMSE

Model Evaluation Measures

Model 1 (LSTM)



Model 2 (GRU)

COMPARISON

05

Comparison in LSTM and GRU

- GRUs are computationally more efficient than LSTM because of less complex structure.

- We calculated MSE and RMSE's for GRUs and LSTM and observed that GRU's has three times less value as compared to LSTM.

- GRUs use less training parameters as compared to LSTM and hence they use less memory and can train and execute faster.

- LSTMs are generally used where long-distance relations modelling is involved.

CONCLUSION

06

Key Take-Away

- GRUs are simpler and thus easier to modify, for example adding new gates in case of additional input to the network. It's just less code in general

- GRUs train faster and perform better than LSTMs on less training data if you are doing stock prediction as per this project

- The GRU controls the flow of information like the LSTM unit, but without having to use a memory unit.

Future Scope

- For better result we can use both the techniques LSTM and GRU to overcome each others drawbacks

- Use of sentimental analysis along with the two above discussed methods, sentimental analysis refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source material

- If the news is positive, then there are good chances that the share prices of that company will go up

- Similarly, if the news article has an overall negative sentiment, the share prices of the company or industry can go down

References

[1] Elliott, Larry. “The Stock Market Turmoil Was All about Good Economic News.” The-Stock-Market-Turmoil-Was-All-about-Good-Economic-News, The Gaurdian, 11 Feb. 2018, www.theguardian.com/business/2018/feb/11/the-stock-market-turmoil-was-all-about-good-economic-news

[2] “National Stock Exchange of India.” Wikipedia, Wikimedia Foundation, 14 Mar. 2018,en.wikipedia.org/wiki/National_Stock_Exchange_of_India.

[3] “HDFC Bank.” Wikipedia, Wikimedia Foundation, 16 Mar. 2018, en.wikipedia.org/wiki/HDFC_Bank.

[4] “Stock Share Price Housing Development Finance Corpltd | Get Quote Hdfc | BSE.” BSE Ltd, www.bseindia.com/stock-share-price/housing-development-finance-corpltd/hdfc/500010/.

[5] “Logistic Regression.” Wikipedia, Wikimedia Foundation, 17 Mar. 2018, en.wikipedia.org/wiki/Logistic_regression.

[6] Haykin, Simon. Neural Networks. 2nd ed., vol. 1 1, Prentice Hall, 1998.

References

[7] Akita, R., Yoshihara, A., Matsubara, T., & Uehara, K. (2016). Deep learning for stock prediction using numerical and textual information. 2016 IEEE/ACIS 15th International Conference on Computer and Information Science, ICIS 2016 - Proceedings. https://doi.org/10.1109/ICIS.2016.7550882

[8] Xu, M., Lan, Y., & Jiang, D. (2016). Unsupervised Learning Part-Based Representation for Stocks Market Prediction. Proceedings - 2015 8th International Symposium on Computational Intelligence and Design, ISCID 2015, 2, 63–66. https://doi.org/10.1109/ISCID.2015.300

[9] Mithani, F., Machchhar, S., & Jasdanwala, F. (2017). A modified BPN approach for stock market prediction. 2016 IEEE International Conference on Computational Intelligence and Computing Research, ICCIC 2016, 0–3. https://doi.org/10.1109/ICCIC.2016.7919718

[10] R. Choudhry and K. Garg, "A Hybrid Machine Learning System for Stock Market Forecasting," World Acad. Sci. Eng. Technol., vol. 2, no. 15, pp. 315–318, 2008.

Project Repository for the Reference

https://github.com/rarpit1994/Stock-Market-Prediction

THANKS

TEAM

- Rupali Wadhwani
- Sourabh Agarwal
- Arpit Rathore
- Harman Bhutani
- Rahul Raj Sharma

QUESTIONS?