
Does Sentiment Analysis of Social Media contradict the Efficient Market Hypothesis?

A presentation by Niclas Joswig

What do you think gets traded on a stock exchange?



Foreign Exchange Market

—

What do you think gets traded on a stock exchange?

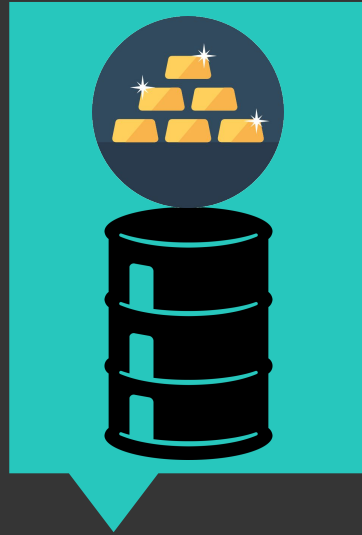


Company Shares



Foreign Exchange Market

What do you think gets traded on a stock exchange?



Commodities



Company Shares

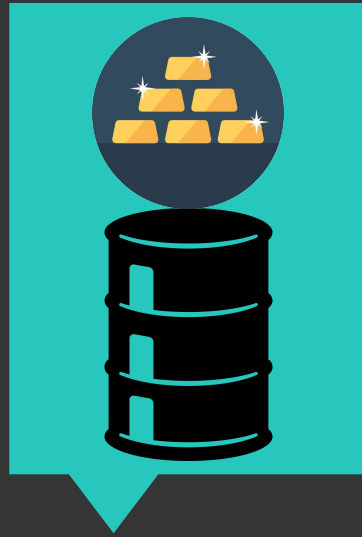


Foreign Exchange Market

What do you think gets traded on a stock exchange?



Indices



Commodities



Company Shares



Foreign Exchange Market

Efficient Market Hypothesis

- The **efficient-market hypothesis** is a theory in financial economics that states that asset prices fully reflect all available information.

A direct implication is that it is impossible to "beat the market" consistently on a risk-adjusted basis since market prices should only react to new information.

x1Down veröffentlicht auf TradingView.com, April 07, 2019 14:09:34 EEST
BATS:AAPL, 1D 197.00 ▲ +1.31 (+0.67%) O:196.45 H:197.10 L:195.93 C:197.00



How do the investors feel about ...

The current
state of the
company

The economic
state of the
country or the
whole world

The branch of
industry the
company is
involved in

```
graph TD; A[The current state of the company] --> D[Social Media]; B[The economic state of the country or the whole world] --> D; C[The branch of industry the company is involved in] --> D;
```

Social Media

Neural Networks

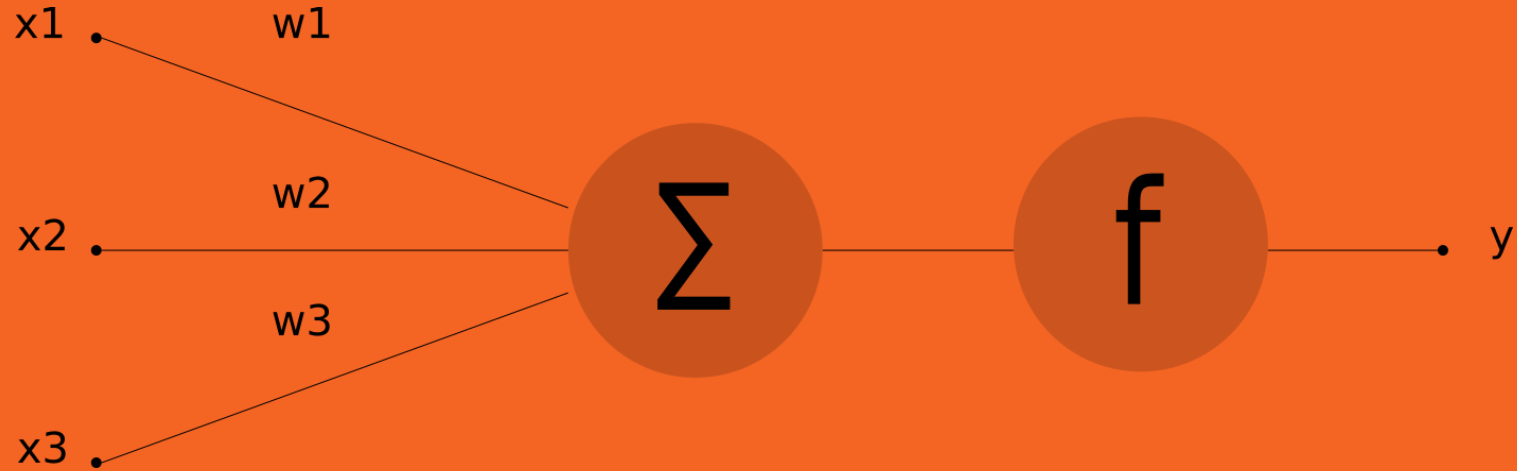
```
graph TD; NN[Neural Networks] --> W2V[Word-2-Vector]; NN --> RN[Recurrent Networks]; RN --> LSTM[LSTM Networks]
```

Word-2-Vector

Recurrent Networks

LSTM Networks

The Neuron



$$y = \delta \sum_{j=0}^n x_j * w_j$$

$$y = \delta(x^T \cdot w)$$

Word-2-Vector

What do we want from the transformation?

Word-2-Vector

What do we want from the transformation?

- Different forms of one word → same vector

Word-2-Vector

What do we want from the transformation?

- Different forms of one word → same vector
- Words with similar meanings → similar vectors

Word-2-Vector

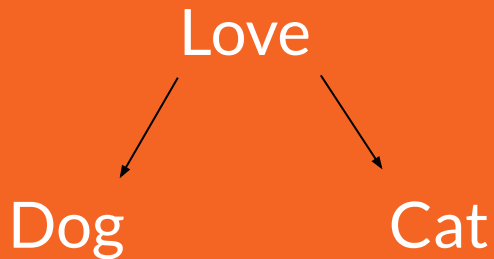
What do we want from the transformation?

- Sparsity
- Different forms of one word \rightarrow same vector
- Words with similar meanings \rightarrow similar vectors
- Perform arithmetic operations on vectors
 \rightarrow Finland - Helsinki = Germany - Berlin

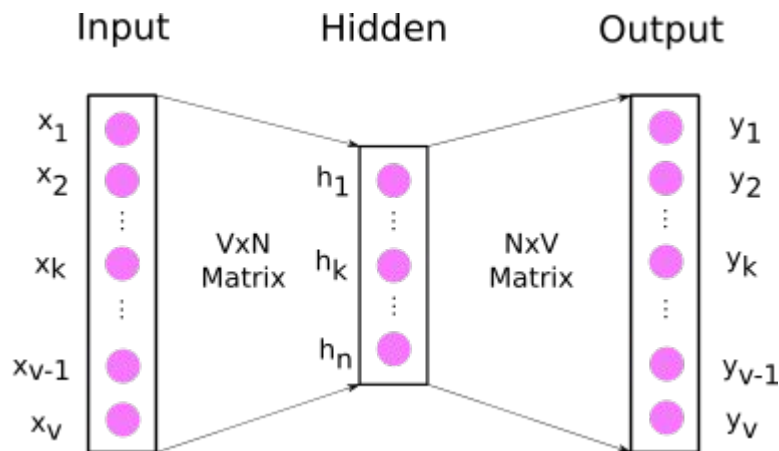
Word-2-Vector

What do we want from the transformation?

- Different forms of one word -> same vector
- Words with similar meanings -> similar vectors
- Perform arithmetic operations on vectors
→ Finland - Helsinki = Germany - Berlin

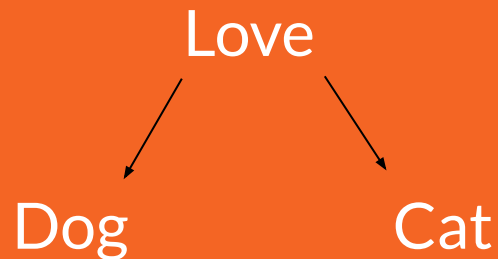


Word-2-Vector

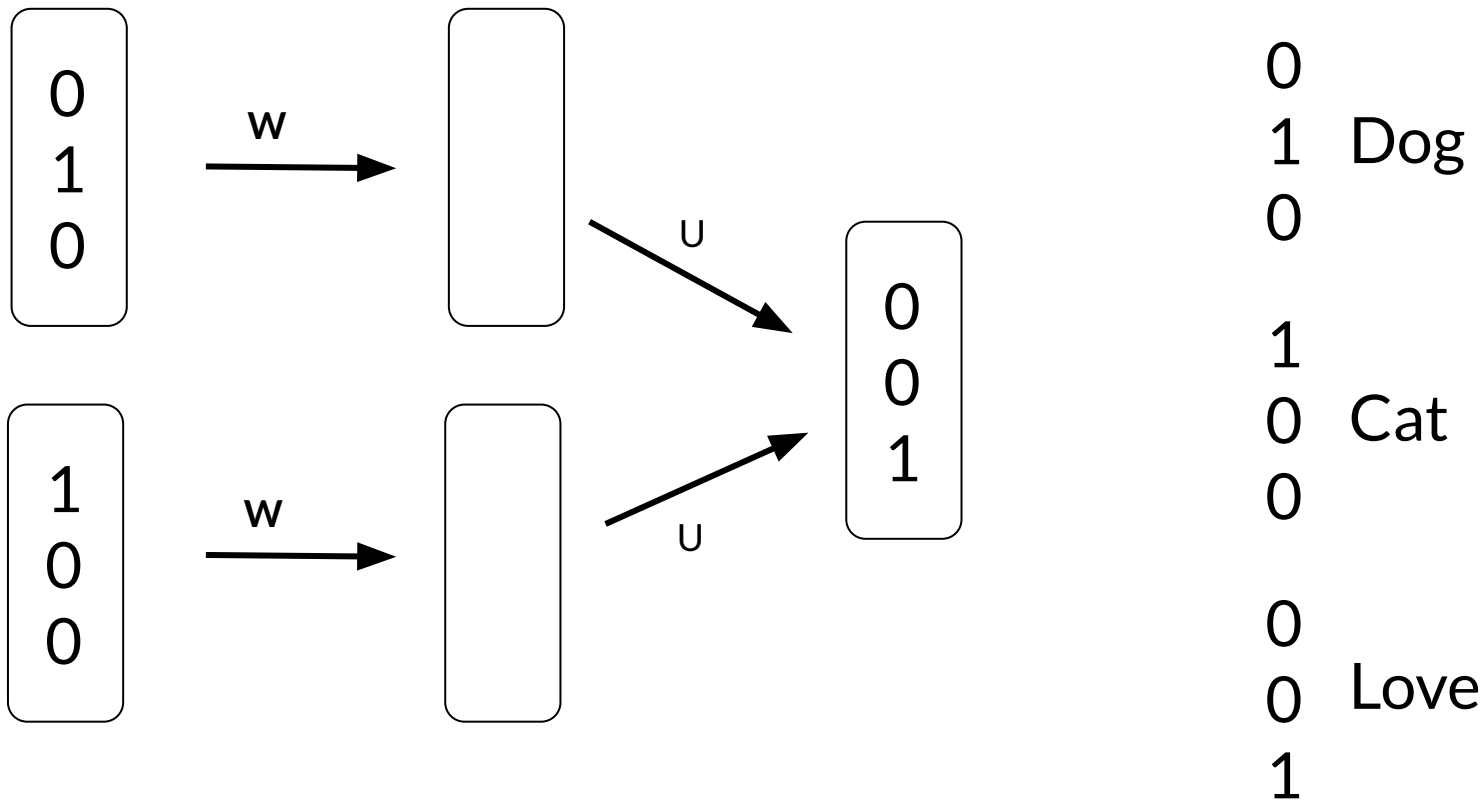


What do we want from the transformation?

- Different forms of one word \rightarrow same vector
- Words with similar meanings \rightarrow similar vectors
- Perform arithmetic operations on vectors
 \rightarrow Finland - Helsinki = Germany - Berlin

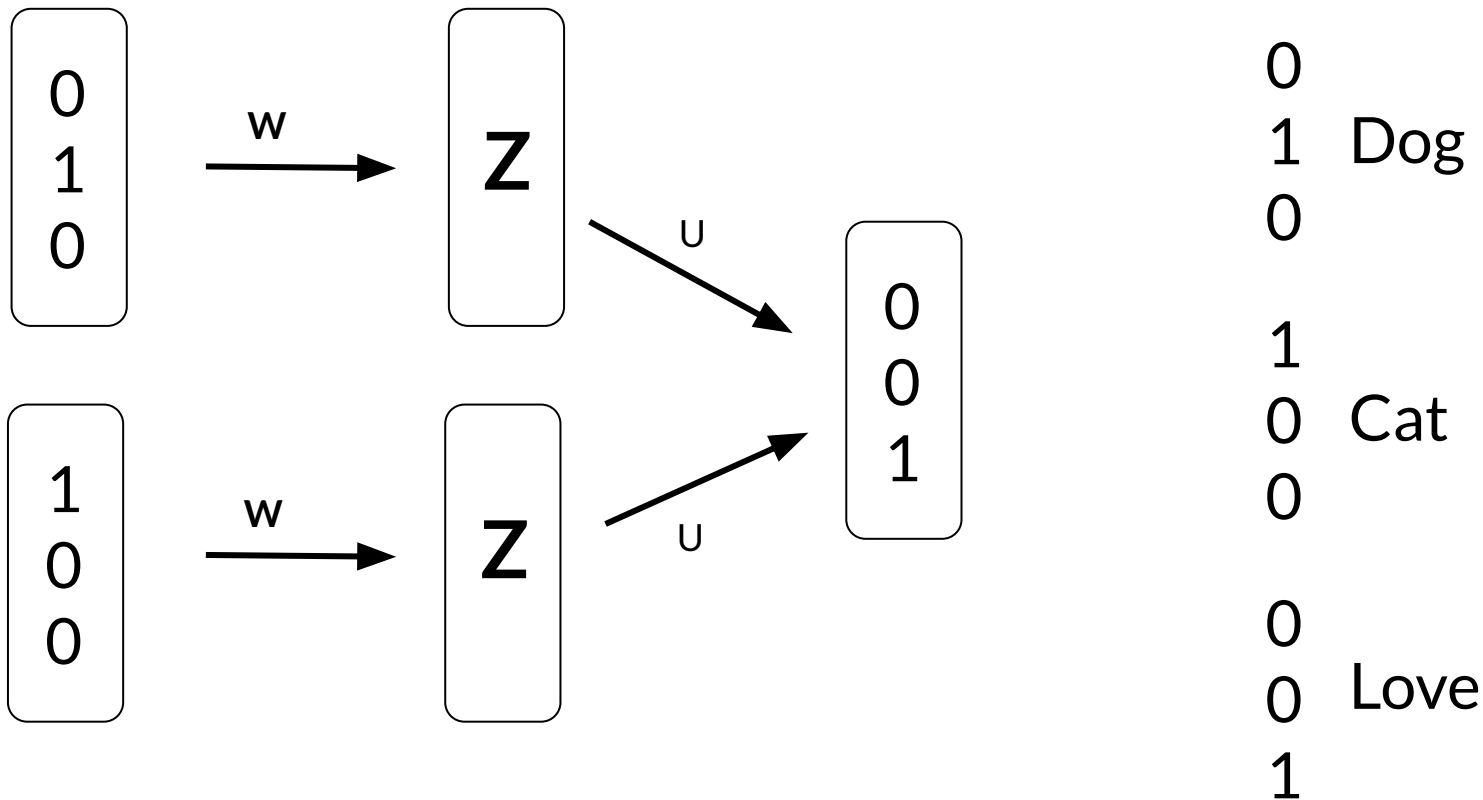


—



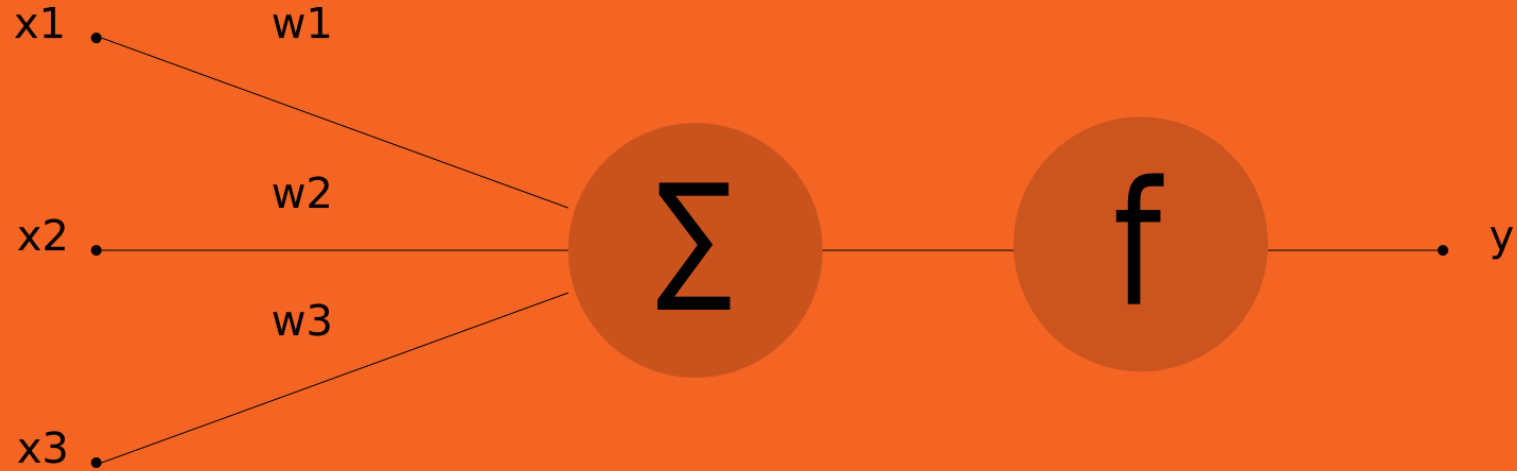
Cat & Dog Word-2-Vector Example

—



Cat & Dog Word-2-Vector Example

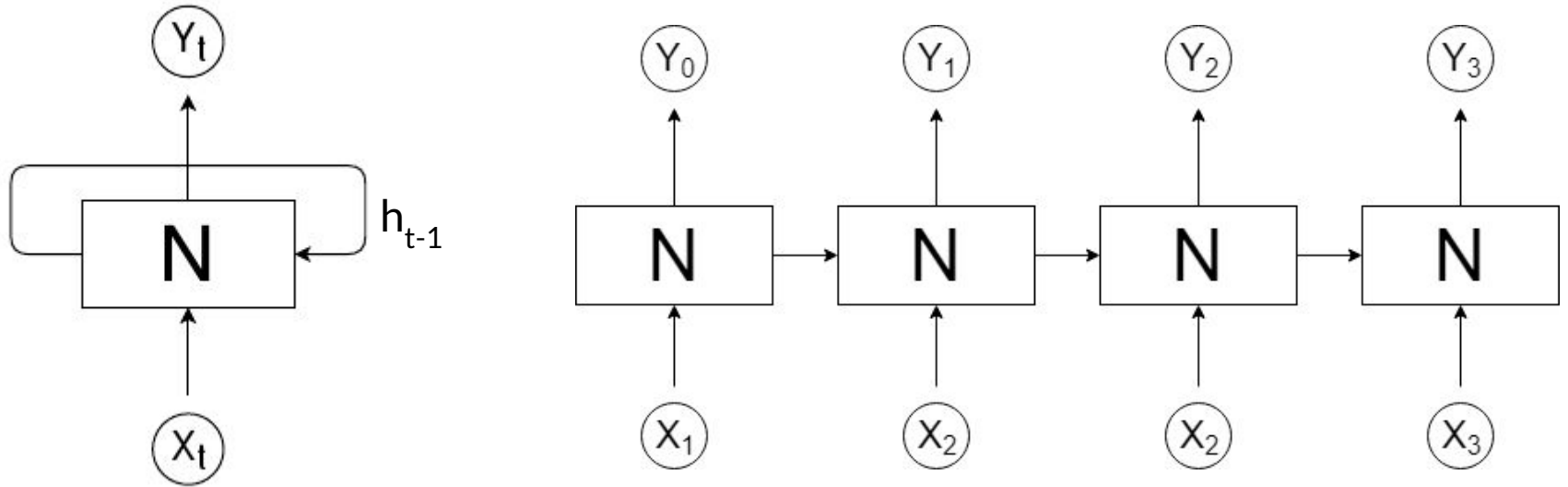
The Neuron



$$y = \delta \sum_{j=0}^n x_j * w_j$$

$$y = \delta(x^T \cdot w)$$

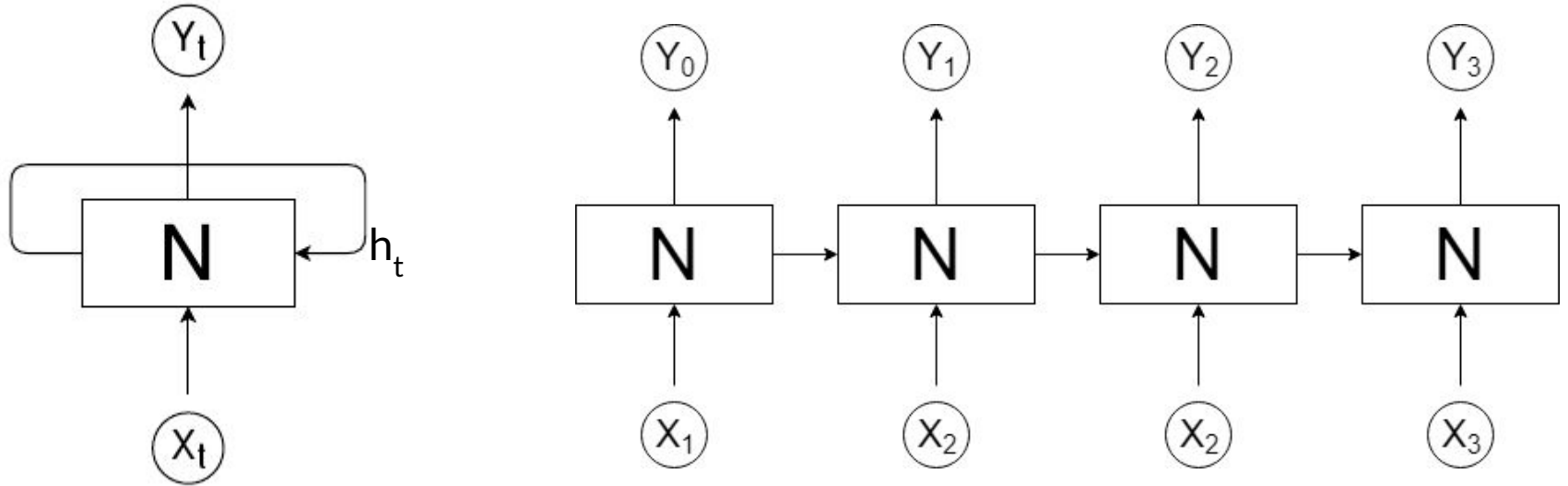
Recurrent Neuron



$$h_t = \delta(W * x_t + U * h_{t-1})$$

$$Y_t = \delta(V * h_t)$$

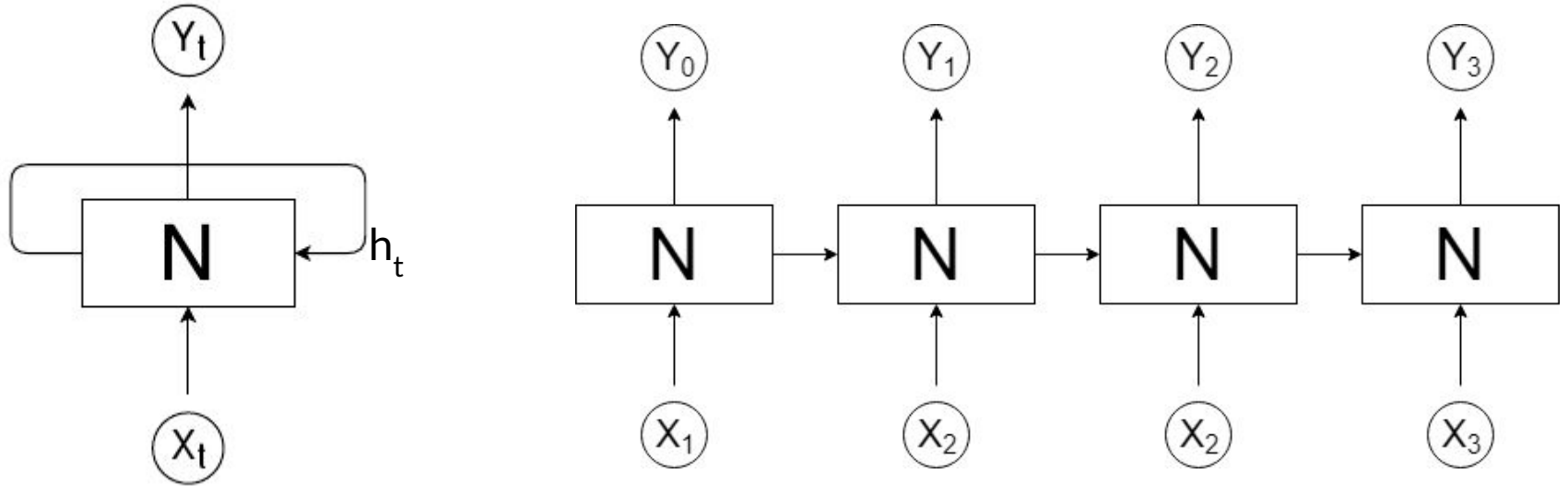
Recurrent Neuron



$$h_t = \delta(W * x_t + U * h_{t-1})$$

$$Y_t = \delta(V * h_t)$$

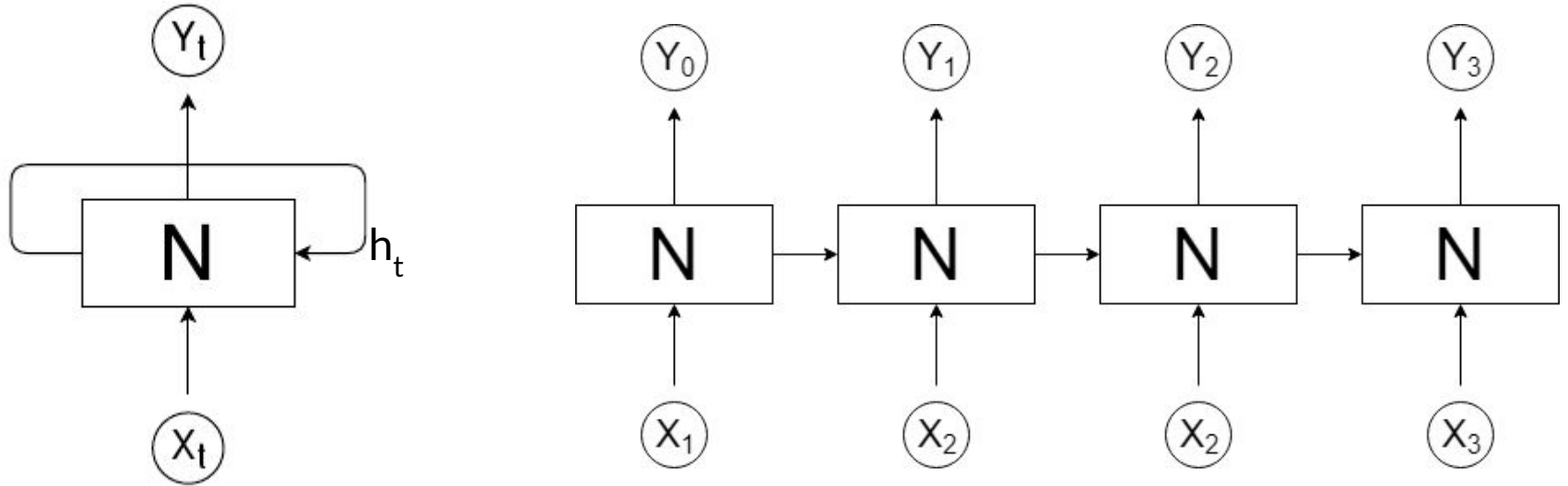
Recurrent Neuron



$$h_t = \delta(W * x_t + U * h_{t-1})$$

$$Y_t = \delta(V * h_t)$$

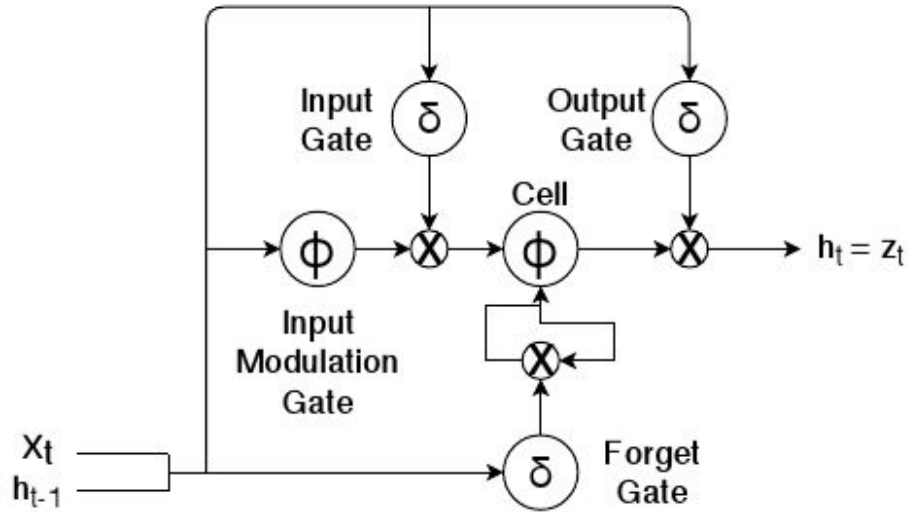
Recurrent Neuron



$$h_t = \delta(W * x_t + U * h_{t-1})$$

$$Y_t = \delta(V * h_t)$$

LSTM Cell



Variables

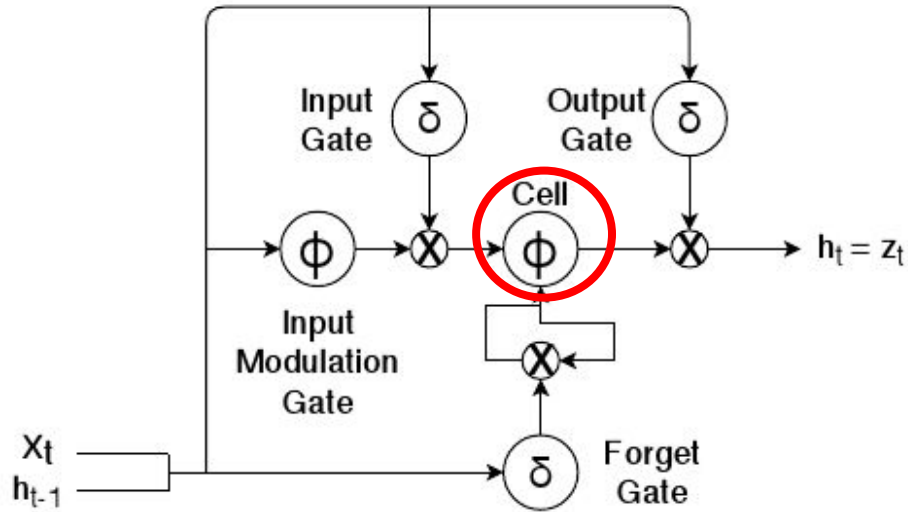
X_t = Input

h_{t-1} = Output from previous step

Z_t = Output

Cell $\hat{=}$ Matrix

LSTM Cell



Variables

X_t = Input

h_{t-1} = Output from previous step

Z_t = Output

Cell $\hat{=}$ Matrix

Two Approaches

Stock Market Prediction Using Neural Networks through News on Online Social Networks by *W. Chen, Y. Zhang, C. Yeo, C. T. Lau and B. Lee*

- Predict exact price of Shenzhen 500 Index
- Base LSTM Model predicts with past Stock data
- Expand model with News Influence Factor
- Evaluate whether News information improve prediction



Predict Effect of Trump's Tweets on Stock Price Milestone by *Tong Yang and Yuxin Yang*

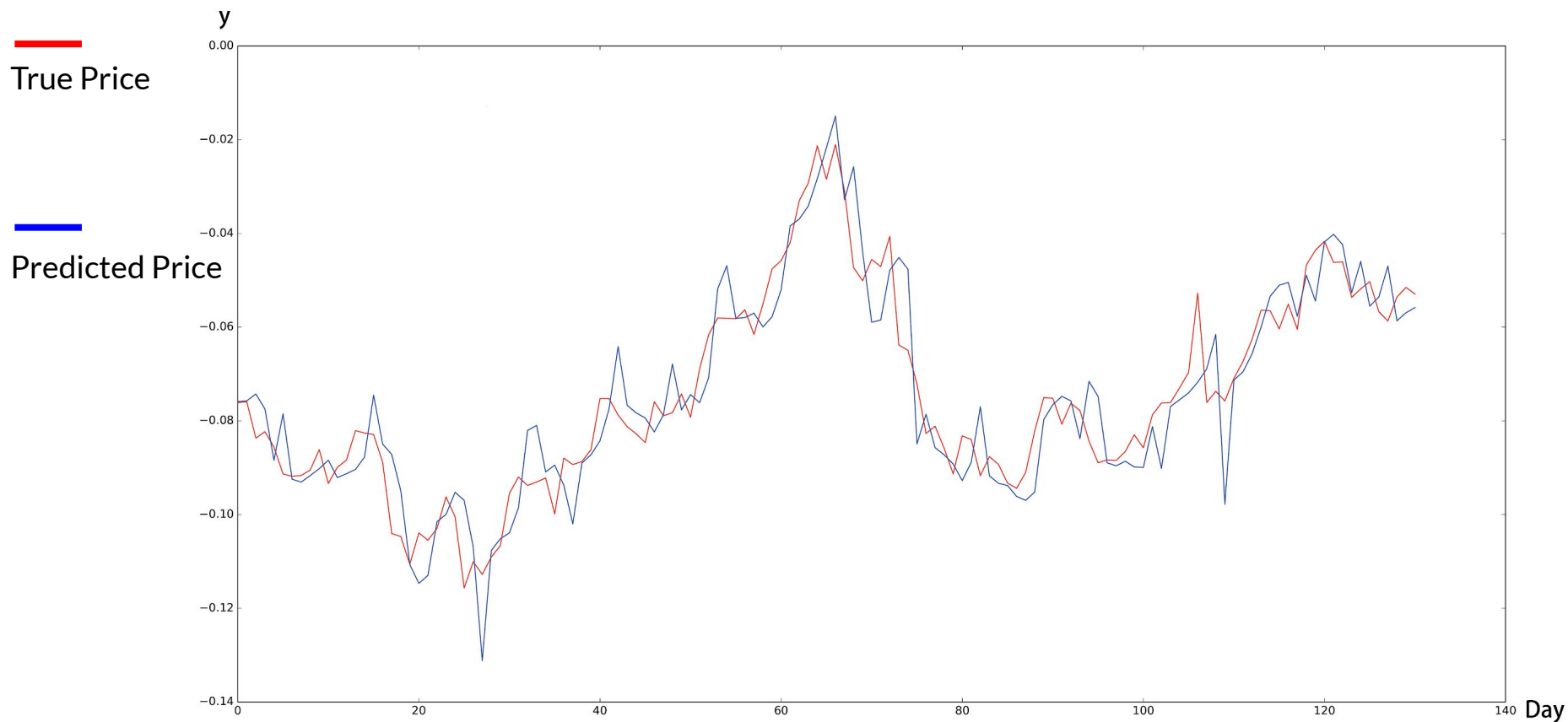
- Use Word-2-Vector on Trump's Tweets
- 3-Class prediction problem (rise, fall, flat)
- Predict S&P 500 Movement with a LSTM Network



Results

1. On Trump dataset: 48% accuracy in a three class prediction problem (5% better than human performance)
2. LSTM Networks tend to overfit fast because of its high complexity structure -> low amount of neurons per Layer
3. Embedding dimension around 200 is performing best
4. Model with News Influence Factor achieved better results than without News Influence Factor

Shenzen 500 Prediction



**Does Sentiment Analysis of
Social Media contradict the
Efficient Market Hypothesis?**

Nope!

Research Concerns

