

Challenges

Machine learning on Time Series Analysis

Solutions

Cheng-Yu Han

2019 / 10 / 25

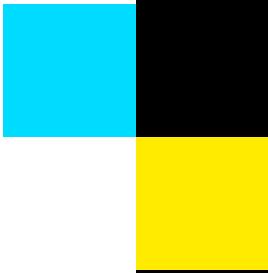
@人工智能與商業應用研討會



Investing

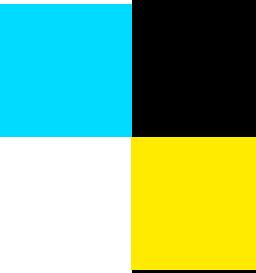
Gambling

Investing



Gambling

Investing



Gambling



Investing

Gambling



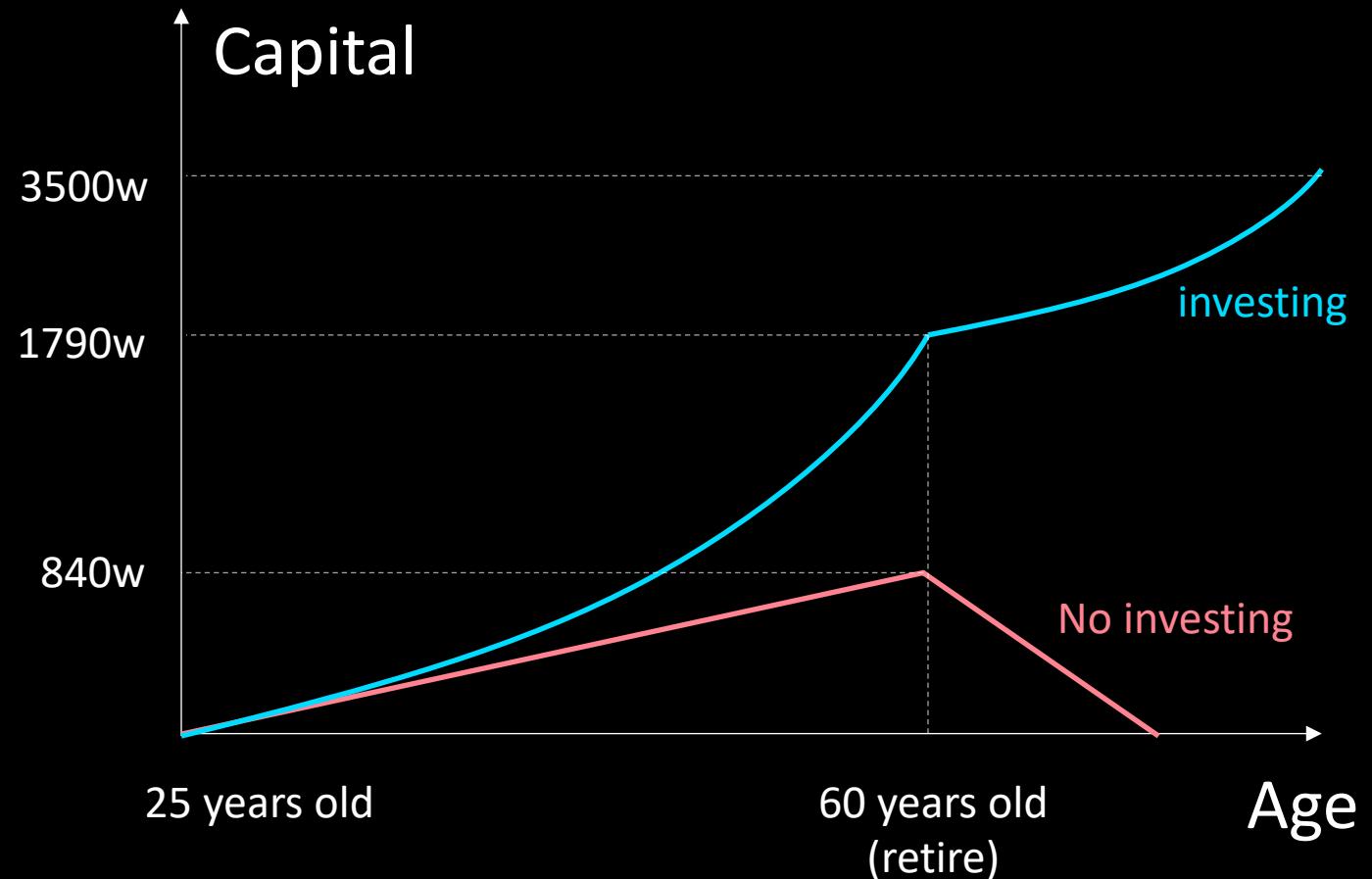


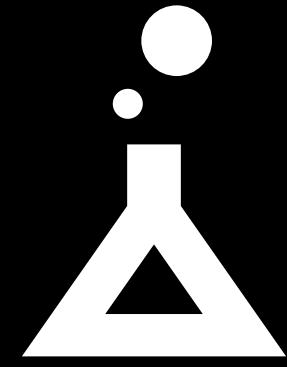


■ Personal Life Financial Plan

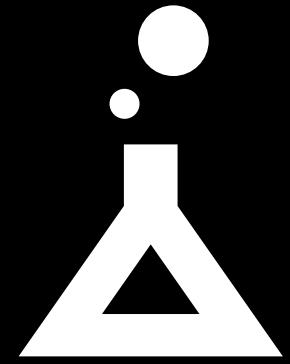
Capital simulation

Income per month.....5w (NTD)
Spent per month.....3w (NTD)
Investment.....70% (of total capital)
Compound interest....5% (each year)





FinLab

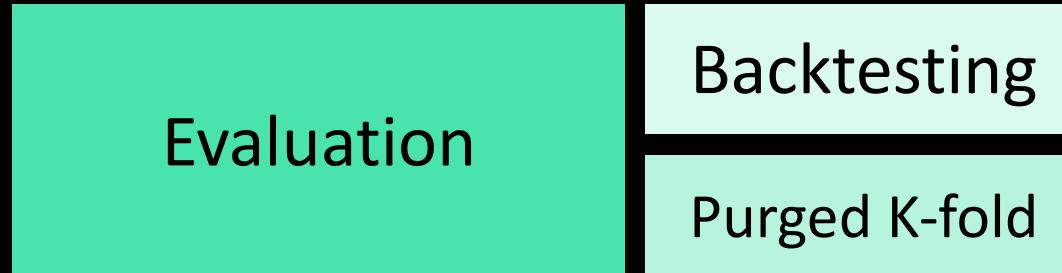
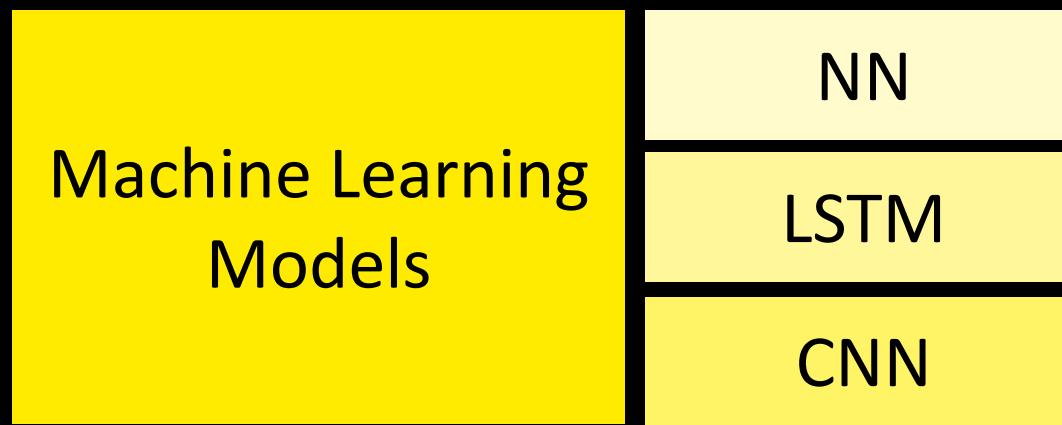


FinLab

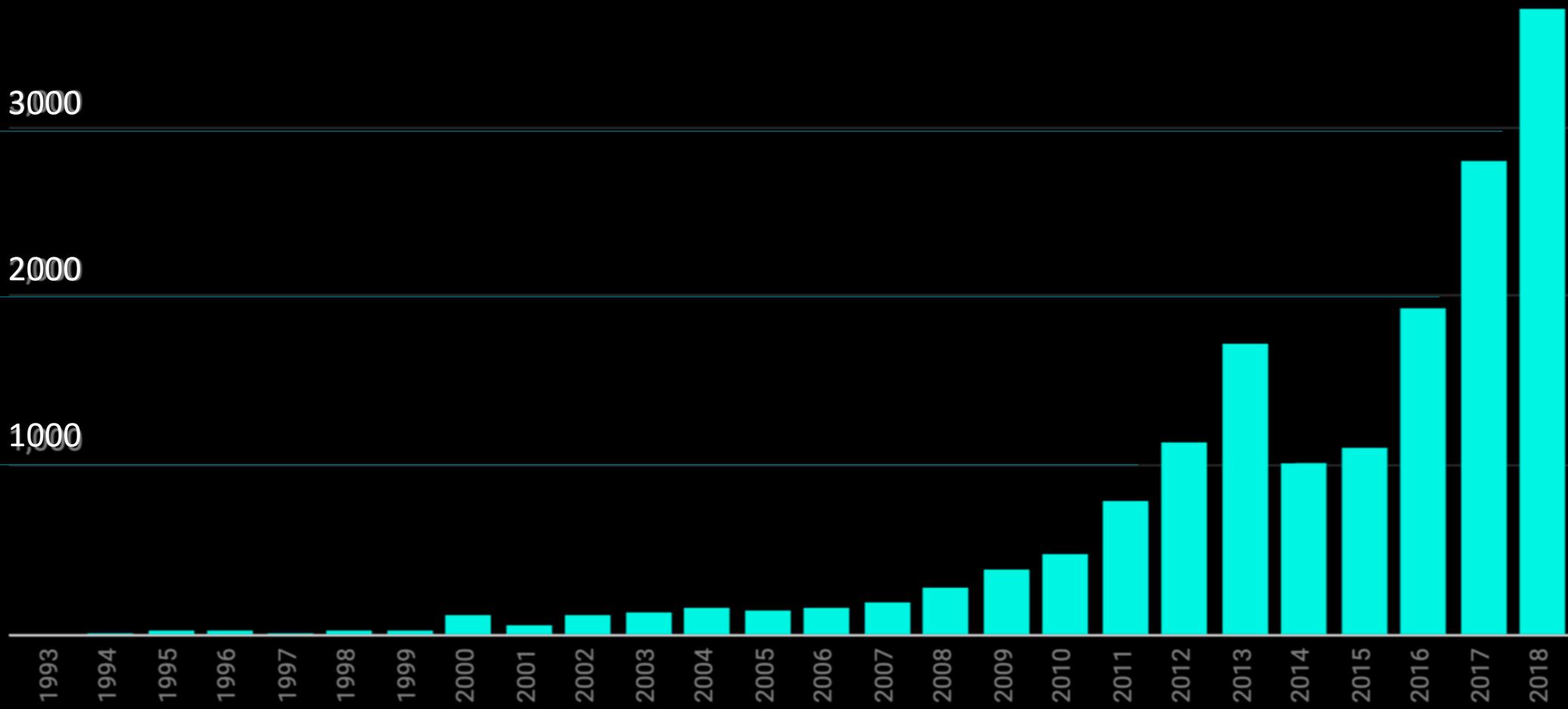


<https://www.finlab.tw/mopcon.html>

■ Outline



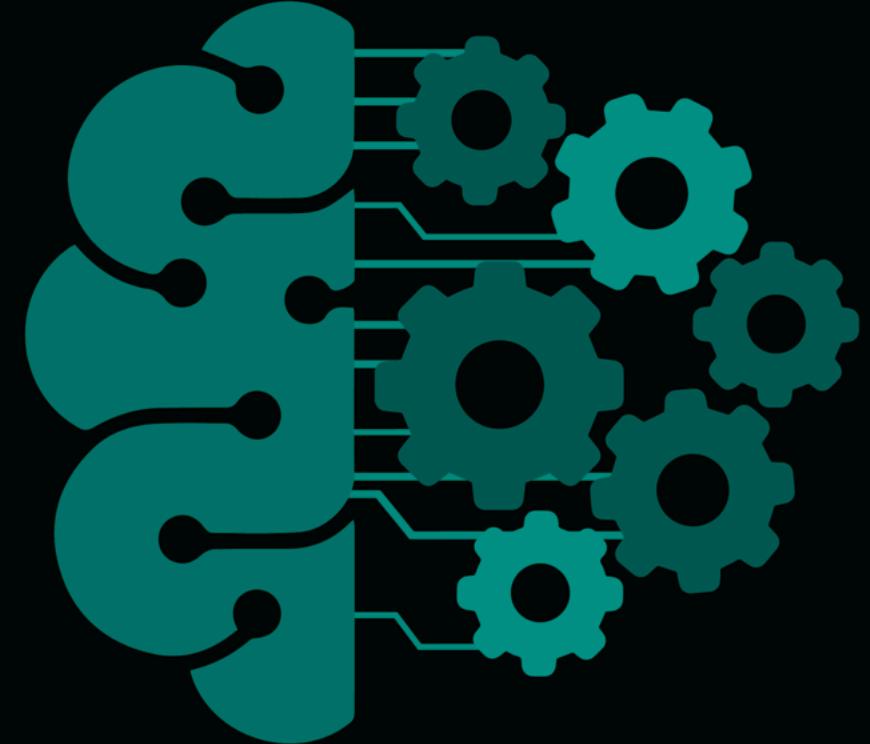
■ Artificial Intelligence papers



All of the papers available in the “artificial intelligence” section (arXiv)

■ Machine Learning

- Mimics the ability to see and hear
- Extract rules automatically from data
- ML spots patterns in high dimensional data



■ Supervised Machine Learning

Training



Testing





Financial Data (Features)

■ Financial Data Structures

Fundamental data

Focusing on creating a portrait of a company

- Useful to combine other data types
- Difficult to confirm data release date
- Missing data is often backfilled
- Consider multiple correction

■ Financial Data Structures

Fundamental data

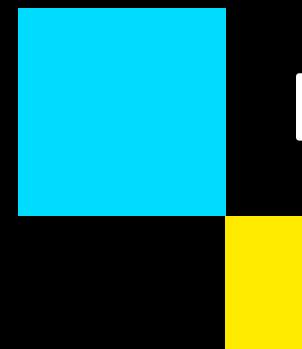
Focusing on creating a portrait of a company

- Useful to combine other data types
- Difficult to confirm data release date
- Missing data is often backfilled
- Consider multiple correction

Trading data

Market participant characteristic footprint
Trading book, price, broker trading summary...etc

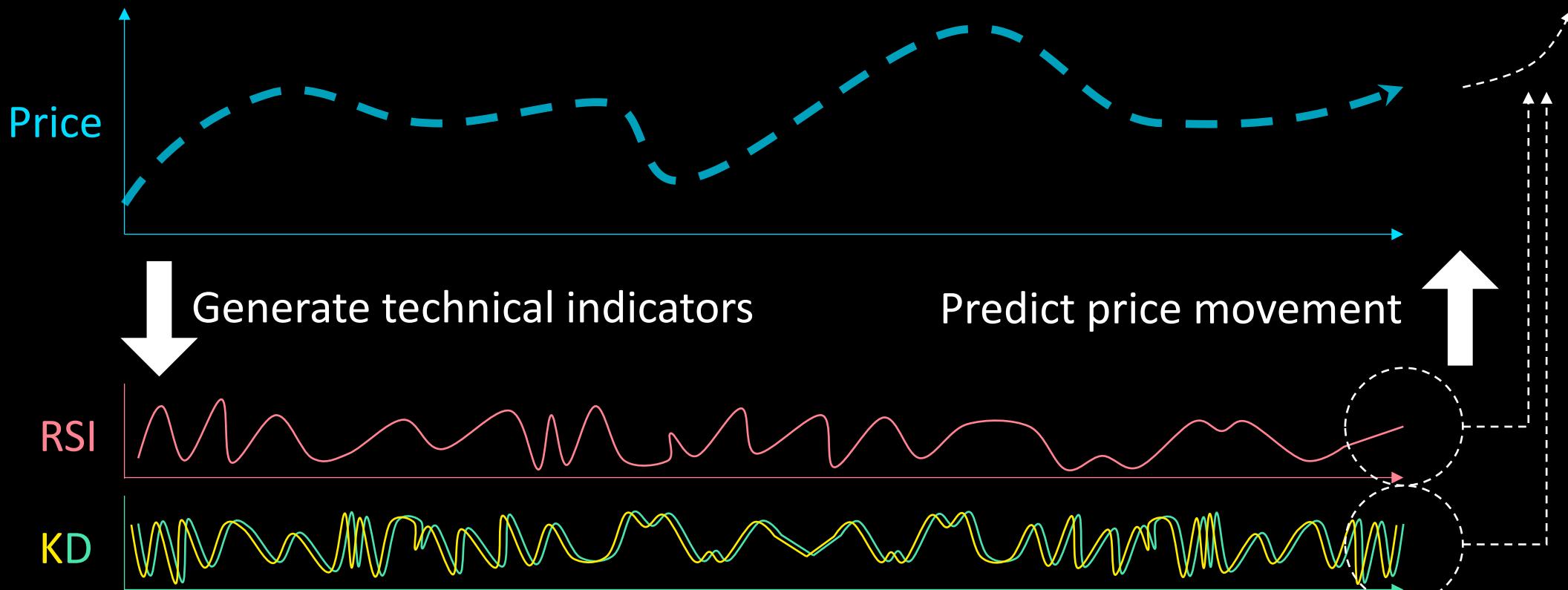
- Data often with timestamp
- Generate extra features (ex: technical indicators)
- Massive amount of data generated in one day
- Some of the data is difficult to obtain

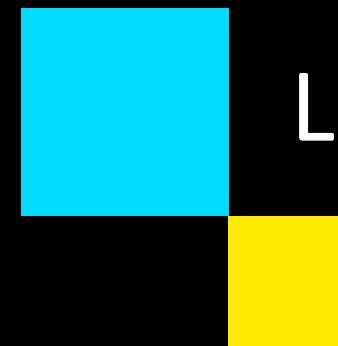


Demo

Creating Technical indicators

Price historical data





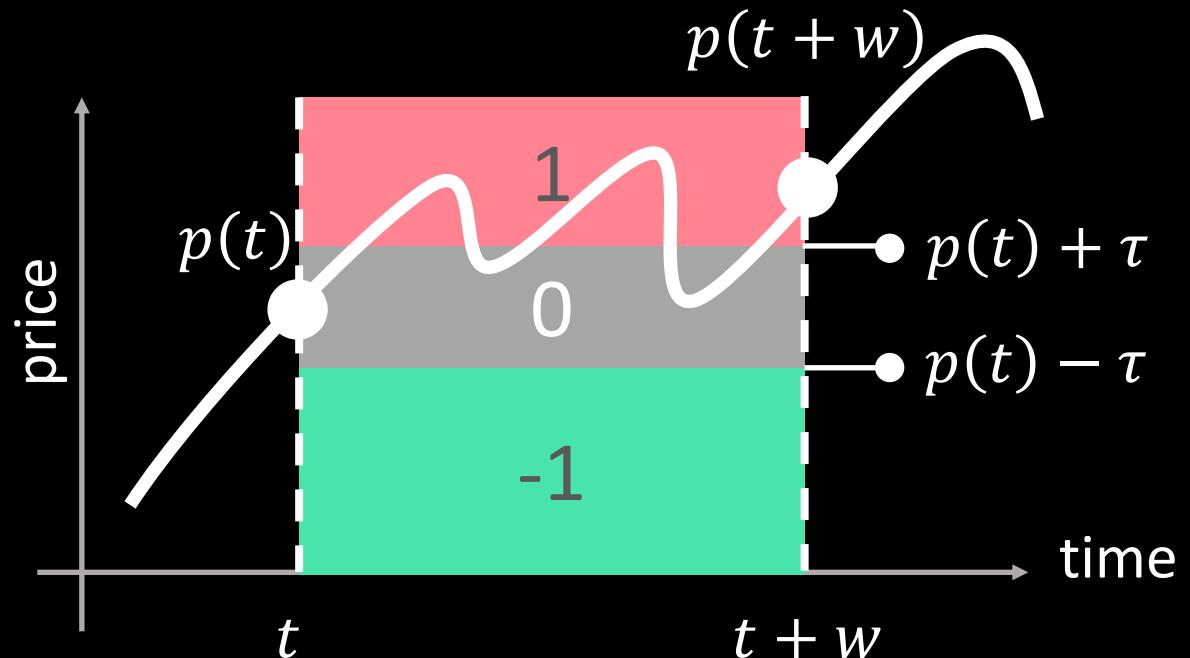
Labeling

■ Challenging of Labeling the data

Fixed time horizon

A popular method in the literature

- τ is a constant
- Do not have stop-loss limits



■ Label Generation Methods

- Triple barrier [Prado 2018]
- Continuous trading signals [Dash 2016]
- Trading Point decision [Chang 2009]

[Prado 2018] Advances in Financial Machine Learning

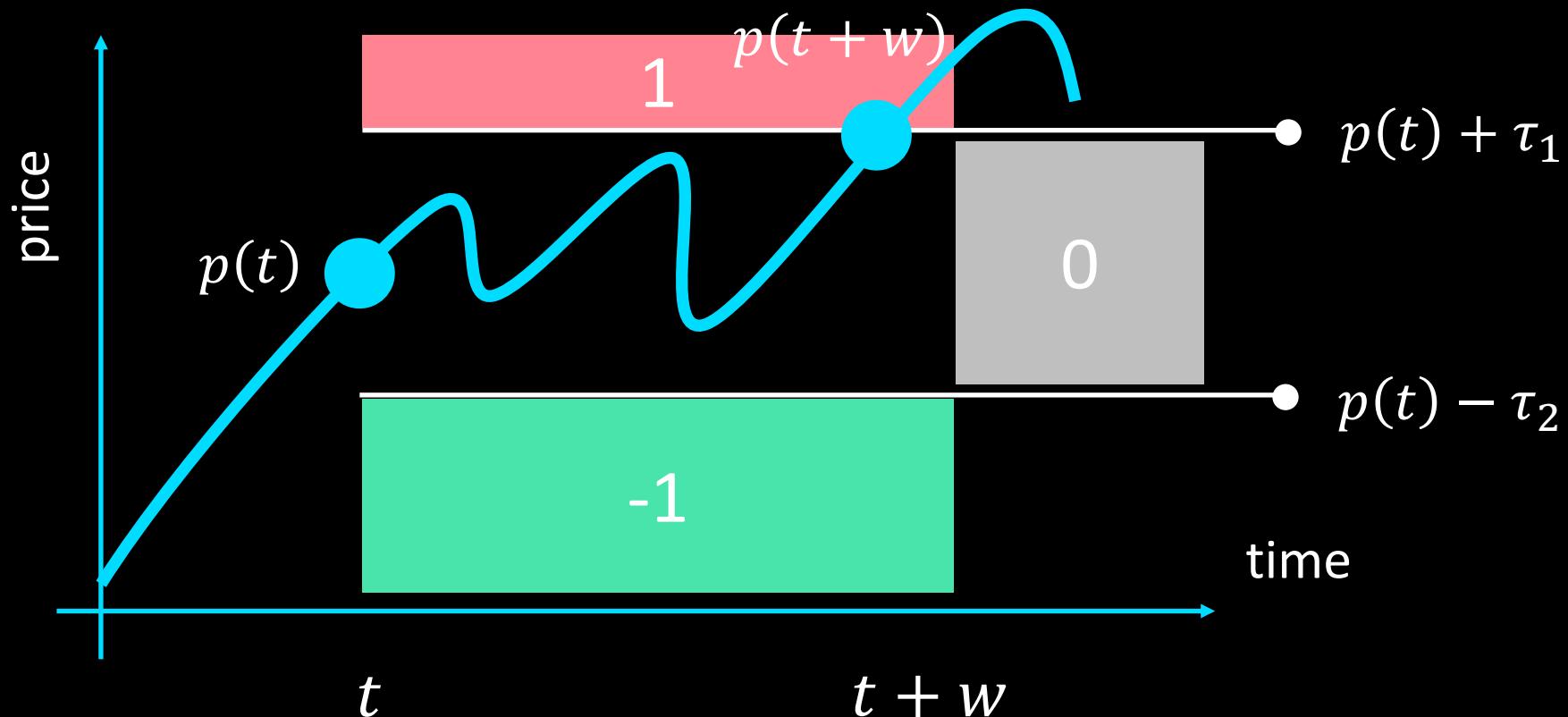
[Tsantekidis 2017] Using Deep Learning to Detect Price Change Indications in Financial Markets

[Dash 2016] A hybrid stock trading framework integrating technical analysis with machine learning techniques

[Chang 2009] Integrating a Piecewise Linear Representation Method and a Neural Network Model for Stock Trading Points Prediction

■ Triple barriers [Prado 2018]

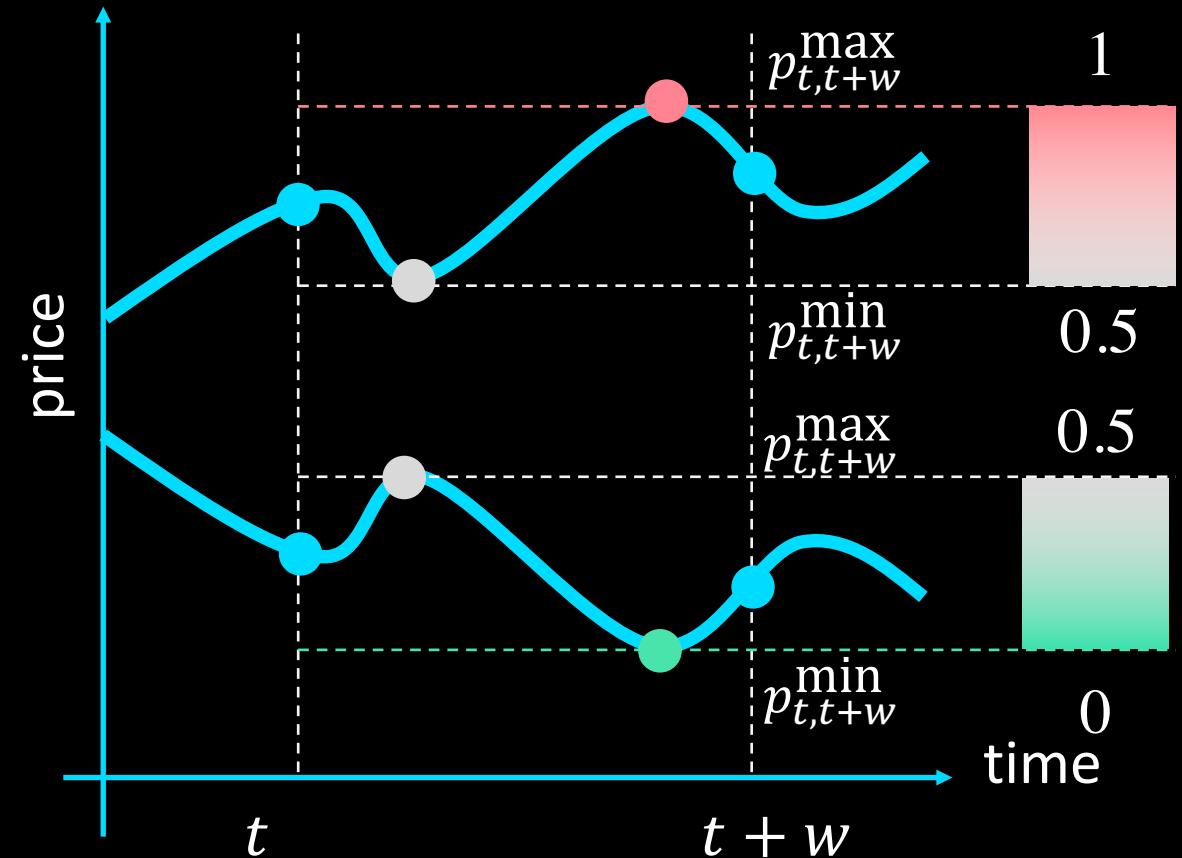
- Horizontal barriers are defined by profit-taking and stop-loss limit
- τ_1 and τ_2 are dynamic according to estimated volatility



■ Continuous trading signals [Dash 2016]

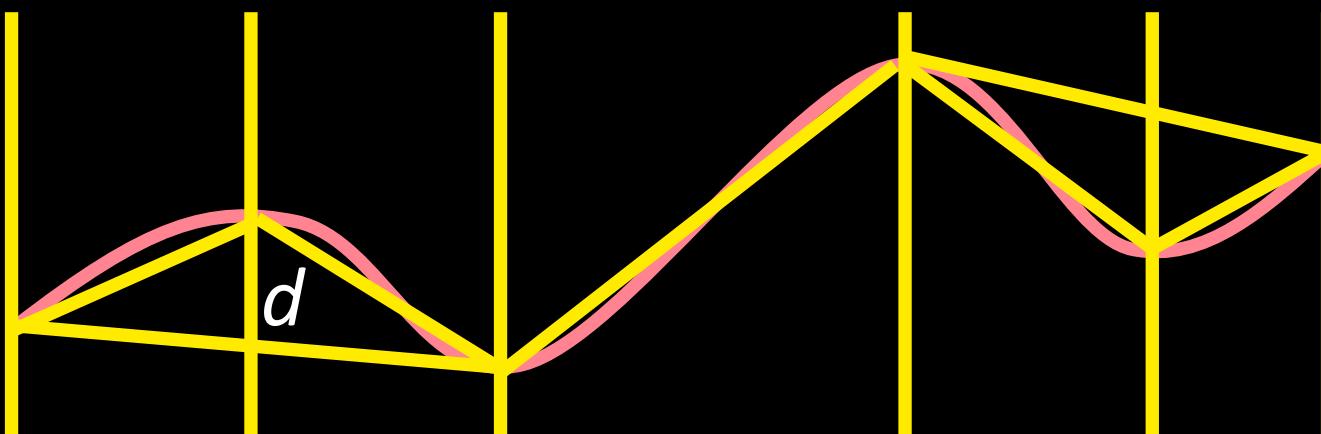
- Using momentum of the stock price
- $y(t)$'s are continuous
- Provides more detailed information

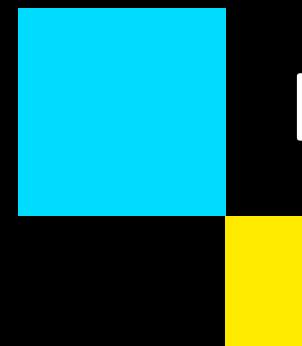
$$y(t) = \begin{cases} \frac{p_{t+w} - p_{t,t+w}^{\min}}{p_{t,t+w}^{\max} - p_{t,t+w}^{\min}} & \text{if } p_{t+w} > p_t \\ 0.5(1 - \frac{p_{t+w} - p_{t,t+w}^{\min}}{p_{t,t+w}^{\max} - p_{t,t+w}^{\min}}) & \text{else} \end{cases}$$



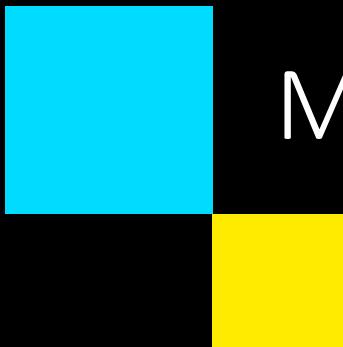
■ Trading point decision

- Find the local minimum and maximum points
- Divide the time series into subsegments
- Threshold value $d \rightarrow$ length of trend





Demo



ML Models

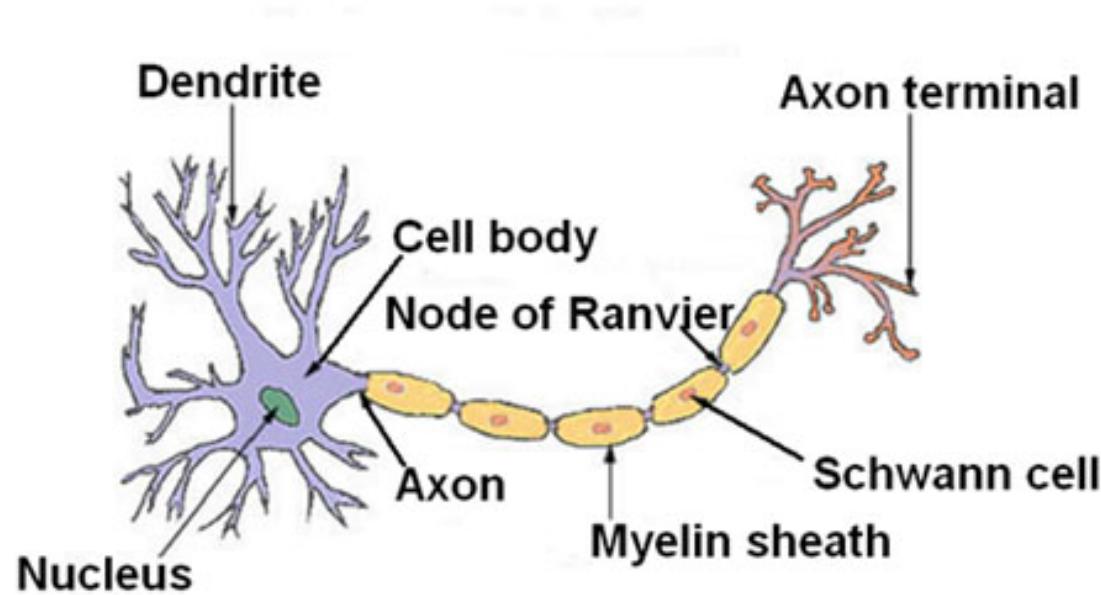


■ Neural Network

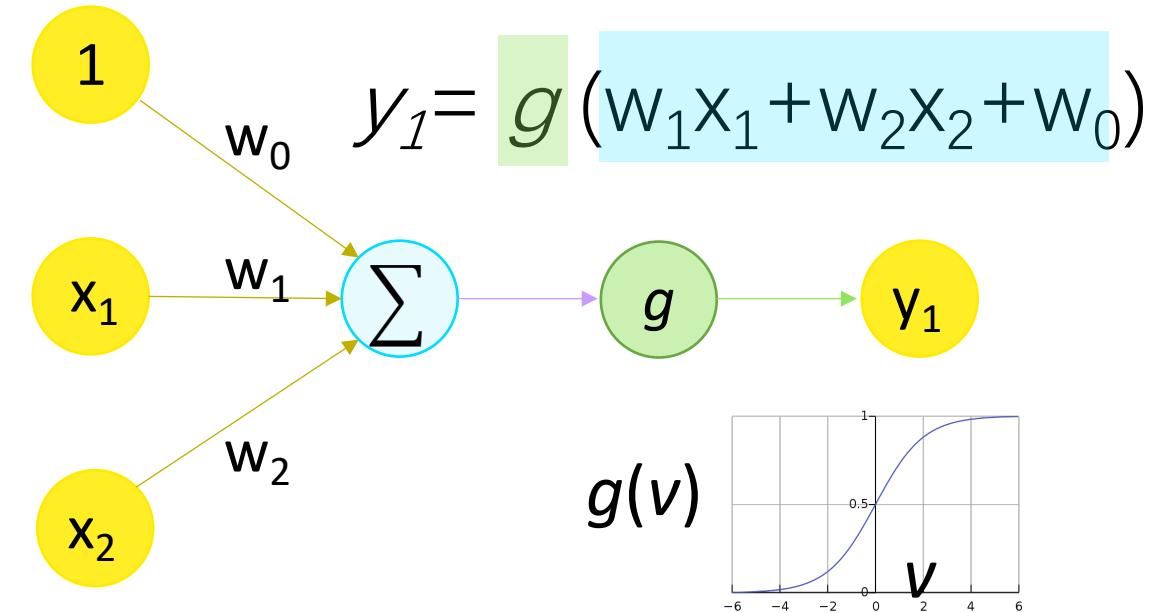
■ Neural Network

- Built to model the human brain
- interpret numeric data through a kind of machine perception

Human neuron structure

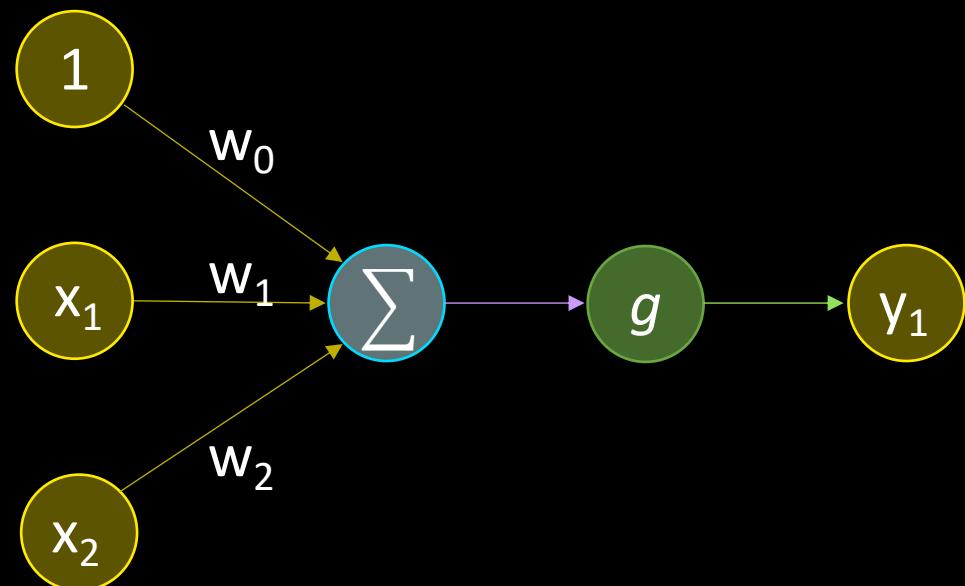


Single neuron model



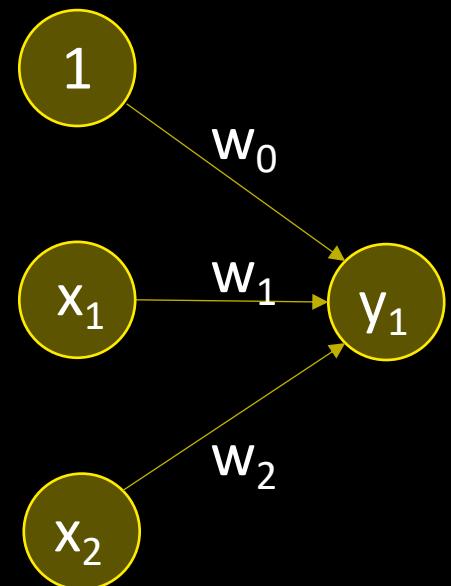
■ Neural Network

Single node in neural network



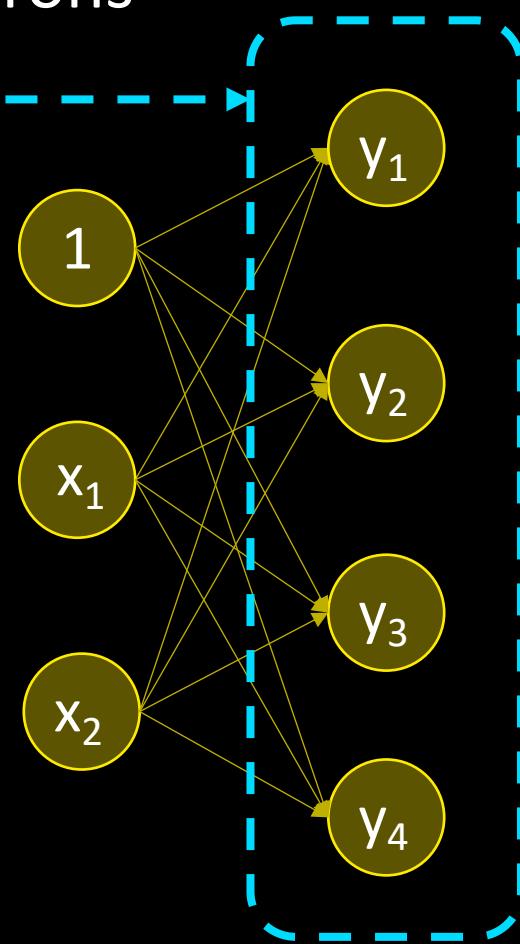
■ Neural Network

Simplified expression



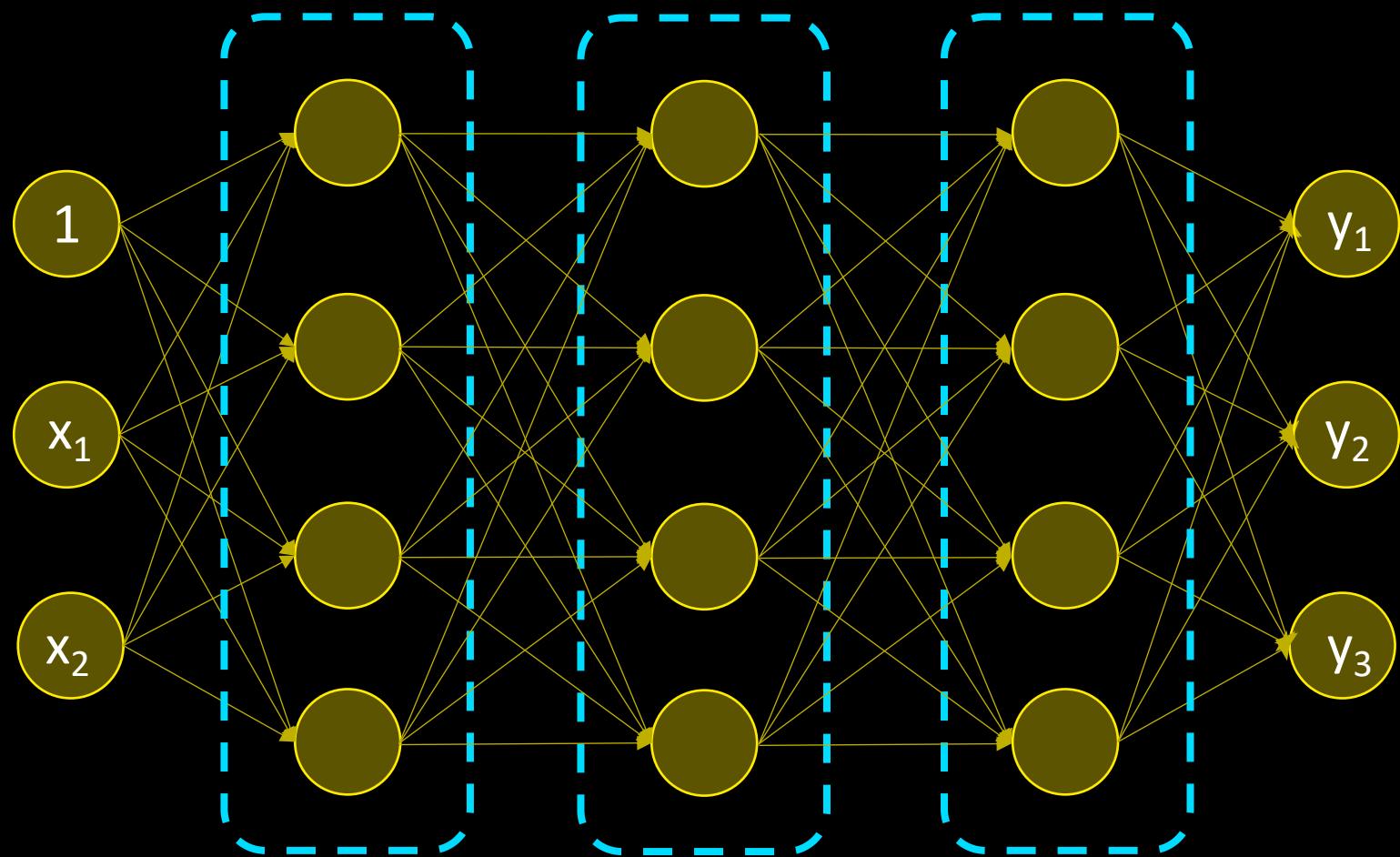
■ Neural Network

A layer contain multiple neurons



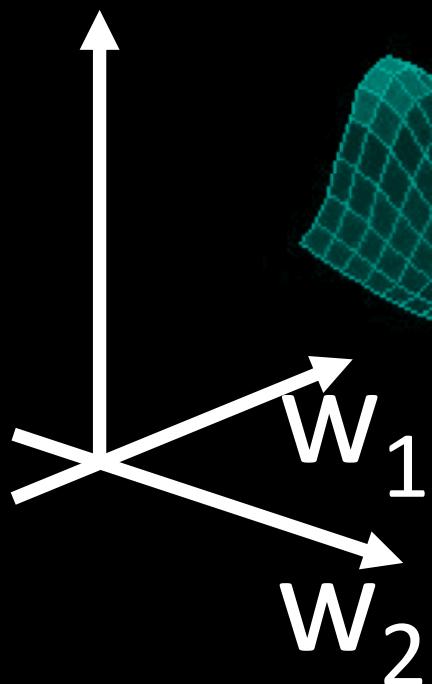
■ Deep Neural Network

Multi-layer deep neural network



■ Neural Network Optimization

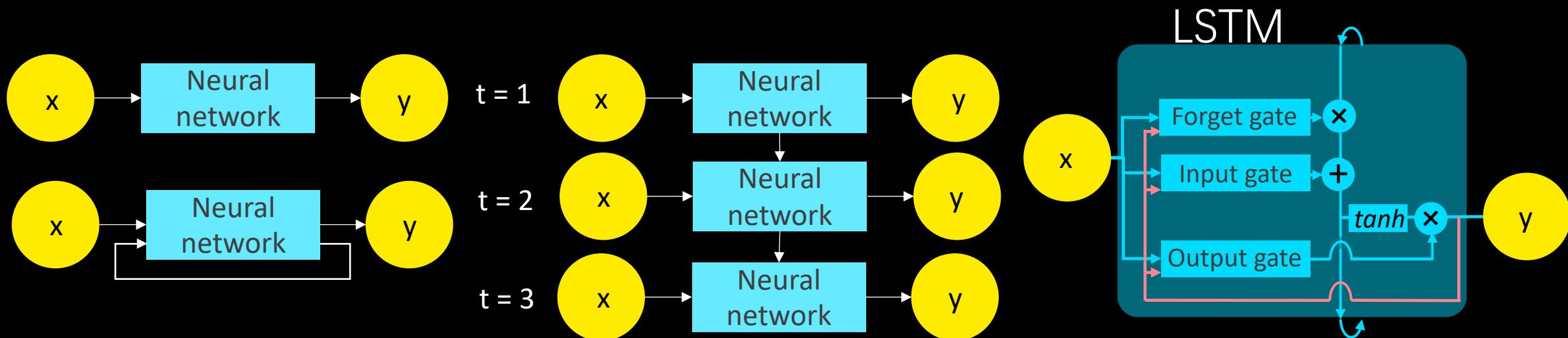
Cost function



■ Long Short Term Memory Neural Network

[Hochreiter 1997]

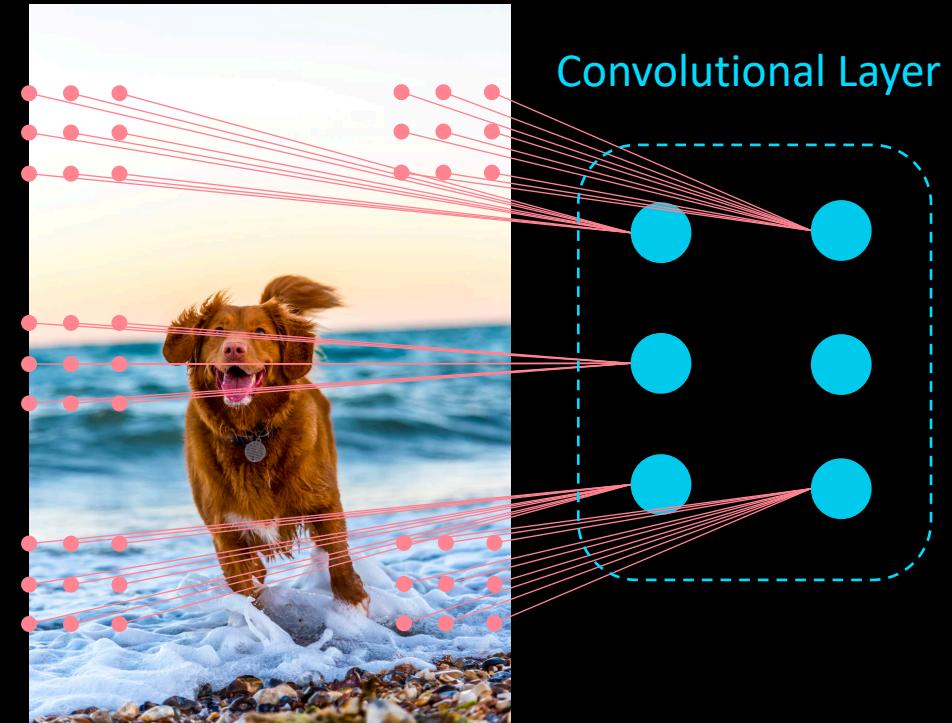
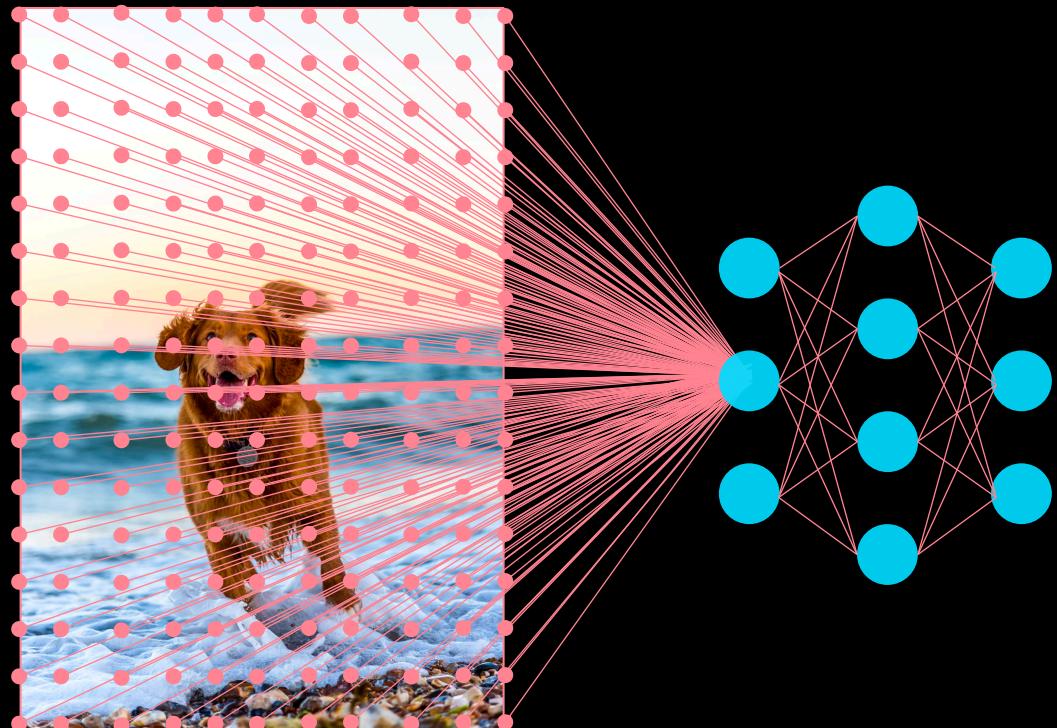
- Can process sequence of data
- LSTM deals with the exploding and vanishing gradient problems



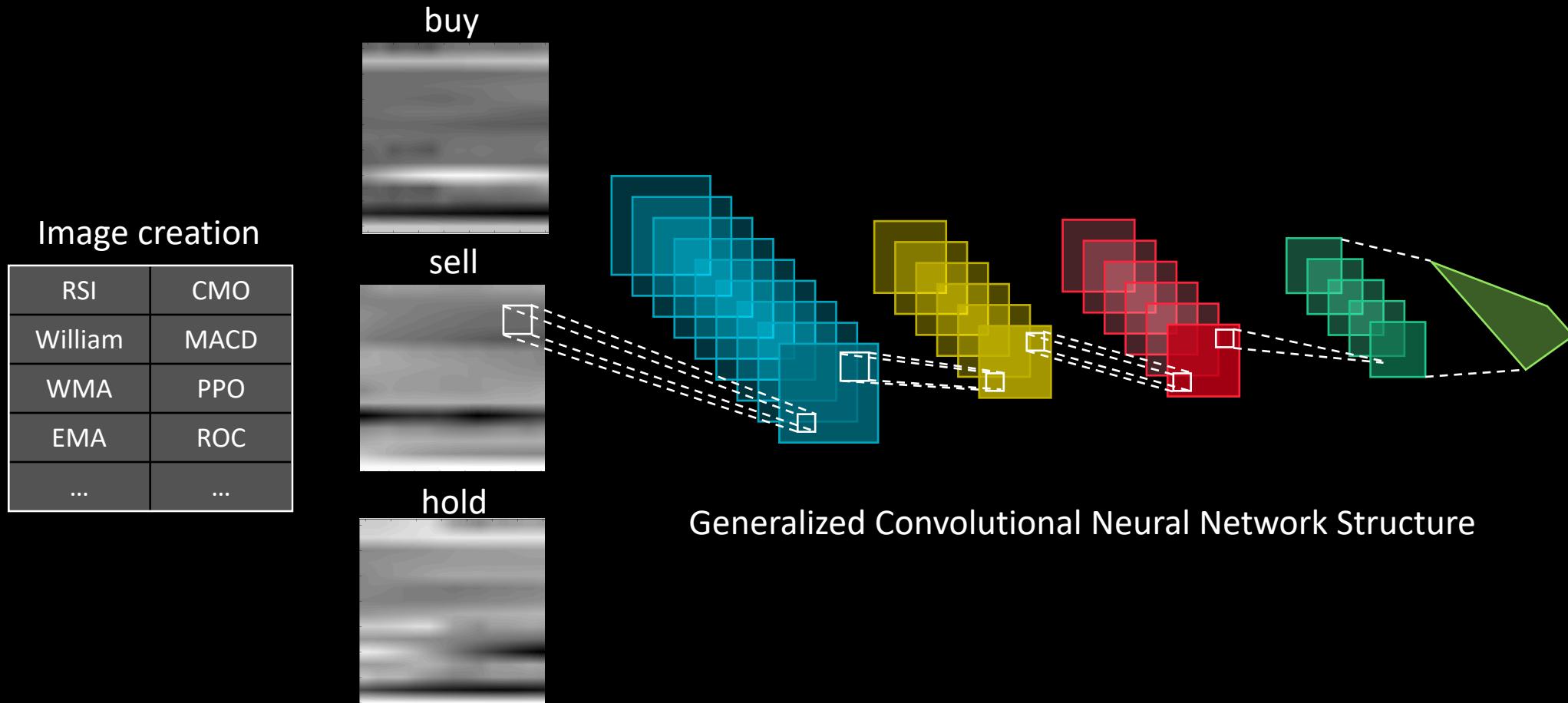
Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9.8 (1997): 1735-1780.

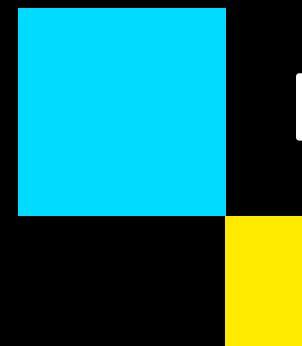
■ Convolutional Neural Network

- Commonly applied to computer visual imagery
- Prevent overfitting



■ Time series to Image conversion Approach [Sezer 2018]

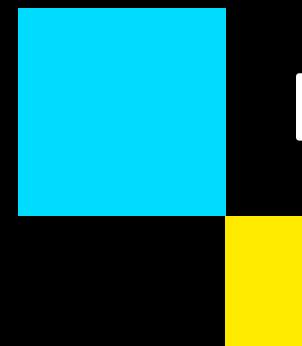




Demo

■ Backtest

- Survivor bias, lookahead bias, transection cost, outlier, overfitting
- Finding the lottery tickets that won the last game
- Solutions
 - Develop model for entire asset or classes
 - Use Bootstrap aggregating
 - Record every backtest conducted
 - Resist the temptation of reusing a failed strategy



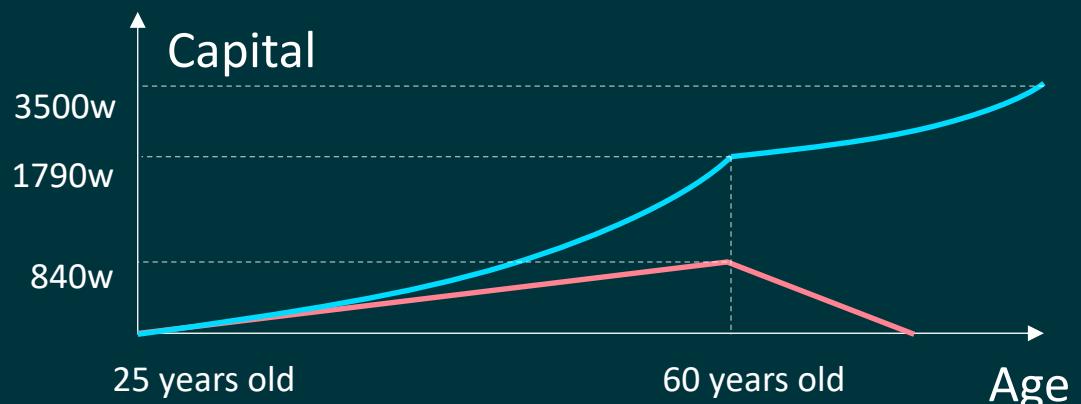
Demo

Conclusion

Investing

Investing or gambling ?

Life Plan



Machine Learning

Financial Data

Features

Machine Learning
Models

NN

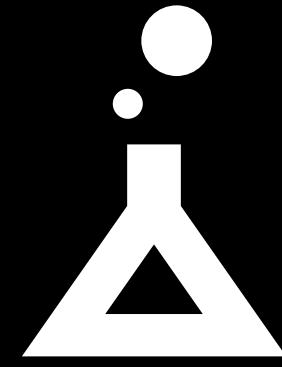
LSTM

CNN

Evaluation

Backtesting

Purged Validation



FinLab