Artificial Neural Network (ANN)

Agenda

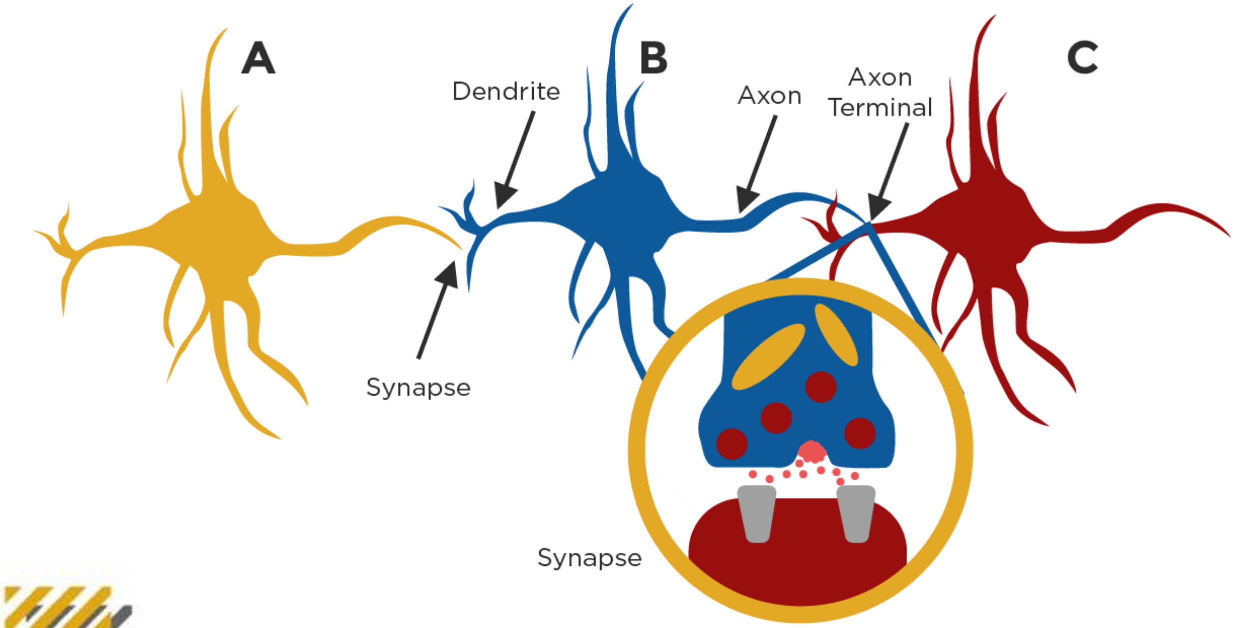
- **Neural Networks Models**
- ANN Advantage & ANN Applications
- Artificial Neural Networks
- Learning Principle for ANN
- Neural Network Properties
- Neural Network Step
- Basic Types of Learning

Neural Networks Models

- Computer modeling scientist think to approach complex system model by computation that is loosely based upon the architecture of the brain.
- Many different models, but all include:
  - Multiple, individual “nodes” or “units” that operate at the same time (in parallel)
  - A network that connects the nodes together
  - Information is stored in a distributed fashion among the links that connect the nodes
  - Learning can occur with gradual changes in connection strength
- Characteristic of human brain : able to remember, count, generalization, adaptation, low consumption energy

Biological Inspiration

Idea : To make the computer more robust, intelligent, and learn, …
Let's model our computer software (and/or hardware) after the brain



- Neurons on the Brain : although heterogeneous, at a low level the brain is composed of neurons
  - A neuron receives input from other neurons (generally thousands) from its synapses
  - Inputs are approximately summed
  - When the input exceeds a threshold the neuron sends an electrical spike that travels that travels from the body, down the axon, to the next neuron(s)

Computation in the Brain

- The brain's neuron network forms a massive parallel information processing system. This contrasts with conventional computers, in which a single processor executes a single series of instructions.
- The time taken for each elementary operation: neurons typically operate at a maximum rate of about 100 Hz, while a conventional CPU carries out hundreds million machine level operations per second. Despite of being built with very slow hardware, the brain has quite remarkable capabilities:
- The performance tends to degrade gracefully under partial damage. In contrast, most programs and engineered systems are brittle: if you remove some arbitrary parts, very likely the whole will cease to function.
- It can learn (reorganize itself) from experience.
- It is possible to do partial recovery from damage if healthy units can learn to take over the functions previously carried out by the damaged areas.
- It performs massively parallel computations extremely efficiently. For example, complex visual perception occurs within less than 100 ms.
- It supports our intelligence and self-awareness. (Nobody knows yet how this occurs.)
- As a discipline of Artificial Intelligence, Neural Networks attempt to bring computers a little closer to the brain's capabilities by imitating certain aspects of information processing in the brain, in a highly simplified way.

Agenda

- Neural Networks Models
- **ANN Advantage & ANN Applications**
- Artificial Neural Networks
- Learning Principle for ANN
- Neural Network Properties
- Neural Network Step
- Basic Types of Learning

ANN Advantage &
ANN Applications

ANN Advantage

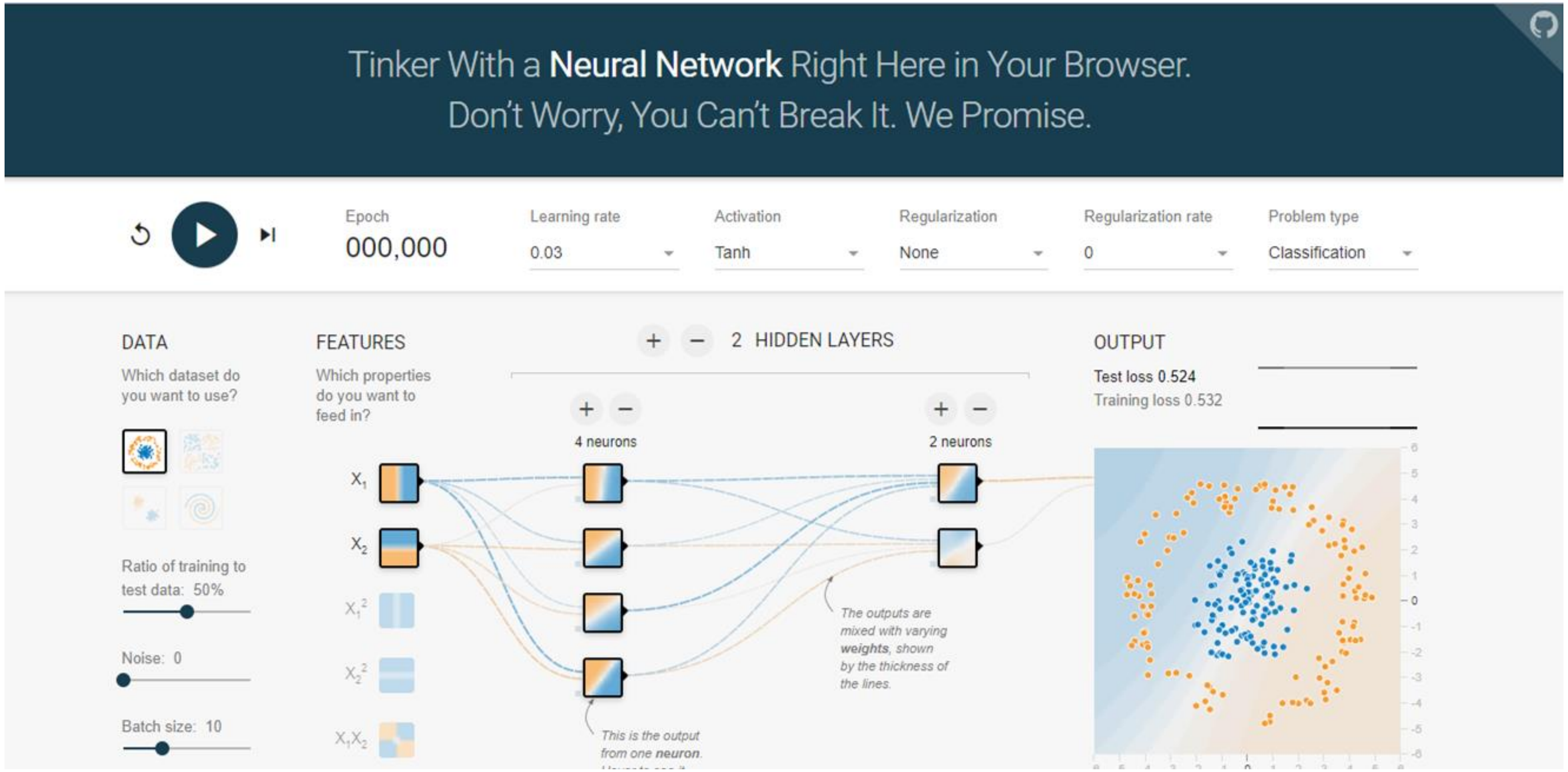
- Able (and faster) to predict based on nonlinear data pattern
- Robust to the missing data

ANN Applications

- Weather forecast, stock price prediction, fraud detection, machine translation, computer vision (deep learning), etc

Tensorflow ANN

https://playground.tensorflow.org



Agenda

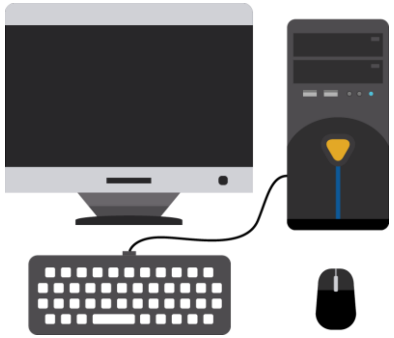
- Neural Networks Models
- ANN Advantage & ANN Applications
- **Artificial Neural Networks**
- Learning Principle for ANN
- Neural Network Properties
- Neural Network Step
- Basic Types of Learning

Typical ANN Tasks

Tasks to be solved by artificial neural networks:

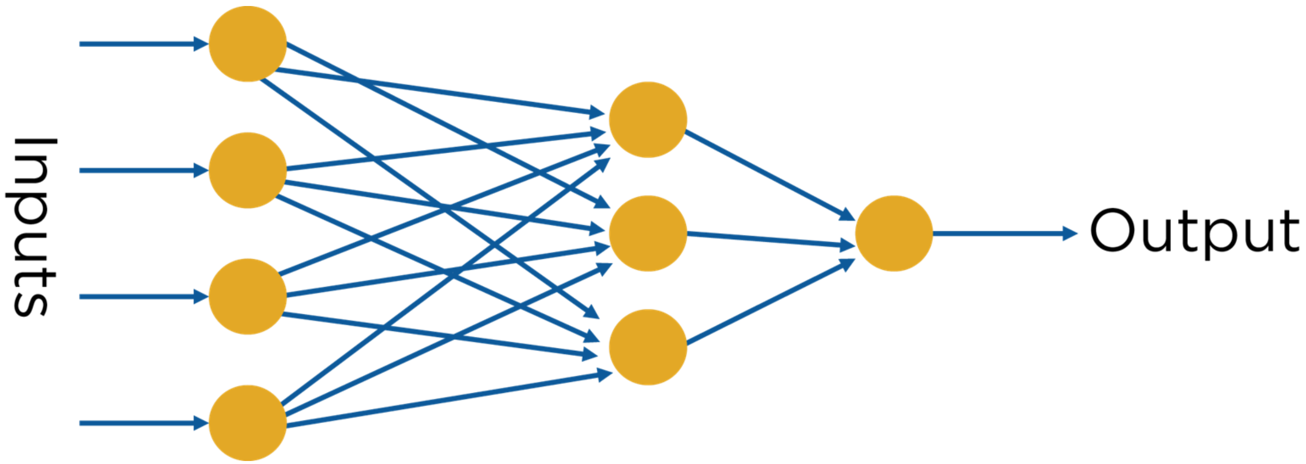
- controlling the movements of a robot based on self-perception and other information (e.g., visual information)
- deciding the category of potential food items (e.g., edible or non-edible) in an artificial world
- recognizing a visual object (e.g., a familiar face);
- predicting where a moving object goes, when a robot wants to catch it.

Brain VS
Computer (Traditional)



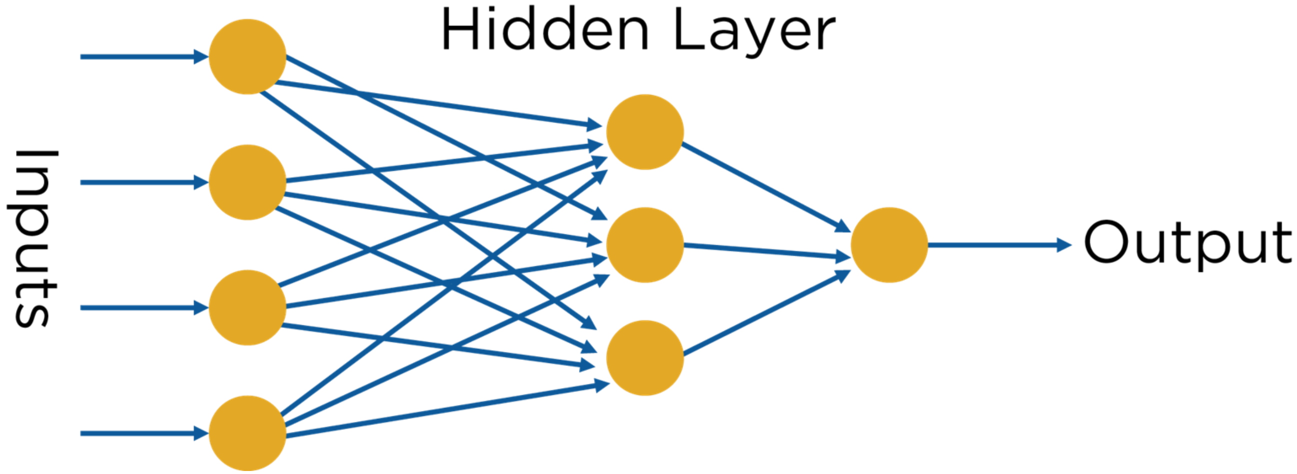
200 billion neurons, 32 trillion synapses	1 billion bytes RAM but trillions of bytes on disk
Element size: 10^{-6} m	Element size: 10^{-9} m
Energy use: 25W	Energy watt: 30-90W (CPU)
Processing speed: 100 Hz	Processing speed: 10^{9} Hz
Parallel, Distributed	Serial, Centralized
Learns: Yes	Learns: Some
Fault Tolerant	Generally not Fault Tolerant
Intelligent/Conscious: Usually	Intelligent/Conscious: Generally No

Artificial Neural Networks



- An artificial neural network is composed of many artificial neurons that are linked together according to a specific network architecture. The objective of the neural network is to transform the inputs into meaningful outputs.
- Types of ANN : Backpropagation (feed-forward), Recurrent Network, Self Organizing Map, Bayesian Networks, etc

Neural Network
Mathematics



$$y_1^1 = f(x_1, w_1^1)$$
$$y_2^1 = f(x_2, w_2^1)$$
$$y_3^1 = f(x_3, w_3^1)$$
$$y_4^1 = f(x_4, w_4^1)$$

$$y^1 = \begin{pmatrix} y_1^1 \\ y_2^1 \\ y_3^1 \\ y_4^1 \end{pmatrix}$$

$$y_1^2 = f(y^1, w_1^2)$$
$$y_2^2 = f(y^1, w_2^2)$$
$$y_3^2 = f(y^1, w_3^2)$$

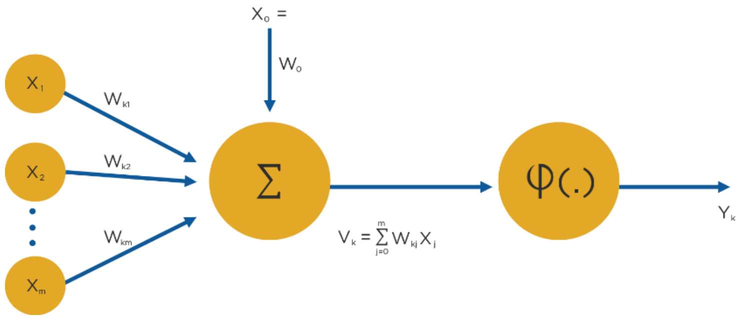
$$y^2 = \begin{pmatrix} y_3^2 \\ y_3^2 \\ y_3^2 \end{pmatrix}$$

$$y_{Out} = f(y^2, w_1^3)$$

Generalized Neuron Models

$$y = \varphi\left(\sum_{j=1}^{n} w_j x_j - u\right)$$

φ = activation function
w = weight for input jth



X_0 and b_k is bias function to increase or decrease input for activation function

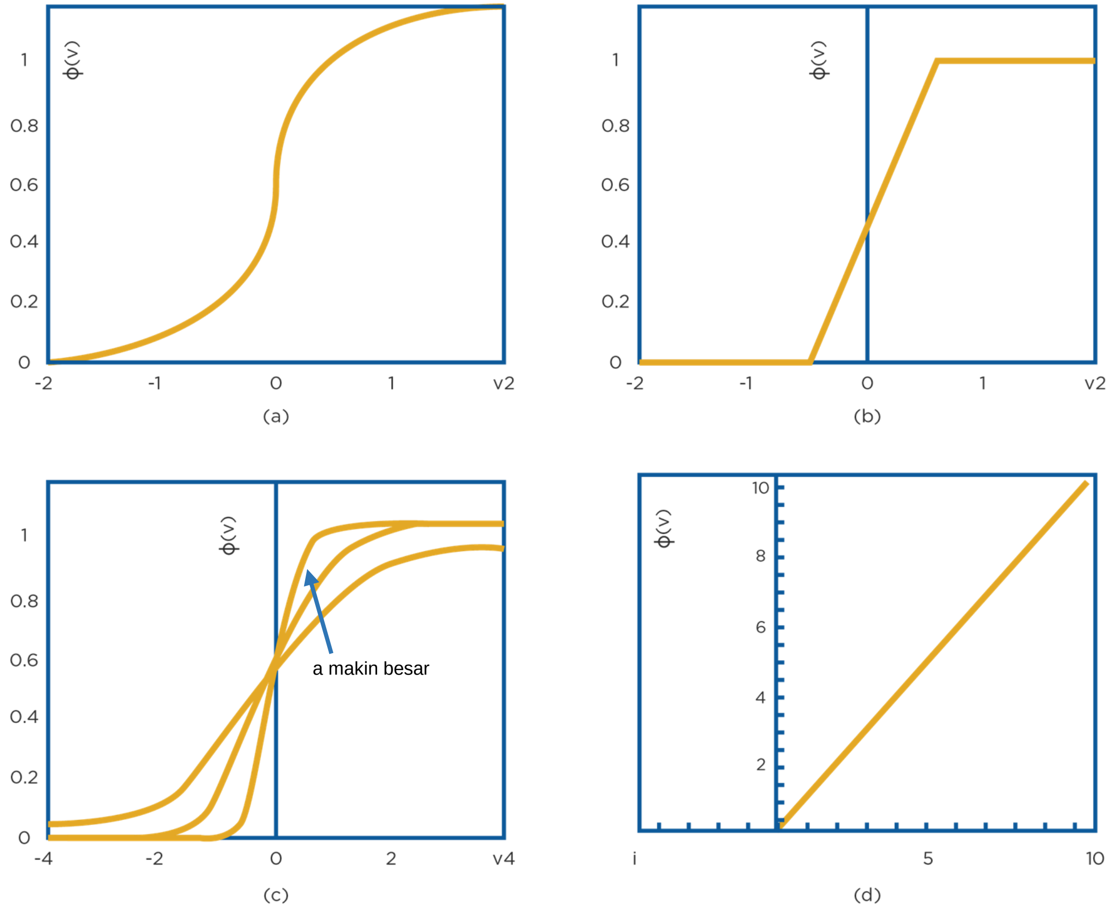
Neuron k can be described as

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \qquad \text{and} \qquad y_k = \varphi(u_k + b_k)$$

*activation function :

- to transform input value to other value according to the function used,
- to enabled ANN handle non linear task,
- without this function, ANN become combination of linear operation

Type of
Activation Function



a. Tanh Function
b. Linear Piecewise
c. Sigmoid Function
d. ReLU Function

Agenda

- Neural Networks Models
- ANN Advantage & ANN Applications
- Artificial Neural Networks
- **Learning Principle for ANN**
- Neural Network Properties
- Neural Network Step
- Basic Types of Learning

Learning as Optimization

- The objective of adapting the responses on the basis of the information received from the environment is to achieve a better state. E.g., the animal likes to eat many energy rich, juicy fruits that make its stomach full, and makes it feel happy.
- In other words, the objective of learning in biological organisms is to optimise the amount of available resources, happiness, or in general to achieve a closer to optimal state.

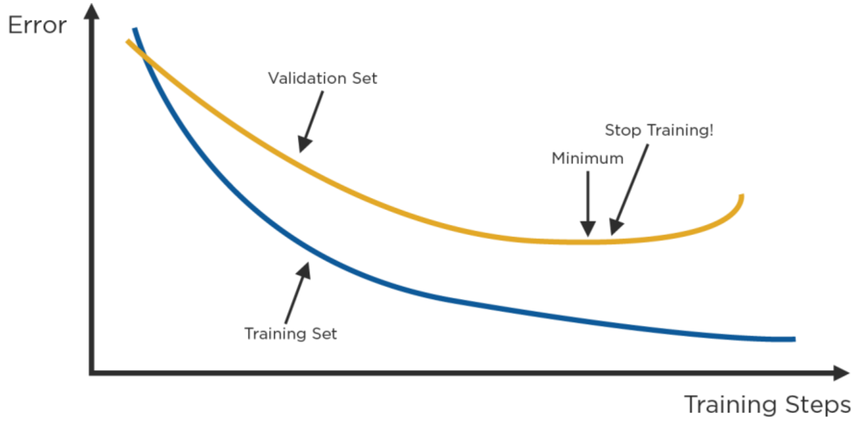
Learning Principle
for ANN

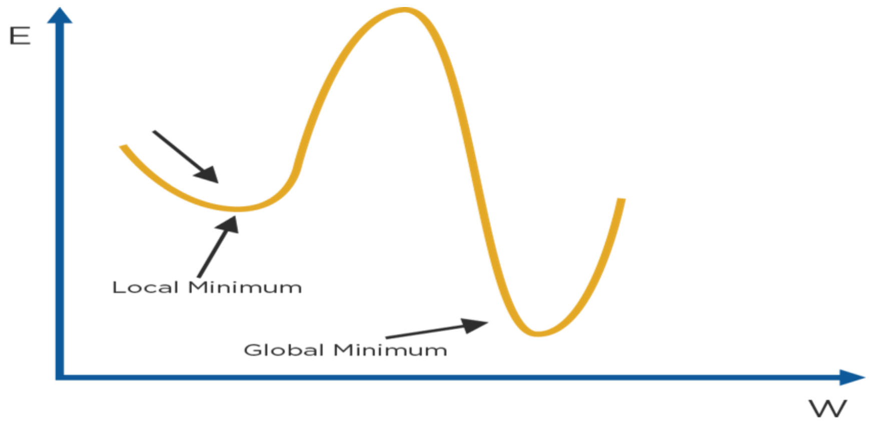
ENERGY MINIMIZATION

We need an appropriate definition of energy for artificial neural networks, and having that we can use mathematical optimisation techniques to find how to change the weights of the synaptic connections between neurons.

ENERGY = measure of task performance error

General Neural Network Steps





1. Split the data into training set and validation (test set)
2. Vary the number of hidden neurons from 1 to 10 in steps 1 or more
3. Train neural network on the training set and measure the performance on the validation set (may be train multiple neural networks to deal with local minimum issue
4. Choose the number of hidden neurons with optimal validation set performance
5. Measure the performance on the independent test set

Some Facts
ANN (1)

- Instead of programming computational system to do specific tasks, teach system how to perform task
- To do this, generate Artificial Intelligence System- AI
- Empirical model which can rapidly and accurately find the patterns buried in data that reflect useful knowledge
- One case of these AI models is neural networks
- AI systems must be adaptive – able to learn from data on a continuous basis
- ANN is non-linear transformation functions similar to logistic regression. So logistic regression is neural network with one neuron
- More complex non-linear patterns requires more hidden layers

Some Facts
ANN (2)

- Neural networks can model very complex patterns and decision boundaries in the data and, as such are very powerful. In fact, they can even model the noise in the training data
- It is a black box process, so it is difficult to explain the usage in everyday applications
- To open up the black box process, we use rule extractions or two-stage modelling
- Rule extraction purpose is to extract if/then classification rules mimicking neural network behavior
- There are two rule extraction: Decompositional and Pedagogical Technique

Agenda

- Neural Networks Models
- ANN Advantage & ANN Applications
- Artificial Neural Networks
- Learning Principle for ANN
- **Neural Network Properties**
- Neural Network Step
- Basic Types of Learning

Neural Network
Properties

1. Epoch : is a measure of the number of times all of the training vectors are used once to update the weight (similar to iteration)
2. Learning Rate : how fast the neural network weight change
3. Activation / Transfer Function : of a node defines the output of that node given an input or set of inputs (non linear function compared to 0 and 1 output of digital circuit)
4. Regularization : is a process of introducing additional information in order to solve an ill-posed problem or to prevent overfitting (L1 and L2 norm)
5. Problem Type : whether it is a classification, regression or else
6. Gradient Descent : is a first-order iterative optimization algorithm for finding the minimum of a function.
7. Regularization Rate

Decompostional
Neural Network Step

Objective : Consider the neural network as a black box (Pedagogical rule extraction) and use neural network predictions as input to a white box analytical techniques such as decision tree

- Train a neural network and prune it as much as possible in terms of connections
- Categorize the hidden unit activation values using clustering
- Extract rules that describe the network outputs in terms of categorized hidden unit activation values
- Extract rules that describe the categorized hidden unit activation values in terms of the network inputs
- Merge the rules obtained in step 3 and 4 to directly relate the inputs to outputs

Example in next slide

Example:
Credit Risk (1)

Customer	Age	Income	Gender	...	Response
Emma	28	1,000	F		No
Will	44	1,500	M		Yes
Dan	30	1,200	M		No
Bob	58	2,400	M		Yes

Customer	Age	Income	Gender	h1	h2	h3	h1	h2	h3	Response
Emma	28	1,000	F	-1.20	2.34	0.66	1	3	2	No
Will	44	1,500	M	0.78	1.22	0.82	2	3	2	Yes
Dan	30	1,200	M	2.1	-0.18	0.16	3	1	2	No
Bob	58	2,400	M	-0.1	0.8	-2.34	1	2	1	Yes

If h1 = 1 and h2 = 3, then response = No

If h2= 2, then response = Yes

If age < 28 income < 1.000, then h1 = 1

If gender = F, then h2 = 3

If age > 34 and income > 1.500, then h2 = 2

If age < 28 and income < 1.000 and gender = F then response = No

If age > 34 and income > 1.500 then response = Yes

Step 1: Start from original data.

Step 2 : Build a neural network (e.g , 3 hidden neurons).

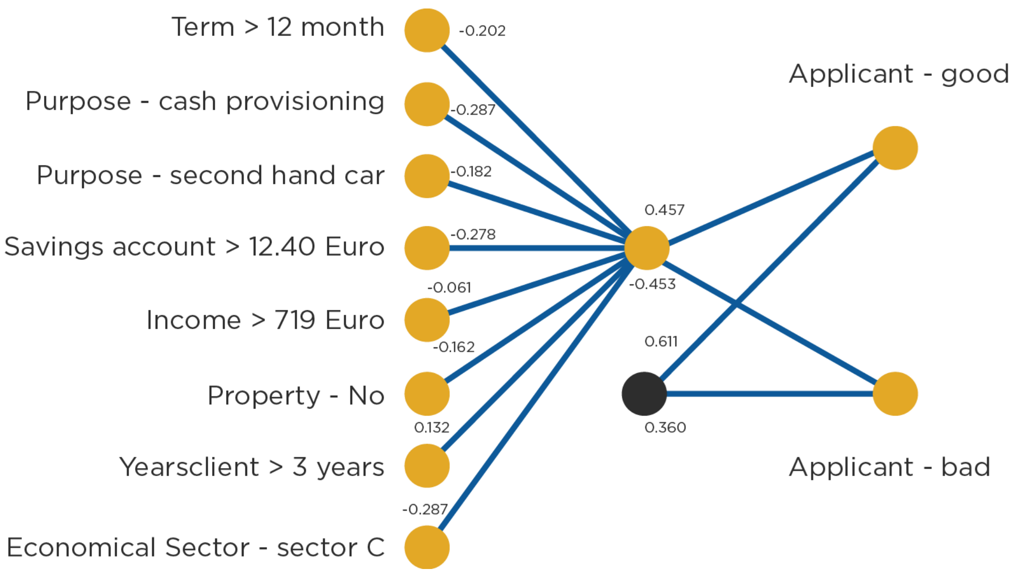
Step 3: Categorize hidden unit activations.

Step 4: Extract rules relating network outputs to categorize hidden units.

Step 5: Extract rules relating categorized hidden units to inputs.

Step 6: Merge both rule sets.

Example:
Credit Risk (2)



- If term > 12 months and purpose = cash provisioning and savings account ≤ 12.40 Euro and years client ≤ 3, then applicant = bad
- If term > 12 months and and purpose = cash provisioning and owns property = no and savings account ≤ 12.40 Euro and years client ≤ 3, then applicant = bad
- If purpose = cash provisioning and income > 719 and owns property = no and savings account ≤ 12.40 Euro and years client ≤ 3, then applicant = bad
- If purpose = secondhand car and income > 719 Euro and owns property = no and savings account ≤ 12.40 Euro and years client ≤ 3, then applicant = bad
- If saving aacounts ≤ 12.40 Euro and economical sector = sector C, then applicant = bad
- Deffault Class: applicant = good

Agenda

- Neural Networks Models
- ANN Advantage & ANN Applications
- Artificial Neural Networks
- Learning Principle for ANN
- Neural Network Properties
- **Pedagogical Neural Network Step**
- Basic Types of Learning

Pedagogical
Neural Network Step

Objective : Decompose the network internal workings by inspecting weight and/or activation values

- The learning data set can be further augmented with artificial data, which then labeled (e.g. classified or predicted) by neural network, so as to further increase the number of observations to make the splitting decisions when building the decision tree.
- Both pedagogical and decompositional approach evaluated in terms of their accuracy, conciseness (e.g, number of rules, number of conditions per rule), and fidelity

		Neural Network Classification	
Rule set classification		Good	Bad
	Good	a	b
	Bad	c	d

Fidelity = (a+d)/(b+c)

Example in next slide

Pedagogical Rule Extraction
Example: Credit Risk (3)

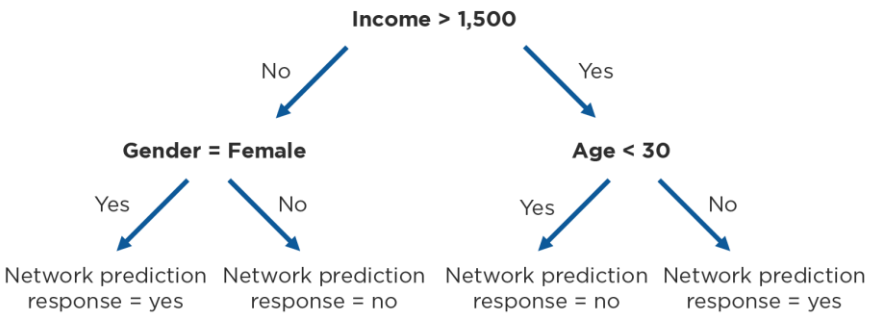
Customer	Age	Income	Gender	...	Response
Emma	28	1,000	F		No
Will	44	1,500	M		Yes
Dan	30	1,200	M		No
Bob	58	2,400	M		Yes

Step 1: Start from original data.

Step 2 : Build a neural network.

Customer	Age	Income	Gender	Network Prediction	Response
Emma	28	1,000	F	No	No
Will	44	1,500	M	Yes	Yes
Dan	30	1,200	M	Yes	No
Bob	58	2,400	M	Yes	Yes

Step 3: Get the network predictions and add them to the data set.



Step 4: Extract rules relating network predictions to original inputs. Generate additional data where necessary.

Two Stage Models

Example: Credit Risk (4)

Customer	Age	Income	Gender	...	Response
Emma	28	1,000	F		No
Will	44	1,500	M		Yes
Dan	30	1,200	M		No
Bob	58	2,400	M		Yes

Step 1: Start from original data.

Customer	Age	Income	Gender	...	Response	Logistic Regression Output
Emma	28	1,000	F		No (=0)	0.44
Will	44	1,500	M		Yes (=1)	0.76
Dan	30	1,200	M		No (=0)	0.18
Bob	58	2,400	M		Yes (=1)	0.88

Step 2 : Build logistic regression model.

Customer	Age	Income	Gender	...	Response	Logistic Regression Output	Error
Emma	28	1,000	F		No (=0)	0.44	-0.44
Will	44	1,500	M		Yes (=1)	0.76	0.24
Dan	30	1,200	M		No (=0)	0.18	-0.18
Bob	58	2,400	M		Yes (=1)	0.88	0.12

Step 3: Calculate errors from logistic regeression model.

Step 4: Build NN predicting errors from logistic regression model.

Customer	Age	Income	Gender	...	Logistic Regression Output	NN Output	Final Output
Bart	28	1,000	F		0.68	-0.32	0.36

Step 5: Score new observations by adding up logistic regression and NN scores.

Agenda

- Neural Networks Models
- ANN Advantage & ANN Applications
- Artificial Neural Networks
- Learning Principle for ANN
- Neural Network Properties
- Neural Network Step
- **Basic Types of Learning**

Basic Types of Learning

ANN

- Supervised Learning

- Unsupervised Learning

Types of Problems

- Mathematical Modeling (Function Approximation)
- Classification
- Clustering
- Forecasting
- Vector Quantization
- Pattern Association
- Control
- Optimization

Mathematical Modeling
(Function Approximation)

- Modeling – often mathematical relationship between two sets of data unknown analytically
- No closed form expression available
- Empirical data available defining output parameters for each input
- Data is often noisy
- Need to construct a mathematical model that correctly generates outputs from inputs
- Approximation carried out using ……………………………
- Network learns ………………………………………….
- Trained Neural Net can be substituted for …………………………..
- Fast computation, reasonable approximation

Classification

- Assignment of objects to specific class
- Given a database of objects and classes of those objects
- Induce ……………………………
- Create a classifier that will ………………..

Clustering

- Grouping together objects similar to one another
- Usually based on some “distance” measurement in object parameter space
- Objects and distance relationships available
- No prior info on classes or groupings
- Objects clustered based on ………………..
- Clustering may precede ………………………
- Similar to statistical k-nearest neighbor clustering method

Forecasting

- Prediction of future events based on history
- Laws underlying behavior of system sometimes hidden; too many related variables to handle
- Trends and regularities often masked by noise
- Prediction system must be able to ……………….
- Time series forecasting – special case of ……….
- Weather, Stock market indices, machine performance

Vector Quantization

- Kohonen classifier – most well known
- Object space divided into several connected regions
- Objects classified based on proximity to regions
- Closest region or node is “winner”
- Form of compression of high dimensional input space
- Successfully used in many geological and environmental classification problems where input object characteristics often unknown

Pattern Association

- Auto-associative systems useful when incoming data is a corrupted version of actual object e.g. face, handwriting
- Corrupt input sample should trigger ……………
- Require a response which …………………
- May require several iterations of repeated modification of input
- Will be discussed under ……………………..

Control

- Manufacturing, Robotic and Industrial machines have complex relationships between input and output variables
- Output variables define state of machine
- Input variables define machine parameters determined by operation conditions, time and human input
- System may be static or dynamic
- Need to map inputs to outputs for stable smooth operation
- Examples include chemical plants, truck backup, robot control

Optimization

- Requirement to improve system performance or costs subject to constraints
- Maximize or Minimize …………………….
- “Terrain” of objective function typically very …………………………………..
- Large number of …………….. affecting objective function (high …………… of problem)
- Design variables often subject to ……………….
- Lots of local ………………………..
- Neural nets can be used to find global optima

Telkom University TÜVRheinland® Precisely Right.

2017 Fifth International Conference on Information and Communication Technology (ICoICT)

Indonesia Infrastructure and Consumer Stock Portfolio Prediction using Artificial Neural Network Backpropagation

S. Prashant Mahasagara[1], Andry Alamsyah[2], Brady Rikumahu[3]
School of Economics and Business
Telkom University, Bandung, Indonesia
[1]mahasagara@students.telkomuniversity.ac.id
[2]andrya@telkomuniversity.ac.id, [3]bradyrikumahu@telkomuniversity.ac.id

***Abstract*— Artificial Neural Network (ANN) method is increasingly popular to build predictive model that generated small error prediction. To have a good model, ANN needs large dataset as an input. ANN backpropagation is a gradient decrease method to minimize the output error squared.**

Stock price movements are suitable with ANN requirement : it is a large data set because stock price is recorded up to every seconds, usually called high frequency data. The implementation of stock price prediction using ANN approach is quite new. The

have certain return on our investment. But certainty is hard to predict.

Stock historical price as big data has the unique pattern. The exploration of pattern is data analytics objective. Generally, data analytics techniques are description, estimation, association, clustering, classification and prediction [3]. With time-series or historical data we can build a predictive model with data analytics to predict the future

***Abstract*— Artificial Neural Network (ANN) method is increasingly popular to build predictive model that generated small error prediction. To have a good model, ANN needs large dataset as an input. ANN backpropagation is a gradient decrease method to minimize the output error squared.**

Stock price movements are suitable with ANN requirement : it is a large data set because stock price is recorded up to every seconds, usually called high frequency data. The implementation of stock price prediction using ANN approach is quite new. The predictive model help investor in building stock portfolio and their decision making process. Buying some stocks in portfolio decrease diversified risk and increases the chance of higher return.

In this paper, we show how to generate prediction model using artificial neural network backpropagation of stock price and forming portfolio with predicted price that bring prediction of the portfolio with the smallest error. The data set we use is historical stock price data from ten different company stocks of infrastructure and consumer sector Indonesia Stock Exchage.

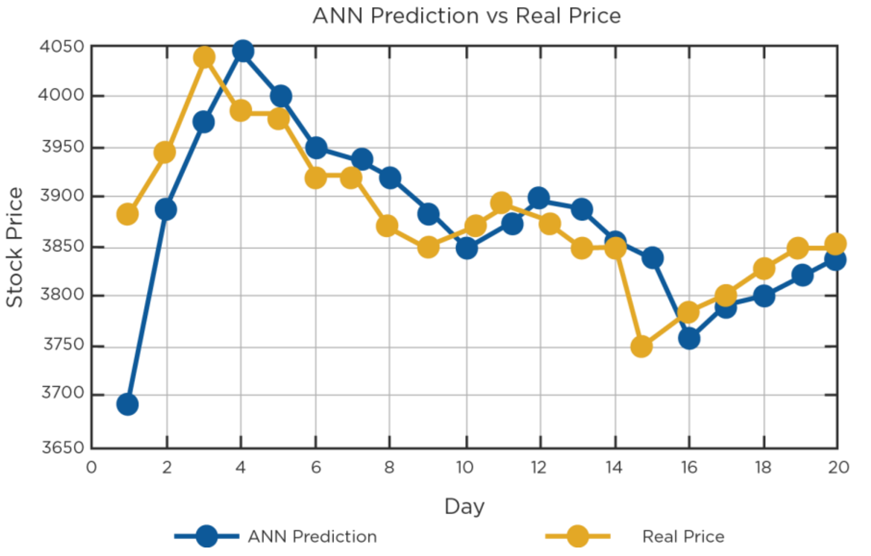
The results is for lower risk condition, ANN predictive model gives higher expected return than the return from real condition, while for higher risk, the return from the real condition is higher than the ANN predictive model.

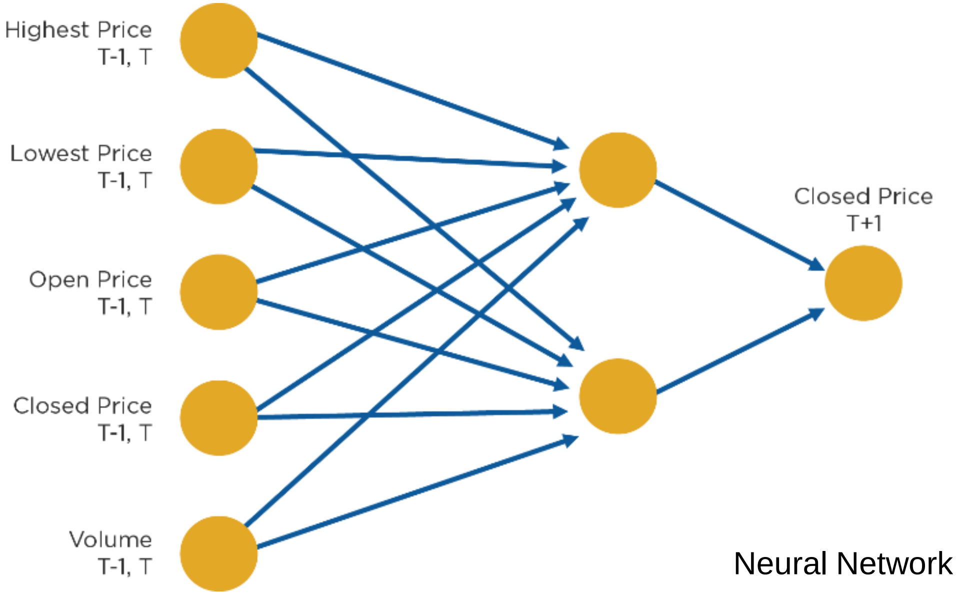
Keywords*— *Stock Portfolio, Artificial Neural Network, Backpropagation

Training Result

The Number of Neuron in Hidden Layer	Error Training
17	0.000944728
20	0.001135016
30	0.001610409
50	0.002910775

Simple Period	Error Training	
	Actual Price	Predicted Price
19/02/2016	3847	3854
22/02/2016	3748	3838





Neural Network Architecture

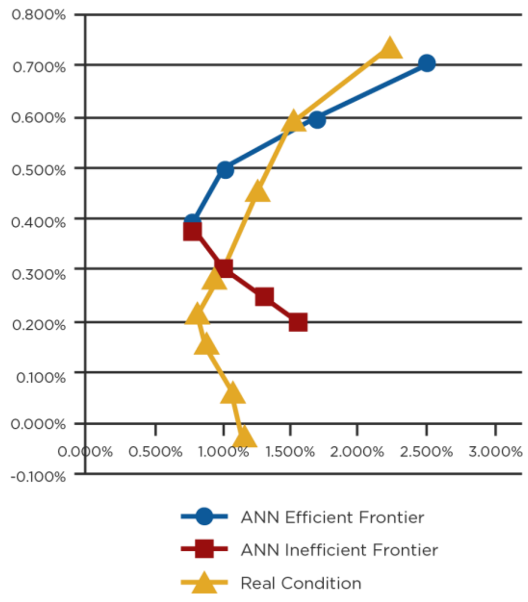
Artificial Neural Network (ANN)

Portfolio	Fund Proportion					Risk	Return
	EXCL	GGRM	ICBP	INDF	PGAS		
A	0.00%	0.00%	0.00%	100.00%	0.00%	2.57%	0.70%
B	0.00%	38.40%	0.00%	48.06%	13.54%	1.80%	0.60%
C	11.64%	42.68%	0.00%	17.97%	27.72%	1.32%	0.50%
D	41.64%	28.62%	2.43%	0.00%	27.31%	0.96%	0.37%
E	47.69%	23.46%	3.79%	0.00%	25.06%	0.97%	0.35%
F	62.82%	10.58%	7.19%	0.00%	19.41%	1.09%	0.30%
G	78.88%	0.00%	10.05%	0.00%	11.07%	1.30%	0.25%
H	100.00%	0.00%	0.00%	0.00%	0.00%	1.59%	0.21%

Eight Portfolio Alternatives Formed Based on Markowitz Portfolio Model

	RMSE (Rupiah)	MAE (Rupiah)	MAPE
EXCL	78	78	1.002%
GGRM	1736	1749	1.402%
ICBP	928	1075	3.478%
INDF	217	248	1.856%
JSMR	200	189	1.684%
KLBF	33	37	1.441%
PGAS	80	65	1.255%
TBIG	153	166	1.407%
TLKM	120	135	2.041%
UNVR	1231	1550	1.861%
Mean	477.7	529.1	1.743%

Error of Prediction



Expected Return

Artificial Neural Network (ANN)

The Future :
Deep Learning

- Deep learning (also known as deep structured learning or hierarchical learning) is part of a broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms. Learning can be supervised, partially supervised or unsupervised.
- Some representations are loosely based on interpretation of information processing and communication patterns in a biological nervous system, such as neural coding that attempts to define a relationship between various stimuli and associated neuronal responses in the brain. Research attempts to create efficient systems to learn these representations from large-scale, unlabeled data sets.
- Deep learning architectures such as deep neural networks, deep belief networks and recurrent neural networks have been applied to fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation and bioinformatics where they produced results comparable to and in some cases superior to human experts.