

Studi Literatur Konsep Dasar Machine Learning Dan Neural Network

Dadang Priyanto¹, Muhammad Zarlis², Herman Mawengkang³, Syahril Efendi⁴

¹Program Pascasarjana Ilmu Komputer, Departemen Ilmu Komputer, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara, Medan, Indonesia.

^{2,3,4} Departemen Ilmu Komputer, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara, Medan, Indonesia
dadangpriyanto@students.usu.ac.id

Abstrak. Kecerdasan buatan merupakan bagian dari ilmu komputer yang pemanfaatannya dapat digunakan dalam berbagai bidang untuk menyelesaikan permasalahan kehidupan manusia. Kecerdasan buatan merupakan program yang dapat meniru pola berpikir manusia. Membuat kecerdasan buatan bisa menggunakan cara Planning, Representasi Pengetahuan, Machine Learning, dan Multi Agent System. Machine Learning (ML) adalah konsep pendekatan dalam artificial intelligence (AI) yang menekankan pada pendekatan matematis untuk melakukan inferensi pada data untuk penyelesaian suatu masalah atau optimisasi yang meniru perilaku manusia. Machine Learning mempunyai dua tujuan utama yaitu Prediksi dan untuk mendapatkan ilmu pengetahuan (knowledge discovery). Dalam pengelolaan data set machine learning mempunyai algoritma yaitu: Supervised Learning, Semi Supervised Learning, Unsupervised Learning dan Reinforcement Learning. Artificial Neural Network (ANN) merupakan salah satu metode dalam Machine Learning (ML) yang menirukan syaraf manusia. Artificial neural network adalah salah satu algoritma supervised learning yang terkenal dan bisa juga digunakan untuk semi-supervised ataupun unsupervised learning. ANN yang bentuknya paling kecil disebut single perceptron yang hanya terdiri dari sebuah neuron. Multilayer perceptron (MLP) yang biasa disebut juga sebagai feedforward neural network, mempunyai beberapa layers, namun secara umum ada tiga layers yaitu input, hidden, dan output layer.

Kata kunci: Kecerdasan buatan, Machine learning, Artificial Neural Network, Multilayer Perceptron

I. PENDAHULUAN

Kecerdasan buatan (artificial intelligence (AI)) sudah akrab ditelinga kita, terutama pada orang yang menekuni bidang Informatika dan ilmu komputer. Kecerdasan buatan merupakan program-program yang memiliki bentuk model matematis (instruksi) yang kita sebut sebagai agen.[4] Banyak permasalahan yang bisa diselesaikan dengan bantuan kecerdasan buatan (AI). Sangat berbeda dengan program biasa yang menghasilkan aksi berdasarkan instruksi. Kecerdasan buatan mempunyai tujuan yaitu untuk menciptakan program yang dapat memprogram (output program adalah sebuah program). Pada dasarnya program adalah automation yang menjalankan suatu instruksi. Seperti program pada umumnya, agen kecerdasan buatan juga menjalankan suatu instruksi. Yang membedakan dari keduanya adalah kemampuan untuk belajar. Adapun belajar yang dimaksudkan adalah mesin mampu belajar apabila program mampu mengupdate parameter.[1] Kecerdasan pada manusia memiliki banyak dimensi, ada unsur Logika, seni, bahasa dan lain-lain, sehingga kecerdasan ini sifatnya multi disiplin ilmu dan sampai saat ini belum ada ilmu yang mampu membuat kecerdasan buatan seperti halnya yang dimiliki oleh manusia. Bentuk program seperti deterministic dianggap sebagai kecerdasan buatan karena mempunyai acting rationally. Untuk membuat kecerdasan buatan bisa menggunakan berbagai cara seperti halnya berikut ini : [4]

1. Planning, yaitu tindakan start state dan goal state, dimana agen harus mampu merencanakan sekuens aksi untuk mengubah start state menjadi goal state. Dalam merencanakan aksi sekuens, ada beberapa kendala (constraints) yang harus dioptimisasi.
2. Representasi pengetahuan, yaitu merepresentasikan suatu pengetahuan dalam bentuk formal. Dengan representasi formal tersebut, dapat melakukan inferensi dengan operasi logika menggunakan bentuk simbolik, seperti simbol logika preposisi, simbol logika orde pertama (first order logic), Fuzzy Logic, abductive reasoning, ontologi, maupun jaringan semantik (semantic web) [2]
3. Machine learning, yaitu suatu teknik untuk melakukan inferensi terhadap data dengan menggunakan pendekatan matematis. Machine learning digunakan untuk membuat model matematis yang merefleksikan pola data.

4. Multi agent sistem, pada dasarnya sistem yang memiliki banyak agen dan berinteraksi satu dengan lainnya untuk menyelesaikan permasalahan. Agen satu mengerjakan suatu hal tertentu, kemudian bekerja bersama untuk menyelesaikan masalah yang lebih besar (yang tidak dapat diselesaikan sendiri).

5. Dan lain-lainya

Representasi pengetahuan dan machine learning kedua dapat digunakan untuk melakukan inferensi. Perbedaan diantara keduanya adalah Inferensi pada bidang keilmuan adalah representasi pengetahuan mencakup tentang bagaimana cara (langkah dan proses) mendapatkan sebuah keputusan. Pada machine learning, inferensi yang dimaksud lebih menitikberatkan ranah hubungan variabel. Pada representasi pengetahuan lebih sulit menerapkan disebabkan kita harus mendefinisikan lebih dahulu term dan aturan-aturan logika. Representasi pengetahuan tidak cocok digunakan dalam pengolahan data yang besar. Machine Learning (ML) adalah representasi data dalam bentuk model matematis karena machine learning merupakan turunan dari Ilmu matematika dan statistic. Machine learning direkomendasikan untuk pengolahan data yang besar karena dengan machine learning model inferensi dapat dilakukan secara otomatis, yang penggunaannya tergantung dari domain permasalahannya.

II. METODOLOGI PENELITIAN

A. *Machine Learning (ML)*

Machine Learning (ML) merupakan konsep dalam artificial intelligence (AI) yang menekankan pada pendekatan matematis untuk melakukan inferensi pada data untuk penyelesaian suatu masalah atau optimisasi yang meniru perilaku kecerdasan manusia. Machine Learning dapat digunakan untuk Prediksi dan untuk memperoleh ilmu pengetahuan (knowledge discovery). Machine learning mengharuskan adanya proses pelatihan, pembelajaran, ataupun training. Oleh sebab itu, ML membutuhkan data untuk dipelajari yang disebut sebagai data set (kumpulan data atau contoh sample) [3]. Machine learning dapat dikategorikan Statistical learning theory yaitu suatu teknik untuk prediksi masa depan dan menyimpulkan atau mendapatkan pengetahuan dari data secara rasional dan non paranormal. Hal ini sesuai dengan konsep intelligent agent, yaitu bertindak berdasarkan lingkungan seperti yang ada dalam data set. Data sample digunakan pada Knowledge discovery untuk membuat model dalam menggeneralisasi "aturan" atau "pola" data sehingga kita dapat menggunakannya untuk mendapatkan informasi atau membuat keputusan.[5,6]

Metode Machine Learning cukup banyak yang dapat digunakan dalam mendapatkan inferensi salah satunya adalah Support Vector Machine (SVM) dan Neural Network (NN). Dalam pengelolaan data set pada ML digunakan algoritma yang dibagi menjadi 4 kategori yaitu : Supervised learning, Semi-supervised learning, Unsupervised learning, Reinforcement learning.

1. *Supervised learning*

Supervised learning adalah pembelajaran yang terarah atau terawasi. Sebagai contoh dalam proses pembelajaran Guru dan siswa, maka guru bertugas mengajar dan mengarahkan sedangkan siswa yang diajar. Manusia berperan sebagai guru, dan mesin berperan sebagai siswa. Pembelajaran metode ini disebut supervised dikarenakan ada yang memberikan contoh jawaban (desired output).

2. *Semi-supervised Learning*

Semi supervised learning hamper sama dengan supervised learning, yang membedakan adalah pada proses pelabelan data. Pada supervised learning, pada contoh diatas ada "guru" yang harus membuat "kunci jawaban" input-output. Sedangkan pada semi-supervised learning tidak ada "kunci jawaban" eksplisit yang harus dibuat guru. Kunci jawaban ini dapat diperoleh secara otomatis, sebagai contoh dari hasil clustering. Pada kategori pembelajaran ini, umumnya kita hanya memiliki sedikit data. Kita kemudian menciptakan data tambahan baik menggunakan supervised maupun unsupervised learning, kemudian membuat model belajar dari data tambahan tersebut.

3. *Unsupervised Learning*

Pada contoh sebelumnya di supervised learning ada guru yang mengajar, kalau pada unsupervised learning tidak ada guru yang mengajar. Perbedaan utamanya adalah pada unsupervised learning tidak ada desired output. Contoh penggunaan unsupervised learning adalah metode clustering dan algoritma yang terkenal adalah K-means. Unsupervised learning adalah mencari sifat-sifat (properties) data.

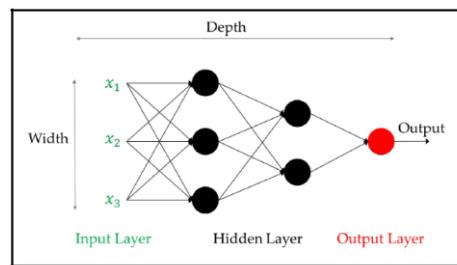
4. *Reinforcement Learning*

Reinforcement Learning (RL) merupakan Paradigma baru dalam Machine Learning. Reinforcement Theory (RT) dalam ilmu psikologi sebagai dasar dalam konsep Reinforcement Learning (RL). Teori ini menjelaskan bagaimana dalam dinamika kehidupan seseorang bisa memilih, menentukan dan mengambil keputusan. Teori ini bisa diterapkan pada situasi tertentu yang seringkali dihadapi manusia. Reinforcement Theory ini menjelaskan bagaimana perilaku manusia itu adalah hasil kompilasi dari

pengalaman-pengalaman yang ia temui/hadapi sebelumnya, (Consequences influence behavior). Contoh pemanfaatan RL adalah machine yang disebut SNARC (Stochastic Neural-Analog Reinforcement Calculator), MENACE (for Matchbox Educable Noughts and Crosses Engine) dan GLEE (Game Learning Expectimaxing Engine). Reinforcement Theory (RT) terdiri dari Empat komponen dasar, yaitu Policy, Reward function Value function dan Model of environment.

B. Neural network (NN)

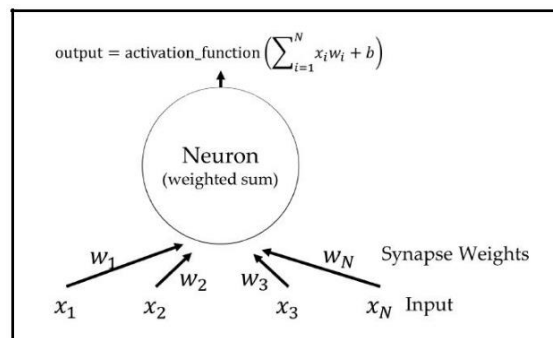
Artificial Neural Network (ANN) merupakan salah satu metode dalam Machine Learning (ML) yang menirukan syaraf manusia. Artificial neural network adalah salah satu algoritma supervised learning yang terkenal dan bisa juga digunakan untuk semi-supervised ataupun unsupervised learning [7, 8, 9, 10, 11, 16]. NN mempunyai lapisan masukan (input layer) dan lapisan keluaran (output layer). Pada setiap lapisan mempunyai satu atau beberapa unit neuron, dan mempunyai sebuah fungsi aktivasi dari unit tersebut untuk menentukan sebuah keluaran. Untuk meningkatkan kemampuan dari NN, dapat ditambahkan lapis tersembunyi atau hidden layer. Data training dapat digunakan untuk melatih NN, semakin banyak data training maka akan semakin baik unjuk kerja dari NN tersebut. Tetapi NN juga mempunyai keterbatasan pada jumlah lapisan, karena semakin banyak jumlah lapisan semakin banyaknya juga jumlah iterasi atau training yang dibutuhkan. Beberapa aplikasi NN antara lain untuk Principal Component Analysis, regresi, klasifikasi citra, dan lain-lain. Secara matematis, ANN ibarat sebuah graf. Selain input layer dan output layer NN juga mempunyai neuron (vertex), serta sinapsis (edge). ANN mempunyai struktur seperti graf, sehingga operasi pada ANN mudah dijabarkan dalam notasi aljabar linear. Seperti terlihat pada gambar (1), Depth (kedalaman) ANN mengindikasikan jumlah layer. Sementara width (lebar) ANN mengindikasikan jumlah unit pada layer.



GAMBAR 1. DEEP NEURAL NETWORK

1. Single Perceptron

ANN yang bentuknya paling kecil disebut single perceptron yang hanya terdiri dari sebuah neuron, seperti terlihat pada Gambar (2) Secara matematis, terdapat fitur vector x yang menjadi input bagi neuron tersebut. Satu fitur vector merepresentasikan suatu data point, event atau instans. Neuron akan memproses masukan (input) ' x ' dengan perhitungan jumlah perkalian antara nilai input dan bobot (synapse weight), yang dilewatkan pada fungsi non-linear [12, 13, 4]. Pada proses training, yang dioptimasi adalah nilai bobot (synapse weight)/(learning parameter), serta terdapat juga bias ' b ' sebagai kontrol tambahan. Keluaran/(output) dari neuron merupakan hasil fungsi aktivasi dari perhitungan jumlah perkalian antara nilai input dengan bobot (synapse weight). Ada beberapa macam fungsi aktivasi, seperti sign function, step function, rectifier dan sigmoid function. Salah satu bentuk tipe sigmoid function diberikan pada persamaan (1). Bila diplot menjadi grafik, fungsi ini memberikan bentuk seperti huruf S.



GAMBAR 2. SINGLE PERCEPTRON

$$\sigma(u) = \frac{1}{1 + e^{-u}} \quad \text{Persamaan (1)}$$

Persamaan fungsi sigmoid pada single perceptron dapat dituliskan kembali menjadi

$$O = f(x, w + b) \quad \text{Persamaan (2)}$$

Dimana o adalah output dan f adalah fungsi non-linear yang dapat diturunkan secara matematis (differentiable non linear function (selanjutnya disebut “fungsi non-linear” saja.)). Prinsip kerja ANN pada dasarnya menggunakan fungsi persamaan non linear, hal inilah yang menjadikan ANN hebat dalam proses prediksi. Untuk melakukan learning parameter (pembelajaran) single perceptron, training bisa dilakukan dengan menggunakan aturan (perceptron training rule). Prosesnya sebagai berikut [4, 12, 13,16]:

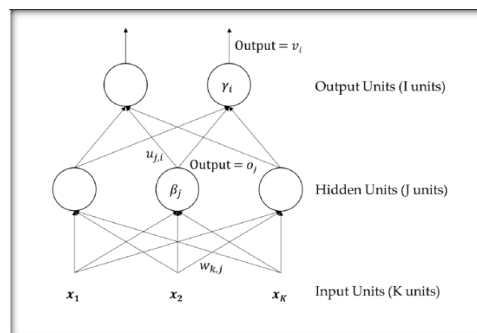
1. Langkah pertama menginisiasi nilai bobot (synapse weights), bisa acak/random ataupun dengan aturan tertentu. [1,11,16]
2. Selanjutnya lewatkan masukan/input pada neuron, kemudian kita akan memperoleh nilai keluaran(output). Langkah ini biasa disebut feedforward.
3. Selanjutnya bandingkan nilai output (actual output) dengan nilai desired output.
4. Apabila nilai keluaran(output) sama dengan desired output, maka tidak perlu melakukan perubahan apapun.
5. Selanjutnya apabila nilai output tidak sesuai dengan desired output, hitung nilai error (loss) kemudian lakukan perubahan terhadap learning parameter (synapse weight).
6. Selanjutnya langkah-langkah tersebut diulang sampai tidak ada perubahan nilai error, dan nilai error kurang dari sama dengan nilai threshold (biasanya mendekati 0), atau sudah mengulang proses latihan sebanyak T kali (threshold). Fungsi error diberikan seperti terlihat pada persamaan (3) (dapat diganti dengan nilai absolute). Perubahan nilai bobot (synapse weight) diberikan seperti pada persamaan (4), dimana y adalah desired output, $o = f(x \cdot w + b)$ menggambarkan actual output untuk x sebagai input, η disebut sebagai learning rate.

$$E(w) = (y - o)^2 \quad \text{Persamaan (3)}$$

$$\Delta w_i = \eta(y - o)x_i \quad \text{Persamaan (4)}$$

2. Multilayer Perceptron

Multilayer perceptron (MLP) biasa disebut sebagai feedforward neural network. MLP mempunyai beberapa layers, setidaknya ada tiga layers yaitu input, hidden, dan output layer seperti terlihat pada Gambar(3). Input layer menerima input (tanpa melakukan operasi apapun), kemudian nilai input (tanpa dilewatkan ke fungsi aktivasi) diberikan ke hidden units seperti di persamaan (5). Pada hidden units, input diproses dilakukan perhitungan dengan fungsi aktivasi untuk setiap neuron, selanjutnya hasilnya diberikan ke layer berikutnya dengan persamaan (6). Keluaran (output) dari input layer akan menjadi nilai input untuk hidden layer. Demikian seterusnya hidden layer akan memberikan hasilnya untuk output layer. Proses ini disebut sebagai feedforward [14, 3]. Berlaku juga untuk artificial neural network apabila memiliki lebih dari tiga layers. Gradient Based Optimization dapat digunakan untuk melakukan optimisasi. MLP adalah gabungan dari banyak fungsi non linear. Gambar (3), masing-masing neuron terkoneksi dengan semua neuron pada layer berikutnya. Struktur ini disebut juga sebagai fully connected.



GAMBAR 3. MULTILAYER PERCEPTRON

$$o_j = \sigma \left(\sum_{k=1}^K \chi^k \omega_{k,j} + \beta_j \right)$$

Persamaan (5)

$$v_i = \sigma \left(\sum_{j=1}^J o_j v_{j,i} + \gamma_i \right) = \sigma \left(\sum_{j=1}^J \sigma \left(\sum_{k=1}^K \chi^k \omega_{k,j} + \beta_j \right) v_{j,i} + \gamma_i \right)$$

Persamaan (6)

Persamaan (5) dan (6) untuk menghitung output pada layer yang berbeda. v, ω adalah learning parameters. β, γ melambangkan noise atau bias. K adalah banyaknya input units dan J adalah banyaknya hidden units. Pada Persamaan (6) dapat dilakukan penyederhanaan sehingga menjadi persamaan (7).

$$\mathbf{v} = \sigma(\mathbf{oU} + \gamma) = \sigma((\sigma(\mathbf{xW} + \beta))\mathbf{U} + \gamma)$$

Persamaan (7)

Untuk melatih MLP, algoritma yang umumnya digunakan adalah backpropagation [15]. Yaitu digunakan untuk memperbaharui bobot parameter (synapse weights) secara bertahap dari output ke input layer. berdasarkan error/loss (output dibandingkan dengan desired output). Intinya adalah mengkoreksi synapse weight dari output layer ke hidden layer, kemudian error tersebut dipropagasi ke layer sebelumnya. Perubahan pada synapse weight suatu layer dipengaruhi oleh perubahan synapse weight pada layer berikutnya. Langkah pertama dengan memberikan pasangan input (x) dan desired output (y) sebagai training data. Agar nilai loss minimal, digunakan algoritma backpropagation dengan prinsip gradient descent yaitu menghitung loss ANN dengan menerapkan **fungsi aktivasi sigmoid**.

Error, untuk MLP diberikan pada persamaan (8) (untuk satu data point), dimana I adalah banyaknya output unit dan \emptyset adalah kumpulan weight matrices (semua parameter pada MLP).

$$E(\theta) = \frac{1}{2} \sum_{i=1}^I (y_i - v_i)^2$$

Persamaan (8)

Proses penurunan untuk melatih MLP, Error/loss diturunkan terhadap tiap learning parameter.

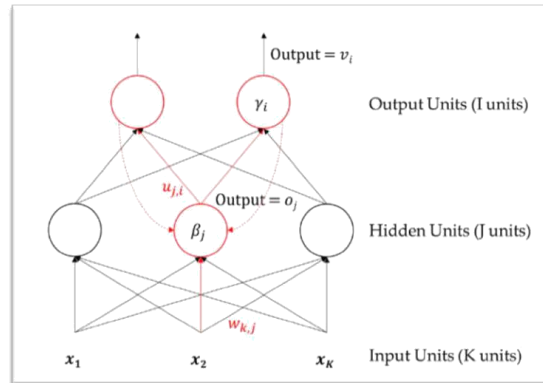
Diferensial $v_{j,i}$ diberikan oleh turunan sigmoid function :

$$\begin{aligned} \frac{\delta E(\theta)}{\delta u_{j,i}} &= (y_i - v_i) \frac{\delta v_i}{\delta u_{j,i}} \\ &= (y_i - v_i) v_i (1 - v_i) o_j \end{aligned}$$

Diferensial $w_{k,j}$ diberikan oleh turunan sigmoid function

$$\begin{aligned} \frac{\delta E(\theta)}{\delta w_{k,j}} &= \sum_{i=1}^I (y_i - v_i) \frac{\delta v_i}{\delta w_{k,j}} \\ &= \sum_{i=1}^I (y_i - v_i) \frac{\delta v_i}{\delta o_j} \frac{\delta o_j}{\delta w_{k,j}} \\ &= \sum_{i=1}^I (y_i - v_i) (v_i (1 - v_i) u_{j,i}) (o_j (1 - o_j) x_k) \end{aligned}$$

Diferensial $w_{k,j}$ memiliki \sum sementara $u_{j,i}$ tidak ada, hal ini disebabkan karena $u_{j,i}$ hanya berkorespondensi dengan satu output neuron. Sementara $w_{k,j}$ berkorespondensi dengan banyak output neuron. Dengan kata lain, nilai $w_{k,j}$ mempengaruhi hasil operasi yang terjadi pada banyak output neuron, sehingga banyak neuron mempropagasi error kembali ke $w_{k,j}$. Ilustrasi diberikan pada Gambar 4.



GAMBAR 4. RINCIAN PENGGUNAAN Σ PADA PENURUNAN DARI HIDDEN KE INPUT LAYER

Metode penurunan serupa dapat juga digunakan untuk menentukan perubahan β dan γ . Jadi proses backpropagation untuk kasus Gambar (3) dapat diberikan seperti pada Gambar (5) dimana η adalah learning rate. Untuk artificial neural network dengan lebih dari 3 layers, bisa diturunkan persamaannya. Secara umum, proses melatih ANN (apapun variasi arsitekturnya) mengikuti framework perceptron training rule.

Input to Hidden Layer

$$o_j = \sigma \left(\sum_{k=1}^K x_k w_{k,j} + \beta_j \right)$$

Hidden to output

$$v_i = \sigma \left(\sum_{j=1}^J o_j u_{j,i} + \gamma_i \right)$$

Output Hidden :

$$\begin{aligned} \delta_i &= (y_i - v_i) v_i (1 - v_i) \\ \Delta u_{j,i} &= -\eta(t) \delta_i o_j \\ \Delta \gamma_i &= -\eta(t) \delta_i \end{aligned}$$

Hidden to Input :

$$\begin{aligned} \phi_j &= \sum_{i=1}^I \delta_i u_{j,i} o_j (1 - o_j) \\ \Delta w_{k,j} &= -\eta(t) \phi_j x_k \\ \Delta \beta_j &= -\eta(t) \phi_j \end{aligned}$$

GAMBAR 5. PROSES LATIHAN MLP MENGGUNAKAN BACKPROPAGATION

IV. KESIMPULAN

Makalah ini memberikan gambaran dasar dalam memahami Artificial Intelligence, Machine Learning, dan Neural Network. Telah diuraikan bahwa kecerdasan buatan merupakan bagian ilmu computer yang berisi program-program yang bersifat matematis dan automation, yang salah satu metode dalam penyelesaiannya bisa menggunakan Machine Learning. Metode Machine Learning yang terkenal dalam mendapatkan inferensi adalah Support Vector Machine (SVM) dan Neural Network (NN). Artificial Neural Network (ANN) merupakan metode dalam Machine Learning (ML) yang menirukan syaraf manusia.

DAFTAR PUSTAKA

- [1] P. Linz. An Introduction to Formal Language and Automata. Jones and Bartlett Publishers, Inc., USA, 2006
- [2] R. Brachman, H. Levesque. Knowledge Representation and Reasoning. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004.
- [3] T. Mitchell. Machine Learning. McGraw-Hill, 1997.
- [4] S. Russel, P. Norvig. Artificial Intelligence: A Modern Approach. Prentice Hall, 1995.
- [5] C. M. Bishop. Pattern Recognition and Machine Learning. Springer, 2006.
- [6] S. Watanabe, H. Nishimori. Fall lecture note on statistical learning theory. Lecture note for Tokyo Institute of Technology, 2016.
- [7] A. Atiya. Learning Algorithms for Neural Network. PhD thesis, California Institute of Technology, 1994.

- [8] T. Mikolov. Statistical Language Models Based on Neural Networks. PhD thesis, Brno University of Technology, 2012.
- [9] K. C. G. Corrado, Thomas Mikolov, Ilya Sutskever and Je_rey Dean. Distributed representations of words and phrases and their compositionality. In Proceedings of CoRR, 2013.
- [10] G. Corrado, Thomas Mikolov, Kai Chen and Je_rey Dean. E_icient estimation of word representations in vector space. In Proceedings of CoRR, 2013.
- [11] K. Yu. Large-scale deep learning at baidu. In Proceedings of the 22nd ACM International Conference on Information and Knowledge Management, pages 2211-2212, 2013.
- [12] M. A. Aizerman, E. A. Braverman, and L. Rozonoer. Theoretical foundations of the potential function method in pattern recognition learning. In Automation and Remote Control., number 25, pages 821{837, 1964.
- [13] F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, pages 65{386, 1958.
- [14] A. Atiya. Learning Algorithms for Neural Network. PhD thesis, California Institute of Technology, 1994.
- [15] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1. chapter Learning Internal Representations by Error Propagation, pages 318{362. MIT Press, Cambridge, MA, USA, 1986.
- [16] J.W.G. Putra, Pengenalan Pembelajaran Mesin dan Deep Learning, Tokiyo Jepang, Desember 2018, <https://www.researchgate.net/publication/323700644>