

Jan Wira Gotama Putra

# Pengenalan Konsep Pembelajaran Mesin dan Deep Learning

Edisi 1.1

July 11, 2018



Untuk Tuhan, Bangsa, dan Almamater



---

## Kata Pengantar

Buku ini ditujukan sebagai bahan pengantar (atau penunjang) mata kuliah *machine learning* untuk mahasiswa di Indonesia, khususnya tingkat sarjana (tidak menutup kemungkinan digunakan untuk tingkat pascasarjana). Buku ini hanya merupakan komplemen, bukan sumber informasi utama. Buku ini memuat materi dasar *machine learning*, yang ditulis sedemikian rupa sehingga pembaca mampu mendapatkan **intuisi**. Pembaca masih harus membaca buku-buku lainnya untuk mendapatkan pemahaman lebih dalam.

Walaupun tidak sempurna, mudah-mudahan buku ini mampu memberi inspirasi. Anggap saja membaca buku ini seperti sedang membaca “*light novel*”. Penulis ingin buku ini bisa menjadi *pointer*; i.e. dengan membaca buku ini, diharapkan kawan-kawan juga mengetahui harus belajar apa (lebih jauhnya) dalam bidang *machine learning*. Setelah membaca buku ini, pembaca diharapkan mampu membaca literatur *machine learning* yang dijelaskan secara matematis (kami memberi rekomendasi bacaan lanjutan).

Di Indonesia, penulis banyak mendengar baik dari teman, junior, senior, dll; suatu pernyataan “kuliah mengajarkan teori saja, praktiknya kurang, dan tidak relevan dengan industri”. Menurut saya di satu sisi itu benar; tapi di sisi lain, karena pemikiran macam itu terkadang kita tidak benar-benar mengerti permasalahan. Ketika mengalami kendala, kita buntu saat mencari solusi karena fondasi yang tidak kokoh. Banyak orang terburu-buru “menggunakan *tools*” karena lebih praktikal. Penulis ingin mengajak saudara/i untuk memahami konsep *machine learning* secara utuh sebelum memanfaatkan.

Buku ini menjelaskan algoritma *machine learning* dari sudut pandang “agak” matematis. Pembaca disarankan sudah memahami/mengambil setidaknya mata kuliah statistika, kalkulus, aljabar linear, pengenalan kecerdasan buatan, dan logika fuzzy. Penulis merasa banyak esensi yang hilang ketika materi *machine learning* hanya dijelaskan secara deskriptif karena itu buku ini ditulis dengan bahasa “agak” matematis. Saat membaca buku ini, disarankan membaca secara runtun. Gaya penulisan buku ini **santai/semiformal** agar lebih mudah dipahami, dengan notasi matematis dibuat seminimal mungkin, mudah-mudahan tanpa mengurangi esensi materi.

Buku ini ditulis menggunakan template monograph (L<sup>A</sup>T<sub>E</sub>X) dari Springer yang dimodifikasi. Dengan demikian, mungkin ada kesalahan pemenggalan kata. Tentunya, buku tidak lepas dari kekurangan, misalnya kesalahan tipografi. Kami sarankan pembaca untuk membaca secara seksama, termasuk menginterpretasikan variabel pada persamaan.

### Petunjuk Penggunaan

Struktur penyajian buku ini dapat dijadikan acuan sebagai struktur kuliah *machine learning* untuk satu semester (bab 1 untuk sesi pertama, dst). Agar dapat memahami materi per bab, bacalah keseluruhan isi bab secara utuh sebelum mempertanyakan isi materi. Penulis sangat menyarankan untuk membahas soal latihan sebagai tambahan materi (bisa juga sebagai PR). Soal latihan ditujukan untuk mengarahkan apa yang harus dibaca/dipahami lebih lanjut.

Pembaca dipersilahkan menyebarkan (*share*) buku ini untuk alasan **NON KOMERSIAL** (pendidikan), tetapi **dimohon kesadarannya untuk tidak menyalin /meniru isi buku ini**. Bila ingin memuat konten diktat ini pada media yang pembaca kelola, dimohon untuk mengontak pengarang terlebih dahulu. Tidak semua istilah bahasa asing diterjemahkan ke Bahasa Indonesia supaya makna sebenarnya tidak hilang (atau penulis tidak tahu versi Bahasa Indonesia yang baku).

Bab lebih awal memuat materi yang relatif lebih “mudah” dipahami dibanding bab berikutnya. Buku ini memberikan contoh dimulai dari contoh sederhana (beserta contoh data). Semakin menuju akhir buku, notasi yang digunakan akan semakin simbolik, beserta contoh yang lebih abstrak. Penulis sangat menyarankan untuk **membaca buku ini secara sekuensial**.

### Kutipan

Buku ini tergolong *self-published work*, tetapi sudah di-*review* oleh beberapa orang. Kami yakin para *reviewers* adalah orang yang berkompeten. Silahkan merujuk buku ini sesuai dengan paduan cara merujuk *self-published work* (apabila diperbolehkan untuk merujuk *self-published work* pada pekerjaan kamu).

### Notasi Penting

Karakter *bold* kapital merepresentasikan matriks ( $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ ). Dimensi matriks ditulis dengan notasi  $N \times M$  dimana  $N$  merepresentasikan banyaknya baris dan  $M$  merepresentasikan banyaknya kolom. Elemen matriks direpresentasikan oleh  $\mathbf{X}_{i,j}$ ,  $\mathbf{X}_{[i,j]}$ , atau  $x_{i,j}$  untuk baris ke- $i$  kolom ke- $j$  (penggunaan akan menyesuaikan konteks pembahasan agar tidak ambigu). Karakter *bold* merepresentasikan vektor ( $\mathbf{x}$ ). Elemen vektor ke- $i$  direpresentasikan oleh  $x_i$  atau  $\mathbf{x}_{[i]}$  tergantung konteks. Ketika penulis menyebutkan vektor, yang dimaksud adalah **vektor baris** (*row vector*, memiliki dimensi  $1 \times N$ , mengadopsi notasi Goldberg [1]). Perhatikan, literatur *machine learning* lainnya mungkin tidak menggunakan notasi *row vector* tetapi *column vector*. Kami

harap pembaca mampu beradaptasi. Simbol “ $\cdot$ ” digunakan untuk melambangkan operator *dot-product*.

Kumpulan data (atau himpunan) direpresentasikan dengan karakter kapital ( $C, Z$ ), dan anggotanya (*data point*, *data entry*) ke- $i$  direpresentasikan dengan karakter  $c_i$ . Perhatikan, elemen vektor dan anggota himpunan bisa memiliki notasi yang sama (himpunan dapat direpresentasikan di komputer sebagai *array*, jadi penggunaan notasi vektor untuk himpunan pada konteks pembicaraan kita adalah tidak salah). Penulis akan menggunakan simbol  $\mathbf{x}_{[i]}$  sebagai elemen vektor apabila ambigu. Fungsi dapat direpresentasikan dengan huruf kapital maupun non-kapital  $f(\dots), E(\dots), G(\dots)$ . Ciri fungsi adalah memiliki parameter! Pada suatu koleksi vektor (himpunan vektor)  $\mathbf{D}$ , vektor ke- $i$  direpresentasikan dengan  $\mathbf{d}_i$ , dan elemen ke- $j$  dari vektor ke- $i$  direpresentasikan dengan  $\mathbf{d}_{i[j]}$ ,  $\mathbf{D}_{i,j}$ , atau  $\mathbf{D}_{[i,j]}$  (karena sekumpulan vektor dapat disusun sebagai matriks).

Karakter non-kapital tanpa *bold* atau indeks ( $a, b, c, x, y, z$ ) merepresentasikan *random variable* (statistik) atau variabel (matematik). Secara umum, saat *random variable* memiliki tertentu, dinotasikan dengan  $x = X$  (nilai tertentu dinotasikan dengan huruf kapital), kecuali disebutkan secara khusus saat pembahasan. Probabilitas direpresentasikan dengan karakter kapital ( $P$ ), dengan karakter non-kapital merepresentasikan probability density ( $p$ ). Penulis yakin pembaca dapat menyesuaikan interpretasi simbol berdasarkan konteks pembahasan. Untuk menginterpretasikan notasi lain, selain yang diberikan pada paduan ini, mohon menyesuaikan dengan ceritera pembahasan.

### Ucapan Terima Kasih

Penulis ingin mengucapkan terima kasih pada Bapak/Ibu/Saudara/i atas kontribusi pada pengembangan dan penulisan buku ini: Adhiguna Surya Kuncoro, Arief Yudha Satria, Candy Olivia Mawalim, Chairuni Aulia Nusapati, Genta Indra Winata, Hayyu Luthfi Hanifah, I Gede Mahendra Darmawiguna, dan Tifani Warnita.

Tokyo, Jepang

Jan Wira Gotama Putra  
<https://wiragotama.github.io/>





---

## Daftar Isi

---

### Bagian I Pengetahuan Dasar

---

<b>1</b>	<b>Pengenalan</b>	3
1.1	Kecerdasan Buatan	3
1.2	Intelligent Agent	6
1.3	Konsep Belajar	8
1.4	Statistical Learning Theory	8
1.5	Training, Development, Testing Set	10
1.6	Supervised Learning	12
1.7	Regresi	14
1.8	Semi-supervised Learning	14
1.9	Unsupervised Learning	15
1.10	Proses Belajar	16
1.11	Tips	17
1.12	Contoh Aplikasi	18
	Soal Latihan	18
<b>2</b>	<b>Fondasi Matematis</b>	19
2.1	Probabilitas	19
2.2	Probability Density Function	21
2.3	Expectation dan Variance	23
2.4	Bayesian Probability	23
2.5	Gaussian Distribution	25
2.6	Teori Keputusan	27
2.7	Teori Informasi	29
	2.7.1 Entropy	29
	2.7.2 Relative Entropy dan Mutual Information	30
2.8	Matriks	31
2.9	Bacaan Lanjutan	32
	Soal Latihan	32

<b>3</b>	<b>Data Analytics</b>	35
3.1	Pengenalan Data Analytics	35
3.2	Nilai Atribut dan Transformasi	37
3.3	Ruang Konsep	38
3.4	Linear Separability	39
3.5	Seleksi Fitur	40
3.6	Classification, Association, Clustering	41
3.7	Mengukur Kinerja	42
3.8	Evaluasi Model	42
3.9	Kategori Jenis Algoritma	44
3.10	Tahapan Analisis	44
	Soal Latihan	45

---

## Bagian II Algoritma Pembelajaran Mesin

---

<b>4</b>	<b>Algoritma Dasar</b>	49
4.1	Naive Bayes	49
4.2	K-means	51
4.3	K-nearest-neighbor	54
	Soal Latihan	54
<b>5</b>	<b>Model Linear</b>	57
5.1	Curve Fitting dan Error Function	57
5.2	Binary Classification	60
5.3	Log-linear Binary Classification	61
5.4	Multi-label Classification	62
5.5	Pembelajaran sebagai Permasalahan Optimisasi	64
5.6	Batasan Model Linear	68
5.7	Overfitting dan Underfitting	69
5.8	Regularization	70
5.9	Transformasi Data	72
5.10	Bacaan Lanjutan	73
	Soal Latihan	73
<b>6</b>	<b>Pohon Keputusan</b>	75
6.1	Inductive Learning	75
6.2	ID3	76
6.3	Isu pada ID3	80
6.4	Pembagian Ruang Konsep	80
	Soal Latihan	81

<b>7</b>	<b>Support Vector Classifier</b>	83
7.1	Maximal Margin Classifier	83
7.2	Support Vector Classifier	88
7.3	Support Vector Machine	89
7.4	Klasifikasi lebih dari dua kelas	90
7.5	Tips	91
	Soal Latihan	91
<b>8</b>	<b>Hidden Markov Model</b>	93
8.1	Probabilistic Reasoning	93
8.2	Generative Model	96
8.3	Part-of-speech Tagging	97
8.4	Hidden Markov Model Tagger	100
8.5	Algoritma Viterbi	102
8.6	Proses Training Hidden Markov Model	104
	Soal Latihan	108
<b>9</b>	<b>Seleksi Fitur dan Metode Evaluasi</b>	109
9.1	Feature Engineering	109
9.2	High Dimensional Data	110
9.3	Feature Selection	110
9.3.1	Subset Selection (Feature Ablation)	111
9.3.2	Shrinkage	112
9.3.3	Principal Components Analysis (Dimension Reduction)	113
9.4	Cross Validation	114
9.5	Replicability, Overclaiming dan Domain Adaptation	116
	Soal Latihan	117
<b>10</b>	<b>Clustering</b>	119
10.1	K-means, Pemilihan Centroid, Kemiripan Data	120
10.2	Hierarchical Clustering	121
10.3	Evaluasi	122
	Soal Latihan	124

---

### Bagian III Neural Networks

---

<b>11</b>	<b>Artificial Neural Network</b>	127
11.1	Definisi	127
11.2	Single Perceptron	128
11.3	Permasalahan XOR	130
11.4	Multilayer Perceptron	132
11.5	Interpretability	135
11.6	Binary Classification	136
11.7	Multi-label Classification	136

#### XIV Daftar Isi

11.8	Deep Neural Network	137
11.9	Tips	140
11.10	Regularization and Dropout	141
11.11	Vanishing and Exploding Gradients	142
11.12	Rangkuman	143
	Soal Latihan	143
<b>12</b>	<b>Autoencoder</b>	<b>145</b>
12.1	Representation Learning	145
12.2	Singular Value Decomposition	147
12.3	Ide Dasar Autoencoder	148
12.4	Representing Context: Word Embedding	151
12.4.1	Vector Space Model	152
12.4.2	Sequential, Time Series dan Compositionality	153
12.4.3	Distributed Word Representation	154
12.4.4	Distributed Sentence Representation	156
12.5	Tips	159
	Soal Latihan	159
<b>13</b>	<b>Arsitektur Neural Network</b>	<b>161</b>
13.1	Convolutional Neural Network	161
13.1.1	Convolution	163
13.1.2	Pooling	164
13.1.3	Rangkuman	166
13.2	Recurrent Neural Network	166
13.3	Part-of-speech Tagging Revisited	171
13.4	Sequence to Sequence	174
13.4.1	Encoder	175
13.4.2	Decoder	176
13.4.3	Beam Search	176
13.4.4	Attention-based Mechanism	178
13.4.5	Variasi Arsitektur Sequence to Sequence	180
13.4.6	Rangkuman	182
13.5	Arsitektur Lainnya	182
13.6	Architecture Ablation	183
	Soal Latihan	183

---

#### Bagian IV Aplikasi dan Topik Tambahan

---

<b>14</b>	<b>Penerapan Pembelajaran Mesin</b>	<b>187</b>
14.1	Sistem Rekomendasi	188
14.1.1	Content-based Filtering	188
14.1.2	Collaborative Filtering	190
14.2	Peringkasan Dokumen	191

14.2.1 Pipelined Approach.....	193
14.2.2 Single-view Approach .....	193
14.3 Konklusi .....	194
14.4 Saran Buku Lanjutan .....	195
Soal Latihan .....	196
<b>Referensi .....</b>	<b>199</b>

