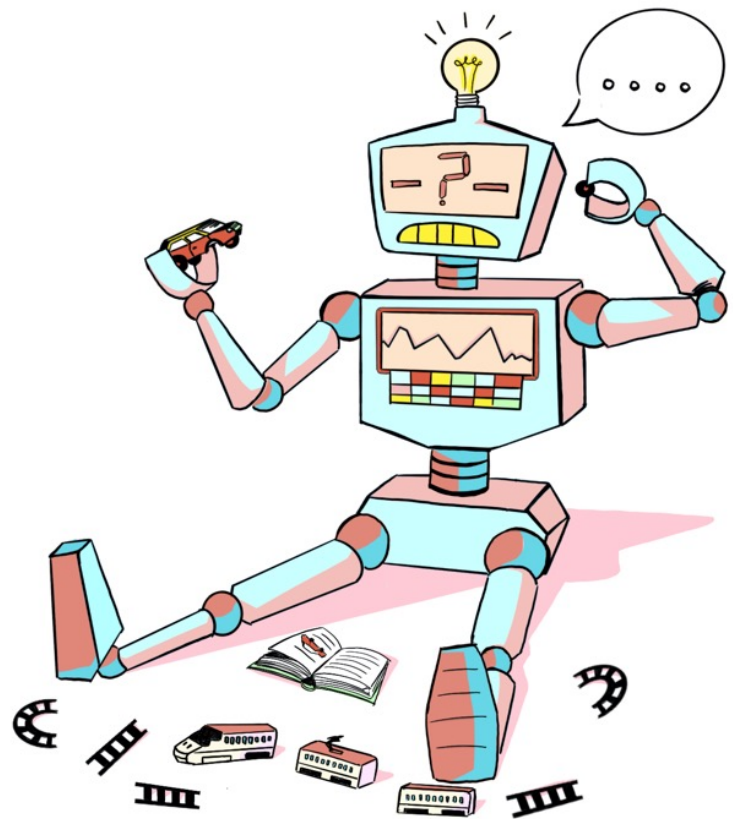


# Pengenalan Konsep Pembelajaran Mesin dan Deep Learning

Edisi 1.4 (17 Agustus 2020)



Jan Wira Gotama Putra



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 komplement, bukan sumber informasi utama. Buku ini memuat materi dasar *machine learning*, yang ditulis sedemikian rupa sehingga pembaca mampu mendapatkan **intuisi**. Materi pada buku ini tidaklah dalam (tapi tidak dangkal); artinya, 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 secara lebih jauh. Setelah membaca buku ini, pembaca diharapkan mampu membaca literatur *machine learning* yang dijelaskan secara lebih matematis ataupun mendalam (kami memberi rekomendasi bacaan lanjutan).

Di Indonesia, penulis banyak mendengar baik dari teman, junior, senior, dll; suatu pernyataan “kuliah mengajari teori saja, praktiknya kurang, dan tidak relevan dengan industri.” Tentu saja pernyataan ini cukup benar, tetapi karena permikiran semacam ini 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. Ada perbedaan yang mendasar antara orang yang hanya mampu menggunakan *tools* dan mengerti konsep secara utuh.

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. Walaupun demikian,

penulis berusaha menggunakan notasi matematis seminimal dan sesederhana mungkin, secukupnya sehingga pembaca mampu mendapatkan intuisi. Saat membaca buku ini, disarankan membaca secara runtun. Gaya penulisan buku ini **santai/semiformal** agar lebih mudah dipahami, 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* yang berdurasi 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 atau 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* (atau mungkin lebih tepat dikatakan sebagai *draft*), tetapi sudah di-*review* oleh beberapa orang. Kami yakin para *reviewer* adalah orang yang berkompeten. Silahkan merujuk buku ini sesuai dengan paduan cara merujuk *self-published work*, apabila memang diperbolehkan untuk merujuk *self-published work* pada pekerjaan pembaca.

### 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 di-

*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 tidaklah 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* dan tanpa indeks, seperti  $(a, b, c, x, y, z)$ , merepresentasikan *random variable* (statistik) atau variabel (matematik). Secara umum, saat *random variable* memiliki nilai 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

Terima kasih pada Bapak/Ibu/Saudara/i atas kontribusi pada 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. Terima kasih pada Natasha Christabelle Santosa atas desain cover.

### Catatan lain

Buku ini adalah *ongoing project*. Versi terakhir dan terakurat dapat diakses pada <https://wiragotama.github.io/>. Buku ini lebih baik dibaca versi full pdf-nya agar pranala bisa di-klik dan gambar memiliki kualitas terbaik.

# Daftar Isi

|  |           |
|--|-----------|
| <b>Bagian I Pengetahuan Dasar</b>                                | <b>1</b>  |
| <b>1 Pengenalan</b>  | <b>3</b>  |
| 1.1 Kecerdasan Buatan . . . . .                                  | 3         |
| 1.2 Intelligent Agent . . . . .                                  | 6         |
| 1.3 Konsep Belajar . . . . .                                     | 8         |
| 1.4 Statistical Learning Theory . . . . .                        | 8         |
| 1.5 Training, Validation, Testing Set . . . . .                  | 11        |
| 1.6 Supervised Learning . . . . .                                | 12        |
| 1.7 Regresi . . . . .  | 15        |
| 1.8 Semi-supervised Learning . . . . .                           | 16        |
| 1.9 Unsupervised Learning . . . . .                              | 16        |
| 1.10 Proses Belajar . . . . .                                    | 18        |
| 1.11 Tips . . . . .  | 18        |
| 1.12 Contoh Aplikasi . . . . .                                   | 19        |
| Soal Latihan . . . . .   | 19        |
| <b>2 Fondasi Matematis</b>                                       | <b>21</b> |
| 2.1 Probabilitas . . . . .                                       | 21        |
| 2.2 Probability Density Function . . . . .                       | 23        |
| 2.3 Expectation dan Variance . . . . .                           | 25        |
| 2.4 Bayesian Probability . . . . .                               | 26        |
| 2.5 Gaussian Distribution . . . . .                              | 27        |
| 2.6 Apakah Karakteristik Sampel Mencerminkan Populasi? . . . . . | 28        |
| 2.7 Teori Keputusan . . . . .                                    | 30        |
| 2.8 Hypothesis Testing . . . . .                                 | 32        |
| 2.9 Teori Informasi . . . . .                                    | 33        |
| 2.10 Matriks . . . . .   | 35        |
| 2.11 Bacaan Lanjutan . . . . .                                   | 36        |
| Soal Latihan . . . . .   | 37        |
| <b>3 Data Analytics</b>  | <b>39</b> |
| 3.1 Pengenalan Data Analytics . . . . .                          | 39        |
| 3.2 Nilai Atribut dan Transformasi . . . . .                     | 41        |
| 3.3 Ruang Konsep . . . . .                                       | 43        |



|   |  |           |
|---|--|-----------|
| 3.4   | Linear Separability . . . . .                          | 43        |
| 3.5   | Seleksi Fitur . . . . .                                | 44        |
| 3.6   | Classification, Association, Clustering . . . . .      | 45        |
| 3.7   | Mengukur Kinerja . . . . .                             | 46        |
| 3.8   | Evaluasi Model . . . . .                               | 47        |
| 3.9   | Kategori Jenis Algoritma . . . . .                     | 48        |
| 3.10  | Tahapan Analisis . . . . .                             | 49        |
|   | Soal Latihan . . . . .                                 | 49        |
| <br><b>Bagian II Algoritma Pembelajaran Mesin</b> |  | <b>51</b> |
| <b>4</b>  | <b>Algoritma Dasar</b>                                 | <b>53</b> |
| 4.1   | Naive Bayes . . . . .                                  | 53        |
| 4.2   | K-means . . . . .                                      | 56        |
| 4.3   | K-nearest-neighbor . . . . .                           | 58        |
|   | Soal Latihan . . . . .                                 | 59        |
| <b>5</b>  | <b>Model Linear</b>                                    | <b>61</b> |
| 5.1   | Curve Fitting dan Error Function . . . . .             | 61        |
| 5.2   | Binary Classification . . . . .                        | 64        |
| 5.3   | Log-linear Binary Classification . . . . .             | 65        |
| 5.4   | Multi-class Classification . . . . .                   | 66        |
| 5.5   | Multi-label Classification . . . . .                   | 70        |
| 5.6   | Pembelajaran sebagai Permasalahan Optimisasi . . . . . | 71        |
| 5.7   | Batasan Model Linear . . . . .                         | 75        |
| 5.8   | Overfitting dan Underfitting . . . . .                 | 76        |
| 5.9   | Regularization . . . . .                               | 78        |
| 5.10  | Transformasi Data . . . . .                            | 79        |
| 5.11  | Bacaan Lanjutan . . . . .                              | 81        |
|   | Soal Latihan . . . . .                                 | 81        |
| <b>6</b>  | <b>Pohon Keputusan</b>                                 | <b>83</b> |
| 6.1   | Inductive Learning . . . . .                           | 83        |
| 6.2   | ID3 . . . . .  | 84        |
| 6.3   | Isu pada ID3 . . . . .                                 | 88        |
| 6.4   | Pembagian Ruang Konsep . . . . .                       | 88        |
|   | Soal Latihan . . . . .                                 | 89        |
| <b>7</b>  | <b>Support Vector Classifier</b>                       | <b>91</b> |
| 7.1   | Maximal Margin Classifier . . . . .                    | 91        |
| 7.2   | Support Vector Classifier . . . . .                    | 96        |
| 7.3   | Support Vector Machine . . . . .                       | 97        |
| 7.4   | Klasifikasi lebih dari dua kelas . . . . .             | 98        |
| 7.5   | Tips . . . . .   | 99        |

|   |            |
|---|------------|
| Soal Latihan . . . . .  | 99         |
| <b>8 Hidden Markov Model</b>                                    | <b>101</b> |
| 8.1 Probabilistic Reasoning . . . . .                           | 101        |
| 8.2 Generative Model . . . . .                                  | 104        |
| 8.3 Part-of-speech Tagging . . . . .                            | 105        |
| 8.4 Hidden Markov Model Tagger . . . . .                        | 108        |
| 8.5 Algoritma Viterbi . . . . .                                 | 111        |
| 8.6 Proses Training Hidden Markov Model . . . . .               | 113        |
| Soal Latihan . . . . .  | 116        |
| <b>9 Seleksi Fitur dan Metode Evaluasi</b>                      | <b>117</b> |
| 9.1 Feature Engineering . . . . .                               | 117        |
| 9.2 High Dimensional Data . . . . .                             | 118        |
| 9.3 Feature Selection . . . . .                                 | 118        |
| 9.4 Evaluasi Kinerja Model . . . . .                            | 122        |
| 9.5 Cross Validation . . . . .                                  | 129        |
| 9.6 Replicability, Overclaiming dan Domain Dependence . . . . . | 132        |
| Soal Latihan . . . . .  | 132        |
| <b>10 Clustering</b>  | <b>135</b> |
| 10.1 K-means, Pemilihan Centroid, Kemiripan Data . . . . .      | 136        |
| 10.2 Hierarchical Clustering . . . . .                          | 137        |
| 10.3 Evaluasi . . . . .   | 139        |
| Soal Latihan . . . . .  | 140        |
| <b>Bagian III Artificial Neural Network</b>                     | <b>141</b> |
| <b>11 Feedforward Neural Network</b>                            | <b>143</b> |
| 11.1 Definisi Artificial Neural Network . . . . .               | 143        |
| 11.2 Single Perceptron . . . . .                                | 144        |
| 11.3 Permasalahan XOR . . . . .                                 | 147        |
| 11.4 Multilayer Perceptron . . . . .                            | 148        |
| 11.5 Interpretability . . . . .                                 | 151        |
| 11.6 Binary Classification . . . . .                            | 153        |
| 11.7 Multi-class Classification . . . . .                       | 154        |
| 11.8 Multi-label Classification . . . . .                       | 154        |
| 11.9 Deep Neural Network . . . . .                              | 155        |
| 11.10 Tips . . . . .  | 158        |
| 11.11 Regularization and Dropout . . . . .                      | 159        |
| 11.12 Vanishing and Exploding Gradients . . . . .               | 160        |
| 11.13 Rangkuman . . . . .                                       | 161        |
| Soal Latihan . . . . .  | 162        |
| <b>12 Autoencoder</b>   | <b>163</b> |

|           |  |            |
|-----------|--|------------|
| 12.1      | Representation Learning . . . . .              | 163        |
| 12.2      | Singular Value Decomposition . . . . .         | 165        |
| 12.3      | Ide Dasar Autoencoder . . . . .                | 166        |
| 12.4      | Resisting Perturbation . . . . .               | 169        |
| 12.5      | Representing Context: Word Embedding . . . . . | 171        |
| 12.6      | Tips . . . . .                                 | 179        |
|           | Soal Latihan . . . . .                         | 179        |
| <b>13</b> | <b>Arsitektur Neural Network</b>               | <b>181</b> |
| 13.1      | Convolutional Neural Network . . . . .         | 181        |
| 13.2      | Recurrent Neural Network . . . . .             | 186        |
| 13.3      | Part-of-speech Tagging Revisited . . . . .     | 191        |
| 13.4      | Sequence to Sequence . . . . .                 | 194        |
| 13.5      | Arsitektur Lainnya . . . . .                   | 203        |
| 13.6      | Architecture Ablation . . . . .                | 203        |
| 13.7      | Transfer Learning . . . . .                    | 204        |
| 13.8      | Multi-task Learning . . . . .                  | 207        |
|           | Soal Latihan . . . . .                         | 211        |
|           | <b>Bagian IV Aplikasi dan Topik Tambahan</b>   | <b>213</b> |
| <b>14</b> | <b>Penerapan Pembelajaran Mesin</b>            | <b>215</b> |
| 14.1      | Sistem Rekomendasi . . . . .                   | 216        |
| 14.2      | Peringkasan Dokumen . . . . .                  | 219        |
| 14.3      | Konklusi . . . . .                             | 222        |
| 14.4      | Saran Buku Lanjutan . . . . .                  | 224        |
|           | Soal Latihan . . . . .                         | 225        |
|           | <b>Referensi</b>                               | <b>227</b> |