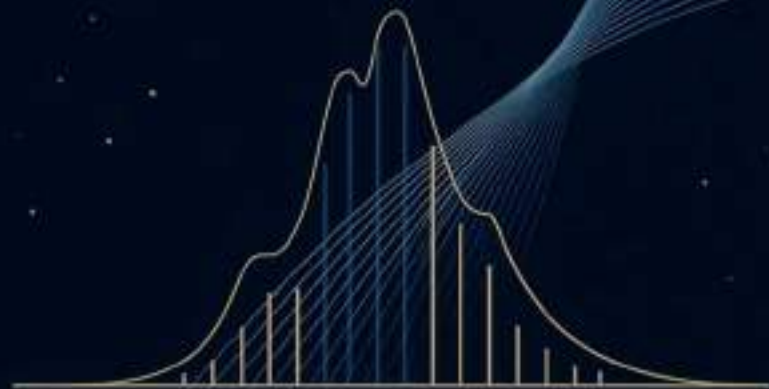


# Machine Learning :

Memahami Multinomial, Distribusi Probabilitas, dan  
Multinomial Naive Bayes



0 1 1 0 0 1 0 1 1 0 0 1  
0 0 1 0 1 1 1 0 1  
0 1 1 0 0 1 0 0 1 0 1

**Muhammad Yusril Helmi Setyawan**  
**Patah Herwanto**  
**Teguh Wiharko**



# Machine Learning

**Memahami Multinomial, Distribusi Probabilitas &  
Multinomial Naive Bayes**

Penulis:

**Muhammad Yusril helmi Setyawan**

**Patah Herwanto**

**Teguh Wiharko**

Editor:

**Fathurrahman Pratama Putra**



## **UU No. 28 Tahun 2014 tentang Hak Cipta**

### **Fungsi dan Sifat Hak Cipta Pasal 4**

Hak Cipta sebagaimana dimaksud dalam Pasal 3 huruf a merupakan hak eksklusif yang terdiri atas hak moral dan hak ekonomi.

### **Pembatasan Pelindungan Pasal 26**

Ketentuan sebagaimana dimaksud dalam Pasal 23, Pasal 24, dan Pasal 25 tidak berlaku terhadap:

- i. penggunaan kutipan singkat Ciptaan dan/atau produk Hak Terkait untuk pelaporan peristiwa aktual yang ditujukan hanya untuk keperluan penyediaan informasi aktual;
- ii. penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk kepentingan penelitian ilmu pengetahuan;
- iii. penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk keperluan pengajaran, kecuali pertunjukan dan fonogram yang telah dilakukan pengumuman sebagai bahan ajar; dan
- iv. penggunaan untuk kepentingan pendidikan dan pengembangan ilmu pengetahuan yang memungkinkan suatu Ciptaan dan/atau produk Hak Terkait dapat digunakan tanpa izin Pelaku Pertunjukan, Produser Fonogram, atau Lembaga Penyiaran.

### **Sanksi Pelanggaran Pasal 113**

1. Setiap orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp100.000.000,00 (seratus juta rupiah).
2. Setiap orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).

# Machine Learning

## Memahami Multinomial, Distribusi Probabilitas & Multinomial Naive Bayes

**Muhammad Yusril helmi Setyawan  
Patah Herwanto  
Teguh Wiharko**

EDITOR :  
**Fathurrahman Pratama Putra**

TATA LETAK:  
**Sari**

DESAIN SAMPUL:  
**Fathurraman Pratama Putra**

SUMBER:  
[www.tdjpublisher.com](http://www.tdjpublisher.com)

ISBN:  
**978-634-7363-01-5**

CETAKAN PERTAMA:  
**Agustus 2025**

UKURAN:  
**vii + 158 Hal; 17,6 cm x 25 cm**

---

Hak Cipta dilindungi Undang-Undang.  
Dilarang menggandakan atau memperbanyak sebagian atau seluruh isi buku ini dalam bentuk  
apa pun tanpa izin tertulis dari penerbit dan penulis.

---

**ANGGOTA IKAPI: 006/NTT/2022**  
**TANGGUH DENARA JAYA PUBLISHER**  
Jl. Timor Raya No. 130 B, Kota Kupang, NTT  
E-mail: [tdj.denarapublisher@gmail.com](mailto:tdj.denarapublisher@gmail.com)  
**081220051382**



# PRAKATA

Kami mengucapkan terima kasih yang tulus kepada semua kontributor yang telah mencurahkan waktu dan upaya mereka untuk proses penulisan buku ini. Buku berjudul "Machine Learning: Memahami Multinomial, Distribusi Probabilitas & Multinomial Naive Bayes" diantisipasi untuk berfungsi sebagai sumber pembelajaran yang efektif, memfasilitasi pemahaman yang lebih dalam tentang konsep dan implementasi pembelajaran mesin, terutama dalam tema multinomial, distribusi probabilitas, dan multinomial Naive Bayes.

Bidang machine learning sangat luas dan rumit, dan terlepas dari banyaknya sumber pembelajaran yang tersedia, konsep dan teknik tertentu tetap menantang untuk dipahami dan diterapkan. Salah satu konsep tersebut adalah multinomial, distribusi probabilitas, dan multinomial Naive Bayes. Tujuan dari buku ini adalah untuk memberikan penjelasan yang mendalam dan sistematis tentang konsep-konsep ini dan untuk menawarkan contoh-contoh praktis dan relevan dari implementasinya.

Buku ini akan memandu pembaca dalam memahami bagaimana distribusi probabilitas digunakan dalam pembelajaran mesin, apa itu Multinomial Naive Bayes, dan bagaimana hal itu dapat diterapkan. Selain itu, buku ini memberikan pemahaman dasar tentang model probabilitas dan bagaimana mereka dapat digunakan untuk membuat prediksi dalam berbagai situasi.

Kami mengakui bahwa buku ini bukannya tanpa kekurangan, dan selalu ada ruang untuk perbaikan. Kami menyambut kritik dan umpan balik yang membangun dari pembaca kami, karena kami sangat bercita-cita untuk meningkatkan kegunaan dan efektivitas buku ini. Meskipun demikian, kami optimis bahwa buku ini akan memberikan kontribusi positif bagi pembaca dan memperkaya pemahaman mereka tentang bidang pembelajaran mesin.

Terakhir, kami berharap buku ini akan menjadi panduan berharga bagi pembaca dalam memahami dan menguasai konsep multinomial, distribusi probabilitas, dan Multinomial Naive Bayes, dan kami berharap dapat membantu dalam mencapai aspirasi dan tujuan mereka di bidang pembelajaran mesin.

Salam hangat dan selamat belajar,

Penulis

# DAFTAR ISI

PRAKATA.....	i
DAFTAR TABEL.....	v
DAFTAR GAMBAR.....	vi
DAFTAR SCRIPT .....	vii
Pendahuluan .....	1
BAB I - Konsep Dasar Probabilitas .....	3
1.1 Probabilitas.....	4
1.2 Prinsip-prinsip dasar.....	5
1.3 Hukum Probabilitas.....	6
BAB II - Variabel Acak dan Distribusi Probabilitas.....	10
2.1 Variabel Acak .....	10
2.2 Distribusi probabilitas .....	12
2.2.1 Probability Mass Function .....	13
2.2.2 Cumulative Distribution Function .....	14
2.2.3 Distribusi Probabilitas Diskrit .....	15
2.2.4 Distribusi Probabilitas Kontinu .....	18
BAB III - Distribusi Multinomial .....	24
3.1 Probabilitas Dari Vektor Jumlah Kejadian .....	24
3.2 Distribusi Multinomial.....	25
BAB IV - Pengantar Machine Learning .....	28
4.1 Machine Learning.....	29
4.2 Penerapan Distribusi Probabilitas Pada Algoritma Machine Learning.....	31
BAB V - Naive Bayes Classifier.....	37
5.1 Cara Kerja Teorema Bayes Untuk Klasifikasi .....	38
5.2 Algoritma Naive Bayes dalam bahasa pemrograman PHP dan Python.....	46



BAB VI - Multinomial Naive Bayes.....	57
6.1 Pola dan Cara Kerja .....	59
6.2 Perbedaan, Kelemahan dan Cara Mengatasi .....	60
6.3 Evaluasi Model Klasifikasi .....	65
BAB VII – Implementasi Python untuk Multinomial Naive Bayes.....	67
7.1 Tujuan dan Gambaran Umum.....	68
7.1.1 Manfaat Implementasi bagi Pemahaman Konsep Klasifikasi Teks .....	68
7.1.2 Tools yang Digunakan .....	70
7.2 Persiapan Lingkungan dan Import Library.....	71
7.3 Dataset dan Preprocessing Teks .....	74
7.3.1 Apa Itu Dataset dalam Konteks NLP? .....	75
7.3.2 Apa Itu Preprocessing Teks?.....	75
7.3.3 Pembuatan Dataset Sederhana (Manual atau CSV).....	76
7.3.4 Penjelasan Fitur (X) dan Label (y).....	78
7.3.5 Tokenisasi dan Representasi Teks dengan Count Vectorizer.....	79
7.3.6 Split Data: Training vs Test .....	81
7.4 Pelatihan Model Multinomial Naive Bayes .....	82
7.4.1 Inisialisasi Model: MultinomialNB() .....	83
7.4.2 Melatih Model dengan Data Pelatihan .....	84
7.5 Pelatihan Model Multinomial Naive Bayes .....	85
7.6 Prediksi dan Evaluasi Model Multinomial Naive Bayes .....	86
7.7 Eksperimen dan Penyesuaian Model.....	88
7.8 Ikhtisar Materi .....	90
BAB VIII – Studi Kasus.....	92
8.1 Implementasi meng-klasifikasikan email .....	92
8.2 Sentiment Analysis pada Dataset Twitter.....	99
8.3 Naive Bayes vs Multinomial Naive Bayes dalam Analisis Sentimen.....	108
8.4 News Classification Using MultinomialNB .....	116
8.5 Studi Kasus: Klasifikasi Sentimen Ulasan Aplikasi Shopee .....	134

8.6 Klasifikasi Harapan Hidup Dunia.....	143
REFERENSI .....	157