

Dasar-Dasar Machine Learning pada Google Colabs



eureka
media aksara
Anggota IKAPI
No. 225/JTE/2021

0858 5343 1992
eurekamediaaksara@gmail.com
Jl. Banjaran RT.20 RW.10
Bojongsari - Purbalingga 53362

ISBN 978-623-120-063-9



9 786231 200839

DASAR-DASAR MACHINE LEARNING PADA GOOGLE COLABS

Dr. Thoyyibah T, S.Kom., M.Kom
Fajar Kurniawan, S.Kom
Dr. Taswanda Taryo, M.Sc



eureka
media aksara

PENERBIT CV.EUREKA MEDIA AKSARA

**DASAR-DASAR MACHINE LEARNING PADA GOOGLE
COLABS**

Penulis : Dr. Thoyyibah T, S.Kom., M.Kom
Fajar Kurniawan, S.Kom
Dr. Taswanda Taryo, M.Sc

Desain Sampul : Ardyan Arya Hayuwaskita

Tata Letak : Herlina Sukma

ISBN : 978-623-120-083-9

Diterbitkan oleh : **EUREKA MEDIA AKSARA, JANUARI 2024**
ANGGOTA IKAPI JAWA TENGAH
NO. 225/JTE/2021

Redaksi:

Jalan Banjaran, Desa Banjaran RT 20 RW 10 Kecamatan Bojongsari
Kabupaten Purbalingga Telp. 0858-5343-1992

Surel : eurekamediaaksara@gmail.com

Cetakan Pertama : 2024

All right reserved

Hak Cipta dilindungi undang-undang

Dilarang memperbanyak atau memindahkan sebagian atau seluruh isi buku ini dalam bentuk apapun dan dengan cara apapun, termasuk memfotokopi, merekam, atau dengan teknik perekaman lainnya tanpa seizin tertulis dari penerbit.

KATA PENGANTAR

Segala puji syukur kami panjatkan kepada Tuhan yang maha Esa, karena atas pertolongan dan limpahan rahmatnya sehingga penulis bisa menyelesaikan buku yang berjudul Peran “Dasar-Dasar Machine Learning pada Google Colabs”. Buku ini disusun secara lengkap dengan tujuan untuk memudahkan para pembaca memahami isi buku ini. Kami menyadari bahwa buku yang ada ditangan pembaca ini masih banyak kekurangan. Maka dari itu kami sangat mengharapkan saran untuk perbaikan buku ini dimasa yang akan datang. Tidak lupa kami mengucapkan terimakasih kepada semua pihak yang telah membantu dalam proses penerbitan buku ini. Semoga buku ini dapat membawa dampak positif bagi para pembaca.

DAFTAR ISI

KATA PENGANTAR.....	iii
DAFTAR ISI.....	iv
DAFTAR GAMBAR.....	vii
BAB 1 DATA ANALITIK DENGAN PYTHON	1
A. Tujuan Pembelajaran.....	1
B. Uraian Materi.....	1
1. Dasar - Dasar Data Analitik.....	1
2. Dasar - Dasar Statistik.....	8
C. Kesimpulan	12
D. Daftar Pustaka	12
BAB 2 EXPLORATORY DATA ANALYSIS (EDA)	14
A. Tujuan Pembelajaran.....	14
B. Uraian Materi.....	14
1. Pengertian Exploratory Data Analysis (EDA).....	14
2. Tujuan Exploratory Data Analysis (EDA)	15
3. Exploratory Data Analysis (EDA) Tools	20
C. Soal.....	23
D. Kesimpulan	27
E. Daftar Pustaka	28
BAB 3 ARTIFICIAL INTELLIGENCE	29
A. Tujuan Pembelajaran:.....	29
B. Uraian Materi.....	29
1. Pengertian Kecerdasan Buatan (Artificial Intelligence)	29
2. Konsep AI	31
3. Tujuan Kecerdasan Buatan	33
4. Ruang Lingkup	33
5. Tugas.....	34
C. Daftar Pustaka	34
BAB 4 MACHINE LEARNING.....	36
A. Tujuan Pembelajaran.....	36
B. Uraian Materi.....	36
1. Pengertian Machine Learning.....	36
2. Kelebihan Machine Learning.....	39
3. Kekurangan Machine Learning.....	40

C. Soal	43
D. Kesimpulan.....	46
E. Daftar Pustaka.....	47
BAB 5 SUPERVISED LEARNING	48
A. Tujuan Pembelajaran:	48
B. Uraian Materi	48
1. Definisi Supervised Learning	48
2. Konsep Supervised Learning	49
3. Tujuan Supervised Learning	51
4. Ruang Lingkup.....	53
5. Tugas	54
C. Daftar Pustaka.....	55
BAB 6 CLASSIFICATION-PROBABILISTIC MODEL (NAIVES BAYES).....	57
A. Tujuan Pembelajaran	57
B. Uraian Materi	57
1. Dasar-Dasar Algoritma Naives Bayes	57
C. Kesimpulan.....	64
D. Daftar Pustaka.....	64
BAB 7 CLASSIFICATION-VECTOR SPACE BASED (KNN AND SVM).....	65
A. Tujuan Pembelajaran	65
B. Uraian Materi	65
C. Soal	76
D. Kesimpulan.....	76
E. Daftar Pustaka.....	77
BAB 8 CLASSIFICATION- NEURAL NETWORK (JST).....	79
A. Tujuan Pembelajaran	79
B. Uraian Materi	79
C. Soal	81
D. Kesimpulan.....	81
D. Daftar Pustaka.....	82
BAB 9 UNSUPERVISED LEARNING.....	84
A. Tujuan Pembelajaran	84
B. Uraian Materi	84
C. Kesimpulan.....	93
D. Soal	94

E. Daftar Pustaka	94
BAB 10 COMMON TEXT MINNING PREPROCESSING.....	96
A. Tujuan Pembelajaran.....	96
B. Uraian Materi.....	96
C. Soal.....	99
D. Kesimpulan	100
E. Daftar Pustaka	101
BAB 11 DOCUMENT VECTOR AND BAG OF WORDS	102
A. Tujuan Pembelajaran:.....	102
B. Uraian Materi.....	102
1. Definisi Document Vector dan Bag of Word	102
2. Kelebihan dan Keterbatasan Document Vactor	107
3. Contoh Penerapan Document Vector	107
4. Kelebihan dan Keterbatasan Bag of Words	109
5. Contoh Penerapan Bag of Words	109
6. Kesimpulan.....	110
7. Tugas.....	111
C. Daftar Pustaka	111
BAB 12 IMAGE PROCESSING	113
A. Tujuan Pembelajaran.....	113
B. Uraian Materi.....	113
C. Soal.....	123
D. Kesimpulan	124
E. Daftar Pustaka	125
TENTANG PENULIS.....	126

DAFTAR GAMBAR

Gambar 9. 1 Melakukan pencarian aplikasi google colab	87
Gambar 9. 2 Cara membuka aplikasi google colab	88
Gambar 9. 3 Tampilan google colab	88
Gambar 9. 4 Menampilkan dataset.....	89
Gambar 9. 5 Mengelompokkan observasi menggunakan kolom numeric dalam data	89
Gambar 9. 6 Memeriksa pusat cluster untuk setiap cluster	89
Gambar 9. 7 Mengklasifikasikan titik dengan model	90
Gambar 9. 8 Mengklasifikasikan model dengan K-Means.....	90
Gambar 9. 9 Model Inersia	90
Gambar 9. 10 Model Inersia 6 cluster	91
Gambar 9. 11 Tampilan Plot Cluster dan Inersia	91
Gambar 9. 12 Bagan plot.....	92
Gambar 12. 1 Grafik Histogram.....	118
Gambar 12. 2 Grafik Histogram Grayscale	118
Gambar 12. 3 Citra Ukuran 8x8	119
Gambar 12. 4 Instalasi Library Python.....	120
Gambar 12. 5 Function Menampilkan Gambar	120
Gambar 12. 6 Function Menampilkan Histogram.....	121
Gambar 12. 7 Function Upload Gambar ke Drive.....	121
Gambar 12. 8 Menampilkan Gambar Asli	122
Gambar 12. 9 Menampilkan Histogram.....	122
Gambar 12. 10 Menampilkan Gambar Skala Abu-abu	123
Gambar 12. 11 Menampilkan Histogram Abu-abu.....	123



DASAR-DASAR MACHINE LEARNING PADA GOOGLE COLABS

Dr. Thoyyibah T, S.Kom., M.Kom

Fajar Kurniawan, S.Kom

Dr. Taswanda Taryo, M.Sc



BAB

1

DATA ANALITIK DENGAN PYTHON

A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar (*basic science*) tentang definisi Topologi Jaringan. Anda harus mampu :

1. Mengetahui dasar - dasar data analitik
2. Mengetahui dasar - dasar statistik

B. Uraian Materi

<p>Tujuan Pembelajaran 1 Pengertian Data Analitik</p>
--

1. Dasar - Dasar Data Analitik

a. Pengertian data analitik

Data analitik adalah suatu tindakan menganalisis data mentah untuk tujuan tertentu. Data analitik saat ini termasuk salah satu keahlian yang banyak dicari organisasi, perusahaan maupun lembaga pemerintahan karena hasilnya bisa dijadikan bahan pertimbangan dalam pengambilan keputusan. Data analitik mengacu pada cara analisis data atau informasi yang telah dikumpulkan begitu banyak dalam satu server (*big data*) yang kemudian hasilnya dapat digunakan untuk mengambil berbagai keputusan strategis.

b) Data Sekunder

Data sekunder adalah data yang didapat tidak secara langsung dari objek penelitian. Peneliti mendapatkan data yang sudah jadi yang dikumpulkan oleh pihak lain dengan berbagai cara atau metode baik secara komersial maupun non komersial. Contohnya adalah pada peneliti yang menggunakan data statistik hasil riset dari surat kabar atau majalah. (Nuryadi, 2017)

C. Kesimpulan

Data analitik adalah suatu tindakan menganalisis data mentah untuk tujuan tertentu. Data analitik saat ini termasuk salah satu keahlian yang banyak dicari organisasi, perusahaan maupun lembaga pemerintahan karena hasilnya bisa dijadikan bahan pertimbangan dalam pengambilan keputusan. Data analitik mengacu pada cara analisis data atau informasi yang telah dikumpulkan begitu banyak dalam satu server (big data) yang kemudian hasilnya dapat digunakan untuk mengambil berbagai keputusan strategis.

Statistik adalah kumpulan fakta yang berbentuk angka-angka yang disusun dalam bentuk daftar atau tabel yang menggambarkan suatu persoalan. Perbedaan dari statistik dan parameter adalah statistik merupakan sembarangan nilai yang menjelaskan nilai dari sampel. Sedangkan parameter merupakan sembarangan nilai yang menjelaskan nilai dari populasi

D. Daftar Pustaka

- Lo, R. (2023). Penggunaan Bahasa Pemrograman Python dalam Menganalisis Hubungan Kualitas Kopi dengan Lokasi Pertanian Kopi. *Jurnal Publikasi Teknik Informatika (JUPTI)*, 101-102.
- Nuryadi, S. M. (2017). *DASAR - DASAR STATISTIK PENELITIAN*. Yogyakarta: Gramasurya.

S.K.W, L. V. (2020). TRANSFORMASI PERAN AUDITOR
DALAM PROSES PENYUSUNAN ANGGARAN
KEGIATAN ORGANISASI KEMAHASISWAAN.
Prosiding ASIC 2020, 50.

BAB 2

EXPLORATORY DATA ANALYSIS (EDA)

A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar (*basic science*) tentang definisi Exploratory Data Analysis. Anda harus mampu :

1. Mengetahui pengertian Exploratory Data Analysis (EDA)
2. Mengetahui jenis-jenis Exploratory Data Analysis (EDA)
3. Mengetahui cara kerja Exploratory Data Analysis (EDA)

B. Uraian Materi

Tujuan Pembelajaran 1

Mengetahui pengertian Exploratory Data Analysis (EDA)

1. Pengertian Exploratory Data Analysis (EDA)

Secara umum, apa itu Exploratory Data Analysis mencakup tentang proses kritis uji investigasi awal pada sebuah data untuk mengidentifikasi pola, menemukan anomali, menguji hipotesis, dan memeriksa asumsi melalui statistik ringkasan dan representasi grafis (visual).

Data specialist menggunakan EDA untuk membedakan kumpulan data apa yang dapat diungkapkan lebih jauh di luar pemodelan data formal atau tugas pengujian hipotesis. Hal ini memungkinkan mereka untuk mendapatkan temuan mendalam tentang variabel dalam kumpulan data dan hubungannya.

E. Daftar Pustaka

Vol. 4 No. 2 (2021): Jurnal Sistem Informasi dan Ilmu Komputer
Prima (JUSIKOM Prima)

<https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v4i2.2475>

DOI: <https://doi.org/10.32520/stmsi.v11i3.1857>

<https://algorit.ma/blog/exploratory-data-analysis-2022/>

BAB 3

ARTIFICIAL INTELLIGENCE

A. Tujuan Pembelajaran:

Setelah mempelajari Bab ini, diharapkan pembaca dapat:

1. Memahami, Mengetahui dan menjelaskan Definisi, Fungsi, Sistematika kecerdasan buatan.
2. Memahami dan menjelaskan aplikasi kecerdasan buatan.
3. Memahami berbagai jenis kecerdasan buatan.

B. Uraian Materi

1. Pengertian Kecerdasan Buatan (Artificial Intelligence)

Kecerdasan buatan merupakan hasil penyatuan bidang elektronika, matematika, dan ilmu komputer yang berupa sebuah sistem yang mampu bertindak layaknya manusia seperti menentukan keputusan, memperkirakan keadaan di masa depan, mengklasifikasikan suatu keadaan, dan berpikir (Prasetyo & Rosyid, 2022).

Kecerdasan Buatan (*Artificial Intelligence - AI*) merujuk pada kemampuan computer atau sistem computer untuk meniru dan mengeksekusi tugas yang biasanya membutuhkan kecerdasan manusia (Pongtambing et al., 2023). Tujuan utama dari kecerdasan buatan adalah memungkinkan computer melakukan tugas yang biasanya membutuhkan pemikiran, analisis, pengambilan keputusan, dan pemecahan masalah manusia (Manongga, 2022).

Menurut para ahli, perkembangan Kecerdasan Buatan akan terus berjalan dengan cepat dan berdampak besar pada

- b. Pada bidang industri: Penggunaan otomasi mesin pada pabrik – pabrik, pengaturan kontrol pada proses produksi dll.
 - c. Pada bidang pertahanan keamanan: Drone pengintai, rudal dengan kendali jarak jauh, penangkal rudal kendali jarak jauh dll.
- Dan pada bidang – bidang lainnya.

5. Tugas

- a. Jelaskan arti AI menurut para ahli.
- b. Jelaskan perbedaan antara system AI dan sistem komputer konvensional.
- c. Berikan contoh aplikasi AI pada dunia saat ini.

C. Daftar Pustaka

- Anggraini, Y. (2023, April 7). Geotimes. *Resiko dan Tantangan Menghadapi Kemajuan Kecerdasan Buatan*.
- Manongga, D. e. (2022). *Dampak Kecerdasan Buatan Bagi Pendidikan*.
- Pabubung, M. R. (2021). *Epistemologi Kecerdasan Buatan (AI) dan Pentingnya Ilmu Etika dalam Pendidikan Interdisipliner*.
- Pongtambing et al., Y. S. (2023). *Peluang dan Tantangan Kecerdasan Buatan Bagi Generasi Muda*.
- Prasetyo, A. A., & Rosyid, H. A. (2022). *Tinjauan Pengaruh Kecerdasan Buatan terhadap Pembelajaran untuk Menghadapi Masa ociety 5.0*.
- Purnamasari, A., & Wijoyo, H. (2023). *Analisis Efektivitas Pembelajaran Bahasa Indonesia Di PERguruan Tinggi Era 5.0*.
- Rahardja, U. (2022). *Masalh Etis dalam Penerapan Sistem Kecerdasan Buatan*.
- Syafi'e, Indrajit, R. E., & Dazki, E. (2022). *Kesiapan Teknologi Informasi Perbankan Hadapi Revolusi Industri Era 4.0*.

Wijoyo, P. &. (2023). *Analisis Efektivitas Pembelajaran Bahasa Indonesia DI Perguruan Tinggi Era 5.0.*

BAB

4

MACHINE LEARNING

A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar (basic science) tentang definisi Machine Learning. Anda harus mampu:

1. Mengetahui pengertian Machine Learning
2. Mengetahui jenis-jenis Machine Learning
3. Mengetahui cara kerja Machine Learning

B. Uraian Materi

Tujuan Pembelajaran 1

Pengertian Machine Learning

1. Pengertian Machine Learning

Machine learning adalah bidang ilmu komputer yang mempelajari bagaimana membuat komputer dapat belajar dari data tanpa diprogram secara eksplisit.

Dalam arti lain, machine learning adalah kemampuan komputer untuk belajar dari data tanpa harus diberi tahu secara eksplisit apa yang harus dilakukan. Komputer dapat belajar dari data untuk membuat prediksi, klasifikasi, atau keputusan.

a. Bagaimana mesin hitung (komputer) dapat belajar?

Mesin hitung (komputer) dapat belajar dengan cara mempelajari data. Data yang diberikan kepada komputer

1. Machine learning adalah bidang ilmu yang mempelajari bagaimana komputer dapat belajar dari data tanpa diprogram secara eksplisit.
2. Proses machine learning terdiri dari dua tahap utama, yaitu fase pelatihan dan fase aplikasi.
3. Machine learning memiliki berbagai macam aplikasi dalam dunia nyata.
4. Machine learning memiliki keterkaitan dengan berbagai bidang lain.

E. Daftar Pustaka

Ahmad, Abu. 2017. "Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning." *Jurnal Teknologi Indonesia* 1-5.

Kusuma, Purba Daru. 2020. *Machine Learning Teori, Program, Dan Studi Kasus*. Bandung: Deepublish.

BAB 5

SUPERVISED LEARNING

A. Tujuan Pembelajaran:

Setelah mempelajari Bab ini, diharapkan pembaca dapat:

1. Memahami, mengetahui dan menjelaskan definisi, konsep supervised learning.
2. Memahami dan menjelaskan tujuan aplikasi supervised learning.
3. Memahami berbagai ruang lingkup supervised learning.

B. Uraian Materi

1. Definisi Supervised Learning

Berikut adalah beberapa definisi supervised learning menurut para ahli:

a. Tom Mitchell

"Sebuah program dikatakan belajar dari pengalaman E terhadap suatu tugas T , dengan ukuran kinerja P , jika kinerjanya dalam tugas T , seperti diukur oleh P , meningkat seiring dengan pengalaman E ." - (Mitchell, 1997)

b. Christopher Bishop

"Pembelajaran supervisi adalah proses di mana model dipelajari dari sekumpulan contoh pelatihan yang berisi pasangan input dan output yang diinginkan, dan tujuan dari model adalah untuk memetakan input ke output." - (Bishop, 2006)

C. Daftar Pustaka

- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, 18(5-6), 602-610.
- Jurafsky, D., & Martin, J. H. (2019). *Speech and Language Processing* (3rd ed.). Prentice Hall.
- Koller, D., & Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques*. MIT Press.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems* 25.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
- Moody, J., & Saffell, M. (2001). Learning to trade via direct reinforcement. *IEEE Transactions on Neural Networks*, 12(4), 875-889.
- Ng, A. (2018). *Machine Learning Yearning*.
- Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, 2(1-2), 1-135.
- Rajkumar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *The New England Journal of Medicine*, 380(14), 1347-1358.

- Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to Recommender Systems Handbook. In *Recommender Systems Handbook* (pp. 1-35). Springer.
- Sahami, M., Dumais, S., Heckerman, D., & Horvitz, E. (1998). A Bayesian approach to filtering junk e-mail. In *Learning for Text Categorization: Papers from the 1998 Workshop*.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT Press.
- Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 71-86.

BAB 6

CLASSIFICATION- PROBABILISTIC MODEL (NAIVES BAYES)

A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan bagaimana kinerja algoritma naives bayes Anda harus mampu :

1. Mengetahui dasar – dasar algoritma naïve bayes
2. Mengetahui dasar – dasar klasifikasi pada algoritma naïve bayes

B. Uraian Materi

Tujuan Pembelajaran 1

Pengertian Algoritma Naïve Bayes

1. Dasar-Dasar Algoritma Naives Bayes

a. Pengertian Algoritma Naives bayes

Algoritma Naive Bayes Classifier merupakan salah satu algoritma yang terdapat

pada teknik klasifikasi. Naïve Bayes Classifier (NBC) merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Selain itu, NBC memiliki komputasi yang mudah, serta memiliki tingkat akurasi tinggi dan error rate yang

C. Kesimpulan

Berdasarkan uraian di atas, dapat ditarik kesimpulan bahwa Naive Bayes adalah metode klasifikasi data berdasarkan faktor-faktor probabilitas. Naïve Bayes Classifier memiliki beberapa kelebihan yaitu, cepat dalam proses perhitungan, algoritma yang sederhana dan akurasi yang tinggi. Namun probabilitas pada Naïve Bayes Classifier tidak bisa mengukur seberapa besar tingkat keakuratan sebuah prediksi, selain itu metode naïve bayes juga memiliki kelemahan pada seleksi atribut yang banyak, yang mengakibatkan memakan banyak waktu.

D. Daftar Pustaka

Ramadhani, Nichyta Dian (2019) : Analisis Kinerja Algoritma Naïve Bayes Classifier Dengan Seleksi Fitur Chi-Square Dan Confix Stripping Stemmer Pada Sistem Klasifikasi Terjemahan Hadits

BAB

7

CLASSIFICATION- VECTOR SPACE BASED (KNN AND SVM)

A. Tujuan Pembelajaran

Tujuan pembelajaran *Classification-Vector Space Based* (KNN and SVM) melibatkan pemahaman dan penerapan konsep serta teknik terkait dengan metode klasifikasi berbasis ruang vektor menggunakan algoritma *K-Nearest Neighbors* (KNN) dan Support Vector Machine (SVM). Anda diharapkan mampu :

1. Memahami konsep dasar klasifikasi dalam ruang vektor.
2. Memahami prinsip kerja algoritma *K-Nearest Neighbors* (KNN) dan Support Vector Machine (SVM) dalam konteks klasifikasi.
3. Mengimplementasikan algoritma KNN dan SVM menggunakan Google Colab.
4. Mampu mengevaluasi performa model klasifikasi menggunakan metrik yang relevan.
5. Mampu memilih model terbaik untuk kasus-kasus tertentu.

B. Uraian Materi

Tujuan Pembelajaran 1

- 1.1 Memahami konsep dasar klasifikasi dalam ruang vektor

Klasifikasi adalah salah satu tugas dalam machine learning yang bertujuan untuk memprediksi kategori atau label dari suatu data berdasarkan informasi yang diberikan pada data yang telah dilatih sebelumnya. Klasifikasi dalam data mining merupakan metode yang umum, suatu objek yang sebelumnya belum diketahui kelas/labelnya dapat melalui metode

- c. Parameter seperti jenis kernel, C , dan γ dapat disesuaikan untuk meningkatkan performa model SVM.
3. Metrik Evaluasi Klasifikasi:
 - a. Akurasi, presisi, recall, dan F1-score adalah metrik evaluasi umum yang digunakan untuk mengukur performa model klasifikasi.
 - b. Pemahaman yang baik terhadap karakteristik dataset dan tujuan bisnis diperlukan untuk memilih metrik yang sesuai.
 - c. Confusion matrix memberikan gambaran yang lebih rinci tentang kinerja model.
 4. Analisis Hasil dan Pemilihan Model Terbaik:
 - a. Analisis hasil evaluasi sangat penting untuk memilih model terbaik.
 - b. Pemilihan model tergantung pada keseimbangan antara precision, recall, akurasi, dan karakteristik khusus dari masalah klasifikasi yang dihadapi.
 - c. ROC curve dan AUC juga dapat memberikan wawasan tambahan untuk model klasifikasi biner

Dengan memahami konsep dasar, implementasi praktis, dan evaluasi performa kedua algoritma klasifikasi ini, diharapkan pembaca dapat lebih percaya diri dalam menggunakan KNN dan SVM untuk menyelesaikan berbagai masalah klasifikasi dalam ruang vector

E. Daftar Pustaka

- Argina, A. M. (2020). Penerapan Metode Klasifikasi K-Nearest Neighbor pada Dataset Penderita Penyakit Diabetes. *Indonesian Journal of Data and Science*, 29-33.
- Leidiyana, A. A. (2020). Pemodelan Klasifikasi Dalam. *J. Tek. Inform. Univ. Muhammadiyah Tangerang vol. ISSN : 254*, 7-14.
- Puspitawati, T. E., Basuki, S., Rahmayanti, V., & Nastiti, S. (2021). Algoritma Support Vector Machine (SVM) Untuk

Identifikasi Komponen Abstrak Pada Jurnal Ilmiah
Berbasis Teknik Klasifikasi. 441–448.

BAB

8

CLASSIFICATION- NEURAL NETWORK (JST)

A. Tujuan Pembelajaran

Setelah mempelajari bab ini, Anda diharapkan mampu:

1. Memahami konsep dasar klasifikasi dengan jaringan saraf tiruan
2. Mampu menerapkan jaringan saraf tiruan untuk klasifikasi

B. Uraian Materi

Tujuan Pembelajaran 1

Memahami konsep dasar klasifikasi dengan jaringan saraf tiruan

Klasifikasi adalah salah satu tugas pembelajaran mesin yang paling umum. Tugas ini bertujuan untuk mengklasifikasikan objek ke dalam kelas yang berbeda. Jaringan saraf tiruan (JST) adalah salah satu model pembelajaran mesin yang paling efektif untuk tugas klasifikasi.

Jaringan Saraf Tiruan untuk Klasifikasi

JST untuk klasifikasi terdiri dari tiga lapisan utama, yaitu:

1. **Lapisan input** menerima data masukan.
2. **Lapisan tersembunyi** memproses data masukan dan menghasilkan output.
3. **Lapisan output** menghasilkan nilai kelas.

3. Output dari lapisan tersembunyi diberikan ke lapisan output.
4. Lapisan output menghasilkan nilai kelas.

JST dapat digunakan untuk mengklasifikasikan objek ke dalam berbagai kelas. Beberapa contoh aplikasi JST dalam klasifikasi adalah:

1. **Klasifikasi gambar**, seperti pengenalan wajah, pengenalan objek, dan klasifikasi citra medis.
2. **Klasifikasi teks**, seperti klasifikasi spam, klasifikasi sentimen, dan klasifikasi bahasa.
3. **Klasifikasi audio**, seperti pengenalan suara, klasifikasi suara, dan klasifikasi musik.
4. **Klasifikasi video**, seperti pengenalan gerakan, klasifikasi perilaku, dan klasifikasi objek bergerak.

JST adalah model pembelajaran mesin yang kuat dan serbaguna yang dapat digunakan untuk berbagai tugas klasifikasi.

D. Daftar Pustaka

- Alkhairi, Putrama, Irfan Sudahri Damanik, dan Agus Perdana Windarto. 2019. "Penerapan Jaringan Saraf Tiruan untuk Mengukur Korelasi Beban Kerja Dosen Terhadap Peningkatan Jumlah Publikasi." *Prosiding SENARIS* 581-601.
- Andriani, Yuli, Anjar Wanto, and Handrizal. 2019. "Jaringan Saraf Tiruan dalam Memprediksi Produksi Kelapa Sawit di PT. KRE Menggunakan Algoritma Levenberg Marquardt." *Prosiding SENARIS* 249-259.
- Muflih, Ghufron Zaida. 2021. "Determination of Backpropagation Neural Network Learning Parameters and Their Effect on the Training Process." *JURISTIK (Jurnal Riset Teknologi Informasi Dan Komputer)* 12-17.
- Muflih, Ghufron Zaida, Sunardi, dan Anton Yudhana. 2019. "Jaringan Saraf Tiruan Backpropagation untuk Prediksi Curah Hujan di Wilayah Kabupaten Wonosobo." *MUST:*

Journal of Mathematics Education, Science and Technology 45-56.

- Putra, Randi Rian. 2019. "IMPLEMENTASI METODE BACKPROPAGATION JARINGAN SARAF TIRUAN DALAM MEMPREDIKSI POLA PENGUNJUNG TERHADAP TRANSAKSI." *Jurnal Teknologi Informasi* 16-20.
- Ramadhan, Aulia Ichwanda, Jaya Tata Hardinata, dan Yuegilion Pranavarna Purba. 2021. "Analisa Jaringan Saraf Tiruan Backpropagation Untuk Memprediksi Prestasi Siswa SMA Muhammadiyah Serbelawan." *BRAHMANA: Jurnal Penerapan Kecerdasan Buatan* 18-26.
- Ridho, Ihda Innar, Anak Agung Gede Bagus Ariana, dan Agus Perdana Windarto. 2023. "Optimasi Fungsi Pembelajaran Jaringan Saraf Tiruan dalam Meningkatkan Akurasi pada Prediksi Ekspor Kopi Menurut Negara Tujuan Utama." *Building of Informatics, Technology and Science (BITS)* 1951-1958.
- Sunardi, Anton Yudhana, dan Ghufron Zaida Muflih. 2020. "Sistem Prediksi Curah Hujan Bulanan Menggunakan Jaringan Saraf Tiruan Backpropagation." *Jurnal Sistem Informasi Bisnis* 155-162.
- Thoriq, Muhammad. 2022. "Peramalan Jumlah Permintaan Produksi Menggunakan Jaringan Saraf Tiruan Algoritma Backpropagation." *Jurnal Informasi Dan Teknologi* 27-32.
- Wanto, Anjar. 2019. "PREDIKSI PRODUKTIVITAS JAGUNG DI INDONESIA SEBAGAI UPAYA ANTISIPASI IMPOR MENGGUNAKAN JARINGAN SARAF TIRUAN BACKPROPAGATION." *SINTECH Journal* 53-62.

BAB 9

UNSUPERVISED LEARNING

A. Tujuan Pembelajaran

1. Memahami konsep dasar dari Unsupervised Learning.
2. Mengetahui jenis-jenis algoritma Unsupervised Learning.
3. Mampu mengimplementasikan Unsupervised Learning menggunakan Google Colab.
4. Memahami aplikasi praktis dari Unsupervised Learning dalam analisis data.

B. Uraian Materi

Tujuan Pembelajaran 1

1.1 Memahami konsep dasar dari Unsupervised Learning

1. Apa itu *Unsupervised Learning*?

Unsupervised Learning atau pembelajaran tak terarah adalah teknik pembelajaran mesin di mana pengguna tidak perlu mengawasi modelnya. Sebaliknya, pembelajaran tak terarah memungkinkan model bekerja sendiri untuk menemukan pola dan informasi yang sebelumnya tidak terdeteksi, terutama yang berkaitan dengan data yang tidak berlabel (greatnusa.com, 2022).

Algoritma *unsupervised learning* memungkinkan pengguna untuk melakukan tugas pemrosesan yang lebih kompleks dibandingkan dengan *supervised learning*. Namun, pembelajaran tak terarah bias lebih tidak terduga dibandingkan dengan metode pembelajaran alami lainnya.

menerapkannya dalam kasus nyata untuk memperkaya pemahaman mereka dalam bidang pembelajaran mesin.

D. Soal

- 1 Jelaskan perbedaan antara Supervised dan Unsupervised Learning.
- 2 Bagaimana algoritma K-Means bekerja dalam Unsupervised Learning?
- 3 Apa fungsi utama dari Principal Component Analysis (PCA)?
- 4 Buatlah sebuah notebook Google Colab untuk melakukan clustering pada dataset tertentu.

E. Daftar Pustaka

- dqlab.id. (2021, Februari 14). <https://dqlab.id/yuk-kenali-jenis-jenis-algoritma-unsupervised-learning>. Retrieved from <https://dqlab.id>: <https://dqlab.id/yuk-kenali-jenis-jenis-algoritma-unsupervised-learning>
- Greatnusa . (2022, februari 24). *unsupervised-learning-adalah*. Retrieved from greatnusa.com: <https://greatnusa.com>
- greatnusa.com. (2022, februari 24). <https://greatnusa.com/artikel/unsupervised-learning-adalah/>. Retrieved from greatnusa.com: <https://greatnusa.com>
- Hakim, R. F. (2022, November 19). <https://medium.com/@986110101/kerja-bareng-colab-2dab324011f0>. Retrieved from <https://medium.com>: <https://medium.com>
- Jovian. (2023, Desember 09). <https://jovian.com/myh0307/sklearn-unsupervised-learning>. Retrieved from <https://jovian.com>: <https://jovian.com>
- Panatagama, A. (2023, Juni 25). <https://terralogiq.com/unsupervised-machine-learning/>. Retrieved from <https://terralogiq.com>: <https://terralogiq.com>

revou.com. (2023, Desember 09).
https://revou.co/kosakata/unsupervised-learning. Retrieved
from [https://revou.co:](https://revou.co:https://revou.co/kosakata/unsupervised-learning)
<https://revou.co/kosakata/unsupervised-learning>

BAB 10

COMMON TEXT MINING PREPROCESSING

A. Tujuan Pembelajaran

Setelah mempelajari bab ini, Anda diharapkan mampu:

1. Menjelaskan pentingnya preprocessing data teks untuk text mining
2. Menerapkan berbagai teknik preprocessing data teks, seperti case folding, stemming, lemmatization, stop words removal, dan normalization

B. Uraian Materi

Tujuan Pembelajaran 1

Menjelaskan pentingnya preprocessing data teks untuk text mining

Data teks adalah data yang terdiri dari teks. Data teks dapat berasal dari berbagai sumber, seperti dokumen, email, situs web, dan media sosial. Text mining adalah proses mengekstrak informasi dari data teks. Sebelum data teks dapat dianalisis dengan teknik text mining, data tersebut harus dipreprocess terlebih dahulu. Preprocessing data teks adalah proses mempersiapkan data teks agar dapat dianalisis dengan lebih efektif.

Tujuan dari preprocessing data teks adalah:

1. Meningkatkan akurasi dan efisiensi proses text mining
2. Menghilangkan noise dari data teks
3. Membuat data teks lebih mudah dianalisis

dipertimbangkan sebelum melakukan preprocessing data teks.

Meskipun terdapat tantangan dan keterbatasan, preprocessing data teks tetap merupakan proses penting yang dapat memberikan manfaat yang signifikan bagi proses text mining.

E. Daftar Pustaka

- Budiman, Ariel Elbert, dan Andreas Widjaja. 2020. "Analisis Pengaruh Teks Preprocessing Terhadap Deteksi Plagiarisme Pada Dokumen Tugas Akhir." *JuTISI* 475-488.
- Harjanta, Aris Tri Jaka. 2015. "Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining." *Jurnal Informatika UPGRIS* 1-9.
- Khomsah, Siti, dan Agus Sasmito Aribowo. 2020. "Model Text-Preprocessing Komentar Youtube Dalam Bahasa Indonesia." *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)* 648 - 654.
- Rahayu, Kharisma, Vindi Fitria, Dhini Septhya, Rahmaddeni, dan Lusiana Efrizoni. 2023. "Klasifikasi Teks untuk Mendeteksi Depresi dan Kecemasan pada Pengguna Twitter Berbasis Machine Learning." *MALCOM: Indonesian Journal of Machine Learning and Computer Science* 108-114.
- Ridwansyah, Tengku. 2022. "Implementasi Text Mining Terhadap Analisis Sentimen Masyarakat Dunia Di Twitter Terhadap Kota Medan Menggunakan K-Fold Cross Validation Dan Naïve Bayes Classifier." *KLIK: Kajian Ilmiah Informatika dan Komputer* 178-185.
- Runimeirati, Abdul Muis, dan Figur Muhammad. 2023. "Pelatihan Text Mining Menggunakan Bahasa Pemrograman Python." *Abdimas Langkanae* 36-46.

BAB 11

DOCUMENT VECTOR AND BAG OF WORDS

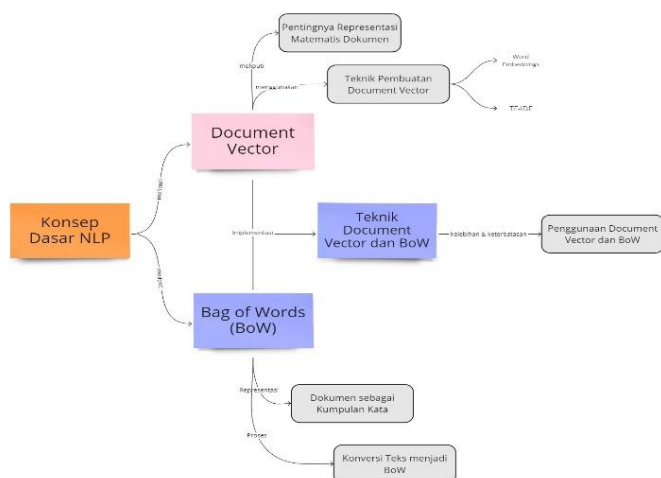
A. Tujuan Pembelajaran:

Setelah mempelajari Bab ini, diharapkan pembaca dapat:

1. Memahami, mengetahui dan menjelaskan definisi, document vector dan bag of words.
2. Memahami kelebihan dan keterbatasan document vector dan bag of words.
3. Memahami contoh penerapan document vector dan bag of words.

B. Uraian Materi

1. Definisi Document Vector dan Bag of Word



miro

efektif, terutama untuk dokumen pendek, meskipun mengabaikan urutan kata.

7. Tugas

- a. Buatlah analisis document vector dari sebuah buku novel online!
- b. Buatlah analisis kata yang sedang viral di social media dengan Teknik Bag of Word!
- c. Buatlah representasi Word Cloud dari soal nomor 1 dan 2!

C. Daftar Pustaka

- Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison-Wesley.
- Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison-Wesley.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, 41(6), 391–407.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding.
- Harris, Z. (1954). Distributional Structure. *Word*, 10(2–3), 146–162.
- Joulin, A., Grave, E., Bojanowski, P., Mikolov, T., & Mikolov, T. (2017). Bag of Tricks for Efficient Text Classification. arXiv preprint arXiv:1607.01759.
- Jurafsky, D., & Martin, J. H. (2019). *Speech and Language Processing* (3rd ed.).
- Jurafsky, D., & Martin, J. H. (2019). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Pearson.

- Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. In International conference on machine learning (pp. 1188-1196).
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. In Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2 (NIPS'13), 3111–3119.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In Advances in neural information processing systems (pp. 3111-3119).
- Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543).
- Salton, G. (1989). A Vector Space Model for Automatic Indexing. *Journal of the American Society for Information Science*, 40(9), 540–557.
- Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513-523.

BAB

12

IMAGE PROCESSING

A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan mengenai pengetahuan dasar tentang hal-hal yang berkaitan dengan *image processing*. Sehingga kita akan mampu:

1. Mengetahui Pengertian *Image Processing*, Elemen Sistem Pemrosesan Citra Digital dan Tujuan Pengolahan Citra
2. Mengetahui Praktek Pemrosesan Citra Digital Dengan Tool Histogram

B. Uraian Materi

Tujuan Pembelajaran 1

Mengetahui Pengertian Image Processing, Elemen Dasar Citra Digital dan Tujuan Pengolahan Citra

1. Pengertian *Image Processing*

Image processing adalah metode untuk memproses atau memanipulasi gambar dua dimensi. Hal ini mencakup berbagai operasi untuk memperbaiki, menganalisis, atau mengubah suatu gambar (Gonzalez, 2018).

Konsep dasar pemrosesan objek pada gambar berasal dari kemampuan indera penglihatan manusia, yang kemudian dihubungkan dengan kemampuan otak manusia. Sejarah pengolahan citra mencatat aplikasinya dalam berbagai bentuk dengan tingkat kesuksesan yang cukup

equalization dan matching, serta normalisasi histogram, dijelaskan sebagai teknik efektif untuk peningkatan kualitas gambar. Pembahasan ditutup dengan langkah-langkah membuat program histogram citra menggunakan Google Colab.

E. Daftar Pustaka

- A. Arora, "Fundamental Steps of Digital Image Processing." Medium, Aug 7, 2019. [Online]. Available: <https://medium.com/futframe-ai/fundamental-steps-of-digital-image-processing-d7518d6bb23c>.
- R. C. Gonzalez and R. E. Woods, "Digital Image Processing" (3rd ed.). Prentice Hall, 2018.
- R. Munir, "Image Histogram." Institut Teknologi Bandung, 2019. [Online]. Available: <https://informatika.stei.itb.ac.id/~rinaldi.munir/Citra/2019-2020/07-Image-Histogram.pdf>.
- S. Ramadani, "Pengembangan Aplikasi Pengolahan Citra untuk Pengelompokan Buah Semangka." Universitas Islam Negeri Sultan Syarif Kasim Riau, 2014.
- W. K. Pratt, "Digital Image Processing: PIKS Inside." John Wiley & Sons, 2007.

TENTANG PENULIS

Dr. Thoyyibah T, S.Kom., M.Kom

Thoyyibah. T. S. Kom. M. Kom. Lulus S1 di Program Studi Teknik Informatika Fakultas Sains dan Teknologi tahun 2011. Lulus S2 di IPB tahun 2014. Lulus S3 di BINUS tahun 2023. Saat ini adalah dosen tetap Universitas Pamulang. Mengampu mata kuliah Komunikasi Data, Jaringan, Automata, Kecerdasan Buatan, Logika Informatika, Machine Learning dll. Aktif menulis artikel di berbagai jurnal ilmiah. Beberapa kali menjadi pemakalah seminar prosiding nasional dan Internasional

Fajar Kurniawan, S.Kom

Lulusan Fakultas Teknik Komputer STMIK Eresha Jakarta (2019) dan saat ini sedang mengejar gelar Magister Komputer di Universitas Pamulang.

Pendidikan:

S1 Teknik Komputer, STMIK Eresha Jakarta (2019)

Sedang menempuh Magister Komputer, Universitas Pamulang

Dr. Taswanda Taryo, M.Sc

Taswanda Taryo. Ybs bekerja di BRIN (dulu BATAN) sejak 1981 dan lulus S1 Fisika Institut Teknologi bandung tahun 1981 dan kemudian melanjutkan sekolah di University of New Brunswick (UNB) dengan gelar Master od Science in Engineering tahun 1991. Akhirnya melanjutkan program S3 di UGM dengan spesialisasi computer engineering pada tahun 2003. Sejak bekerja di BATAN, pernah menjadi salah satu pimpinan BATAN tahun 2008-2016. Ybs bergabung dengan Program Pasca Sarjana UNPAM tahun 2010 dan sekarang sebagai dosen senior Magister Teknik Informatika UNPAM.