

PROPOSAL PROYEK DESAIN INOVASI DATA SCIENCE

Analisis Pengaruh Pola Hidup terhadap Pencapaian Indeks Prestasi Mahasiswa Menggunakan Algoritma Naive Bayes



Kelompok : 28
Anggota Kelompok:

- 1. Muhammad Rakha Ervinpermana Putra - 255150200111069**
- 2. Fahrizal Dwi Candra - 255150207111073**
- 3. Michael Juanio Patrick - 255150201111037**

DEPARTEMEN TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BRAWIJAYA
2025

DAFTAR ISI

DAFTAR ISI.....	2
ABSTRAK.....	3
BAB I	
PENDAHULUAN.....	4
1.1 Latar Belakang.....	4
1.2 Rumusan Masalah.....	4
1.3 Tujuan.....	4
1.4 Manfaat.....	5
BAB II	
TINJAUAN PUSTAKA.....	5
2.1 Pola Hidup Mahasiswa.....	5
2.2 Indeks Prestasi (IP).....	6
2.3 Hubungan Pola Hidup dan Prestasi Akademik.....	6
2.4 Teknologi dan Metode yang Digunakan.....	6
BAB III	
METODOLOGI DAN SOLUSI.....	7
3.1 Metodologi Perancangan.....	7
3.2 Solusi.....	8
BAB IV	
HIPOTESIS HASIL.....	9
4.1 Prediksi Keluaran Utama.....	9
4.2 Pencapaian Tujuan.....	9
4.3 Kesesuaian dengan Kajian Pustaka.....	9
DAFTAR PUSTAKA.....	9
LAMPIRAN.....	11

ABSTRAK

Mahasiswa seringkali menghadapi tantangan dalam mengelola pola hidup secara mandiri, yang berdampak pada kebiasaan seperti kurang tidur dan pola makan tidak teratur. Fenomena ini berpotensi menurunkan konsentrasi dan performa akademik, yang selaras dengan isu pada *Sustainable Development Goals (SDGs)* poin ke-3 (Kesehatan dan Kesejahteraan) dan ke-4 (Pendidikan Berkualitas). Penelitian ini bertujuan untuk menganalisis pengaruh pola hidup terhadap Indeks Prestasi (IP) mahasiswa, mengidentifikasi faktor-faktor gaya hidup yang paling berpengaruh, dan merancang solusi berbasis data. Dengan menggunakan pendekatan kuantitatif, data dikumpulkan melalui metode survei untuk mendapatkan informasi pola hidup (data primer) dan data akademik mahasiswa (data sekunder). Solusi yang diusulkan adalah membangun sistem klasifikasi cerdas menggunakan algoritma Naive Bayes untuk memprediksi kategori IP mahasiswa (Rendah, Sedang, Tinggi) berdasarkan variabel-variabel seperti durasi tidur, frekuensi olahraga, dan kebiasaan makan. Diharapkan model yang dihasilkan mampu memprediksi IP dengan tingkat akurasi yang tinggi dan mengonfirmasi hipotesis bahwa pola hidup sehat memiliki hubungan positif yang signifikan dengan prestasi akademik. Hasil penelitian ini diharapkan dapat memberikan rekomendasi praktis bagi mahasiswa dan menjadi landasan bagi institusi pendidikan untuk mengembangkan program pendukung kesejahteraan mahasiswa.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Menjadi mahasiswa bukan hanya tentang mengikuti perkuliahan dan mendapatkan nilai, tetapi juga tentang kemampuan untuk mengatur diri dan waktu secara mandiri. Banyak mahasiswa baru yang mengalami perubahan drastis ketika mulai kuliah mereka harus beradaptasi dengan jadwal padat, tugas yang menumpuk, serta kegiatan organisasi atau pekerjaan sambilan. Akibatnya, tidak sedikit mahasiswa yang mengalami kelelahan fisik maupun mental.

Masalah seperti begadang untuk menyelesaikan tugas, makan tidak teratur, atau kurang tidur menjadi hal yang umum. Padahal, kebiasaan tersebut bisa memengaruhi kemampuan otak dalam menerima pelajaran dan menurunkan daya konsentrasi. Misalnya, penelitian dari *Sleep Foundation* (2023) menunjukkan bahwa seseorang yang tidur kurang dari 6 jam per malam berpotensi mengalami penurunan fokus hingga 40% dibandingkan orang yang tidur cukup.

Fenomena ini juga terlihat di kalangan mahasiswa Indonesia. Berdasarkan survei kecil yang dilakukan di beberapa universitas, Sejumlah 206 mahasiswa (67%) mengalami penundaan pada jenis tugas membaca buku atau referensi yang berkaitan dengan tugas akademik yang diwajibkan. Kondisi ini menunjukkan bahwa gaya hidup bukan hanya soal kebiasaan sehari-hari, tetapi juga faktor penting yang dapat menentukan hasil belajar.

Penelitian ini sejalan dengan tujuan Sustainable Development Goals (SDGs) poin ke-3 (*Good Health and Well-being*) dan poin ke-4 (*Quality Education*), karena keduanya menekankan pentingnya kesejahteraan fisik dan mental untuk mendukung pendidikan berkualitas.

1.2 Rumusan Masalah

- Bagaimana pengaruh pola hidup mahasiswa terhadap IP atau nilai akademik mereka?
- Faktor-faktor pola hidup apa yang paling berpengaruh terhadap prestasi belajar mahasiswa?
- Bagaimana mahasiswa dapat membentuk pola hidup yang lebih baik untuk meningkatkan prestasi akademik?

1.3 Tujuan

- Mengetahui hubungan antara pola hidup mahasiswa dengan nilai akademik.
- Mengidentifikasi faktor yang memengaruhi IP.
- Memberikan gambaran dan rekomendasi pola hidup ideal bagi mahasiswa agar prestasi tetap optimal.

1.4 Manfaat

- Secara Teoritis: Menambah referensi dalam bidang psikologi pendidikan dan perilaku mahasiswa terkait gaya hidup dan kinerja akademik.
- Secara Praktis: Memberi wawasan kepada mahasiswa tentang pentingnya pola hidup seimbang untuk mendukung pencapaian akademik.
- Bagi Kampus: Dapat menjadi dasar bagi lembaga untuk menyusun program pendampingan atau kegiatan penunjang kesejahteraan mahasiswa.

BAB II

TINJAUAN PUSTAKA

2.1 Pola Hidup Mahasiswa

Pola hidup adalah kumpulan kebiasaan sehari-hari seseorang yang mencerminkan cara mereka mengatur waktu, energi, dan prioritas. Pada mahasiswa, pola hidup meliputi cara belajar, jadwal tidur, konsumsi makanan, aktivitas sosial, hingga manajemen stres.

Pola hidup yang sehat mencakup tidur cukup (7–8 jam per hari), olahraga rutin, konsumsi makanan bergizi, serta menjaga keseimbangan antara waktu kuliah dan hiburan. Pola hidup yang tidak teratur, seperti sering begadang, pola makan tidak teratur, dan kurang olahraga, dapat menurunkan kemampuan konsentrasi serta semangat belajar mahasiswa.

2.2 Indeks Prestasi (IP)

Indeks Prestasi (IP) adalah ukuran capaian akademik mahasiswa selama satu semester atau keseluruhan studi. IP menggambarkan tingkat keberhasilan mahasiswa dalam memahami materi kuliah, biasanya dinyatakan dalam skala 0,00–4,00.

Nilai IP dipengaruhi oleh banyak faktor, antara lain kemampuan kognitif, motivasi belajar, lingkungan sosial, serta pola hidup sehari-hari. Mahasiswa dengan gaya hidup yang seimbang cenderung memiliki fokus, daya tahan belajar, dan produktivitas akademik yang lebih baik.

2.3 Hubungan Pola Hidup dan Prestasi Akademik

Beberapa penelitian terdahulu menunjukkan adanya hubungan positif antara pola hidup sehat dengan prestasi akademik. Mahasiswa yang memiliki pola tidur cukup dan asupan gizi seimbang cenderung memiliki fokus serta daya ingat yang lebih baik. Sebaliknya, mahasiswa yang sering begadang, jarang sarapan, atau mengalami stres berlebih biasanya mengalami penurunan motivasi dan nilai akademik.

Dengan demikian, pola hidup menjadi salah satu faktor penting yang dapat memengaruhi pencapaian Indeks Prestasi (IP) mahasiswa. Analisis hubungan ini dapat dilakukan menggunakan pendekatan berbasis data, sehingga hasilnya lebih objektif dan terukur.

2.4 Teknologi dan Metode yang Digunakan

Dalam penelitian ini digunakan metode klasifikasi untuk menganalisis pengaruh pola hidup terhadap pencapaian Indeks Prestasi mahasiswa. Metode klasifikasi digunakan untuk mengelompokkan mahasiswa ke dalam kategori tertentu, seperti IP tinggi, IP sedang, dan IP rendah, berdasarkan variabel pola hidup seperti jam tidur, frekuensi olahraga, kebiasaan makan, serta intensitas belajar.

Algoritma yang digunakan adalah Naive Bayes, yaitu salah satu algoritma klasifikasi yang berbasis pada probabilitas (teorema Bayes). Naive Bayes bekerja dengan menghitung kemungkinan suatu data masuk ke dalam kelas tertentu berdasarkan nilai-nilai fitur yang dimilikinya. Kelebihan algoritma Naive Bayes antara lain:

1. Sederhana dan cepat dalam proses perhitungan.
2. Akurat untuk data survei atau kuesioner, seperti data pola hidup mahasiswa.
3. Cocok untuk jumlah data yang tidak terlalu besar dan memiliki variabel kategori (misalnya “sering olahraga”, “jarang makan pagi”, “tidur cukup”).

Dengan menggunakan algoritma Naive Bayes, penelitian ini dapat memprediksi kategori IP mahasiswa berdasarkan pola hidup yang mereka jalani. Hasil dari klasifikasi ini diharapkan dapat memberikan gambaran seberapa besar pengaruh kebiasaan hidup terhadap prestasi akademik mahasiswa.

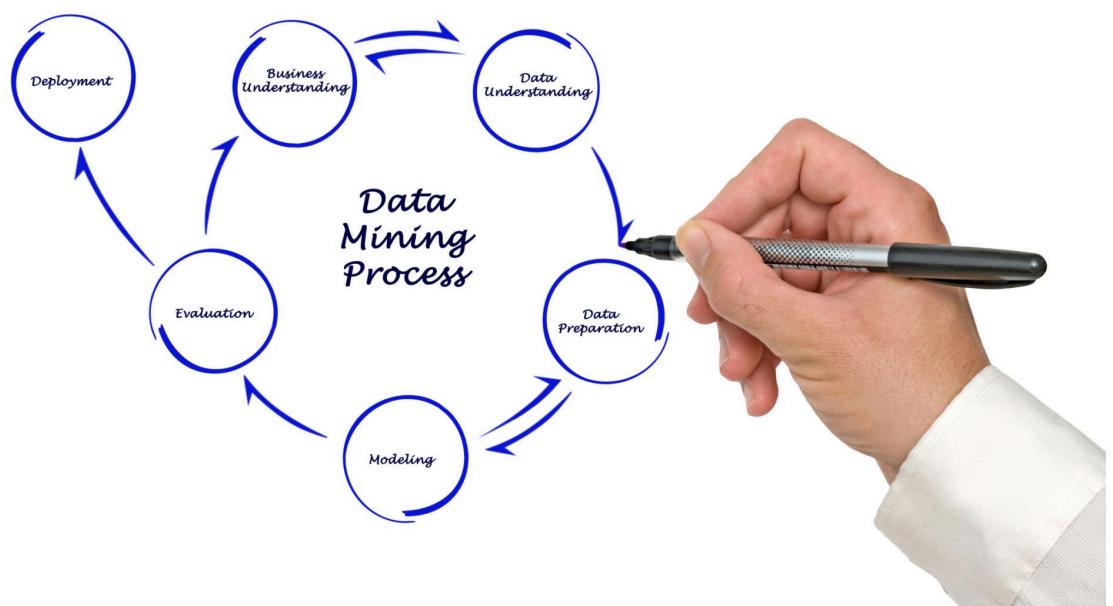
BAB III

METODOLOGI DAN SOLUSI

3.1 Metodologi Perancangan

Hal yang perlu dijelaskan :

- **Pendekatan Penelitian** Penelitian ini menggunakan pendekatan kuantitatif dengan metode *survei* dan *analisis data sekunder*. Data dikumpulkan melalui kuesioner yang disebar kepada mahasiswa untuk mendapatkan informasi mengenai pola hidup mereka, seperti durasi tidur, frekuensi olahraga, kebiasaan makan, dan intensitas belajar. Selain itu, data akademik (Indeks Prestasi) mahasiswa juga akan dikumpulkan sebagai data sekunder untuk dianalisis.
- **Tahapan Perancangan** Proses perancangan sistem prediksi ini mengikuti alur kerja *data mining* yang sistematis, seperti digambarkan dalam diagram alur berikut:



- **Tools yang Digunakan**

- **Software:** Proses analisis data dan implementasi algoritma Naive Bayes akan menggunakan bahasa pemrograman **Python** dengan bantuan pustaka (*library*) seperti **Pandas** untuk manipulasi data, **Scikit-learn** untuk penerapan model *machine learning*, dan **Matplotlib** atau **Seaborn** untuk visualisasi data.
- **Hardware:** Komputer atau laptop dengan spesifikasi standar yang mampu menjalankan lingkungan pengembangan Python.

3.2 Solusi

- **Penjelasan Solusi Utama** Solusi yang diusulkan adalah membangun sebuah **sistem klasifikasi cerdas** menggunakan algoritma **Naive Bayes** untuk memprediksi kategori Indeks Prestasi (IP) mahasiswa (misalnya: Rendah, Sedang, Tinggi) berdasarkan data

pola hidup mereka. Sistem ini akan mengolah data seperti jam tidur, kebiasaan sarapan, frekuensi olahraga, dan manajemen stres untuk memberikan prediksi yang akurat.

- **Cara Kerja Solusi**

1. **Input:** Mahasiswa atau pihak kampus memasukkan data mengenai kebiasaan hidup sehari-hari ke dalam sistem.
2. **Proses:** Algoritma Naive Bayes yang sudah dilatih akan menghitung probabilitas (kemungkinan) dari data input tersebut untuk masuk ke dalam setiap kategori IP (Rendah, Sedang, atau Tinggi).
3. **Output:** Sistem akan menampilkan hasil prediksi berupa kategori IP yang paling mungkin dicapai oleh mahasiswa tersebut berdasarkan pola hidupnya. Sebagai contoh, jika seorang mahasiswa memiliki data "tidur kurang dari 5 jam", "jarang olahraga", dan "sering melewatkkan sarapan", sistem mungkin akan memprediksi bahwa mahasiswa tersebut berisiko memiliki IP dalam kategori "Rendah".

- **Manfaat Solusi**

- **Bagi Mahasiswa:** Memberikan peringatan dini (*early warning*) agar dapat memperbaiki pola hidup untuk meningkatkan performa akademik.
- **Bagi Kampus:** Membantu dosen pembimbing akademik atau konselor dalam mengidentifikasi mahasiswa yang berisiko mengalami penurunan prestasi, sehingga intervensi dapat dilakukan lebih awal.
- **Dampak Positif:** Mendorong terciptanya lingkungan kampus yang lebih sadar akan pentingnya kesehatan fisik dan mental sesuai dengan tujuan SDGs ke-3 (*Good Health and Well-being*) dan ke-4 (*Quality Education*).

- **Batasan Solusi**

- Model prediksi ini bergantung pada kualitas dan kejujuran data yang diisi oleh mahasiswa melalui kuesioner. Data yang tidak akurat akan menghasilkan prediksi yang tidak valid.
- Prediksi yang dihasilkan bersifat probabilistik (berdasarkan kemungkinan) dan bukan merupakan kepastian mutlak, karena prestasi akademik juga dipengaruhi oleh faktor lain di luar pola hidup (misalnya: kemampuan kognitif awal, kondisi sosial-ekonomi).
- Model ini dirancang khusus untuk konteks mahasiswa di lingkungan yang diteliti dan mungkin memerlukan penyesuaian jika diterapkan di populasi mahasiswa yang berbeda.

BAB IV

HIPOTESIS HASIL

Berdasarkan metodologi dan solusi yang telah dirancang, berikut adalah hipotesis atau dugaan hasil yang diharapkan dari penelitian ini:

4.1 Prediksi Keluaran Utama

Sistem klasifikasi berbasis algoritma Naive Bayes yang dirancang diperkirakan dapat berjalan dengan baik dan mampu mengelompokkan mahasiswa ke dalam kategori IP "Rendah", "Sedang", dan "Tinggi" dengan tingkat akurasi di atas 70%. Faktor-faktor seperti durasi tidur, kebiasaan sarapan, dan frekuensi belajar mandiri diprediksi menjadi variabel yang paling berpengaruh signifikan terhadap hasil prediksi IP.

4.2 Pencapaian Tujuan

Tujuan penelitian yang telah ditetapkan pada Bab I diperkirakan akan tercapai. Penelitian ini akan berhasil menunjukkan adanya hubungan yang kuat dan terukur antara pola hidup dengan nilai akademik. Selain itu, model yang dihasilkan dapat memberikan rekomendasi konkret mengenai aspek pola hidup mana yang perlu diperbaiki oleh mahasiswa untuk mengoptimalkan prestasi mereka.

4.3 Kesesuaian dengan Kajian Pustaka

Hasil penelitian ini dihipotesiskan akan sejalan dan memperkuat temuan-temuan dari penelitian sebelumnya yang telah dibahas pada Bab II. Secara spesifik, hasil akan mengonfirmasi bahwa kebiasaan hidup yang tidak sehat (seperti kurang tidur dan pola makan buruk) berkorelasi negatif dengan pencapaian akademik. Dengan demikian, penelitian ini akan memberikan bukti empiris tambahan yang mendukung teori bahwa kesejahteraan mahasiswa adalah fondasi penting untuk pendidikan berkualitas.

DAFTAR PUSTAKA

- Arifianto, F., & Rahayu, N. (2022). *Educational data mining for student academic prediction using K-Means clustering and Naïve Bayes classifier*. Jurnal Pilar Nusa Mandiri, 18(1), 45–54. Retrieved from <https://ejournal.nusamandiri.ac.id/index.php/pilar/article/view/1432>
- Brain Health. (n.d.). *How lack of sleep impacts cognitive performance and focus*. Retrieved October 16, 2025, from <https://brain.health/blog/articles/how-lack-of-sleep-impacts-cognitive-performance-and-focus/>
- National Sleep Foundation. (2023). *2023 Sleep in America Poll Report*. Retrieved from https://www.thensf.org/wp-content/uploads/2024/04/2023-Sleep-in-America-Poll-Report_FIN_AL.pdf
- Putra, D. H., & Susanti, E. (2023). *Classification models for academic performance: A comparative study of Naïve Bayes and Random Forest algorithms in analyzing University of Lampung student grades*. Jurnal Teknik Informatika (JUTIF), 4(3), 230–239. Retrieved from <https://jutif.if.unsoed.ac.id/index.php/jurnal/article/view/2066>
- Sari, D. P., & Rahmawati, I. (2017). *Academic procrastination (conduct academic delay) students Faculty of Engineering Surabaya State University*. Jurnal BK UNESA, 7(2), 1–7. Universitas Negeri Surabaya. Retrieved from <https://ejournal.unesa.ac.id/index.php/jurnal-bk-unesa/article/view/18984/17335>
- Sleep Foundation. (2023). *Sleep deprivation and cognitive performance: What research shows*. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC2656292/>
- Widodo, R., & Prasetyo, D. (2022). *Optimasi model prediksi kelulusan mahasiswa menggunakan algoritma Naïve Bayes*. Buletin Teknologi, 4(2), 101–108. Retrieved from <https://jurnal.kdi.or.id/index.php/bt/article/view/2825>
- Sleep Foundation. (2023). *Sleep deprivation and cognitive performance: What research shows*. Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC2656292/>
- Daffa', M. R. (2024). *Prediksi prestasi akademik mahasiswa menggunakan algoritma Naive Bayes (Studi kasus: Mahasiswa Teknik Informatika angkatan 2020 UIN Maulana Malik Ibrahim Malang) [Skripsi]*, UIN Maulana Malik Ibrahim Malang. Etheses UIN Malang. Retrieved from <https://etheses.uin-malang.ac.id/79064/>
- Deb, B., & Ahmed, K. U. (2024). Utilizing nutritional and lifestyle data for predicting student academic performance: A machine learning approach. *International Journal of Information and Communication Technology*, 10(2), 233–242. Retrieved from https://www.researchgate.net/publication/383142885_UTILIZING_NUTRITIONAL_AND_LIFESTYLE_DATA_FOR_PREDICTING_STUDENT_ACADEMIC_PERFORMANCE_A_MACHINE_LEARNING_APPROACH

Khan, A., Al Mamun, M. A., & Rahman, M. M. (2023). Predicting student academic performance: A machine learning analysis of study habits and lifestyle factors. *International Journal of Advanced Research*, 11(11), 707-717. Retrieved from <https://doi.org/10.21474/IJAR01/18023>

LAMPIRAN

Wajib melampirkan :

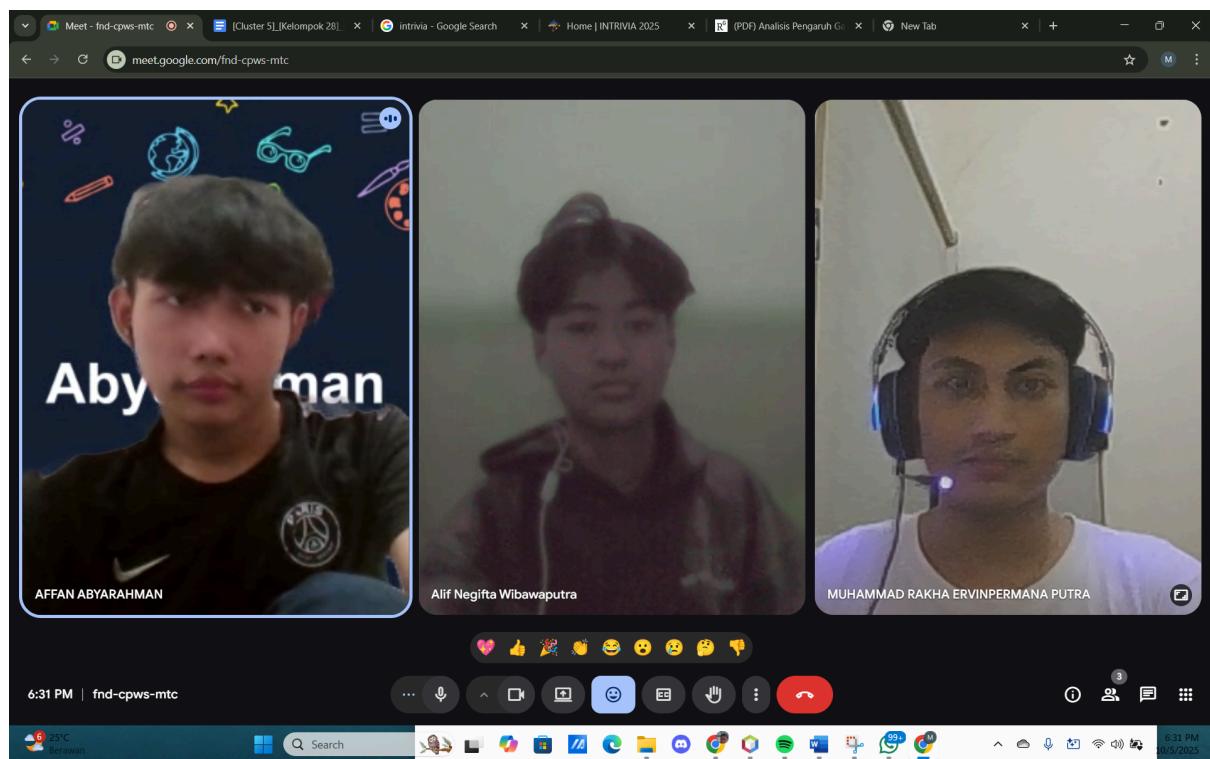
- List Pembagian kerja kelompok

BAB 1-2 : Muhammad Rakha Ervinpermana Putra

BAB 3-4 : Farizal Dwi Candra

Pitch Deck : Michael

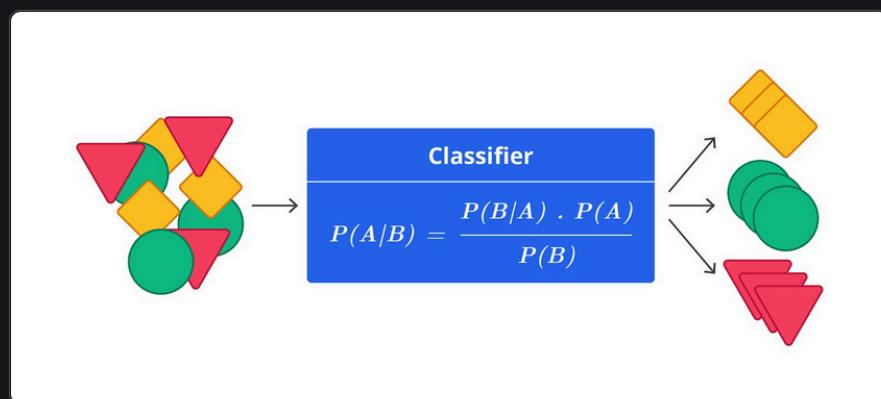




Naive Bayes

Naive Bayes adalah algoritma klasifikasi berbasis probabilitas yang berdasarkan pada Teorema Bayes, dengan asumsi bahwa fitur-fitur dalam data bersifat independen satu sama lain. Nama "naive" (naif) merujuk pada asumsi independensi ini, yang sering kali tidak realistik. Namun, dalam praktiknya dapat menghasilkan model yang efektif. Naive Bayes menggunakan **prinsip probabilitas** untuk memprediksi kelas dari data baru berdasarkan pengamatan fitur yang ada.

Secara matematis, Naive Bayes bekerja dengan menghitung kemungkinan suatu data termasuk dalam kelas tertentu berdasarkan dua faktor: kemungkinan awal dari setiap kelas (*probabilitas prior*) dan kemungkinan fitur dalam data jika kelas tersebut benar (*probabilitas likelihood*).



Rumus untuk menghitung probabilitas posterior $P(A|B)$ dalam konteks Teorema Bayes sebagai berikut.

untuk Pemula

Cari

$$P(A|B) = \frac{P(B/A) \cdot P(A)}{P(B)}$$

Artinya:

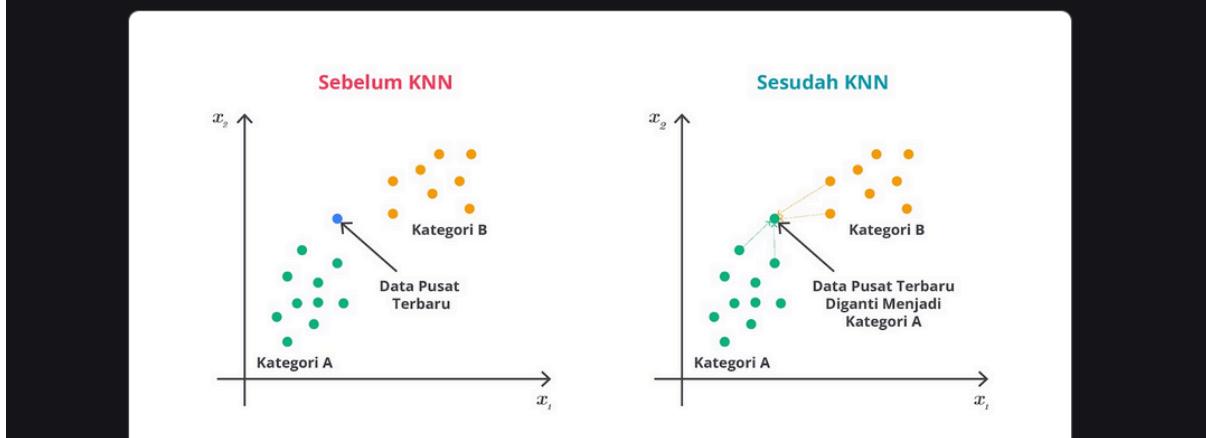
- $P(A|B)$ adalah probabilitas posterior, yaitu probabilitas bahwa peristiwa A terjadi mengingat bahwa peristiwa B telah terjadi.
- $P(B/A)$ adalah probabilitas likelihood, yaitu probabilitas bahwa peristiwa B terjadi jika peristiwa A sudah terjadi.
- $P(A)$ adalah probabilitas prior, yaitu probabilitas awal dari peristiwa A sebelum mempertimbangkan peristiwa B .
- $P(B)$ adalah probabilitas total dari peristiwa B , yang sering dianggap konstan dapat membandingkan probabilitas posterior untuk berbagai nilai A .

Setelah menghitung kemungkinan untuk setiap kelas, model ini memilih kelas dengan kemungkinan tertinggi sebagai hasil klasifikasi. Naive Bayes bisa digunakan untuk berbagai jenis data, seperti teks, gambar, atau tabel dan sering dipakai dalam aplikasi, seperti filter spam, analisis sentimen, serta pengenalan pola.

K-Nearest Neighbors (KNN)

Algoritma K-Nearest Neighbors (KNN) adalah metode supervised learning yang digunakan untuk mengatasi masalah klasifikasi dan regresi. Evelyn Fix dan Joseph Hodges mengembangkan algoritma ini pada tahun 1951 yang kemudian diperluas oleh Thomas Cover. KNN merupakan salah satu algoritma klasifikasi yang paling sederhana dan intuitif dalam machine learning.

Algoritma ini digunakan untuk mengklasifikasikan data baru berdasarkan kedekatannya dengan data yang sudah diberi label dalam dataset pelatihan. KNN sering digunakan karena kemudahannya dalam pemahaman dan implementasi meskipun pada praktiknya, ia dapat menjadi sangat efektif untuk berbagai masalah klasifikasi.



Decision Tree

Decision Tree adalah algoritma machine learning yang sering digunakan dalam tugas klasifikasi dan regresi. Struktur dari algoritma ini mirip dengan bentuk pohon dengan setiap cabang mewakili keputusan atau percabangan dari data berdasarkan fitur-fitur yang ada.

Struktur dasar dari Decision Tree melibatkan tiga komponen utama, yaitu akar (*root node*), *node (decision node)*, dan daun (*leaf node*). Root node mewakili seluruh dataset dan menjadi titik awal untuk pemisahan data. Node-node di sepanjang cabang pohon mewakili keputusan yang diambil berdasarkan fitur tertentu, sedangkan *leaf node* adalah hasil akhir dari proses klasifikasi atau regresi, seperti label kelas atau nilai numerik.

Struktur Dasar dari Decision Tree

