

# MACHINE LEARNING

(UAS)



Nama : Anggi Nur Fitri  
Npm : 41155055190057  
Kelas : Teknik Informatika B

TEKNIK INFORMATIKA  
UNIVERSITAS LANGLANGBUANA  
2022/2023

**Bagian 1 (40%)** Jawab Pertanyaan berikut:

1. Apa itu *Linear dan Logistic Regresion* dan apa gunanya?

regresi sering digunakan dalam penelitian kuantitatif. Regresi dilakukan dalam pengujian pengaruh, biasanya menguji antara pengaruh variable independen terhadap variable dependen (Sekaran & Bougie, 2016). Perbedaan antara regresi linier dengan regresi logistik terletak pada jenis variabel dependennya. Regresi linier digunakan apabila variabel dependennya numerik sedangkan regresi logistik digunakan pada data yang dependennya berbentuk kategori yang dikotom

Kegunaan analisis regresi adalah untuk mengetahui variabel-variabel kunci yang memiliki pengaruh terhadap suatu variabel bergantung, pemodelan, serta pendugaan (estimation) atau peramalan (forecasting). Selain itu, masih ada beberapa kegunaan lainnya, yakni:

1. Membuat estimasi rata-rata dan nilai variabel tergantung dengan didasarkan pada nilai variabel bebas.
2. Untuk menguji hipotesis karakteristik dependensi.
3. Meramalkan nilai rata-rata variabel bebas yang didasari nilai variabel bebas diluar jangkauan sample.

2. Apa itu *Support Vector Machine* dan apa gunanya?

Support Vector Machine (SVM) merupakan salah satu algoritma machine learning dengan pendekatan berbasis supervised learning yang dapat digunakan untuk masalah klasifikasi dan regresi. Digunakan untuk menganalisis data dan mengurutkannya ke dalam salah satu dari dua kategori.

3. Apa itu *K-Nearest Neighbor* dan apa gunanya?

Algoritma ***k-Nearest Neighbor*** adalah algoritma *supervised learning* dimana hasil dari instance yang baru diklasifikasikan berdasarkan mayoritas dari kategori **k**-tetangga terdekat.

Tujuan dari algoritma ini adalah untuk mengklasifikasikan obyek baru berdasarkan atribut dan *sample-sample* dari *training data*.

#### 4. Apa itu Naive Bayes dan apa gunanya?

Naive bayes merupakan metode pengklasifikasian berdasarkan probabilitas sederhana dan dirancang agar dapat dipergunakan dengan asumsi antar variabel penjelas saling bebas (independen). Pada algoritma ini pembelajaran lebih ditekankan pada pengestimasian probabilitas.

Algoritma naive bayes memiliki banyak manfaat sehingga banyak digunakan di berbagai aspek kehidupan. Antara lain sebagai berikut:

- Real time prediction
- Multiclass prediction, disini naive bayes dapat memprediksi probabilitas beberapa kelas variabel target
- Text classification, selain itu naive bayes juga sering digunakan untuk membuat text classification karena memiliki tingkat keberhasilan yang lebih tinggi dibandingkan dengan algoritma lain
- Recommendation system, naive bayes juga digunakan untuk proyek data mining untuk menyaring informasi yang tidak terlihat dan memprediksi apakah pengguna menginginkan sumber daya yang diberikan atau tidak.

#### 5. Apa itu *Decision Tree* dan apa gunanya?

Decision tree adalah algoritma machine learning yang menggunakan seperangkat aturan untuk membuat keputusan dengan struktur seperti pohon yang memodelkan kemungkinan hasil, biaya sumber daya, utilitas dan kemungkinan konsekuensi atau resiko.

Kegunaannya decision tree memiliki kemampuan untuk mem-*break down* proses pengambilan keputusan yang kompleks menjadi lebih simple, sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan. Selain itu juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target.

6. Apa itu Random Forest dan apa gunanya?

Random Forest adalah algoritma machine learning yang menggabungkan keluaran dari beberapa decision tree untuk mencapai satu hasil. Sesuai namanya, Forest atau 'hutan' dibentuk dari banyak tree (pohon) yang diperoleh melalui proses bagging atau bootstrap aggregating.

Kegunaannya :

- Random Forest digunakan di perbankan untuk memprediksi kelayakan kredit dari seorang debitur atau pemohon pinjaman. Algoritma ini membantu kreditur atau lembaga pemberi pinjaman membuat keputusan yang baik apakah akan memberikan pinjaman kepada debitur atau tidak. Bank juga menggunakan algoritma Random Forest untuk mendeteksi penipuan.
- Pekerja di bidang kesehatan menggunakan Random Forest untuk mendiagnosis pasien. Pasien didiagnosis dengan menilai riwayat medis mereka sebelumnya. Rekam medis masa lalu ditinjau untuk menetapkan dosis yang tepat untuk pasien.
- Analisis keuangan bisa menggunakan Random Forest untuk mengidentifikasi pasar potensial untuk saham. Algoritma ini juga memungkinkan bagi seorang trader untuk mengetahui kondisi dan perilaku sebuah saham.
- Melalui algoritma Random Forest, vendor e-commerce dapat memprediksi preferensi pelanggan berdasarkan perilaku konsumsi masa lalu.

7. Apa itu *K-Means* dan apa gunanya?

**K-Means Clustering** adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pemodelan unsupervised learning dan menggunakan metode yang mengelompokkan data berbagai partisi. Kegunaannya untuk meminimalisasikan fungsi objective yang telah di set dalam proses clustering. Tujuan tersebut dilakukan dengan cara meminimalikan variasi data yang ada didalam cluster dan memaksimalkan variasi data yang ada di cluster lainnya.

8. Apa itu *Agglomerate Clustering* dan apa gunanya?

**Agglomerative** (metode penggabungan) adalah strategi pengelompokan hirarki yang dimulai dengan setiap objek dalam satu cluster yang terpisah kemudian membentuk cluster yang semakin membesar. Jadi, banyaknya cluster awal adalah sama dengan banyaknya objek. .... (Supranto, 2004)

Kegunaannya dapat memanfaatkan data produksi dan data penjualan perusahaan untuk melihat pola dendrogram dan pengelompokan produk-produk yang paling menguntungkan

9. Apa itu *Apriori Algorithm* dan apa gunanya?

Algoritma Apriori adalah salah satu algoritma pada data mining untuk mencari frequent item/itemset pada transaksional database

Algoritma apriori berguna untuk menentukan frequent itemset pada sekumpulan data, banyak digunakan pada data transaksi atau biasa disebut market basket, misalnya sebuah swalayan memiliki market basket, dengan adanya algoritma apriori, pemilik swalayan dapat mengetahui pola pembelian seorang konsumen

10. Apa itu *Self Organizing Map* dan apa gunanya?

*Self-organizing maps* (SOM) merupakan suatu jenis *artificial neural network* yang dilatih dengan metode *unsupervised learning*. Jaringan ini mampu menghasilkan sebuah representasi terpisah atas ruang *input* sampel pelatihan dengan dimensi rendah (biasanya dua dimensi).

Tujuannya untuk melakukan visualisasi data dengan cara mengurangi dimensi data melalui penggunaan *self-organizing neural networks* sehingga manusia dapat mengerti *high-dimensional* data yang dipetakan dalam bentuk *low-dimensional* data.

## Pemetaan warna dengan Scattered Plot :

