

Nama : Charderra Eka Bagas Sanjaya
Nim : 20210801088
Prodi : Teknik Informatika
UTS Machine Learning

Essay :

1. Pembelajaran mesin (ML) adalah cabang dari kecerdasan buatan (AI) yang berfokus pada pembuatan algoritma dan model statistik. Ini memungkinkan komputer untuk belajar dan meningkatkan kinerja dari pengalaman tanpa harus diprogram untuk setiap tugas.

Cara Pembelajaran Mesin Berfungsi:

- Pengumpulan Data, Untuk pembelajaran mesin, data dalam jumlah besar dan berkualitas tinggi diperlukan. Data tersebut dapat berupa angka, gambar, teks, atau jenis data lainnya.
 - Pemilihan Fitur, Fitur adalah atribut atau karakteristik data yang relevan yang akan digunakan untuk pembelajaran. Pemilihan Algoritma: Ada banyak jenis algoritma ML, termasuk regresi, pohon keputusan, dan jaringan saraf tiruan, dan pemilihan algoritma yang tepat tergantung pada jenis masalah yang ingin diselesaikan.
2. Untuk kegiatan sehari-hari:
 - 1. Mencegah spam email:
 - Mengapa: Kami menerima banyak email setiap hari, kebanyakan spam. Memfilter email spam secara manual sangat sulit dan memakan waktu.
 - Manfaat: Filter spam email berbasis ML dapat mengidentifikasi pola dalam email spam, seperti kata kunci atau alamat pengirim yang mencurigakan, sehingga menghemat waktu dan tenaga dengan mengidentifikasi dan memindahkan email spam ke folder yang berbeda.
 - 2. Saran Produk/Konten:
 - Mengapa: Banyak barang dan konten tersedia di platform e-commerce atau layanan streaming. Secara manual, menemukan yang sesuai dengan minat kita sangat sulit dan memakan waktu.
 - Manfaat: ML dapat melihat riwayat tontonan dan pembelian kami, serta preferensi pengguna lain yang sebanding. Dengan demikian, ML dapat memberikan rekomendasi produk atau konten yang relevan dengan minat kita, sehingga memudahkan kita menemukan hal-hal yang kita sukai.

3. Beberapa faktor dapat menentukan kategori taksonomi yang digunakan dalam penerapan machine learning (ML), seperti

1. Berdasarkan Pendekatan Pembelajaran:

- Supervised Learning (Pembelajaran Terarah), Metode ini melatih model pembelajaran mesin dengan data yang sudah memiliki label atau tujuan yang jelas. Tujuannya adalah agar model dapat belajar dari data dan membuat prediksi yang akurat untuk data yang belum dilabelkan. Regresi linear, pohon keputusan, dan Support Vector Machine (SVM) adalah beberapa contoh algoritma pembelajaran yang diawasi.
- Unsupervised Learning (Pembelajaran Tak Terarah), Metode ini menggunakan data yang tidak memiliki label atau tujuan untuk melatih model pembelajaran mesin. Tujuannya adalah agar model memiliki kemampuan untuk menemukan struktur atau pola yang tersembunyi dalam data yang dikumpulkan. Clustering (K-Means, DBSCAN) dan Principal Component Analysis (PCA) adalah dua contoh algoritma pembelajaran tidak diawasi.
- Pembelajaran Semi-Terarah (Semi-Terarah), Teknik ini merupakan kombinasi dari pembelajaran yang diawasi dan tidak diawasi, di mana model dilatih dengan sebagian data berlabel dan sebagian data tidak berlabel. Tujuannya adalah untuk menggunakan kedua jenis data untuk meningkatkan kinerja model.
- Pembelajaran Penguatan (Reinforcement Learning), Dalam pendekatan ini, model belajar melalui interaksi dengan lingkungannya dan menerima umpan balik dalam bentuk penghargaan atau hukuman. Tujuannya adalah agar model dapat mengetahui apa yang harus dilakukan untuk mencapai tujuan tertentu.

2. Berdasarkan pekerjaan yang telah diselesaikan:

- Klasifikasi adalah proses mengkategorikan data ke dalam kelas atau kategori yang telah ditentukan. Salah satu contohnya adalah mengklasifikasikan email spam sebagai spam atau tidak spam.
- Regresi adalah pekerjaan memprediksi nilai numerik kontinu. Contoh: membuat perkiraan harga rumah berdasarkan luas tanah dan jumlah kamar.
- Clustering adalah proses mengelompokkan data ke dalam kelompok-kelompok yang memiliki karakteristik yang mirip. Sebagai contoh, mengelompokkan pelanggan berdasarkan cara mereka membeli sesuatu.
- Mengidentifikasi Anomali: Tugas mengidentifikasi data yang tidak biasa atau menyimpang dari pola normal. Untuk ilustrasi, identifikasi penipuan kartu kredit.

Studi Kasus

1. **Source Code:**

```
import pandas as pd

# Dataset
data = {
    'Hari': ['Senin', 'Selasa', 'Rabu', 'Kamis', 'Jumat', 'Sabtu', 'Minggu'],
    'Datang': [2, 3, 4, 1, 2, 5, 2],
    'Biaya': [30000*2, 35000*3, 25000*4, 15000*1, 20000*2, 30000*5,
35000*2],
    'Mahasiswa': ['Ani', 'Budi', 'Jono', 'Lono', 'Joni', 'Ani', 'Budi']
}

# Buat DataFrame
df = pd.DataFrame(data)

# Hitung mean kolom `Datang`, bulatkan hasilnya menjadi 2 desimal
rata_rata_datang = round(df['Datang'].mean(), 2)

# Cari baris dimana `Biaya` maksimum terjadi, lalu ambil nilai pada kolom
`Hari` dari baris tersebut
hari_biaya_tertinggi = df.loc[df['Biaya'].idxmax(), 'Hari']

# Filter DataFrame yang `Biaya` lebih besar dari 110000, lalu dapatkan nilai
unik dari kolom `Hari`
hari_biaya_lebih_110000 = df[df['Biaya'] > 110000]['Hari'].unique().tolist()

# Hitung frekuensi setiap nilai unik di kolom `Mahasiswa` dan dapatkan nilai
teratas
mahasiswa_paling_banyak_datang = df['Mahasiswa'].value_counts().index[0]

# Filter DataFrame dimana `Hari` sama dengan 'Minggu', lalu dapatkan nilai
unik dari kolom `Mahasiswa`
datang_hari_minggu = df[df['Hari'] == 'Minggu']['Mahasiswa'].unique().tolist()

# Dapatkan nilai maksimum dan minimum kolom `Biaya`
biaya_tertinggi = df['Biaya'].max()
biaya_terendah = df['Biaya'].min()

# Hitung frekuensi setiap nilai unik pada kolom `Datang`, lalu dapatkan nilai
teratas dan terbawah
frekuensi_datang_tertinggi = df['Datang'].value_counts().index[4]
frekuensi_datang_terendah = df['Datang'].value_counts().index[3]

# Cetak hasilnya
print("Rata-rata mahasiswa datang pada minggu ini:", rata_rata_datang)
print("Hari dengan biaya tertinggi:", hari_biaya_tertinggi)
```

```

print("Hari dengan biaya lebih dari 110000:", hari_biaya_lebih_110000)
print("Mahasiswa yang paling banyak datang ke kampus:",
mahasiswa_paling_banyak_datang)
print("Mahasiswa yang datang pada hari Minggu:", datang_hari_minggu)
print("Biaya tertinggi:", biaya_tertinggi)
print("Biaya terendah:", biaya_terendah)
print("Frekuensi datang tertinggi:", frekuensi_datang_tertinggi)
print("Frekuensi datang terendah:", frekuensi_datang_terendah)

```

Output:

```

Rata-rata mahasiswa datang pada minggu ini: 2.71
Hari dengan biaya tertinggi: Sabtu
Hari dengan biaya lebih dari 110000: ['Sabtu']
Mahasiswa yang paling banyak datang ke kampus: Ani
Mahasiswa yang datang pada hari Minggu: ['Budi']
Biaya tertinggi: 150000
Biaya terendah: 15000
Frekuensi datang tertinggi: 5
Frekuensi datang terendah: 1

```

2. Source Code:

```

import pandas as pd
from plotnine import ggplot, aes, geom_bar, geom_text, theme_bw

fakultas = ["Bisnis", "D3 Perhotelan", "ICT", "Ilmu Komunikasi", "Seni dan
Desain"]
jumlah_mahasiswa = [260, 28, 284, 465, 735]
akreditasi = ["A","A","B","A","A"]

# Create DataFrame (Corrected)
info_mahasiswa = pd.DataFrame({"fakultas": fakultas, "jumlah_mahasiswa":
jumlah_mahasiswa, "akreditasi": akreditasi})

# Create Plot
gambar = (
    ggplot(info_mahasiswa, aes(x="fakultas", y="jumlah_mahasiswa",
fill="fakultas"))
)

gambar = gambar + geom_bar(width=1, stat="identity")

# Show the plot
print(gambar)

```

Output :

