

**FAKULTAS ILMU KOMPUTER – UPN “VETERAN” JAWA TIMUR****SOAL UJIAN TENGAH SEMESTER GANJIL TA. 2019/2020**

Nama : Fadlila Agustina
NPM : 21083010050

MATA KULIAH : PENGANTAR SAINS DATA
SMT / PROGRAM STUDI: I / SAINS DATA
HARI / TANGGAL : SENIN, 18 OKTOBER 2021
WAKTU : 120 MENIT
SIFAT : TERBUKA
DOSEN : Sugiarto, S.Kom., M.Kom
Kartika Maulida Hindrayani S.Kom, M.Kom
KELAS / PRARALEL : B

PETUNJUK

- ✓ **Lembar Jawaban** : Tuliskan Nama, NPM, Paralel Kelas, di sebelah kiri atas lembar jawaban.
- ✓ **Lembar Soal** : Dikumpulkan kembali. Tuliskan Nama dan NPM di pojok kiri atas.

SOAL :

1. Jelaskan sesuai dengan pemahaman anda mengenai daur hidup (life cycle) Data Science dan berikan contoh di masing-masing tahap!
2. Jelaskan mengenai perbedaan peran data engineer, data scientist, data analyst dan kaitan peran tersebut dalam Big Data dengan data sains!
3. Tantangan apa saja yang dihadapi oleh seorang pekerja di bidang data saat ini di Indonesia? Jelaskan menurut pendapat anda!
4. Pada sebuah perusahaan perbankan yang bonafit, perekaman data sudah dilakukan dalam setiap prosesnya. Pihak manajemen perbankan ingin melakukan segmentasi customer yang memiliki rekening di perbankan. Segmentasi dilakukan agar penjualan produk perbankan seperti kredit, asuransi, program tabungan, dll dapat lebih tepat sasaran dan menjalin hubungan dengan pelanggan. Menurut anda, apa saja fitur-fitur penting yang harus dilibatkan ke dalam dataset. Sebutkan dan jelaskan beserta contoh sampelnya (paling sedikit 5 sampel).
5. Project charter merupakan tahap awal dalam Data Science untuk mengidentifikasi dan mengklarifikasi detail proyek yang akan dikerjakan kepada client.

Studi kasus : Kementerian Perhubungan menghadapi berbagai permasalahan dalam sektor transportasi, baik transportasi darat, laut, dan udara. Permasalahan tersebut dapat

| Acuan : | Soal ini dibuat oleh : | Ditinjau dan Divalidasi oleh : |
|--|---|---|
| 1. RPS 2. CPMK Pengantar Sains Data | Dosen Pengampu Sugiarto, S.Kom., M.Kom Kartika Maulida Hindrayani S.Kom, M.Kom | Koordinator Mata Kuliah (Nama Dosen Pengampu) |

**FAKULTAS ILMU KOMPUTER – UPN “VETERAN” JAWA TIMUR****SOAL UJIAN TENGAH SEMESTER GANJIL TA. 2019/2020**

diatasi dengan kebijakan pemerintah dan kepatuhan rakyat dalam sektor transportasi. Data Engineer dan Data Scientist Kementerian Perhubungan diminta untuk melakukan pengumpulan data dari sektor transportasi dan pengolahan data untuk mendapatkan sebuah insight. Tuangkan studi kasus tersebut dalam project charter!

1. Definisikan research goal
2. Konteks dan misi proyek
3. Apa yang akan Anda analisis
4. Apa saja resources (sumber daya) yang harus disiapkan
5. Pastikan bahwa rencana proyek Anda mempunyai nilai kontribusi dan dapat dicapai
6. Buat indikator ketercapaian proyek Anda
7. Buat Timeline

~ Kerjakanlah soal secara jujur, mandiri, semoga sukses~
Selamat mengerjakan

| Acuan : | Soal ini dibuat oleh : | Ditinjau dan Divalidasi oleh : |
|---|---|--|
| <ol style="list-style-type: none">1. RPS2. CPMK Pengantar Sains Data | Dosen Pengampu Sugiarto, S.Kom., M.Kom Kartika Maulida Hindrayani S.Kom, M.Kom | Koordinator Mata Kuliah (Nama Dosen Pengampu) |



Nama : Fadlila Agustina
NPM : 21083010050
Kelas Paralel : B

LEMBAR JAWABAN UJIAN TENGAH SEMESTER GANJIL

1. Daur Hidup Data Science :

➤ **Business Understanding**

Pemahaman tentang bisnis, mengidentifikasi topik dan permasalahan yang terjadi di sekitar kita.
Contoh : di bidang keuangan dan perekonomian ada kasus seperti investasi bodong.

➤ **Data Mining**

Proses pengambilan atau pengumpulan data penting dari suatu data besar yang berasal dari primer dan sekunder.

Contoh : gaji, jenis investasi yang diinginkan, nama, asal, dll.

➤ **Data Cleaning**

Memeriksa data dan menghapus data yang tidak diperlukan.

Contoh : menghapus asal dan beberapa data diri yang kurang dibutuhkan.

➤ **Data Exploration**

Pengelompokan hasil dari data cleaning.

Contoh : kelompok data diri, kelompok keuangan, kelompok pengertian tentang investasi.

➤ **Feature Engineering**

Menyeleksi sekiranya data yang terpenting dan lebih menguntungkan dalam suatu project.

Contoh : gaji yang diterima, presentase yang ingin diinvestasikan.

➤ **Predictive Modeling**

Melakukan prediksi model atau tipe machine learning yang cocok digunakan dalam suatu project.

Contoh : supervised learning, unsupervised learning, regression model, association rules.

➤ **Data Visualization**

Presentasi dari sebuah data agar mudah dipahami dan dicerna oleh otak manusia.

Contoh : dashboard visualization, business intelligence, grafik atau plot.

2. Perbedaan peran data engineer, data scientist, data analyst, dan kaitan peran tersebut dalam Big Data dengan data sains.

- **Data Engineer** : generalis serbaguna yang menggunakan ilmu computer untuk membantu memproses kumpulan data besar. Data engineer biasanya fokus pada pengkodean, pembersihan kumpulan data, dan penerapan permintaan yang datang dari data scientist.
- **Data Scientist** : dapat mengambil project data science dari ujung ke ujung, membantu menyimpan data dalam jumlah besar, membuat pemodelan prediktif, dan menyajikan temuan.
- **Data Analyst** : membantu orang-orang dari seluruh perusahaan memahami kueri (suatu bahasa computer yang digunakan untuk melakukan permintaan terhadap basis data dan sistem informasi) tertentu dengan bagan.

3. Tantangan yang dihadapi oleh seorang pekerja di bidang data saat ini di Indonesia? Jelaskan menurut pendapat anda!

- Tantangan yang dihadapi oleh seorang pekerja di bidang data saat ini di Indonesia adalah kurangnya tenaga profesional yang memahami data terlebih sebuah big data, pendapatan atau gaji yang diterima di dalam negeri tidak sebanding dengan pekerjaan yang dilakukan atau gajinya lebih besar jika bekerja di luar negeri, sehingga banyak WNI yang lebih memilih bekerja di luar negeri daripada di dalam negeri.

4. Sebutkan dan jelaskan beserta contoh sampelnya (paling sedikit 5 sampel).



- Fitur-fitur yang harus dilibatkan dalam sebuah dataset, yaitu jenis pekerjaan, perkiraan pendapatan atau gaji yang didapatkan, dan presentase penyesihan gaji.
- a) Jenis pekerjaan :
- Pegawai Negeri Sipil (camat)
 - Tenaga Kesehatan (perawat)
 - Pengusaha (wiraswasta)
 - Pegawai Swasta (buruh)
 - Jasa Hukum (pengacara)
- b) Perkiraan pendapatan yang didapatkan
- Pegawai Negeri Sipil (camat) : 3 – 4 juta
 - Tenaga Kesehatan (perawat) : 4 – 7 juta
 - Pengusaha (wiraswasta) : 3 – 5 juta
 - Pegawai Swasta (buruh) : 3 – 4 juta
 - Jasa Hukum (pengacara) : 4 – 6 juta
- c) Presentase penyesihan gaji untuk asuransi dll
- 1 – 2 juta : 3 – 5 %
 - 2 – 3 juta : 5 – 7 %
 - 4 – 5 juta : 7 – 8 %
 - 6 – 8 juta : 10 – 11 %
 - 6 – 10 juta : 12 – 14 %

5. 1. Research Goal

Transportasi memiliki peran strategis dalam proses pembangunan yang tidak hanya dapat mendorong pembangunan daerah, tetapi juga mendukung pembangunan ekonomi nasional. Tanpa adanya transportasi sebagai sarana penunjang tidak dapat diharapkan tercapainya hasil yang memuaskan dalam usaha pengembangan ekonomi dari suatu negara. Transportasi akan memberikan kontribusi manfaat yang besar terhadap pertumbuhan ekonomi. Dilihat dari pentingnya peran transportasi dalam suatu daerah, maka diperlukan manajemen pengelolaan transportasi kota yang baik agar dapat mempermudah mobilitas masyarakat perkotaan dalam melakukan aktivitas.

2. Konteks dan Misi

Kelengkapan sarana dan prasarana transportasi perkotaan yang baik akan mampu memberikan pelayanan kepada masyarakat dalam melakukan mobilitas. Terdapat beragam moda transportasi di perkotaan. Lalu lintas dan angkutan jalan sebagai bagian dari sistem transportasi nasional harus dikembangkan potensi dan perannya untuk mewujudkan keamanan, keselamatan, ketertiban, dan kelancaran berlalu lintas dan angkutan jalan dalam rangka mendukung pembangunan ekonomi dan wilayah.

Misi project adalah menganalisa sistem transportasi darat di perkotaan khususnya di wilayah kota Surabaya.

3. Analisis

- Pertumbuhan kebutuhan transportasi sering tidak seimbang dengan penyedia prasarana.
- Kualitas dan jumlah angkutan umum yang belum memadai sarana dan prasarana, jaringan jalan dan jaringan trayek, terminal, sistem pengendalian dan pelayanan angkutan umum belum berhasil ditata secara konseptual.
- Penggunaan kendaraan pribadi yang kurang efisien.



4. Resources (sumber daya)

Pengumpulan data-data, baik data sekunder maupun primer dengan menggunakan metode library research.

5. Nilai Kontribusi

Dengan menganalisis studi kasus tersebut dapat membantu pemerintah mengatasi kemacetan di wilayah perkotaan dan membantu pemerintah menumbuhkan perekonomian karena transportasi merupakan sektor yang sangat penting bagi pemerintah.

6. Indikator Ketercapaian

- Masyarakat dapat memahami penggunaan transportasi dengan baik 40 %
- Kemacetan teratasi menjadi 30 %
- Pertumbuhan ekonomi meningkat 15 %
- Perkembangan sarana transportasi 15 %

7. Timeline

- Mulai project (1 November 2021)
- Identifikasi permasalahan (10 November 2021)
- Pengumpulan data dan analisis data (15 November 2021)
- Pengujian awal project (20 November 2021)
- Pengembangan project (30 November 2021)
- Pengembangan dan perbaikan project (5 Desember 2021)
- Penutupan project (15 Desember 2021)