**AI/ML Engineer Challenge**

**Deadline : 4 Agustus 2024, 00.00**

**1. LLM Adaptability :** **Membuat Soal Pilihan Ganda Menggunakan API LLM dari Cakra.ai**

**Deskripsi:** Anda diminta untuk membuat aplikasi yang dapat menghasilkan 5 pertanyaan dengan pilihan ganda (ABCD) dalam bahasa Indonesia berdasarkan dokumen yang diberikan. Anda akan menggunakan API LLM dari cakra.ai untuk membantu menghasilkan pertanyaan-pertanyaan tersebut.

**Persiapan:**

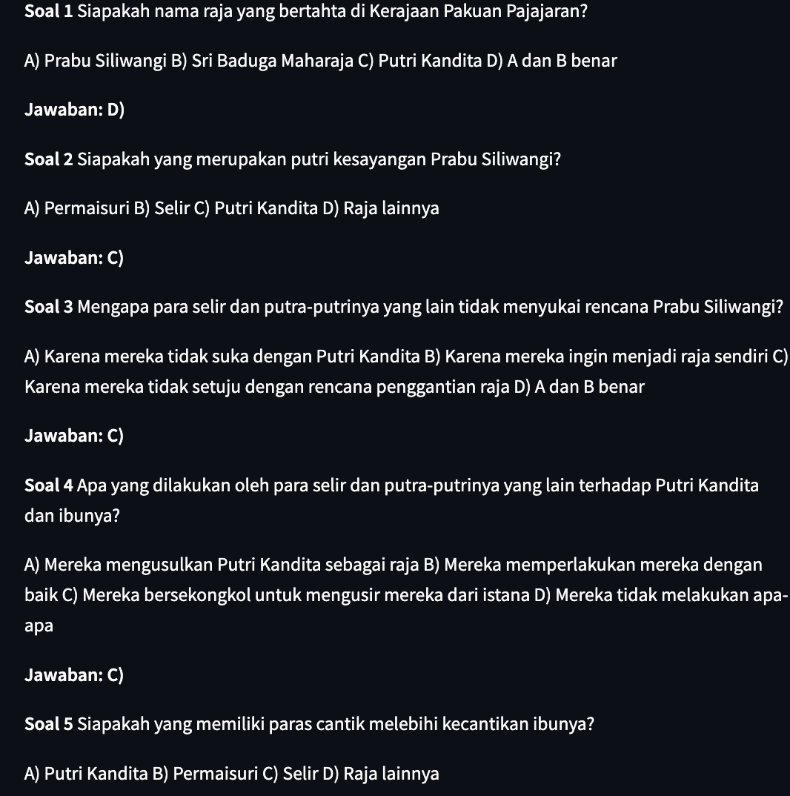
* **Download Buku:** Download buku yang dibutuhkan untuk dibuat menjadi soal pilihan ganda <https://drive.google.com/file/d/1MlpSUiOp8Auestg9Yiwce6ulWwfWOlNS/view?usp=sharing> .
* **Login:** Login pada <https://saas.cakra.ai> untuk mendapatkan key akses API dari cakra.ai.
  + **Akses:**
    - **Username:** trial\_sabrina
    - **Password:** Qwesabrina123
  + Dapatkan akses API key untuk mengonsumsi API LLM cakra.ai pada menu "My Service".
  + A screenshot of a service

    Description automatically generated
* **Pelajari Dokumentasi:** Pelajari dokumentasi API cakra.ai untuk memahami cara mengonsumsi API LLM.
  + Pilih service API LLM pada service list untuk melihat dokumentasi API LLM cakra.ai.
  + A screenshot of a service

    Description automatically generated
  + Pelajari dokumentasi API LLM cakra.ai di <https://saas.cakra.ai/service/detail/1>.
  + A screenshot of a computer

    Description automatically generated

**Membuat Aplikasi:**

* **Bahasa Pemrograman:** Buat aplikasi dalam bahasa pemrograman Python yang akan melakukan hal berikut:
  + **Load Dokumen:** Load dokumen yang diberikan. (5 poin)
  + **Konsumsi API:** Konsumsi API LLM dari cakra.ai untuk mendapatkan output dari API LLM tersebut. (30 poin)
  + **Output:** berikan output menjadi 5 pertanyaan dengan pilihan ganda ABCD dalam bahasa Indonesia. (50 poin)  
    contoh output :  
    
  + **Simple UI:** Berikan simple UI pada streamlit, gradio, flask, etc untuk inference programnya. (15 poin)

2. **Data Science Challenge: Memprediksi Kualitas Jeruk Berdasarkan Dataset**

**Deskripsi:** Anda diminta untuk membuat model prediktif yang dapat memprediksi kualitas jeruk berdasarkan dataset yang diberikan. Dataset ini berisi informasi mengenai berbagai atribut yang mempengaruhi kualitas jeruk.

**Persiapan:**

* **Download Dataset:** Download dataset yang dibutuhkan dari Kaggle.
  + <https://www.kaggle.com/datasets/shruthiiiee/orange-quality>

**Langkah-Langkah:**

1. **Preprocessing Data (20 poin):**
   * Lakukan preprocessing data yang diperlukan, seperti menangani missing values, normalisasi, encoding categorical variables, dll.
2. **Feature Engineering (10 poin):**
   * Buat fitur-fitur tambahan jika diperlukan untuk meningkatkan performa model.
3. **Membangun Model (30 poin):**
   * Bagi dataset menjadi data latih dan data uji.
   * Bangun model prediktif menggunakan algoritma machine learning yang sesuai (misalnya, linear regression, decision tree, random forest, dll).
   * Latih model dengan data latih dan evaluasi performa model dengan data uji.
4. **Evaluasi Model (20 poin) :**
   * Gunakan metrik evaluasi yang sesuai (misalnya, R-squared, Mean Absolute Error, Mean Squared Error) untuk mengukur performa model.
   * Lakukan tuning hyperparameter jika diperlukan untuk meningkatkan performa model.
5. **Model Inference (20 Poin):**
   * Berikan inference model nya menggunakan streamlit, gradio, flask, etc

**Pengiriman:**  
Unggah kode repositori Anda di GitHub, BitBucket, atau layanan serupa lainnya dan pastikan bahwa repositori tersebut dapat diakses.

* **Deploy Aplikasi:** Deploy aplikasi Anda pada platform seperti Huggingface space, GitHub Pages, Streamlit, dll. Sertakan URL dari repositori dan situs web yang live.
* **Ringkasan Proyek:** Masukkan ringkasan singkat (1-2 paragraf) tentang proyek Anda, menekankan fitur inovatif atau pertimbangan penting dalam proses pengembangan.

**What We Look For:**  
Kami mencari solusi yang menunjukkan:

* Pemahaman yang kuat tentang integrasi API dan kemampuan untuk memanfaatkan teknologi LLM.
* Kemampuan dalam menangani teks bahasa Indonesia secara efektif.
* Kemampuan dalam menangani data dengan metode preprocessing yang tepat.
* Keahlian dalam memilih dan membangun model machine learning yang sesuai.
* Dokumentasi dan penjelasan kode yang baik.