**Soal LKS Bidang Lomba**

**Artificial Intelligence (AI)**

**Tingkat Kabupaten Sidoarjo**

**2025**

1. **Deskripsi Kompetisi**

Sekolah Menengah Kejuruan (SMK) memiliki peran penting dalam membekali siswa dengan keterampilan yang siap digunakan di dunia industri. Oleh karena itu, pemahaman terhadap **kepuasan siswa mengenai berbagai aspek sekolah**, seperti kualitas pengajaran, fasilitas, kesiapan industri, serta sistem pembelajaran digital, sangat penting untuk meningkatkan mutu pendidikan vokasi.

Dalam lomba ini, peserta ditantang untuk mengembangkan **sistem kecerdasan buatan (AI) yang dapat menganalisis sentimen siswa SMK** berdasarkan data survei kepuasan mereka. Sistem ini harus mampu **melakukan pembersihan data, menerapkan metode pembobotan kata, melakukan klasifikasi sentimen, serta menampilkan hasil dalam bentuk antarmuka grafis (GUI)**.

Peserta akan diberikan **dataset** yang berisi tanggapan siswa terhadap berbagai aspek sekolah, termasuk tetapi tidak terbatas pada:

* Kualitas pengajaran dan kurikulum
* Ketersediaan fasilitas dan infrastruktur
* Kesiapan sekolah dalam mendukung praktik industri
* Kepuasan terhadap sistem pembelajaran digital
* Interaksi dengan tenaga pendidik dan bimbingan karir

Tantangan utama dalam kompetisi ini adalah **memproses data teks dalam bahasa Indonesia**, membangun **model klasifikasi sentimen**, serta mengembangkan **antarmuka pengguna (GUI) interaktif** untuk menyajikan hasil analisis dalam format yang dapat digunakan oleh pemangku kepentingan sekolah.

1. **Tugas Peserta**

Peserta diminta untuk:

1. **Melakukan Data Preprocessing**  
   Membersihkan data teks dari karakter tidak relevan, melakukan tokenisasi, serta melakukan normalisasi teks.
2. **Menggunakan TF-IDF untuk Pembobotan Kata**  
   Menerapkan metode **Term Frequency-Inverse Document Frequency (TF-IDF)** sebagai representasi numerik dari teks.
3. **Melatih Model Klasifikasi**  
   Mengembangkan model machine learning atau deep learning yang dapat mengkategorikan sentimen siswa menjadi **positif atau negatif**.
4. **Mengembangkan Antarmuka Pengguna (GUI)**  
   Membuat aplikasi berbasis GUI yang memungkinkan pengguna untuk memasukkan teks atau mengunggah file survei baru untuk mendapatkan analisis sentimen secara otomatis.
5. **Persyaratan dan Batasan**

* Bahasa pemrograman utama: Python
* Perpustakaan yang diperbolehkan:
  + 1. Pandas
  + 2. NumpyMatplotlib
  + 3. Seaborn
  + 4. Scikit learn
  + 5. Nltk/sastrawi
  + 6. Re
  + 7. Joblib
  + 8. Streamlit
  + 9. Gradio
  + 10. Imblearn
  + 11. wordcloud
  + 12. pickle
  + 13. PyQt
* Dataset: Disediakan dalam format CSV
* Output yang diharapkan: Label sentimen (Positif, Negatif) dengan visualisasi statistik