

Proposal Probabilitas
Analisis Tingkat Pengangguran di Indonesia



Di susun oleh :
Muhammad Ma'rufil Kurhi
(241552010002)

Dosen pengampu : hendri karisma s.kom m,t

Fakultas ilmu komputer
Jurusan teknik informatika
Sekolah tinggi manajemen ilmu komputer tazkia
(stmik tazkia)

1. Latar Belakang

Pengangguran merupakan salah satu masalah ekonomi utama yang dihadapi banyak negara, termasuk Indonesia. Tingkat pengangguran yang tinggi dapat berdampak negatif pada pertumbuhan ekonomi, kesejahteraan sosial, dan stabilitas politik. Memahami tren dan faktor-faktor yang memengaruhi tingkat pengangguran sangat penting untuk merumuskan kebijakan yang efektif.

Proposal ini bertujuan untuk menganalisis data tingkat pengangguran di Indonesia dari tahun 2011 hingga 2022. Analisis ini akan menggunakan dua pendekatan utama: visualisasi data untuk memahami tren historis dan pemodelan regresi untuk memprediksi tingkat pengangguran.

2. Tujuan

Tujuan dari proposal ini adalah:

1. **Memvisualisasikan tren historis tingkat pengangguran** di Indonesia dari tahun 2011 hingga 2022.
2. **Menganalisis hubungan antara jumlah penduduk bekerja dan jumlah pengangguran** dengan tingkat pengangguran.
3. **Membangun model regresi** untuk memprediksi tingkat pengangguran.
4. **Mengevaluasi performa model** yang dibangun.

3. Metodologi

Metodologi yang digunakan dalam analisis ini terdiri dari beberapa langkah, sebagaimana yang sudah Anda terapkan dalam program Python yang Anda unggah.

3.1. Pengumpulan Data

Data yang digunakan berasal dari file '**data pengangguran.xlsx**' yang berisi:

- **Tanggal:** Periode data (bulanan/tahunan).
- **Jumlah Penduduk Bekerja:** Jumlah total penduduk yang memiliki pekerjaan.
- **Jumlah Pengangguran:** Jumlah total penduduk yang tidak memiliki pekerjaan tetapi sedang mencari pekerjaan.

3.2. Pra-pemrosesan Data

Data akan diolah menggunakan pustaka pandas di Python. Langkah-langkahnya meliputi:

1. **Menghitung Tingkat Pengangguran:** Tingkat pengangguran dihitung dengan formula:

$$\text{Tingkat Pengangguran(\%)} = \frac{\text{Jumlah Pengangguran}}{\text{Jumlah Penduduk Bekerja}} \times 100$$

2. **Konversi Tanggal:** Kolom Tanggal akan diubah menjadi format datetime untuk mempermudah analisis berdasarkan tahun.

3.3. Visualisasi Data

Visualisasi akan dilakukan menggunakan pustaka matplotlib. Sebuah bar chart akan dibuat untuk menunjukkan tingkat pengangguran per tahun, dilengkapi dengan garis tren untuk memudahkan interpretasi. Ini memungkinkan kita melihat pola atau kecenderungan penurunan atau kenaikan tingkat pengangguran dari tahun ke tahun.

3.4. Pemodelan Regresi

Pemodelan akan menggunakan **Stochastic Gradient Descent (SGD) Regressor** dari pustaka scikit-learn.

1. **Pembagian Data:** Data akan dibagi menjadi data latih (training) dan data uji (testing) dengan perbandingan **80:20**. Data latih digunakan untuk melatih model, sedangkan data uji digunakan untuk menguji performa model.
2. **Variabel:**
 - **Variabel Independen (X):** Jumlah Pengangguran dan Jumlah Penduduk Bekerja.
 - **Variabel Dependen (y):** Tingkat Pengangguran (%).
3. **Pelatihan Model:** Model SGDRegressor akan dilatih menggunakan data latih.
4. **Evaluasi Model:** Performa model akan dievaluasi menggunakan metrik berikut:
 - **Mean Squared Error (MSE):** Mengukur rata-rata kuadrat dari error, menunjukkan seberapa dekat prediksi model dengan nilai sebenarnya.
 - **R2 Score:** Mengukur seberapa baik model dapat menjelaskan variabilitas dalam variabel dependen. Nilai mendekati 1 menunjukkan model yang sangat baik.

4. Hasil yang Diharapkan

- **Grafik tren pengangguran** yang menunjukkan pergerakan tingkat pengangguran dari tahun 2011 hingga 2022.
- **Hasil evaluasi model** (nilai MSE dan R2) yang menunjukkan kemampuan model dalam memprediksi tingkat pengangguran.
- **Kesimpulan dan rekomendasi** berdasarkan hasil analisis, seperti apakah tingkat pengangguran cenderung menurun atau naik, dan apakah model yang dibuat dapat digunakan untuk prediksi di masa depan.