**LAPORAN TUGAS**

**ANALISIS DATA PREDIKSI CUACA MENGGUNAKAN POWER BI**

**Sebuah gambar berisi teks, lambang, simbol, lingkaran

Deskripsi dibuat secara otomatis**

**DISUSUN OLEH :**

**M. HILMI FAJRUL ALAM 210411100175**

**PRODI TEKNIK INFORMATIKA**

**JURUSAN TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS TRUNOJOYO MADURA**

**2023**

1. Analisis Data

Pada data prediksi cuaca ini terdapat lima fitur yaitu date (tanggal), precipitation (pengendapan), temp max (suhu maksimal), temp min (suhu minimal), wind (angin) dan satu label yaitu weather (cuaca). Proses selanjutnya adalah cleanning data dengan menyeleksi fitur fitur yang memang diperlukan dan yang tidak diperlukan. Setelah mencari literatur terkait bawasannya fitur date (tanggal) tidak berpengaruh terhadap hasil sehingga fitur ini dapat dihapus. Data prediksi cuaca ini diperoleh dari Kaggle berikut merupakan link datanya <https://www.kaggle.com/datasets/ananthr1/weather-prediction>

Sebuah gambar berisi teks, cuplikan layar, software, nomor

Deskripsi dibuat secara otomatis

1. Visualisasi Data

Berikut adalah visualisasi data fitur yang ada pada data prediksi cuaca

1. Precipitation

Precipitation atau presipitasi merupakan proses di mana uap air yang ada pada atmosfer berubah bentuk menjadi cair (air) ataupun padat (salju) dan jatuh ke permukaan bumi. Cuaca berkaitan erat dengan banyak dan jenis presipitasi yang jatuh.

Sebuah gambar berisi teks, cuplikan layar, garis, Font

Deskripsi dibuat secara otomatis

1. Temp max

Temperatur maksimal merupakan suhu maksimal yang diukur pada kurun waktu tertentu. Suhu maksimal merupakan salah satu faktor penentu cuaca. Kareana suhu maksimal dapat mengindikasikan cuaca panas atau dingin.

Sebuah gambar berisi teks, cuplikan layar, Plot, diagram

Deskripsi dibuat secara otomatis

1. Temp min

Temperatur minimal merupakan suhu minimal yang diukur pada kurun waktu tertentu. Sama halnya dengan suhu maksimal suhu minimal merupakan salah satu faktor penentu cuaca. Kareana suhu minimal dapat mengindikasikan cuaca panas atau dingin.

Sebuah gambar berisi teks, cuplikan layar, garis, Font

Deskripsi dibuat secara otomatis

1. Wind

Wind atau angin merupakan salah satu faktor penting untuk memprediksi cuaca. Karena angin dapat membawa perubahan masa udara yang bisa mempengaruhi suhu serta kelembaban. Selain itu angin juga dapat berpengaruh terhadap pembentukan awan, pemodelan pergerakan badai, dan prediksi arus udara atas.

Sebuah gambar berisi teks, cuplikan layar, Plot, diagram

Deskripsi dibuat secara otomatis

1. Missing Value

Pada dataset ini tidak terdapat missing value sehingga kitab isa langsung ketahap selanjutnya.

1. Normalisasi Data

Normalisasi data adalah proses untuk memastikan bahwa beberapa variabel memiliki rentang nilai yang sama, tanpa ada variabel yang terlalu besar atau terlalu kecil untuk memudahkan analisis statistik. Terdapat beberapa teknik normalisasi data tetapi pada data ini digunakak teknik min-max scaling yang mengubah nilai dalam data dengan rentang angka 0 sampai dengan 1 menggunakan rumus yang mana x merupakan nilai asli, min merupakan nilai minimum dalam data dan max merupakan nilai maksimum dalam data. Dan ini merupakan data yang sudah dinormalisasi

Sebuah gambar berisi teks, cuplikan layar, nomor, Tidak bisa

Deskripsi dibuat secara otomatis

1. Balancing data

Balancing data adalah proses menciptakan keseimbangan atau pemerataan antara berbagai kategori atau kelas kumpulan data. Ketidakseimbangan sering terjadi ketika satu jenis atau kelas memiliki jumlah sampel yang jauh lebih besar dibandingkan jenis atau kelas lainnya. Ketidakseimbangan ini dapat membuat model pembelajaran mesin kurang mampu mengidentifikasi kelas minoritas. Untuk menyeimbangkan data pada dataset ini digunakan Python dengan teknik RandomOverSampler.

Sebuah gambar berisi teks, cuplikan layar, Font, nomor

Deskripsi dibuat secara otomatis

Gambar diatas adalah gambar data yang belum diseimbangkan. Dapat dilihat pada gambar tesebut bahwa jumlah data tidak seimbang pada setiap kategori sehingga perlu diseimbangkan.

Sebuah gambar berisi teks, cuplikan layar, Font, nomor

Deskripsi dibuat secara otomatisSebuah gambar berisi teks, cuplikan layar, nomor, Font

Deskripsi dibuat secara otomatis

Sebuah gambar berisi teks, cuplikan layar, Font, software

Deskripsi dibuat secara otomatis

Setelah dilakukan proses balancing data dapat dilihat bahwa jumlah data tiap kategori yang awalnya tidak seimbang menjadi seimbang dengan jumlah 641 tiap masing masing kategori.