

1 Pendahuluan

1.1 Latar Belakang

Polusi udara adalah pencemaran yang terjadi sebagai akibat interaksi kompleks antara lingkungan alam dan lingkungan antropogenik. Berbagai bahan kimia yang lepas ke atmosfer menyebabkan pencampuran udara dengan bahan kimia tersebut. Partikel yang bercampur dengan udara dapat mencapai tingkat berbahaya baik di dalam ruangan atau di luar ruangan. Polutan penyebab polusi udara dapat berasal dari asap, metana, karbon dioksida, jamur, dan lainnya. Polutan dapat berasal dari aktivitas manusia seperti asap yang dihasilkan dari kendaraan, kegiatan pabrik, penggunaan AC dan lain-lain. Negara-negara berkembang dan berpenduduk padat melihat polusi udara telah menjadi salah satu masalah serius. Efek negatif bagi kesehatan yang diakibatkan oleh polusi udara menjadi intensif yang diperhatikan dengan serius. Terpapar polusi udara dapat meningkatkan risiko penyakit [1]. Menurut WHO, kematian dini akibat polusi udara dapat mencapai 7 juta per tahun. Tingkat polusi udara dapat diketahui melalui kualitas udara dimana dipengaruhi oleh beberapa unsur seperti PM, SO₂, NO, NO₂, O₃ [2].

Menyadari betapa serius masalah polusi udara tersebut, mendorong penelitian [3] menggunakan metode Naïve Bayes. Naïve Bayes merupakan metode pengklasifikasian peluang sederhana dengan asumsi antar variabel penjelas saling bebas (independen). Dari hasil penelitian tersebut, mendapatkan hasil akurasi sebesar 97,08%. Sedangkan pada suatu penelitian [4] yang melakukan deteksi jantung dengan metode yang sama, mendapatkan hasil klasifikasi dengan akurasi sebesar 74%. Namun, penelitian ini tidak membahas mengenai studi polusi udara dan tidak melakukan training data dalam ukuran yang jauh lebih besar. Dalam melakukan pemetaan peta prediksi, penelitian mengenai persebaran polusi udara [5], memprediksi nilai polutan. Namun, pada penelitian tersebut tidak mengembangkan peta distribusi persebaran polusi udara. Berdasarkan penelitian di atas, dengan mempertimbangkan kelebihan dan kekurangannya, dalam penelitian ini, menggunakan metode Naïve Bayes dalam pembuatan peta prediksi klasifikasi sebaran polusi udara. Dengan adanya penelitian ini, diharapkan dapat menghasilkan peta prediksi yang akurat sebagai bentuk kontribusi peneliti terhadap penanganan atauantisipasi polusi udara.

1.2 Batasan Masalah

Dalam penelitian ini, batasan masalah yang dihadapi yaitu fokus wilayah yang digunakan adalah kota Jakarta. Pemilihan Jakarta sebagai fokus wilayah untuk diteliti karena merupakan kota dengan kepadatan penduduk tertinggi beserta aktivitas kegiatan yang dilakukan. Faktor tersebut menjadi salah satu sebab penelitian dilakukan. Data yang digunakan adalah data indeks kualitas udara ambien berdasarkan Indeks Standar Pencemaran Udara (ISPU) kota Jakarta dari tahun 2016-2021 dengan jumlah data sebanyak lebih dari 5000 baris data. Selain itu, digunakan data lain seperti data iklim, jumlah kendaraan, jumlah pohon dan kepadatan penduduk untuk membentuk model prediksi klasifikasi pada tahun 2022. Berdasarkan model yang telah dibentuk, sistem yang dibangun adalah membentuk peta prediksi polusi udara dengan pemetaan kategori indeks polusi. Sistem menampilkan bentuk peta Jakarta dengan pembagian berdasarkan kelurahan dan ditandai dengan warna untuk setiap kategori polusi udara yang berbeda. Namun, sistem hanya dapat menampilkan kategori polusi udara untuk satu kelurahan dalam setiap satu kali prediksi dilakukan. Karena kebutuhan komputasi dan manajemen memori computer yang terbatas, sehingga tidak dapat dilakukan pemetaan prediksi secara serentak dengan menggunakan semua kelurahan yang ada.

1.3 Tujuan

Tujuan dari penelitian ini adalah membangun sistem peta prediksi klasifikasi polusi udara dengan menerapkan metode *Naïve Bayes* berbasis website. Dengan menggunakan metode tersebut, ingin diketahui bagaimana bentuk implementasi untuk membangun model klasifikasi dan prediksi berdasarkan data yang digunakan sekaligus mengetahui bagaimana performansi *Naïve Bayes* pada pengembangan peta sebaran polusi udara.

1.4 Organisasi Tulisan

Susunan penelitian yang dilakukan dibangun berdasarkan ketentuan organisasi tulisan yaitu pendahuluan, studi terkait, sistem yang dibangun, evaluasi dan kesimpulan. Pendahuluan merupakan bagian dimana dijelaskan latar belakang yang ada, masalah-masalah yang akan dihadapi beserta batasan dan tujuannya. Studi Terkait menjelaskan mengenai studi yang dilakukan berdasarkan sumber seperti jurnal dan sebagainya sebagai landasan pemilihan metode, membangun solusi dan pembentukan sistem yang dibangun. Pada bagian berikutnya, merupakan kerangka pembentukan sistem berdasarkan literasi pada bagian

sebelumnya. Evaluasi merupakan tahapan atau scenario pengujian untuk mendapatkan hasil yang sesuai dengan tujuan. Terakhir adalah kesimpulan, yaitu rangkuman dari semua skenario pengujian yang dilakukan dengan menggunakan sistem sebagai hasil akhir dan sebagai pengembangan untuk penelitian selanjutnya.