

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Seiring meningkatnya perkembangan industri, ekonomi juga urbanisasi di Indonesia sejak abad 21 (Syamil dan Nusantara, 2023) (Associate, 2021), masalah polusi udara menjadi semakin serius yang berdampak pada lingkungan ekologis dan menarik perhatian dunia (Wang dan Shen, 2016). Dampak buruk pada makhluk hidup menjadi suatu masalah akibat adanya pencemaran udara (Satra dan Rachman, 2016). Dampak tersebut dapat mempengaruhi penurunan kualitas udara, sering terjadi di daerah kota besar di Indonesia. Adapun faktor-faktor yang mengakibatkan udara tercemar yaitu adanya aktivitas transportasi, kegiatan industri, dan limbah yang di buang ke udara (Cahyono, 2016). Pencemaran udara yang berlebihan dapat berpotensi negatif untuk Kesehatan hingga mengakibatkan kematian (Iqbal dan Hermanto, 2017). Selain itu, faktor alam juga dapat mencemari kualitas udara antara lain aktifitas dari gunung vulkanik. Sedangkan faktor manusia yang dapat mencemari kualitas udara antara lain adanya aktifitas kendaraan/transportasi di jalan raya dan aktifitas industri yang berkembang sejalan dengan pertumbuhan penduduk (Agista, dkk, 2020). Dan karbon monoksida yang cukup tinggi, sehingga menjadi faktor penyebab gas rumah kaca yang dapat mempengaruhi kualitas udara (Handayani, 2020).

Telah dilakukan penelitian menggunakan metode prediksi kualitas udara sehingga dianggap dapat memberikan kontribusi yang bermanfaat bagi para pakar lingkungan dan kesehatan untuk mengidentifikasi potensi risiko kesehatan yang disebabkan oleh polusi udara. *Machine Learning* merupakan salah satu teknologi informasi yang dapat mendukung langkah-langkah klasifikasi tersebut. (Gladkova dan Saychenko, 2022; Li, dkk 2023). *Machine learning* adalah cabang *Artificial Intelligence* yang didasari oleh algoritma dan model statistik pada sistem komputer yang mampu belajar secara otomatis, berkat teknologi ini dapat membuat klasifikasi, prediksi, atau keputusan secara otomatis untuk melakukan tugas tersebut (Su , dkk, 2022). Mengatur dan menyelesaikan masalah polusi udara

merupakan proses jangka panjang. Prediksi kualitas udara bisa membantu pencegahan kerusakan udara akibat pencemaran. Oleh sebab itu, perlu adanya klasifikasi kualitas udara dalam peramalan yang akan terjadi pada waktu yang tepat, agar departemen pemerintah dan masyarakat dapat mengambil tindakan perlindungan dan mencegah insiden polusi yang serius. Seperti contoh, beberapa pabrik di Indonesia telah ditutup sementara karena penyebab polusi udara (Deny, 2023)

Monitoring kualitas udara yang berkelanjutan diperlukan untuk pemantauan yang efektif dan pengambilan langkah-langkah mitigasi yang tepat untuk menanggulangi pencemaran udara yang berlebihan. Untuk itu dibutuhkan klasifikasi udara yang akurat dan efisien. Algoritma *machine learning* (ML) dapat membantu dalam memproses prediksi kualitas udara secara real-time, mengidentifikasi pola polusi udara, dan memberikan informasi yang akurat untuk pengambilan keputusan yang cepat dan responsif. Penelitian prediksi kualitas udara menggunakan *machine learning* dengan memanfaatkan data historis dan algoritme pembelajaran mesin memungkinkan prediksi akurat mengenai tingkat kualitas udara perkotaan di masa depan dalam skala global. Oleh karena itu, algoritma pembelajaran mesin dapat secara efektif secara efektif memperkirakan tingkat AQI (Ravindiran, dkk, 2023)

Algoritma *Random forest* (RF) dan K-Nearest Neighbors (KNN) digunakan untuk memprediksi kinerja termal pemanas udara surya (SAH) dengan panel penyerap bersirip bentuk C. dengan hasil RF mendapatkan akurasi yang lebih tinggi dan cocok untuk memprediksi hal tersebut (Saravanan, dkk, 2023)

Penerapan *ensemble learning* telah diterapkan untuk membangun sistem klasifikasi topik pada tweet menggunakan ekspansi fitur metode Word2Vec dan *Random forest*. Pada ekspansi fitur diterapkan menggunakan 3 corpus Word2Vec (tweet, berita dan gabungan antara tweet dan berita) Serta, terdapat tiga variasi ekspansi fitur, yakni top 1, top 5, dan top 10, yang diterapkan dalam mencari model terbaik, sehingga mendapatkan model terbaik menggunakan fitur top 5 dengan nilai akurasi 99,49% dan nilai F1-Score yang juga meningkat pada model fitur top 5 sebesar 0,9949 menggunakan kamus kata gabungan dari data berita dan tweet. Pada penelitian ini, model ekspansi fitur ini berhasil meningkatkan nilai akurasi untuk

metode klasifikasi *Random forest* yang sebelumnya hanya mendapat nilai akurasi 98.44% dan nilai F1-Score sebesar 0.9842 (Ramli dan Sibaroni, 2022)

Metode *ensemble* juga telah digunakan pada penelitian (Nurmasani dan Pristyanto, 2021), (Ngo, dkk, 2022) yang hasil nya menunjukkan bahwa ensemble dapat digunakan sebagai solusi untuk kasus prediksi. Berdasarkan pengujian yang dilakukan menunjukkan bahwa ensemble bisa mengoptimalkan akurasi dibandingkan dengan prediksi lain nya tunggal lainnya.

Perkiraan kualitas udara otomatis menggunakan AQP-EDLMRA teknik telah diperkenalkan dalam penelitian ini. Yang disajikan Teknik AQP-EDLMRA menggunakan model *ensemble voting* dengan menggunakan tiga metode *deep learning* (DL) yaitu long short term memory(LSTM), belief network (DBN) Simple Additive Weighting (SAW). Dalam model pemungutan suara ansambel, data baru bisa diklasifikasikan berdasarkan suara tertimbang dari prediksi mereka hasil. Untuk menyesuaikan nilai *hyperparameter* DL metode, MRA dieksploitasi dalam pekerjaan ini. Itu evaluasi eksperimental pendekatan AQP-EDLMRA adalah diuji menggunakan serangkaian data kualitas udara dan hasil perbandingan yang komprehensif menunjukkan bahwa Teknik AQP-EDLMRA dengan menggunakan *ensemble voting* telah mencapai peramalan yang lebih baik. (Sivanesh, dkk, 2023)

Dari rangkuman penelitian yang telah dijelaskan sebelumnya, dapat dikatakan bahwa penggunaan *ensemble learning* cenderung memberikan hasil prediksi yang lebih baik dibandingkan dengan menggunakan *single* algoritma. Penelitian ini mencakup serangkaian proses prediksi dengan menggunakan metode *ensemble learning* dengan menggabungkan algoritma *K-neares neighbor*, *Random forest* dan *Support Vector Machine*. Dengan adanya penggabungan algoritma ini diharapkan menghasilkan model terbaik untuk lebih optimal pada proses prediksi, pengambilan keputusan dalam mempertahankan kualitas udara yang jauh lebih efektif dan lebih efisien.

## 1.2 Tujuan Penelitian

1. Menggabungkan algoritma *Machine learning* menggunakan metode *ensemble learning* yang digunakan untuk prediksi kualitas udara agar dapat memberikan hasil prediksi yang lebih baik.
2. Mengetahui tingkat error menggunakan matrix MSE dan RMSE hasil prediksi kualitas udara dengan menggabungkan *k-Nearest Neighbor (KNN)*, *Support Vector Machine (SVM)* dan *Random forest (RF)* yang dioptimalkan.

## 1.3 Manfaat penelitian

Untuk pendidikan :

1. Menambah pengetahuan dan wawasan terkait penggunaan algoritma *machine learning* khususnya *ensemble learning* dalam memecahkan masalah yang kompleks, seperti prediksi kualitas udara.
2. Penelitian ini dapat dijadikan sebagai sumber pembelajaran yang berharga bagi mahasiswa atau pelajar yang ingin mempelajari algoritma *machine learning* dan penggunaannya dalam prediksi kualitas udara.
3. Memberikan kesempatan bagi mahasiswa atau peneliti muda untuk mengembangkan penelitian lanjutan terkait prediksi kualitas udara dengan memanfaatkan algoritma *machine learning*.

Untuk Masyarakat :

- 1 Dengan menggunakan teknologi *machine learning*, dapat dilakukan prediksi kualitas udara yang lebih akurat dan detail. Hal ini memungkinkan masyarakat untuk mengetahui kualitas udara di sekitar mereka dengan lebih tepat dan lebih cepat. Informasi ini dapat membantu masyarakat untuk mengambil tindakan yang tepat untuk menjaga kesehatan mereka dan lingkungan sekitar, dengan demikian masyarakat dapat mengetahui kualitas udara lebih akurat.
- 2 Dengan memiliki informasi yang lebih akurat tentang kualitas udara di sekitar mereka, masyarakat dapat lebih sadar akan pentingnya menjaga kualitas udara yang baik untuk kesehatan mereka dan lingkungan sekitar. Hal ini dapat mendorong masyarakat untuk mengambil tindakan untuk mengurangi polusi udara dan mempromosikan kebijakan yang lebih ramah lingkungan. Sehingga dapat meningkatkan kesadaran masyarakat akan kualitas udara.