

BAB I

USULAN GAGASAN

1.1 Deskripsi Umum Masalah

Permasalahan utama yang dihadapi Indonesia salah satunya adalah polusi lingkungan berupa polusi udara, terutama di daerah kota besar. Peningkatan jumlah penduduk menyebabkan peningkatan pada proses pembakaran pada bidang industri. Selain di bidang industri, meningkatnya penduduk Indonesia juga menyebabkan peningkatan penggunaan bahan bakar untuk kebutuhan transportasi masyarakat. Bidang industri dan transportasi menjadi salah satu ancaman permasalahan lingkungan di Indonesia, karena dapat menghasilkan gas dan polutan yang mencemari udara [1]. Paparan polutan udara jangka panjang dapat menyebabkan berbagai masalah kesehatan yang menyerang saluran pernapasan. Selain merugikan kesehatan, polusi udara juga memberi dampak negatif pada lingkungan dan iklim. Gas dan polutan berbahaya bagi lingkungan salah satunya adalah karbon dioksida (CO_2) dan partikulat tersuspensi di udara ($\text{PM}_{2.5}$). Berdasarkan Peraturan Menteri Kesehatan Nomor 2 tahun 2023, $\text{PM}_{2.5}$ memiliki batas konsentrasi yang lebih ketat dibandingkan polutan lainnya pada periode pengukuran 24 jam. Data baku mutu menunjukkan bahwa $\text{PM}_{2.5}$ lebih berbahaya karena partikel ini dapat menyebabkan berbagai masalah kesehatan yang serius. Sementara itu, CO_2 dianggap polutan berbahaya karena peningkatan konsentrasinya di atmosfer mempercepat efek rumah kaca [2].

$\text{PM}_{2.5}$ merupakan suatu partikulat di udara yang berdiameter kurang dari $2,5 \mu\text{m}$ yang berasal dari emisi langsung proses pembakaran. $\text{PM}_{2.5}$ juga merupakan pembentukan sekunder dari reaksi kimia atmosfer, sehingga umumnya mengandung lebih banyak spesies organik daripada PM_{10} [3]. Indikator lain dari polutan berbahaya adalah karbon dioksida (CO_2). CO_2 masuk ke atmosfer melalui berbagai proses, termasuk pembakaran bahan bakar fosil seperti batu bara, gas alam, dan minyak, serta limbah padat. CO_2 secara alami berada di atmosfer sebagai bagian dari siklus karbon bumi, yang melibatkan peredaran alami CO_2 di atmosfer, lautan, tanah, tumbuhan, dan hewan [4].

Pemantauan gas polutan CO_2 dan $\text{PM}_{2.5}$ diperlukan untuk memberikan informasi kualitas udara. Sistem pemantauan kualitas udara menjadi upaya dalam memantau pencemaran udara di lingkungan dengan menggunakan sensor SKU:SEN0219 untuk CO_2 dan sensor SKU:SEN0460 untuk $\text{PM}_{2.5}$. Pemantauan telah dilakukan di beberapa stasiun yang dilengkapi dengan sensor untuk mengukur tingkat $\text{PM}_{2.5}$ dan CO_2 . Stasiun tersebut terletak di TULT, GKU, dan DELI. Pemantauan ini dilakukan agar masyarakat dapat melakukan tindakan

pencegahan. Analisis sistem pemantauan kualitas udara menjadi suatu tantangan yang harus dihadapi. Data yang dikirim oleh hardware sering kali mengalami kendala, seperti tidak masuk ke dalam *database* atau terjadi keterlambatan dalam pengiriman data. Berikut hasil pengiriman data *raw* ke *database* pada periode September - Oktober 2024.

Tabel 1.1 Data *raw* periode September-Oktober 2024 pada tiga stasiun pengukuran.

Stasiun/Kategori Data	Deli	GKU	TULT
Total data <i>raw</i> seharusnya	43920	43920	43920
Total data <i>raw</i> terkirim	39055	40195	32499
Persentase ketersediaan data	88.9%	91.5%	74%

Tabel 1.1 menunjukkan ketersediaan data *raw* yang didapat pada *database*. Persentase ketersediaan data yang paling rendah adalah stasiun TULT. Kondisi ini dapat disebabkan oleh gangguan dari alat ukur atau masalah transmisi. Jika data dari *hardware* tidak terkirim ke dalam *database*, hal ini akan berdampak negatif terhadap performa dan kualitas data. Data yang tidak lengkap dapat mempengaruhi analisis dan validasi data yang dibutuhkan. Berdasarkan analisis tersebut, perlu adanya pengembangan pada sistem agar dapat mendeteksi terjadinya data hilang.

Selain itu, untuk mendukung pemantauan kualitas udara, telah dikembangkan sistem prediksi menggunakan algoritma *Long Short-Term Memory* (LSTM). LSTM adalah salah satu jenis arsitektur *Recurrent Neural Network* (RNN) yang sering digunakan dalam masalah-masalah yang berkaitan dengan *deep learning*. Kinerja sistem prediksi ini dievaluasi menggunakan metrik *Root Mean Square Error* (RMSE) pada tiga stasiun pemantauan udara yang berbeda, yaitu GKU, TULT, dan DELI. Hasil evaluasi menunjukkan bahwa kinerja model dalam memprediksi parameter CO₂ dan PM_{2.5} bervariasi di setiap stasiun pengukuran. Beberapa stasiun menunjukkan tingkat kesalahan prediksi yang lebih tinggi untuk kedua parameter, sementara stasiun lainnya memiliki kesalahan prediksi yang lebih rendah, terutama untuk parameter PM_{2.5}. Variasi ini menunjukkan bahwa model tersebut masih memerlukan evaluasi dan penyempurnaan lebih lanjut untuk meningkatkan akurasi prediksi. Untuk mencapai tujuan ini, analisis lebih lanjut terhadap model prediksi diperlukan, serta peningkatan melalui optimasi model atau pendekatan kombinasi atau *hybrid* LSTM dengan model lain yang dapat meningkatkan keakuratan hasil dan performansi sistem prediksi.

Di sisi lain, tindakan pencegahan penting dilakukan oleh masyarakat agar dapat mengurangi paparan polusi. Dalam menghadapi polusi udara, sistem peringatan dini memegang peranan penting dalam memberikan informasi yang tepat mengenai kondisi udara. Sistem peringatan dini ini berfokus pada pemberitahuan tingkat keparahan CO₂ dan PM_{2.5} yang tersebar di udara. Apabila CO₂ dan PM_{2.5} terdeteksi melebihi baku mutu yang telah ditetapkan, sistem akan memberikan notifikasi kepada masyarakat agar dapat mengambil langkah perlindungan, seperti menggunakan masker saat keluar rumah. Notifikasi tersebut akan disampaikan melalui media sosial, sehingga informasi dapat tersebar luas dan cepat diterima oleh masyarakat. Dengan adanya pemberitahuan yang cepat dan tepat, masyarakat dapat segera mengambil tindakan preventif untuk melindungi diri dari dampak buruk polusi udara [5].

1.2 Analisis Masalah

1.2.1 Aspek Kesehatan

Kualitas udara yang semakin memburuk membawa dampak signifikan bagi kesehatan manusia, baik dalam jangka pendek maupun jangka panjang. Dalam jangka pendek, paparan polusi udara dapat menyebabkan iritasi pada mata dan tenggorokan. Sedangkan dalam jangka panjang, polusi udara dapat mengakibatkan berbagai penyakit serius seperti gangguan fungsi paru, penyakit pernapasan, kanker, kardiovaskular, aterosklerosis, dan bahkan kematian dini [6]. Dilansir dari *World Health Organization* (WHO) pada penelitian tahun 2019, polusi udara luar ruangan, baik di kota maupun di desa, diperkirakan mengakibatkan 4,2 juta kematian dini secara global setiap tahun. Kematian ini terjadi karena paparan partikel halus yang dapat memicu penyakit kardiovaskular, menyerang saluran pernapasan serta kanker. Sekitar 37% kematian dini akibat polusi udara luar ruangan disebabkan oleh penyakit jantung iskemik dan stroke, 18% karena penyakit paru obstruktif kronis, 23% karena infeksi saluran pernapasan akut bagian bawah, dan 11% terkait dengan kanker pada saluran pernapasan [7]. Dari data tersebut, sistem pemantauan kualitas udara menjadi sangat penting dalam memberikan informasi kepada masyarakat mengenai tingkat polusi udara, sehingga langkah pencegahan dan penanggulangan dapat diambil untuk mengurangi dampak negatif yang ditimbulkan bagi kesehatan.

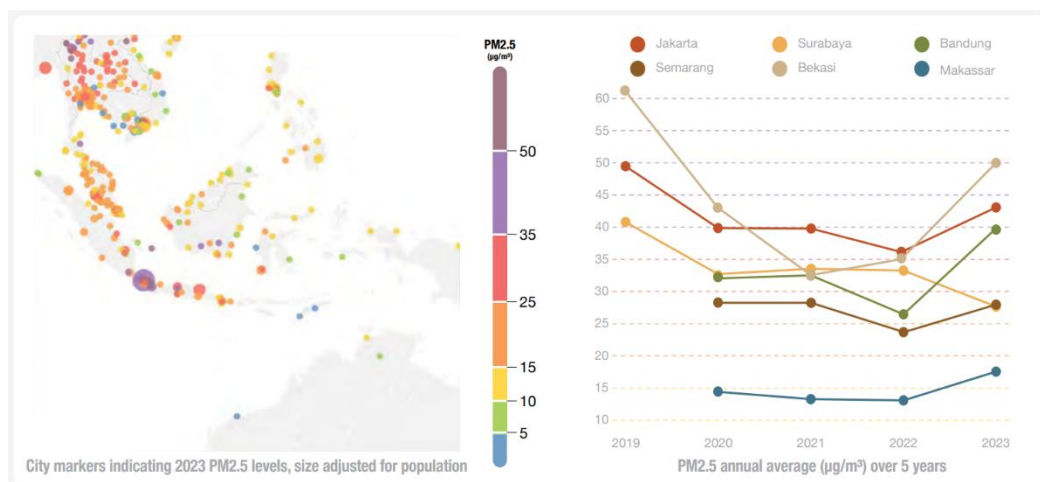
1.2.2 Aspek Keberlanjutan

Sistem pemantauan kualitas udara akan terus berkembang karena pemantauan dilakukan dalam jangka panjang. Sistem terintegrasi dengan berbagai parameter kualitas udara, sehingga memungkinkan untuk melakukan pemantauan dari berbagai stasiun

pengukuran secara bersamaan. Sensor yang bekerja dalam sistem pemantauan kualitas udara berfokus pada sensor CO₂ dan PM_{2.5}. Data yang dihasilkan pada sensor akan tersimpan pada *database* untuk dilakukan pengolahan. Data yang telah diolah akan ditampilkan pada *website* dan media sosial yang memberikan visualisasi data secara representatif yang dapat diakses oleh publik. Sistem memungkinkan pengumpulan data secara terus menerus mengenai kondisi kualitas udara, sehingga pengembangan sistem akan terus dilakukan secara berkelanjutan sebagai upaya pemantauan bagi lingkungan.

1.2.3 Aspek Lingkungan

Polusi udara dapat menimbulkan masalah serius bagi lingkungan. Masalah yang ditimbulkan seperti pemanasan global, penipisan lapisan ozon, dan sebagainya. Pemanasan global terjadi akibat kadar CO₂ yang tinggi di lapisan atmosfer. Kadar CO₂ yang tinggi dapat menghalangi pantulan panas dari bumi ke atmosfer sehingga permukaan bumi menjadi lebih panas. Selain gas CO₂, PM_{2.5} juga dapat menimbulkan masalah bagi lingkungan. PM_{2.5} yang dihasilkan dari proses pembakaran dapat menurunkan kualitas udara dan menyebabkan terbentuknya kabut tebal. Kondisi ini dapat memperburuk polusi, terutama dalam kondisi cuaca yang stagnan [8].



Gambar 1.1 Grafik PM_{2.5} di Indonesia (2019-2023)

Dilansir dari *World Air Quality Report 2023* IQAir konsentrasi PM_{2.5} rata-rata di Indonesia mengalami peningkatan 20% pada tahun 2023 menjadi 37.1 µg/m³ [9]. Pada gambar grafik di atas, perbandingan konsentrasi dilakukan pada 6 kota berbeda di Indonesia, antara lain Jakarta, Semarang, Surabaya, Bekasi, Bandung, dan Makassar. Dapat dilihat pada grafik bahwa Kota Bandung menempati peringkat ketiga kota di Indonesia dengan tingkat PM_{2.5} tertinggi di Indonesia.

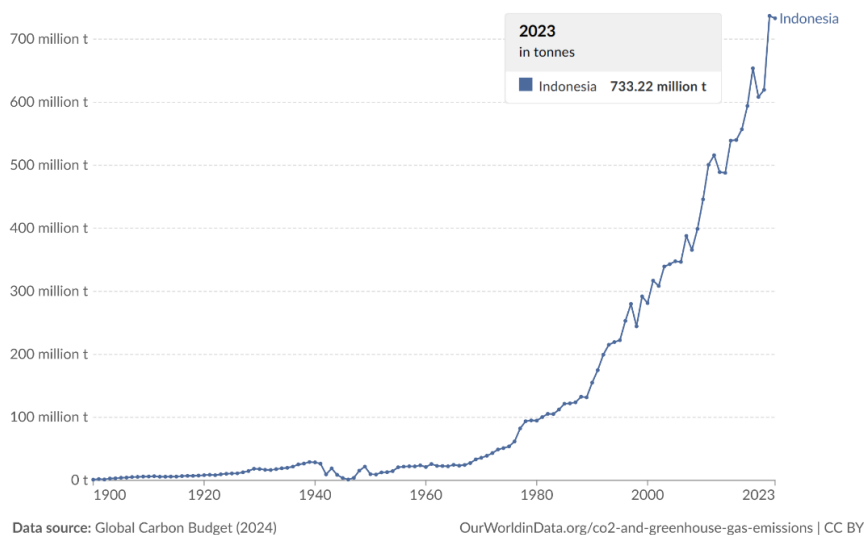
City	2023	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	2022
Jakarta	43.8	20.6	19.2	31.2	28.4	49.8	48.7	55.3	58.3	53.0	61.1	53.0	44.3	36.2
Surabaya	27.6	25.3	18.4	24.1	16.4	19.9	21.2	31.3	31.2	26.2	36.9	39.8	39.7	34.4
Bandung	39.6	23.0	17.8	31.2	30.2	43.3	43.1	37.3	52.7	42.5	53.1	49.8	49.1	26.1
Semarang	28.1	20.8	18.6	23.4	24.1	31.6	24.7	20.0	37.2	31.4	38.2	35.4	30.4	24.3
Bogor	49.9	34.1	22.9	39.6	38.0	71.5	76.9	76.6	60.4	51.4	46.8	47.2	36.6	35.4
Makassar	17.4	10.4	9.5	11.3	11.7	18.4	17.7	20.5	22.6	27.1	26.8	24.8	11.9	13.2

Gambar 1.2 Data PM_{2.5} untuk 6 kota besar di Indonesia (2022-2023)

Dari perbandingan konsentrasi PM_{2.5} di atas, Kota Bandung mengalami peningkatan konsentrasi rata-rata PM_{2.5} yang cukup tinggi dari tahun 2022 hingga 2023. Konsentrasi PM_{2.5} di kota Bandung bernilai 26.1 µg/m³ pada tahun 2022 mengalami peningkatan menjadi 39.6 µg/m³ pada tahun 2023. Jika dilihat dari data konsentrasi rata-rata bulanan tahun 2023, bulan Oktober memiliki rata-rata konsentrasi PM_{2.5} tertinggi dibandingkan bulan lainnya. Hal ini disebabkan oleh pergeseran awal musim hujan sehingga konsentrasi PM_{2.5} terdeteksi sangat tinggi dibandingkan dari tahun sebelumnya pada bulan yang sama.

Annual CO₂ emissions

Carbon dioxide (CO₂) emissions from fossil fuels and industry¹. Land-use change is not included.



1. Fossil emissions: Fossil emissions measure the quantity of carbon dioxide (CO₂) emitted from the burning of fossil fuels, and directly from industrial processes such as cement and steel production. Fossil CO₂ includes emissions from coal, oil, gas, flaring, cement, steel, and other industrial processes. Fossil emissions do not include land use change, deforestation, soils, or vegetation.

Gambar 1.3 Grafik CO₂ di Indonesia (1900-2023)

Sumber polusi udara di Indonesia banyak berasal dari pertambangan batu bara, kebakaran hutan, dan pembukaan lahan dengan penebangan dan pembakaran di Sumatera dan Kalimantan untuk pengembangan pertanian. Biasanya, polusi di kota-kota besar

banyak diakibatkan oleh emisi gas buang CO₂ dari sistem transportasi. Dapat dilihat pada grafik di atas yang dilansir dari *Global Carbon Budget* tahun 2024, emisi gas CO₂ di Indonesia mencapai 733,22 juta ton pada tahun 2023. Meskipun jumlah ini lebih sedikit dibandingkan pada tahun 2022 yang mencapai 737,07 juta ton, namun emisi gas CO₂ di Indonesia cukup tinggi dibandingkan dengan beberapa negara lain [10].

1.3 Analisis Solusi yang Ada

Solusi untuk sistem pemantauan kualitas udara telah memiliki beberapa pengembangan, seperti adanya *website* yang menampilkan informasi kualitas udara. Saat ini, terdapat dua pengembangan solusi, yaitu bidang *hardware* dan *software*. Berfokus pada bidang *software*, solusi yang dikembangkan telah memberikan informasi yang tepat sesuai dengan kebutuhan masyarakat. Namun, performansi sistem masih memerlukan analisis dan pengembangan lebih lanjut untuk memperbaiki stabilitas sistem dalam menjalankan program yang telah dikembangkan.

1.3.1 Validasi Data

Validasi data merupakan sistem yang sudah ada saat ini dan telah dikembangkan oleh peneliti sebelumnya. Sistem validasi data saat ini menggunakan metode *IT Automation* untuk memvalidasi data secara otomatis, memastikan bahwa data yang dikumpulkan, disimpan, atau diproses sesuai dengan standar dan aturan yang berlaku. Validasi data bertujuan untuk memastikan integritas, akurasi, konsistensi, dan keandalan data. Setiap jenis sensor membutuhkan metode validasi yang berbeda bergantung pada karakteristik sensor tersebut. Tujuan validasi data adalah untuk mengurangi kesalahan yang sering terjadi selama proses pengiriman data dari sensor. Proses validasi data dapat mencakup berbagai langkah seperti menghilangkan data yang dianggap sebagai *outlier*, memeriksa kelengkapan data, menghapus nilai NaN, dan menghapus data yang berada di luar batas yang telah ditentukan oleh sensor.

Metode validasi data yang telah dikembangkan saat ini telah berjalan secara otomatis, sehingga dapat mengurangi adanya intervensi manual. Namun, hasil data validasi yang telah dirata-ratakan terkadang tidak mencapai ketersediaan 75 % sehingga menyebabkan menurunnya kualitas data. Hal ini disebabkan oleh adanya data hilang pada data *raw*.

1.3.2 Prediksi Konsentrasi PM_{2.5} dan CO₂

Prediksi data merupakan langkah membuat proses perkiraan atau hasil yang mungkin terjadi di masa yang akan datang dengan mengamati pola yang ada pada data masa lalu. Prediksi data yang dilakukan saat ini berfokus pada prediksi konsentrasi PM_{2.5} dan CO₂.

Prediksi data dapat dilakukan menggunakan *machine learning*, terutama ketika data tersebut terkait dengan deret waktu. *Machine learning* memerlukan suatu model untuk menjadi landasan prediksi data yang akan digunakan di masa mendatang. Model ini memungkinkan kita untuk menghasilkan *output* yang tepat berdasarkan data yang ada saat ini. Singkatnya, model ini dibentuk melalui pengumpulan data, analisis untuk memahami pola dan tren, lalu dilatih, sehingga menjadi sebuah model. Ada banyak jenis model prediktif yang dapat dibuat dengan menggunakan berbagai algoritma, seperti LSTM, *Naïve Bayes*, dan *Decision Tree*. Model prediktif yang dipilih pada sistem prediksi yang sudah ada saat ini adalah LSTM.

LSTM adalah salah satu algoritma yang sangat populer untuk memprediksi dan mengklasifikasi data deret waktu. Setelah membangun model prediksi menggunakan metode LSTM, penting untuk menguji dan mengevaluasi kinerja model guna memastikan keandalan dan efisiensi sistem yang telah dikembangkan. Evaluasi ini melibatkan penggunaan *quality metrics*, yaitu metrik atau indikator yang digunakan untuk menilai seberapa baik model dapat memprediksi data. *Quality metrics* yang umumnya digunakan mencakup MAE (*Mean Absolute Error*), MSE (*Mean Squared Error*), RMSE (*Root Mean Squared Error*), dan R^2 Score (Koefisien Determinasi). Penggunaan metrik-metrik ini membantu dalam menilai tingkat ketepatan dan akurasi prediksi model terhadap parameter yang diukur. Hasil nilai *quality metrics* dapat dilihat pada **Tabel 1.2**.

Tabel 1.2 *Quality Metrics* Sistem Prediksi

Stasiun/Quality Metrics	GKU		TULT		DELI	
	PM _{2.5}	CO ₂	PM _{2.5}	CO ₂	PM _{2.5}	CO ₂
R ²	-0,342	-0.425	0,58	-2.36	0.0984	-0.706
MSE	128279,85	39242.58	38,125	16936.21	1120	5525.54
MRSE	258,161	198.09	6,174	130.13	33.566	74.33
MAE	298,714	169.24	18	94.97	24.5	42.32

Berdasarkan evaluasi model, terlihat bahwa model prediksi TULT (PM_{2.5}) memiliki performa terbaik dengan R² sebesar 0,58 dan kesalahan yang rendah (MSE: 38,125, RMSE: 6,174, MAE: 18). Namun, model untuk GKU (PM_{2.5}), TULT (CO₂), dan DELI (PM_{2.5}) menunjukkan performa yang buruk dengan R² yang sangat rendah atau negatif, serta kesalahan yang besar. Adanya nilai *R-square* (R²) negatif mengindikasikan bahwa model yang dievaluasi memiliki kinerja yang lebih buruk dibandingkan dengan model

dasar [11]. R^2 bisa menjadi negatif ketika model tidak sesuai dengan data dengan baik, seperti ketika hasil model tidak lebih baik daripada menggunakan garis horizontal sebagai referensi. Di sisi lain, nilai R^2 positif menandakan bahwa model memiliki kinerja yang lebih baik dibandingkan dengan model dasar [12]. Hal ini menunjukkan perlunya perbaikan signifikan dalam model-model tersebut.

1.3.3 Visualisasi Data Berbasis *Website*

Visualisasi data dapat memberikan pemahaman informasi yang lebih komprehensif dan representatif dari data yang dihasilkan. Melalui proses ini, data yang disajikan dapat berupa grafis yang menarik, tabel, *heatmap*, diagram, bagan, dan peta. Visualisasi data bertujuan untuk menyampaikan informasi dengan cara yang lebih mudah dipahami secara visual, memungkinkan pengguna untuk mengenali pola, hubungan dan wawasan yang mungkin sulit terlihat ketika data disajikan dalam format daftar biasa.

Salah satu platform yang dapat digunakan untuk menampilkan visualisasi data yaitu *website*. *Website* merupakan sebuah halaman daring tempat menyimpan berbagai informasi dan menyajikannya agar dapat diakses oleh semua orang selama terhubung dengan internet. *Website* menjadi salah satu sarana untuk menyajikan informasi kualitas udara kepada masyarakat. Melalui platform *website*, masyarakat bisa mendapatkan akses informasi yang berkaitan dengan data hasil pengukuran $PM_{2.5}$ dan CO_2 yang sudah dilakukan proses validasi. Data validasi akan direpresentasikan dalam bentuk grafik pada *website* untuk membantu pengguna dalam mendapatkan informasi kualitas udara secara jelas.

1.3.4 *Early Warning System*

Early Warning System (EWS) merupakan sistem peringatan dini yang memberikan peringatan terhadap potensi bahaya dari parameter kualitas udara. EWS memberikan notifikasi peringatan yang bertujuan untuk memberi informasi pada pengguna mengenai tingkat keparahan kualitas udara yang terukur. EWS akan membandingkan data prediksi yang sudah diperoleh dengan baku mutu kualitas udara ambien. Saat ini, media penyampaian informasi EWS yang telah dikembangkan adalah menggunakan *email*. Ketika hasil prediksi melebihi baku mutu, sistem akan memberikan notifikasi berupa hasil data kuantitatif tingkat kualitas udara secara detail kepada pengguna.

Salah satu kelebihan penggunaan *email* adalah mudah diakses dari berbagai perangkat. Tingkat ketersediaan dan aksesibilitas tinggi memungkinkan notifikasi *email* dapat diterima secara *real-time*. Namun, akses pengguna *email* masih lebih rendah

dibandingkan dengan akses pengguna jejaring sosial. Dilansir dari Data Digital 2024 di Indonesia pada bulan Januari, tingkat pengguna *email* hanya sebesar 62.8% dibandingkan data pengguna jejaring sosial sebesar 98.4% [13]. Dari data statistik tersebut, tidak semua pengguna mengakses *email* untuk dapat memperoleh informasi, sehingga memungkinkan notifikasi terlewat dan tidak tersampaikan pada pengguna.

1.4 Analisis Solusi yang Diajukan

1.4.1 Pengembangan Fitur Pendukung pada *Website* untuk Optimalisasi *Monitoring* Kualitas Udara

Fitur tambahan ini dikembangkan untuk meningkatkan kemampuan sistem dalam memantau kualitas udara secara lebih menyeluruh dan responsif. Salah satu fitur baru adalah halaman *monitoring* kualitas udara yang menampilkan data dari seluruh stasiun (*All Stations*), memungkinkan pengguna untuk mengakses informasi secara lebih lengkap. Selain itu, fitur *windrose* yang mengikuti skala Beaufort ditambahkan untuk memvisualisasikan arah dan kecepatan angin, sehingga membantu pengguna memahami pola angin di berbagai lokasi dengan lebih baik.

Fitur lain yang ditambahkan adalah visualisasi sistem *flagging*, yang memudahkan pengguna dalam mengidentifikasi data yang bermasalah, seperti data yang hilang atau tidak valid. Dengan adanya fitur ini, admin dapat dengan mudah melihat status data yang memerlukan tindak lanjut. Data *history flagging* juga disediakan untuk menganalisis kejadian *flagging* sebelumnya, yang berguna dalam evaluasi dan perbaikan sistem. Fitur notifikasi *flagging* yang terintegrasi pada *dashboard* memberikan peringatan langsung saat ditemukan data bermasalah, sehingga admin dapat segera mengambil langkah yang diperlukan.

1.4.2 Sistem *Flagging*

Sistem *flagging* adalah sistem yang berfungsi untuk menandai dan mendeteksi ketika terjadi data hilang dalam suatu periode tertentu. Jika sistem mendeteksi adanya data hilang, maka sistem akan memberikan peringatan kepada teknisi sebagai indikator bahwa telah terjadi data hilang. Sistem *flagging* juga berperan penting dalam pemantauan kualitas data sehingga teknisi dapat segera mengambil tindakan untuk menganalisis penyebab dan melakukan perbaikan yang diperlukan. Sistem *flagging* yang akan dikembangkan menggunakan metode *sliding window*. Dari kedua metode tersebut akan dipilih berdasarkan perbandingan kelebihan dan kekurangannya. Adapun perbandingan dari masing - masing metode sistem *flagging* dapat dilihat pada **Tabel 1.3**.

Tabel 1.3 Perbandingan Metode Sistem *Flagging*

No.	Metode	Kelebihan	Kekurangan
1.	<i>Threshold-Based</i>	Mudah diaplikasikan karena hanya memerlukan pengecekan terhadap jumlah kehilangan data [14].	Tidak ada analisis terkait penyebab atau pola data yang hilang.
2.	<i>Sliding Windows</i>	Memungkinkan pemantauan secara terus-menerus dalam jendela tertentu sehingga mendeteksi kehilangan data secara langsung [15].	Anomali kecil dari satu <i>window</i> waktu mungkin terdeteksi sebagai kehilangan data.

Berdasarkan perbandingan, metode yang akan dipilih untuk sistem *flagging* adalah *sliding windows*. Metode ini dipilih karena dapat mengecek data hilang dalam rentang waktu tertentu sesuai dengan ukuran *window* yang ditentukan. Jika terdapat data tidak mengirim pada waktu yang ditentukan, sistem ini akan memberikan peringatan. Selain itu, *sliding windows* memiliki algoritma sederhana sehingga komputasinya tidak berat pada sistem.

Notifikasi pada sistem *flagging* diperlukan untuk memastikan kepada teknisi untuk segera menangani masalah yang terdeteksi oleh sistem secara cepat. Kehilangan data yang tidak ditindaklanjuti secara tepat waktu dapat menyebabkan gangguan operasional. Oleh karena itu, notifikasi memberikan keunggulan dalam efisiensi komunikasi karena teknisi dapat menerima peringatan tanpa perlu memantau sistem sepanjang waktu. Pemilihan platform yang tepat sangat penting untuk memastikan informasi sesuai kebutuhan pengguna. Berikut beberapa media sosial yang dibandingkan sebagai platform sistem *flagging*.

Tabel 1.4 Perbandingan Notifikasi Media Sosial

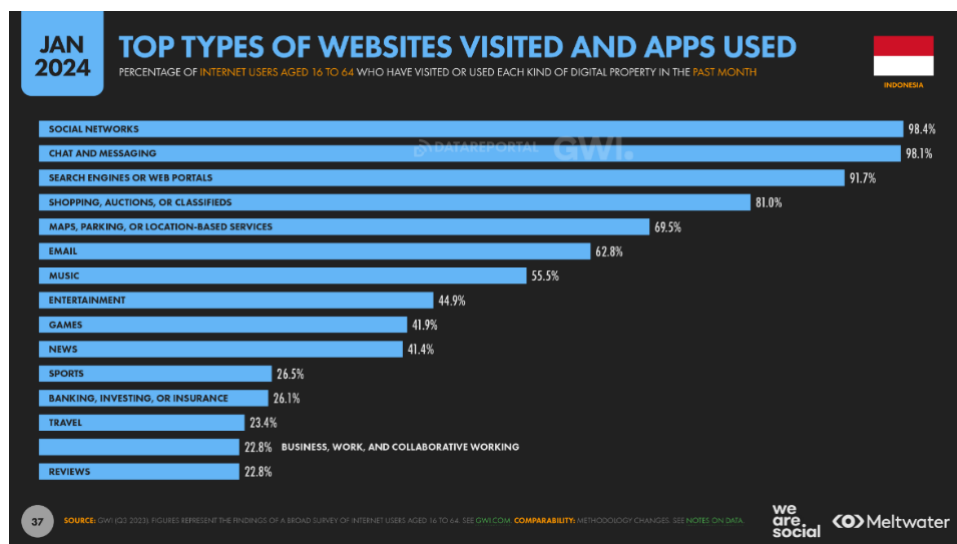
Media Sosial	Kelebihan	Kekurangan
SMS	• Respon cepat • Kemudahan implementasi sehingga tidak perlu aplikasi pihak ketiga [16]	• Tergantung pada jaringan SMS • Biaya operasional tinggi • Pesan terbatas
WhatsApp	• Kemudahan komunikasi • Aksesibilitas tinggi [17]	• API berbayar

<i>Email</i>	• Biaya operasional rendah • Dokumentasi otomatis [18]	• Respons lebih lambat
--------------	---	--

Dari perbandingan media sosial di atas, *email* dijadikan notifikasi yang sesuai untuk sistem *flagging* karena sifatnya formal dan terorganisir. *Email* memungkinkan pengiriman notifikasi yang terstruktur dan profesional. Selain itu, notifikasi yang dikirimkan melalui *email* otomatis terdokumentasi dan dapat disimpan dalam waktu lama yang memberikan kemudahan akses bagi pengguna untuk meninjau kembali informasi penting.

1.4.3 Early Warning System via WhatsApp

Pengembangan sistem peringatan dini (*early warning system*) dilakukan dengan mengubah platform notifikasi dari *email* menjadi media sosial. Dengan adanya media sosial, pengguna dapat berbagi dan memperoleh informasi dengan lebih cepat dan mudah. Berdasarkan data statistik digital pada bulan Januari 2024 di Indonesia, pengguna *email* hanya sebanyak 62.8% dari jumlah populasi pengguna internet di Indonesia. Jumlah ini jauh lebih rendah dibandingkan dengan pengguna jejaring sosial yang mencapai 98.4% [13].



Gambar 1.4 Data Pengguna Aplikasi dan Website Januari 2024

Banyaknya pengguna jejaring sosial membuat platform media sosial menjadi alternatif usulan solusi yang dapat dikembangkan. Hal ini karena media sosial marak digunakan oleh masyarakat, baik untuk sarana komunikasi, sarana hiburan, sarana edukasi, maupun sarana informasi. Karena banyaknya pengguna media sosial, maka informasi yang didapatkan dapat tersebar dengan cepat dan informatif, sehingga media

sosial menjadi salah satu solusi pengembangan *early warning system*. Adapun pilihan media sosial yang akan digunakan adalah SMS, Telegram, dan WhatsApp. Perbandingan antara media SMS, Telegram, dan WhatsApp dijelaskan dalam tabel berikut.

Tabel 1.5 Perbandingan Platform Media Sosial

Media Sosial	Kelebihan	Kekurangan
SMS	• Tidak perlu menggunakan internet sehingga dapat menjangkau area dengan sinyal yang lemah. • SMS dapat mencapai semua ponsel. • Tidak mendukung pengiriman multimedia dan pesan panjang terbagi menjadi beberapa pesan.	• Berbayar. • Waktu untuk menerima pesan 10 detik. • IMEI harus didaftarkan dan membutuhkan waktu lama.
Telegram	• Gratis. • Waktu untuk menerima pesan hanya butuh 5 detik [19]. • Bot harus didaftarkan tidak butuh waktu lama [19].	• Memerlukan koneksi internet. • Pengguna yang terbatas.
WhatsApp	• Banyak pengguna. • Menyediakan API untuk integrasi dengan sistem lain.	• API berbayar. • Memerlukan koneksi internet.

Berdasarkan perbandingan di atas media sosial yang akan dipilih menjadi platform untuk notifikasi pengguna adalah WhatsApp. WhatsApp dipilih karena merupakan platform notifikasi yang cukup banyak digunakan pengguna media sosial di Indonesia. Menurut Statista, Indonesia menempati posisi ketiga di dunia untuk jumlah pengguna WhatsApp terbanyak. Pada Juni 2021, jumlah pengguna WhatsApp di Indonesia telah meningkat menjadi 84,8 juta [20]. Pada platform WhatsApp, terdapat fitur WhatsApp Business yang dapat mengirimkan pesan *broadcast* atau pesan pemberitahuan ke beberapa kontak dalam satu kali pengiriman pesan. Integrasi antara *website* Biru-Langit dengan WhatsApp dilakukan dengan menetapkan nomor khusus sehingga semua notifikasi peringatan mengenai kualitas parameter udara dapat disampaikan melalui nomor tersebut.

1.4.4 Evaluasi Sistem Prediksi

Sistem prediksi bertujuan untuk memberikan informasi tentang kualitas parameter udara setelah pengukuran hari ini. Sistem ini menggunakan data historis untuk memperkirakan atau memprediksi kondisi di masa mendatang berdasarkan tren yang telah diidentifikasi. Dalam sistem prediksi yang telah dikembangkan saat ini, model prediksi berbasis *machine learning* diterapkan untuk menganalisis data deret waktu. Model prediksi yang dipilih adalah LSTM (*Long Short-Term Memory*), yang merupakan pengembangan dari RNN (*Recurrent Neural Network*). LSTM dirancang untuk mengatasi ketergantungan jangka panjang, yang merupakan kelemahan utama dari RNN. Selain itu, LSTM adalah salah satu algoritma yang paling populer untuk prediksi dan klasifikasi data yang berhubungan dengan waktu.

Untuk memastikan keandalan dan efisiensi sistem yang telah dibangun, diperlukan pengujian performa. Pengujian ini penting untuk memastikan keandalan dan akurasi prediksi yang dihasilkan oleh sistem. Dengan pengujian performa, dapat membantu memahami kemampuan model dalam memprediksi kualitas udara dan melakukan penyesuaian yang diperlukan untuk meningkatkan akurasi. Hal ini juga membantu dalam menentukan parameter-parameter penting yang berpengaruh dalam prediksi, sehingga model dapat terus disempurnakan seiring dengan bertambahnya data dan perubahan kondisi lingkungan. Pengujian performa tidak hanya akan menilai efektivitas model tetapi juga memungkinkan perbaikan dan peningkatan sistem lebih lanjut.

Berdasarkan hasil evaluasi performa sistem prediksi yang ada, diperlukan peningkatan performa algoritma *machine learning*. Untuk mengatasi ini, terdapat beberapa metode prediktif yang dapat digunakan untuk peningkatan performa yaitu *decision tree*, *naive bayes*, LSTM, dan CNN-LSTM. Dari keempat model prediktif tersebut akan dipilih berdasarkan perbandingan kelebihan dan kekurangannya. Adapun perbandingan dari masing-masing model prediktif dapat dilihat pada **Tabel 1.6**.

Tabel 1.6 Perbandingan Model Prediktif

No.	Metode	Kelebihan	Kekurangan
1.	<i>Decision Tree</i>	Data lebih akurat, meningkatkan efisiensi komputasi, dan menghindari hilangnya atribut kontinu.	Percabangan bisa saja kosong dan percabangan tidak signifikan.

2.	<i>Naïve Bayes</i>	Kinerja <i>naive bayes</i> masih tetap unggul ketika pengujian dilakukan pada tipe data kategori, semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas, dan menghasilkan performa baik.	Sangat sensitif pada fitur yang terlalu banyak sehingga menghasilkan akurasi lebih rendah dan ukuran dari vektor fitur yang dihasilkan cukup besar dan membutuhkan teknik untuk memperkecil ukuran vektor tersebut.
3.	LSTM	Dapat menangani masalah <i>vanishing gradient</i> dan menyimpan informasi dengan jangka waktu yang panjang.	Waktu pelatihan yang lama dan kombinasi parameter yang kompleks.
4.	CNN-LSTM	Efektif dalam mengekstraksi fitur-fitur penting dari data spasial dan menangani ketergantungan jangka panjang dalam data sekuensial. Model ini mampu mengelola data kompleks seperti kualitas udara yang berubah-ubah dan memberikan invarian translasi, sehingga mengenali pola yang sama di lokasi berbeda.	Komputasi yang cukup kompleks, waktu pelatihan yang lama, dan kebutuhan akan data yang besar.

Berdasarkan perbandingan di atas, model prediktif yang akan dipilih untuk pengembangan prediksi adalah kombinasi model LSTM (*Long Short-Term Memory*) dengan model CNN (*Convolutional Neural Network*). Dengan menggabungkan model hybrid CNN-LSTM (*Convolutional Neural Network - Long Short-Term Memory*), maka model ini dapat memanfaatkan keunggulan CNN dalam mengenali pola fitur-fitur penting dari data secara spesifik dan keunggulan LSTM dalam menangani ketergantungan jangka panjang dalam data sekuensial. Kombinasi ini memungkinkan model untuk secara efektif mengelola data yang kompleks dan tidak terstruktur, memberikan hasil prediksi yang lebih akurat dan handal.