

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Salah satu kebutuhan mendasar masyarakat modern adalah energi listrik yang akan meningkat seiring dengan perkembangan teknologi, pertumbuhan populasi, kemajuan ekonomi dan perubahan gaya hidup. Tentunya jumlah bahan bakar berbasis fosil yang tersedia itu terbatas, selain penggunaan bahan bakar fosil ini mengeluarkan gas berbahaya seperti Karbon Dioksida, tetapi juga suatu saat nanti akan habis [4]. Pada 2022, total emisi Karbon Dioksida (CO₂) di Indonesia mencapai sekitar 696,7 juta ton, di mana 42,6% berasal dari pembangkitan tenaga listrik dengan peningkatan rata-rata 4,1% per tahunnya [5].

Emisi karbon dioksida telah lama menjadi salah satu permasalahan utama sektor pembangkit listrik karena peningkatan kebutuhan energi yang semakin tinggi menyebabkan penipisan bahan bakar fosil dan peningkatan polusi udara [6]. Karbon Dioksida merupakan penyebab utama *global warming* dan perubahan iklim [7], maka itu salah satu kunci dalam menanggulangi efek ini adalah dengan mengurangi penggunaan listrik. Tetapi menurut Neng Shen (2020), untuk mengurangi emisi Karbon akan diperlukan pengendalian konsumsi bahan bakar fosil dan hal ini dapat memengaruhi perkembangan ekonomi suatu negara [8]. Sebagai solusi, *Carbon Trading* digunakan. *Carbon Trading* adalah suatu alat atau kebijakan yang mendorong pengurangan emisi Karbon yang dihasilkan oleh perusahaan dengan menggunakan sistem perdagangan [9]. Pada dasarnya *Carbon Trading* bekerja dengan prinsip "*cap and trade*" yaitu sebuah perusahaan akan memiliki batas kuota jumlah Karbon emisi yang diperbolehkan dan jika kuota atau *cap* tersebut dilewati, maka akan dikenakan denda [10]. Metode pembayaran denda ini bisa dilakukan dengan membayar secara langsung kepada pemerintah atau dengan metode kedua yaitu "*trade*". Perusahaan yang memerlukan lebih banyak kuota emisi Karbon, dapat membeli kuota tersebut di pasar Karbon. Sebaliknya jika sebuah perusahaan tidak menggunakan seluruh kuota yang dimiliki, kuota tersebut dapat dijual pada pasar Karbon. Metode ini telah terbukti efektif dalam mengurangi emisi karbon, di antaranya sistem perdagangan emisi Karbon Eropa menjadi yang paling sukses [11]. Maka itu pentingnya untuk sebuah perusahaan atau entitas lainnya dapat memprediksi jumlah Emisi Karbon yang akan dihasilkannya agar dapat menyusun

sebuah strategi untuk mengurangi emisi Karbon.

Dalam beberapa tahun terakhir, kemajuan teknologi khususnya di bidang kecerdasan artifisial seperti *machine learning*, telah memberikan peluang baru dalam menganalisis dan memprediksi konsumsi energi listrik dan penggunaan algoritma *machine learning*, seperti regresi, *neural network*, *decision trees*, dan lain-lain, dapat membantu proses prediksi atau identifikasi pola dan tren yang sulit atau bahkan tidak dapat di deteksi oleh metode analisis tradisional [12]. *Electricity demand forecasting* memiliki peran yang penting dalam optimalisasi dan manajemen penggunaan listrik [13], dengan mengetahui jumlah energi listrik yang akan di konsumsi, pemilik rumah tangga dapat lebih memahami kapan dan dimana sebagian besar energi listrik di konsumsi dan dengan pengetahuan tersebut pemilik rumah tangga dapat menyusun sebuah strategi untuk mengurangi penggunaan energi tersebut. Sehingga strategi tersebut tidak hanya biaya listrik akan turun, tetapi juga jumlah emisi Karbon yang dihasilkan juga akan turun.

Untuk melakukan prediksi-prediksi tersebut, diperlukannya sebuah model *machine learning*. Salah satu algoritma umum yang digunakan dalam *machine learning* adalah *decision tree*. *Decision tree* adalah sebuah metode klasifikasi yang menggunakan prinsip sebuah *tree* yaitu sebuah kumpulan aturan dan kondisi untuk membentuk suatu keputusan[14]. Adapun beberapa penelitian lainnya yang meneliti juga mengenai penggunaan algoritma *decision tree*, seperti Charbuty et al. (2021) meneliti dan membahas secara mendalam mengenai cara kerja, performa, tipe *decision tree*, dan kelebihan kekurangan [15]. Lalu Cui Xiwen et al. (2021) meneliti mengenai *carbon emission forecasting* di China dengan menggunakan *gradient boosting decision tree* untuk memprediksi total emisi Karbon dalam tingkat nasional berdasarkan total *GDP*, impor ekspor, pertumbuhan urbanisasi, penggunaan energi milik negara China [16]. Lalu Jinli Wei et al. (2022) meneliti mengenai kalkulasi emisi Karbon yang dihasilkan oleh transportasi umum dengan menggunakan algoritma *decision tree* dan penelitian ini berhasil membuat sebuah model yang dapat memprediksi emisi karbon transportasi umum dengan tingkat akurasi yaitu 92.37% [17].

Tetapi *decision tree* sendiri memiliki beberapa limitasi seperti pertama kurang stabil karena sedikit perubahan dalam *training data* dapat mengubah total struktur *tree*-nya, kedua karena *local decisions decision tree* cenderung hanya fokus pada atribut terbaik dari setiap keputusan, sehingga menghasilkan *tree* yang kurang optimal atau cenderung *overfitting*, ketiga karena metode *local search tree* yang memiliki pengaruh terbesar dalam klasifikasi, bisa tersembunyi di bawah *tree* yang

lemah, sehingga *tree* dengan pengaruh terbesar memiliki kesempatan tinggi untuk dilewatkan dan masih banyak lagi [18].

Untuk menanggulangi kekurangan tersebut diperlukannya sebuah algoritma yang bernama *ensemble*. *Ensemble* adalah suatu algoritma yang dimana proses *training* atau pembangunan model tidak hanya meliputi satu model melainkan banyak model yang akan digabungkan menjadi satu untuk menghasilkan model yang kuat [19]. Salah satu metode *ensemble* yang populer dan akan digunakan dalam penelitian ini adalah *boosting*. Dengan menggunakan metode *boosting*, model dapat menggabungkan *decision tree* yang lemah-lemah menjadi satu *decision tree* yang kuat, sehingga akurasi model dalam memprediksi emisi Karbon dapat meningkat dan juga kuat atau tahan terhadap *overfitting* [19].

Salah satu algoritma *boosting* yang populer adalah *XGBoost Regression*, huruf 'X' dalam *XGBoost* memiliki arti ekstrem, disebut ekstrem karena berbeda dengan *gradient boosting* pada biasanya, *XGBoost* merupakan algoritma *gradient boosting* yang lebih ter-regularisasi [20]. Regularisasi adalah salah satu metode yang diterapkan untuk mengurangi *overfitting* dari model regresi yang dibentuk berdasarkan data *training*, dengan menggunakan regularisasi model akan memiliki sistem penalti kompleksitas pada fungsi akhir yang akan mengurangi kesempatan *overfitting* [21]. Menurut Tianqi Chen (2016), *XGBoost* telah dibuktikan memiliki performa prediksi yang sangat efektif, hal ini dapat dilihat dari hasil kompetisi *machine learning* yang diselenggarakan oleh *Kaggle*, yang dimana dalam kompetisi itu diantara 29 *winning solution* 17 diantaranya menggunakan *XGBoost* [22]. Adapun beberapa penelitian yang meneliti juga mengenai penggunaan *XGBoost*, seperti Ruoran Wang et al. (2022) menyatakan bahwa *XGBoost* memiliki performa yang lebih baik dibanding algoritma *regression* lainnya dalam memprediksi tingkat kematian korban cedera otak, penelitian ini menemukan bahwa *XGBoost* memiliki akurasi yang lebih tinggi yaitu 95.5% dibanding *logistic regression* yaitu 70% [23]. Lalu J.Avanijaa et al. (2021) juga meneliti penggunaan *XGBoost Regression* dalam memprediksi harga rumah, yang dimana penelitian ini berhasil membuat sebuah model dengan tingkat akurasi rata-rata 95.2% atau tingkat *error* rata-rata 4.8% [24].

Berdasarkan penjelasan di atas, tujuan utama penelitian ini adalah mengimplementasi dan mengevaluasi performa algoritma *XGBoost Regression* dalam memprediksi penggunaan energi listrik atau emisi Karbon sebuah rumah tangga, dengan harapan hasil prediksi tersebut dapat digunakan untuk meningkatkan efisiensi penggunaan listrik dan mengurangi emisi Karbon.

1.2 Rumusan Masalah

Berdasarkan latar belakang dan permasalahan yang telah dijelaskan, maka rumusan masalah dalam penelitian ini adalah

1. Bagaimana cara mengimplementasi algoritma *XGBoost regression* untuk prediksi emisi Karbon hasil penggunaan listrik rumah tangga?
2. Bagaimana nilai RMSE, R-Squared, MAE, dan MAPE untuk model prediksi emisi Karbon hasil penggunaan listrik rumah tangga menggunakan algoritma *XGBoost regression*?

1.3 Batasan Permasalahan

Terdapat beberapa batasan masalah yang ditetapkan untuk menghindari penelitian yang terlalu luas, berikut adalah batasan masalah tersebut:

1. *Dataset* yang digunakan pada penelitian ini berasal dari kaggle yang merupakan penelitian "*Data driven prediction models of energy use of appliances in a low-energy house*" oleh Luis M. (2017) [25] terkait penggunaan listrik sebuah rumah tangga yang berada di Belgia pada tahun 2016, dengan jumlah dataset 19735.
2. Fitur yang digunakan untuk melakukan prediksi adalah suhu ruangan, kelembapan ruangan, penggunaan listrik lampu.
3. Ratio Train Test Split yang akan digunakan adalah 75% Training dan 25% Validation.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah.

1. Untuk mengimplementasi algoritma *XGBoost regression* untuk mendapatkan hasil prediksi emisi Karbon hasil penggunaan listrik rumah tangga.
2. Untuk mengevaluasi nilai RMSE, R-Squared, MAE, dan MAPE dari model prediksi emisi Karbon hasil penggunaan listrik rumah tangga menggunakan algoritma *XGBoost regression*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah mendapatkan pengetahuan dan pemahaman mengenai faktor apa saja yang dapat memengaruhi jumlah emisi Karbon hasil penggunaan listrik rumah tangga. Penelitian ini juga memiliki manfaat yaitu membangun sebuah *machine learning model* yang dapat memprediksi emisi Karbon hasil penggunaan listrik rumah tangga.

1.6 Sistematika Penulisan

Berisikan uraian singkat mengenai struktur isi penulisan laporan penelitian, dimulai dari Pendahuluan hingga Simpulan dan Saran. Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN

Bab ini terdiri atas enam bagian, yaitu latar belakang masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan laporan.

- Bab 2 LANDASAN TEORI

Bab ini terdiri atas penjelasan mengenai teori-teori dan pengetahuan dasar yang diperlukan dan digunakan untuk menjalankan dan mendukung penelitian ini.

- Bab 3 METODOLOGI PENELITIAN

Bab ini terdiri atas penjelasan mengenai tahap-tahap yang akan dilakukan dari awal hingga akhir penelitian guna mencapai suatu kesimpulan atau jawaban dari penelitian ini.

- Bab 4 HASIL DAN DISKUSI

Bab ini terdiri atas penjelasan mengenai hasil implementasi algoritma *XGBoost regression* disertai dengan gambar, potongan kode, ataupun *graph*.

- Bab 5 KESIMPULAN DAN SARAN

Bab ini terdiri atas kesimpulan yang didapatkan dari hasil penelitian yang sudah dilakukan dan saran untuk meningkatkan pengembangan penelitian selanjutnya.