JURNAL REKAYASA INFRASTRUKTUR
ISSN : 2460-335X
E-ISSN : 2686-6587
Vol 9 No 2, November 2023

OPTIMASI MIX DESIGN BETON MELALUI TEKNOLOGI MACHINE LEARNING

Adji Putra Abriantoro[1*], J. Rajes Khana[2]
[1]Program Studi Teknik Sipil, Universitas 17 Agustus 1945 Jakarta, DKI Jakarta
adji.putra@uta45jakarta.ac.id

ABSTRACT

The concrete mix design is a crucial step in ensuring the quality of concrete used in construction projects. Traditional mix design methods rely on trial and error, which can be time-consuming and escalate construction costs. In recent years, machine learning technology has been developed to predict concrete properties and optimize mix designs for high-quality concrete. This study aims to explore the application of machine learning in high-quality concrete mix design, considering theoretical conditions, methods, and related research. The focus of this research is to test the use of machine learning techniques in predicting concrete properties and optimizing mix designs for high-quality concrete. The theory and concepts of machine learning will be applied to concrete mix design datasets and relevant properties. The methods employed will include data pre-processing, feature selection, and model training and evaluation. The performance of the machine learning model will be compared with traditional mix design methods to determine its effectiveness. Furthermore, the results and benefits of this study will demonstrate the potential advantages of using machine learning in determining high-performance concrete mix designs. By accurately predicting concrete properties and optimizing mix designs, construction projects can be completed more efficiently and with higher quality. This technology also has the potential to reduce costs associated with trial and error methods and minimize the environmental impact of concrete production. The success of this study will pave the way for further research and development in the field of machine learning for high-performance concrete mix designs.

Keywords: accuracy, concrete, efficiency, machine learning, mix design

ABSTRAK

Mix design beton memegang peranan penting dalam menjamin kualitas beton yang diterapkan dalam proyek konstruksi. Pendekatan tradisional dalam *mix design* mengandalkan *trial and error* yang memakan waktu serta meningkatkan biaya pembangunan. Namun, belakangan ini, kemajuan *teknologi machine learning* telah ditujukan untuk memprediksi sifat-sifat beton serta mengoptimalkan desain campuran pada beton berkualitas tinggi. Penelitian ini bertujuan untuk mengeksplorasi penerapan *machine learning* dalam desain campuran beton mutu tinggi dengan mempertimbangkan teori, metode, dan riset yang relevan. Fokusnya adalah pada penggunaan teknologi *machine learning* untuk memprediksi sifat-sifat beton dan mengoptimalkan *mix design* untuk beton mutu tinggi. Konsep dan teori *machine learning* akan diterapkan pada dataset desain campuran beton dan sifat-sifat yang terkait. Langkah-langkahnya mencakup pra-pemrosesan data, pemilihan fitur, serta pelatihan dan evaluasi model. Performa model machine learning akan dibandingkan dengan metode desain campuran tradisional untuk menilai efektivitasnya. Selain itu, hasil dan manfaat dari penelitian ini akan menggambarkan potensi kegunaan machine learning dalam menentukan desain campuran beton mutu tinggi. Dengan kemampuan memprediksi sifat-sifat beton secara akurat dan melakukan optimalisasi desain campuran, proyek konstruksi dapat diselesaikan secara lebih efisien dan dengan standar kualitas yang lebih tinggi. Tak hanya itu, teknologi ini juga berpotensi mengurangi biaya yang terkait dengan uji coba berulang serta menurunkan dampak lingkungan dari produksi beton. Keberhasilan penelitian ini akan membuka pintu bagi eksplorasi dan pengembangan lebih lanjut dalam ranah *machine learning* untuk *mix design* beton mutu tinggi.

Kata kunci: Akurasi, Beton, Efisiensi, Machine learning, Mix design

1. PENDAHULUAN

Mix design beton merupakan suatu proses yang kompleks dan multistage di mana penulis berusaha untuk menemukan komposisi bahan yang terbaik untuk menciptakan beton yang memiliki performa unggul. Desain campuran dan kekuatan beton adalah faktor kritis dalam desain bangunan beton bertulang, mix design beton yang baik dapat membantu memastikan bahwa beton mencapai kekuatan dan daya tahan yang diinginkan yang sangat penting dalam konstruksi bangunan beton bertulang[1][2]. Proses ini melibatkan kajian mendalam terhadap berbagai faktor seperti jenis dan proporsi agregat, campuran air semen, aditif, dan lainnya, dengan tujuan untuk mencapai kualitas dan kekuatan beton yang optimal. Beberapa penelitian secara independen telah menunjukkan bahwa perkembangan kekuatan beton ditentukan tidak hanya oleh rasio air/campuran (w/c ratio), tetapi juga dipengaruhi oleh kandungan bahan-bahan lain[3]. Langkah awal dalam mix design beton melibatkan pemahaman mendalam tentang karakteristik material-material yang akan digunakan, serta evaluasi terhadap persyaratan teknis proyek yang bersangkutan. Selanjutnya, berbagai parameter seperti kuat tekan, workability, tahan terhadap erosi, dan lain-lain, harus diperhitungkan secara cermat untuk memastikan bahwa desain campuran beton memenuhi standar yang direncanakan. Tujuan utama dari mix desgn beton adalah untuk memperkirakan komposisi kuantitatif dan proporsi yang tepat dari komponen campuran beton. Dalam pembuatan beton digunakan komposisi campuran yang direncanakan akan mencapai mutu beton yang terbaik. Mutu beton antaranya yang paling signifikan adalah kekuatan tekan. Mix desgn beton memiliki peran penting dalam pemenuhan renana mutu beton baik kekuatan beton maupun daya tahannya[4].

Mostafaei et.al (2023) membahas tentang dampak lingkungan dari desain campuran beton dan pentingnya mengoptimalkan desain campuran untuk mengurangi dampak negatif serta menyoroti keterbatasan metode manual tradisional dalam mix design beton [5]–[7]. Metode mix design beton secara konvensional masih memiliki beberapa kekurangan. Kami ingin memperkenalkan cara untuk merancang campuran beton berdasarkan persamaan matematika yang dikembangkan oleh algoritma machine learning. Hasil pendektan analisa machine learning ini dapat menjadi alat untuk perkiraan kasar dari mutu beton yang dihasilkan. alam beberapa dekade terakhir, kebutuhan akan alat kekuatan yang efektif dirasakan, untuk memprediksi kinerja beton, terutama kekuatan tekanannya. Beberapa alat statistik telah diterapkan dalam beberapa dekade terakhir. Sangat populer di antaranya adalah persamaan regresi linear, multilinear, dan nonlinear[8]–[10]. Selain itu, perkembangan teknologi machine learning belakangan ini mendapatkan perhatian yang signifikan, dan prediksi untuk masa depan teknologi ini semakin menjanjikan. Machine learning merupakan cabang dari kecerdasan buatan yang memungkinkan komputer untuk belajar dari data dan pengalaman, serta dapat melakukan tugas-tugas kompleks tanpa harus secara eksplisit diprogram [11].

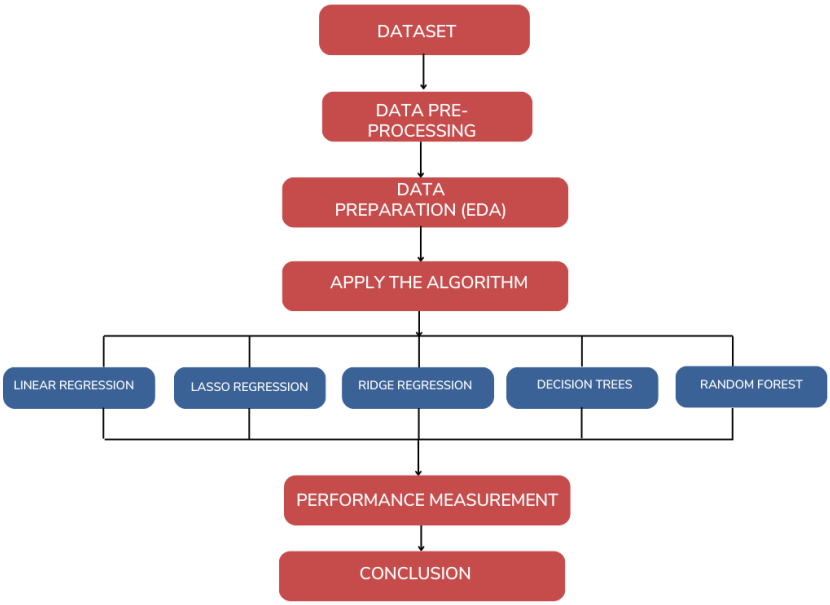
Kamath et al. (2022) menemukan bahwa Decision regression tree, Random forest regression, Lasso regression, Ridge regression, dan Multiple linear regression tampil paling baik dalam sebuah studi perbandingan lima model regresi Machine Learning (ML)[12]. Dua set data dari masa lalu dianalisis menggunakan lima model regresi ML yang berbeda, dan dua dataset dari laboratorium mutakhir digunakan. Kinerja model dievaluasi menggunakan koefisien determinasi (R2), mean absolute error, root-mean-squared error, dan mean absolute percentage error. Model-model tersebut kemudian dibandingkan secara statistik dengan penelitian-penelitian sebelumnya tentang topik tersebut. Terakhir, analisis sensitivitas dan pengujian parametrik digunakan untuk menilai kinerja prediktor dan memberikan wawasan tentang dampak variabel-variabel tersebut. Kehadiran teknologi ini telah membawa perubahan fundamental dalam berbagai bidang, termasuk dalam desain campuran beton. Dalam konteks desain campuran beton, machine learning dapat digunakan untuk menganalisis data besar terkait performa beton, properti bahan, dan kondisi lingkungan, untuk kemudian memberikan rekomendasi komposisi campuran yang optimal. Hal

ini memberikan keunggulan dalam efisiensi dan akurasi desain, serta dapat menghemat waktu dan sumber daya.

2. METODE PENELITIAN

a. Dataset

Dataset yang digunakan pada penelitian ini didapatkan dari UC Irvine Machine Learning Repository's yang menyediakan dataset yang diperlukan. Estimasi kekuatan tekan beton memiliki delapan variabel antaralain agregat halus, agregat kasar, air, umur beton, superplastisizer, dan komponen limbah, seperti slag dan abu terbang. Korelasi antara distribusi variabel dan parameter yang digunakan untuk menghitung kekuatan tekan digambarkan dalam Gambar 1.



Gambar 1. Study Model

Dalam dunia machine learning, terdapat sejumlah besar parameter input yang mempengaruhi hasil yang diinginkan. Parameter-parameter ini berperan penting dalam menentukan kualitas dan akurasi model yang dihasilkan. Salah satu contoh aplikasi dari teknik ini adalah dalam prediksi kuat tekan beton. Beton adalah material konstruksi yang kompleks dengan berbagai komponen, termasuk semen, air, agregat kasar, agregat halus dan aditif. Semua komponen ini mempengaruhi karakteristik dan sifat akhir beton, termasuk kuat tekannya. Oleh karena itu, pemilihan dan pengaturan proporsi komponen ini sangat penting. Selain itu, umur beton juga mempengaruhi kuat tekan beton. Beton yang lebih tua biasanya memiliki kuat tekan yang lebih tinggi dibandingkan dengan beton yang lebih muda. Oleh karena itu, umur beton juga menjadi salah satu parameter input dalam model machine learning.

Output dari model ini adalah kuat tekan beton. Kuat tekan adalah ukuran kekuatan material untuk menahan beban kompresi. Dalam konteks beton, kuat tekan adalah indikator utama kualitas dan keandalan material tersebut. Namun, perlu diingat bahwa output model sangat sensitif terhadap

jumlah titik data dan faktor-faktor yang dimasukkan ke dalamnya. Oleh karena itu, penting untuk memiliki data yang cukup dan relevan untuk mendapatkan hasil yang akurat dan dapat diandalkan. Dalam penelitian ini, digunakan data penelitian yang mencakup 1030 pengamatan. Data ini digunakan untuk memberikan perkiraan mix design untuk beton mutu tinggi. Dengan menggunakan teknik machine learning, peneliti dapat membuat prediksi yang akurat tentang kuat tekan beton berdasarkan berbagai parameter input.

b. Data Preprocessing

Pre-data processing adalah tahap krusial dalam rangka memastikan keakuratan, kehandalan, dan integritas data yang akan digunakan dalam penelitian yang membahas efektivitas penerapan Machine Learning (ML) untuk menentukan rencana mix design pada beton mutu tinggi. Proses ini melibatkan langkah-langkah seperti pengumpulan data yang cermat, identifikasi potensi outliers, serta deteksi dan penanggulangan nilai yang hilang atau inkonsisten. Dalam konteks karakteristik material beton, umur, dan kuat tekan beton, pengelompokan dan penyaringan data menjadi penting untuk memastikan hanya data yang relevan dan berkualitas tinggi yang akan digunakan dalam analisis selanjutnya. Penghapusan data yang tidak konsisten atau tidak relevan juga merupakan bagian dari pre-data processing guna meminimalkan bias dan mengoptimalkan kualitas dataset.

Selain itu, tahap pre-data processing juga mencakup normalisasi atau standarisasi data, memastikan bahwa variabel-variabel dengan skala yang berbeda dapat diolah secara konsisten. Hal ini membantu meminimalisir potensi distorsi hasil akibat dominasi variabel dengan skala besar, yang mungkin menghambat kinerja algoritma ML. Pengaturan data dengan standar yang tepat juga merupakan langkah strategis untuk memaksimalkan kinerja model ML yang akan dikembangkan. Oleh karena itu, pre-data processing bukan hanya sekadar proses persiapan data, tetapi juga sebuah strategi yang cerdik untuk mengoptimalkan validitas dan efektivitas model ML yang direncanakan, dengan menghasilkan hasil yang tepat dan informatif dalam perancangan mix beton mutu tinggi.

c. Data Preparation (Explortaoory Data Analysis / EDA)

Data preparation, atau yang dikenal dengan istilah Exploratory Data Analysis (EDA), merupakan tahap yang sangat kritis dalam rangkaian penelitian yang membahas efektivitas penerapan Machine Learning (ML) dalam menentukan rencana mix design pada beton mutu tinggi. Proses EDA ini melibatkan analisis mendalam terhadap dataset yang telah terkumpul, yang mencakup identifikasi dan pemahaman pola data, distribusi variabel, serta korelasi antar variabel yang kompleks. Kami melakukan eksplorasi yang cermat terhadap hubungan antara parameter-parameter terkait material beton, umur, dan kuat tekan beton, memungkinkan penulis untuk memetakan interaksi yang kompleks di antara variabel tersebut. Hasil dari analisis ini memberikan wawasan mendalam terkait karakteristik esensial yang menjadi landasan dalam perancangan mix beton mutu tinggi. Selain itu, melalui EDA, penulis dapat mengidentifikasi dengan tepat anomali atau outlier yang dapat mempengaruhi hasil analisis ML, sehingga memastikan keakuratan dan kehandalan model yang akan dikembangkan. Dengan demikian, seluruh proses EDA ini menjadi tonggak penting dalam memastikan bahwa dataset yang digunakan untuk membangun model ML memiliki kualitas yang optimal, menghasilkan rencana mix design beton mutu tinggi yang akurat dan dapat diandalkan.

d. Apply the Algorithm

Tahapan ini berfokus pada penerapan berbagai algoritma machine learning untuk membangun model yang dapat membantu dalam menentukan rencana mix design beton mutu tinggi yang optimal, selain itu langkah penting ini mendalaminya dengan mengimplementasikan serangkaian model algoritma Machine Learning yang memiliki relevansi yang erat dengan konteks penelitian. Penerapan rangkaian algoritma ini, yang secara cermat dipilih sesuai dengan tujuan penelitian yang memiliki relevansi kuat dengan topik penelitian. Pemilihan algoritma yang sesuai dengan konteks penelitian dan tujuan spesifik juga menjadi faktor kunci dalam memberikan kontribusi yang berarti dalam penelitian ini. Misalnya, penelitian ini dapat mempertimbangkan berbagai faktor seperti komposisi bahan yang mempengaruhi beton mutu tinggi. Dengan memadukan dan menganalisis informasi yang berasal dari berbagai model, diharapkan akan diperoleh wawasan yang lebih dalam dan terperinci mengenai cara memaksimalkan kualitas beton mutu tinggi, seiring dengan efisiensi penggunaan sumber daya dan biaya produksi. Ini akan mendukung industri konstruksi dalam mencapai tujuan berkelanjutan dan optimal, yang mengarah pada penggunaan beton berkualitas tinggi untuk proyek-proyek infrastruktur dan bangunan yang lebih baik.

e. Performance measurement

Sebagai indikator seberapa baik kinerja suatu model, penelitian ini menggunakan koefisien korelasi (R2), root-mean-squared error (RMSE), dan mean absolute percentage error (MAPE). Menurut penelitian Fee et al. (2022), representasi matematikanya adalah sebagai berikut:

$$R^2 = \frac{\sum \left(y_a - y'_a\right)\left(y_p - y'_p\right)}{\sqrt{\sum \left(y_a - y'_a\right)^2 \sum \left(y_p - y'_p\right)^2}}$$

Root Mean Square Error (RMSE) merupakan ukuran seberapa baik model dapat memprediksi data. RMSE dihitung dengan mengambil akar kuadrat dari rata-rata perbedaan kuadrat antara nilai yang diprediksi dan nilai sebenarnya. Dalam rumus ini:

$$\text{RMSE} = \frac{\sqrt{\left(y_a - y_p\right)^2 + \left(y_a - y_p\right)^2 \ldots + \left(y_a - y_p\right)^2}}{n}.$$

RMSE memberikan ukuran rata-rata seberapa jauh prediksi dari nilai sebenarnya. Semakin rendah nilai RMSE, semakin baik model dalam memprediksi data (Singla etal, 2022).

Sedangkan Mean Square Error (MSE) adalah metrik evaluasi yang digunakan dalam statistik dan machine learning untuk mengukur seberapa akurat model regresi dalam memprediksi nilai numerik, dalam rumus sebagai berikut:

$$\text{MSE} = \frac{1}{N}\sum_{j=1}^{N}\left(y_j - \check{y}_j\right)^2$$

MSE dihitung dengan menjumlahkan perbedaan kuadrat antara nilai prediksi dan nilai sebenarnya, kemudian membaginya dengan jumlah total observasi1. Semakin rendah nilai MSE, semakin baik model dalam memprediksi data[13], [14].

Lalu Mean Absolute Error (MAE) adalah metrik yang digunakan untuk mengevaluasi kinerja model regresi yang erupakan ukuran statistik yang digunakan untuk mengukur sejauh mana prediksi model berbeda dari nilai sebenarnya. MAE dihitung dengan mencari rata-rata dari selisih absolut antara nilai prediksi dan nilai sebenarnya. MAE memberikan gambaran tentang seberapa

besar rata-rata kesalahan prediksi model terhadap nilai sebenarnya. Semakin rendah nilai MAE, semakin baik model dalam memprediksi data[15].

Dengan rumus sebagai berikut:

$$\mathrm{MAE} = \frac{1}{N}\sum_{j=1}^{N}\left|y_j - \check{y}_j\right|.$$

Dimana:

R2 adalah koefisien determinasi.

ya dan yp adalah nilai aktual dan prediksi.

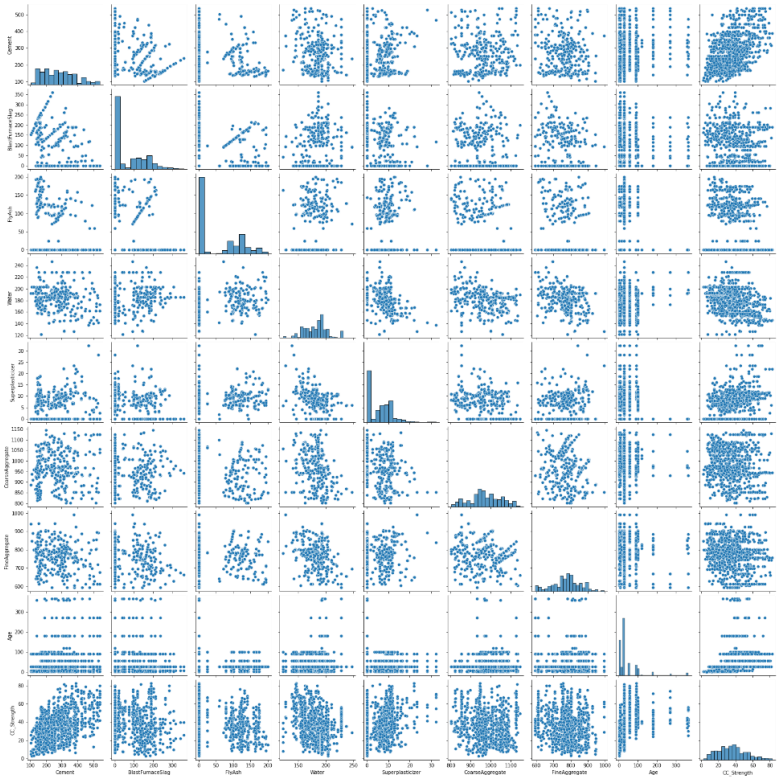
y'a dan y'p adalah rata-rata nilai aktual dan prediksi.

3. HASIL DAN PEMBAHASAN

3.1 Exploratory Data Analysis (EDA)

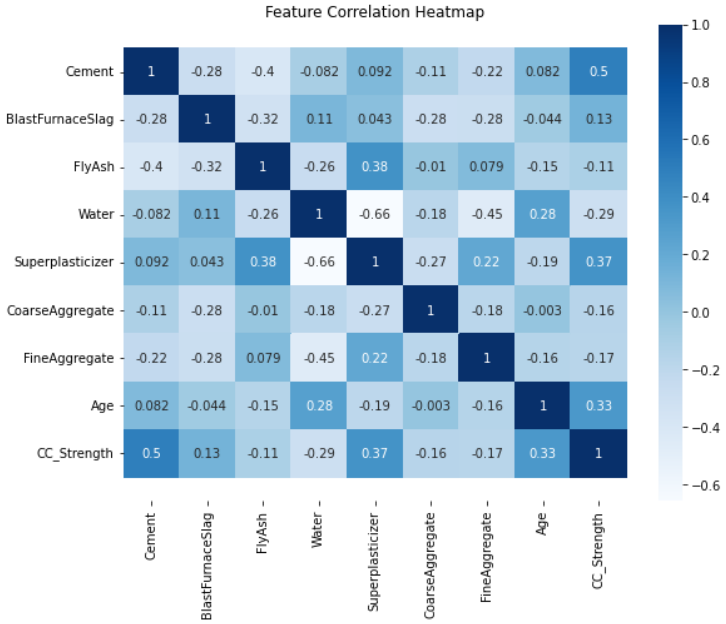
Hasil pemeriksaan hubungan antar variabel dalam suatu dataset menggunakan grafik pasangan (pairplot) dari library Seaborn. Dalam konteks analisis data, korelasi adalah ukuran sejauh mana dua variabel berubah bersama. Jika dua variabel memiliki korelasi tinggi, itu berarti jika satu variabel berubah, variabel lainnya cenderung berubah juga. Sebaliknya, jika dua variabel memiliki korelasi rendah, itu berarti perubahan dalam satu variabel tidak selalu diikuti oleh perubahan dalam variabel lainnya.

Dalam hal ini, pairplot digunakan untuk memeriksa apakah ada korelasi tinggi antara variabel independen dalam dataset. Jika ada korelasi tinggi antara beberapa variabel independen, itu bisa menjadi masalah dalam beberapa jenis analisis data dan machine learning, seperti regresi linier. Oleh karena itu, penting untuk memeriksa korelasi antar variabel sebelum melakukan analisis lebih lanjut.



Gambar 2. Pairplot

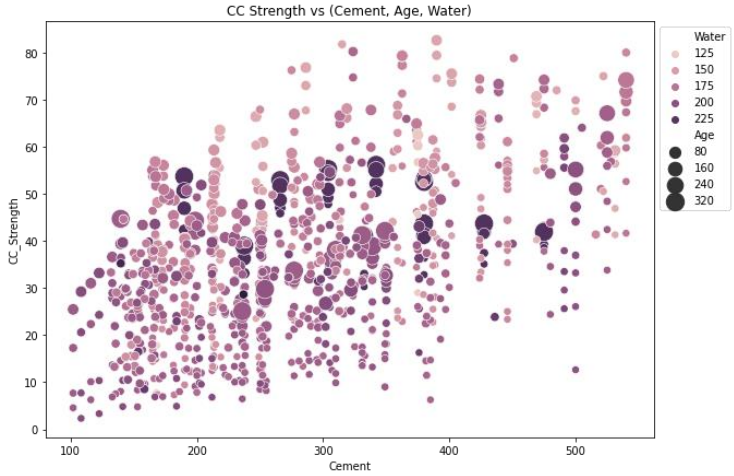
Dari hasil analisa pairplot tidak tampak adanya korelasi tinggi antara variabel independen (fitur). Hal ini dapat dikonfirmasi lebih lanjut dengan membuat plot korelasi Pearson. Tabel korelasi ini memberikan gambaran tentang sejauh mana hubungan antara dua variabel. Tabel korelasi adalah matriks persegi yang menunjukkan koefisien korelasi antara setiap pasangan variabel dalam dataset. Nilai dalam tabel berkisar antara -1 dan 1. Sedangkan koefisien korelasi adalah ukuran sejauh mana dua variabel berubah bersama. Jika dua variabel memiliki korelasi positif, itu berarti jika satu variabel meningkat, variabel lainnya cenderung meningkat juga. Sebaliknya, jika dua variabel memiliki korelasi negatif, itu berarti jika satu variabel meningkat, variabel lainnya cenderung menurun. Nilai 0 menunjukkan tidak ada korelasi antara dua variabel. Nilai yang mendekati 1 atau -1 menunjukkan korelasi yang kuat. Dalam konteks ini, tabel menunjukkan bahwa ada beberapa fitur yang memiliki korelasi kuat satu sama lain. Dengan memahami korelasi antara fitur, kita dapat membuat model prediktif yang lebih baik dan menghindari masalah seperti multikolinearitas dalam model regresi. Tabel korelaso Pearson sebgai berikut:



Gambar 3. Feature Correlation Heatmap

Adapun dari hasil analisa tabel Pearson diketahui bahwa tidak ada korelasi yang tinggi antara kuat tekan beton dan fitur-fitur lainnya kecuali Semen, yang memang seharusnya demikian untuk mendapatkan kekuatan yang lebih besar. Usia dan Superplastisizer adalah dua fitur lain yang memiliki korelasi yang kuat dengan kuat tekan beton. Superplastisizer tampaknya memiliki korelasi negatif yang tinggi dengan Air, dan korelasi positif dengan Fly Ash dan Agregat Halus.

Hal ini berarti bahwa semen adalah komponen utama yang menentukan kuat tekan beton, karena semen berfungsi sebagai bahan pengikat yang menyatukan agregat-agregat lainnya. Usia juga berpengaruh, karena semakin lama beton mengeras, semakin kuat pula ikatan antara semen dan agregat. Superplastisizer adalah bahan tambahan yang digunakan untuk meningkatkan kelenturan dan kemudahan pengecoran beton tanpa menambahkan Air. Air berperan sebagai bahan pengaktif Semen, tetapi jika terlalu banyak, dapat mengurangi kuat tekan beton. Fly ash dan agregat halus adalah bahan-bahan pengganti atau pelengkap semen yang dapat meningkatkan kuat tekan beton dengan cara mengisi ruang-ruang kosong (pori) antara agregat kasar dan mengurangi porositas beton. Selanjutnya kita dapat menganalisis korelasi-korelasi ini secara visual dengan membuat plot hubungan-hubungan tersebut sebagai berikut:



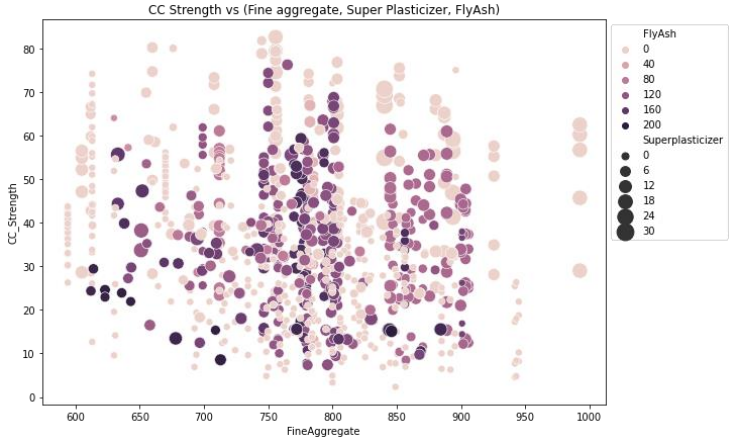
Gambar 4. Plot Hubungan (Semen, Umur, Air)

Kekuatan tekan meningkat seiring dengan jumlah semen. Kekuatan tekan juga meningkat seiring dengan usia beton. Penggunaan semen yang rendah memerlukan lebih banyak semen untuk mencapai kekuatan yang lebih tinggi. Semakin banyak semen yang digunakan, semakin banyak air yang dibutuhkan. Kekuatan beton meningkat ketika digunakan lebih sedikit air dalam praktiknya.

Dalam hal ini, beberapa faktor dapat mempengaruhi kekuatan tekan beton.

1. Jumlah Semen: Semen adalah komponen utama dalam campuran beton yang berfungsi sebagai pengikat. Oleh karena itu, semakin banyak semen yang digunakan dalam campuran, semakin kuat ikatan yang terbentuk antara agregat, sehingga meningkatkan kekuatan tekan beton.
2. Usia: Beton membutuhkan waktu untuk mengeras dan mencapai kekuatan penuhnya. Proses ini dikenal sebagai pengerasan beton. Semakin lama beton mengeras, semakin kuat strukturnya, sehingga kekuatan tekannya meningkat.
3. Kebutuhan Semen pada Usia Muda: Beton yang masih muda atau baru dicor memerlukan lebih banyak semen untuk mencapai kekuatan yang lebih tinggi. Ini karena semen baru mulai bereaksi dengan air dan proses pengerasan belum selesai.
4. Kebutuhan Air pada Semen: Semen yang lebih banyak cenderung memerlukan lebih banyak air dibandingkan dengan penggunaan semen yang sedikit. Ini mungkin disebabkan oleh penyerapan air oleh semen selama proses pengerasan.
5. Penggunaan Air dalam Pengerjaan Beton: Air berperan penting dalam pembuatan campuran beton karena air diperlukan untuk reaksi kimia pada hidrasi. Namun, penggunaan air yang berlebihan dapat mengurangi kekuatan tekan beton karena dapat menciptakan porositas dalam struktur beton setelah air menguap.

Selanjutnya melihat hubungan antara parameter degan variable agregat halus, superplasticizer dan juga fly ash sebagai berikut:



Gambar 5. Plot Hubungan (Agregat halus, SP, dan Flyash)

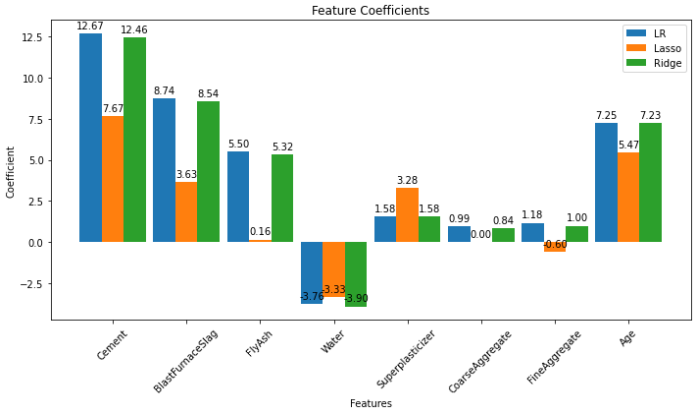
Seiring peningkatan fly ash, kekuatan tekan menurun. Kekuatan tekan juga menurun seiring peningkatan jumlah air. Namun kekuatan tekan beton meningkat seiring peningkatan superplastisizer dan penggunaan lebih sedikit air.
Dalam konteks campuran beton, fly ash adalah bahan tambahan atau bahan substitusi semen yang biasanya digunakan untuk meningkatkan kekuatan dan daya tahan beton. Namun, jika digunakan dalam jumlah yang berlebihan, fly ash dapat mengurangi kekuatan tekan beton karena sifatnya yang lebih ringan dan kurang padat dibandingkan dengan semen.
Lalu superplastisizer adalah bahan kimia yang ditambahkan ke dalam campuran beton untuk meningkatkan kelenturan dan kemudahan pengecoran (*workability*) tanpa perlu menambahkan lebih banyak air. Oleh karena itu, peningkatan superplastisizer dapat meningkatkan kekuatan tekan beton. Sedangkan penggunaan agregat halus seperti pasir berfungsi untuk mengisi ruang kosong antara agregat kasar seperti kerikil dalam campuran beton. Oleh karena itu, penggunaan lebih banyak agregat halus dapat meningkatkan kepadatan dan kekuatan tekan beton itu sendiri.

Kita membuat kesimpulan dari EDA dengan mengamati plot titik-titik, namun ada interaksi non-linier yang tersembunyi antara fitur-fitur yang tidak dapat kita lihat secara visual. Interaksi ini dapat mempengaruhi hasil analisis kita. Kita dapat memahami plot secara visual seperti yang ditunjukkan, namun kita juga dapat menggunakan fitur plotting baris-demi-baris dan kolom-demi-kolom dari seaborn untuk melakukan analisis lebih mendalam, namun kita masih belum mampu untuk melacak semua korelasi ini dengan sendirinya. Korelasi ini dapat memberikan informasi penting tentang hubungan antara fitur-fitur. Untuk alasan ini, kita dapat menggunakan Machine Learning untuk menangkap hubungan-hubungan ini dan memberikan wawasan yang lebih baik tentang masalah ini. Machine Learning adalah metode yang dapat belajar dari data dan menemukan pola-pola yang tersembunyi di dalamnya

3.2 Estimasi Kualitas Model
Regresi linear adalah metode yang digunakan untuk membangun model prediksi dengan mengasumsikan bahwa hubungan antara variabel independen dan variabel dependen adalah linear. Dalam regresi linear, terdapat tiga jenis yaitu regresi linear tanpa regularisasi, regresi linear dengan regularisasi L1 (Lasso Regression), dan regresi linear dengan regularisasi L2 (Ridge Regression). Ketiga jenis regresi linear ini dievaluasi dengan membandingkan Root Mean Squared Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE), dan R2 Score.

Model	RMSE	MSE	MAE	R2
Linear Regression	10,29	105,78	8,23	0,57
Lasso Regression	10.68	114,13	8,66	0,54
Ridge Regression	10,29	105,86	8,24	0,57

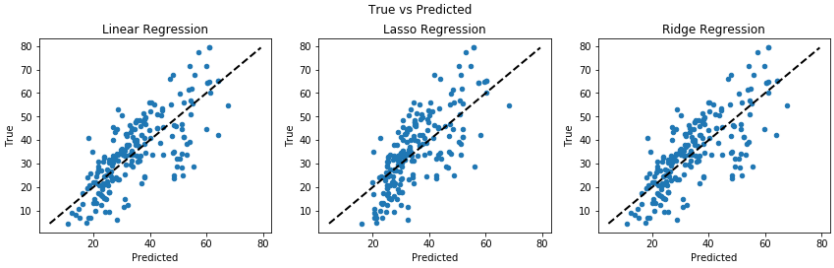
Hasil evaluasi menunjukkan bahwa ketiga jenis regresi linear memiliki performa yang hampir sama. Namun, Lasso Regression dapat mengurangi kompleksitas model dengan menjaga koefisien sekecil mungkin. Sedangkan Linear Regression dan Ridge Regression memiliki koefisien yang hampir sama. Dalam memilih jenis regresi linear yang tepat, perlu diperhatikan bahwa Lasso Regression dapat mengurangi performa model.



Gambar 6. Feature Correlation

Dari hasil analisa plotting koefisien menunjukan bahwa Regresi Lasso adalah teknik yang efektif dalam mengurangi kompleksitas model dengan menjaga koefisien sekecil mungkin. Dengan melakukan ini, Regresi Lasso dapat membantu dalam menyelesaikan masalah overfitting dengan meminimalkan koefisien yang tidak penting. Dengan demikian, model dapat lebih mudah diinterpretasikan dan lebih stabil.
Selain itu, koefisien dengan Regresi Linier dan Regresi Ridge hampir sama. Ini menunjukkan bahwa Regresi Lasso dapat mempertahankan kinerja model sambil mengurangi kompleksitasnya. Meskipun demikian, penting untuk dicatat bahwa Regresi Lasso memiliki keunggulan dalam hal seleksi fitur, karena dapat menghilangkan fitur yang tidak relevan dengan menetapkan koefisiennya menjadi nol.

Selanjutnya melakukan analisis plotting prediksi dari ketiga model tersebut sebagai berikut:

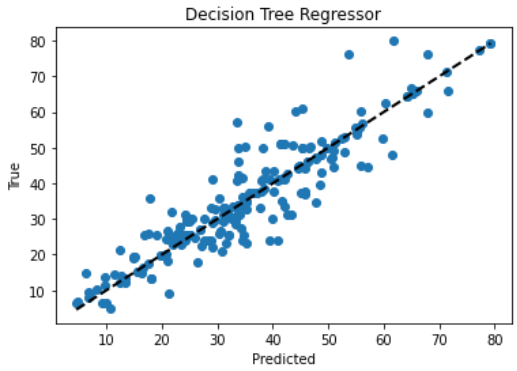


Gambar 7. Plotting Analisis

Ini adalah serangkaian tiga plot titik yang membandingkan nilai prediksi dan nilai sebenarnya dari variabel target. Plot pertama adalah untuk Regresi Linier, plot kedua untuk Regresi Lasso, dan plot ketiga untuk Regresi Ridge. Sumbu x mewakili nilai prediksi dan sumbu y mewakili nilai sebenarnya. Semakin dekat titik-titik dengan garis hitam, semakin kecil kesalahannya.
Jika kita amati grafik antara nilai prediksi dan nilai aktual dari variabel target, kita dapat menyimpulkan bahwa Regresi Linier dan Regresi Ridge berkinerja baik karena prediksinya lebih dekat dengan nilai sebenarnya. Sementara itu, Regresi Lasso mengurangi kompleksitas model dengan biaya penurunan kinerja dalam kasus ini. (Semakin dekat titik-titik dengan garis hitam, semakin kecil kesalahannya.)

Selain regresi linear, terdapat algoritma lain yang dapat digunakan dalam machine learning yang digunakan, yaitu Decision Trees an Random Forest Regressor. Decision Trees adalah algoritma yang membangun model prediksi dalam bentuk pohon keputusan.

Model	RMSE	MSE	MAE	R2
Decision Tree	6,44	41,47	4,33	0,83



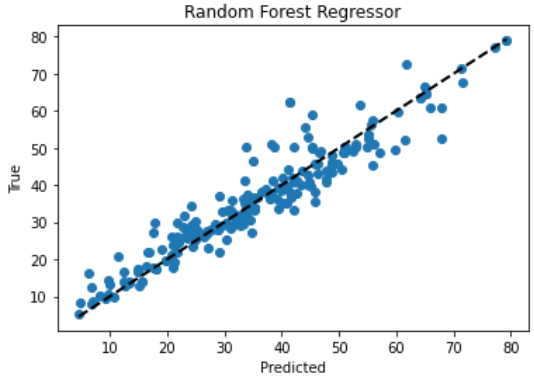
Gamabr 8. Decision Tree

Nilai Root Mean Squared Error (RMSE) telah turun dari 10.29 menjadi 7.31, sehingga Decision Trees telah meningkatkan kinerja secara signifikan. Penurunan ini menunjukkan bahwa model ini dapat memprediksi dengan lebih akurat dibandingkan sebelumnya. Ini adalah peningkatan yang signifikan, mengingat bahwa RMSE adalah ukuran yang sangat sensitif terhadap perubahan dalam prediksi model. Semakin rendah nilai RMSE, semakin baik model kita dalam memprediksi data yang belum pernah dilihat sebelumnya.

Peningkatan ini juga dapat diamati dalam plot. Dalam plot, kita dapat melihat bahwa lebih banyak titik berada di garis, yang menunjukkan bahwa prediksi kita lebih dekat dengan nilai sebenarnya. Garis ini mewakili tempat di mana prediksi kita sama dengan nilai sebenarnya. Oleh karena itu, semakin banyak titik yang berada di garis, semakin akurat prediksi tersebut. Dengan demikian, kita dapat menyimpulkan bahwa Decision Trees telah berhasil meningkatkan kinerja model secara signifikan dibandingkan tiga model regresi yang sudah dibahas sebelumnya.

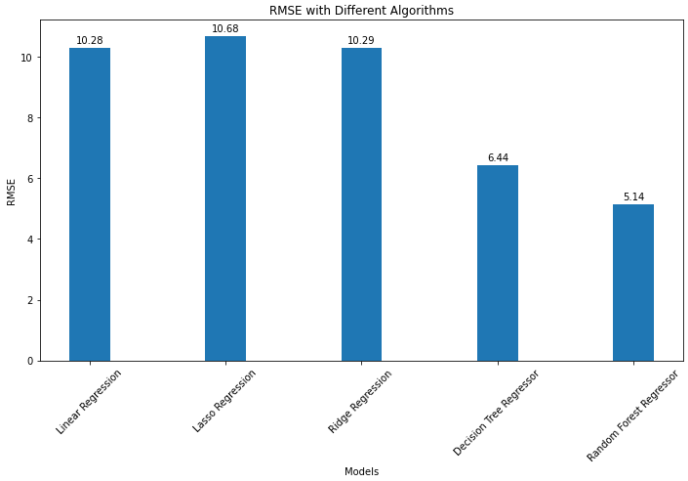
Selanjutnya algoritma yang digunakan terakhir adalah Random Forest Regressor. Random forest regressor adalah algoritma yang menghasilkan banyak pohon keputusan secara acak dan menggabungkan hasil prediksi dari setiap pohon untuk menghasilkan prediksi akhir.

Model	RMSE	MSE	MAE	R2
Rando Forest Regressor	5,41	26,45	3,56	0,89



Gambar 9. Random Forest Regressor

Ini adalah grafik scatter plot yang menunjukkan hasil dari metode Random Forest Regressor. Garis regresi ditampilkan dalam warna hitam. Data ditampilkan dalam warna biru. Sumbu x menunjukkan nilai yang diprediksi, sedangkan sumbu y menunjukkan nilai sebenarnya. Terdapat beberapa data yang menyimpang dari garis regresi, namun secara umum data menyebar dengan baik di sekitar garis regresi, yang menunjukkan bahwa model kita bekerja dengan baik. Selain itu, nilai Root Mean Squared Error (RMSE) dengan Random Forest Regressor sekarang adalah 5.11. Hal ini menunjukkan bahwa metode ensemble seperti Random Forest Regressor dapat meningkatkan kinerja model dengan menggabungkan prediksi dari beberapa pohon keputusan. Dengan demikian, model dapat memanfaatkan kekuatan kolektif dari pohon-pohon tersebut untuk membuat prediksi yang lebih akurat. Hasil evaluasi menunjukkan bahwa Random Forest Regressor memiliki performa yang terbaik dibandingkan dengan empat algoritma yang telah dibahas sebelumnya.



Gambar 10. RMSE 5 Algoritma Model

Dalam memilih algoritma yang tepat untuk memprediksi kekuatan beton, dapat dilakukan perbandingan antara performa dari setiap algoritma. Hasil perbandingan menunjukkan bahwa Random Forest Regressor adalah algoritma yang paling baik dalam memprediksi kekuatan beton

4. KESIMPULAN

Berdasarkan analisis yang telah dilakukan, dapat disimpulkan bahwa model Random Forest Regressor berkinerja terbaik dalam memprediksi Kekuatan Tekan Beton. Model ini dilatih dan dievaluasi menggunakan data yang telah diproses dan distandarisasi. Hasil menunjukkan bahwa model ini mencapai RMSE terendah dan Skor R2 tertinggi, menunjukkan bahwa model mampu menangkap pola kompleks dalam data dan membuat prediksi yang akurat. Secara keseluruhan, analisis ini memberikan wawasan berharga bagi praktisi dan dapat digunakan untuk perencanaan mix design beton.

DAFTAR PUSTAKA

[1] S. Pushkar, “Life-Cycle Assessment of the Substitution of Sand with Coal Bottom Ash in Concrete: Two Concrete Design Methods,” *Applied Sciences*, vol. 9, no. 17, p. 3620, Sep. 2019, doi: 10.3390/app9173620.

[2] A. H. Hossein, H. AzariJafari, and R. Khoshnazar, “The role of performance metrics in comparative LCA of concrete mixtures incorporating solid wastes: A critical review and

guideline proposal," *Waste Management*, vol. 140, pp. 40–54, Mar. 2022, doi: 10.1016/j.wasman.2022.01.010.

[3] ACI, "Fly Ash Concrete Mix Design and the Water-Cement Ratio Law," *ACI Mater J*, vol. 91, no. 4, 1994, doi: 10.14359/4050.

[4] C. Argiz, A. Moragues, and E. Menéndez, "Use of ground coal bottom ash as cement constituent in concretes exposed to chloride environments," *J Clean Prod*, vol. 170, pp. 25–33, Jan. 2018, doi: 10.1016/j.jclepro.2017.09.117.

[5] H. Mostafaei, B. Badarloo, N. F. Chamasemani, M. A. Rostampour, and P. Lehner, "Investigating the Effects of Concrete Mix Design on the Environmental Impacts of Reinforced Concrete Structures," *Buildings*, vol. 13, no. 5, p. 1313, May 2023, doi: 10.3390/buildings13051313.

[6] A. P. Abriantoro, "PENGARUH FLY ASH SEBAGAI SUBSTITUSI SEMEN TERHADAP DURASI INITIAL SETTING TIME, FLOWABILITY DAN KUAT TEKAN UMUR 1 HARI BETON SELF-COMPACTING CONCRETE (SCC) DENGAN PENAMBAHAN 0,15% CITRIC ACID," *JURNAL KAJIAN TEKNIK SIPIL*, vol. 7, no. 2, pp. 32–43, Jun. 2023, doi: 10.52447/jkts.v7i2.6852.

[7] A. Putra Abriantoro and E. Susilowati, "EKSPERIMEN SUBSTITUSI SERBUK KACA PADA BINDER UNTUK MENINGKATKAN KUAT TEKAN BETON SELF-COMPACTING CONCRETE (SCC) PADA USIA 1 HARI".

[8] S. Bhanja and B. Sengupta, "Investigations on the compressive strength of silica fume concrete using statistical methods," *Cem Concr Res*, vol. 32, no. 9, pp. 1391–1394, Sep. 2002, doi: 10.1016/S0008-8846(02)00787-1.

[9] T. Kim, J. M. Davis, M. T. Ley, S. Kang, and P. Amrollahi, "Fly ash particle characterization for predicting concrete compressive strength," *Constr Build Mater*, vol. 165, pp. 560–571, Mar. 2018, doi: 10.1016/j.conbuildmat.2018.01.059.

[10] H. Naderpour and M. Mirrashid, "Estimating the compressive strength of eco-friendly concrete incorporating recycled coarse aggregate using neuro-fuzzy approach," *J Clean Prod*, vol. 265, p. 121886, Aug. 2020, doi: 10.1016/j.jclepro.2020.121886.

[11] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput Sci*, vol. 2, no. 3, p. 160, May 2021, doi: 10.1007/s42979-021-00592-x.

[12] M. V. Kamath, S. Prashanth, M. Kumar, and A. Tantri, "Machine-Learning-Algorithm to predict the High-Performance concrete compressive strength using multiple data," *Journal of Engineering, Design and Technology*, Feb. 2022, doi: 10.1108/JEDT-11-2021-0637.

[13] H. M. R. Hasan, A. S. A. Rabby, M. T. Islam, and S. A. Hossain, "Machine Learning Algorithm for Student's Performance Prediction," in *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, IEEE, Jul. 2019, pp. 1–7. doi: 10.1109/ICCCNT45670.2019.8944629.

[14] P. Marcelino, M. de Lurdes Antunes, E. Fortunato, and M. C. Gomes, "Machine learning approach for pavement performance prediction," *International Journal of Pavement Engineering*, vol. 22, no. 3, pp. 341–354, Feb. 2021, doi: 10.1080/10298436.2019.1609673.

[15] D. Nemirovsky, T. Arkose, N. Markovic, M. Nemirovsky, O. Unsal, and A. Cristal, "A Machine Learning Approach for Performance Prediction and Scheduling on Heterogeneous CPUs," in *2017 29th International Symposium on Computer Architecture*

and High Performance Computing (SBAC-PAD), IEEE, Oct. 2017, pp. 121–128. doi: 10.1109/SBAC-PAD.2017.23.