

## Pemilihan Algoritma Machine Learning Yang Optimal Untuk Prediksi Sifat Mekanik Aluminium

<sup>(1)\*</sup>Desmarita Leni

<sup>(1)</sup>Teknik Mesin, Fakultas Teknik, Universitas Muhammadiyah Sumatera Barat.

Jl. Pasir Jambak No.4, Pasie Nan Tigo, Kec. Koto Tengah, Kota Padang, Sumatera Barat 25586, Indonesia

\*Email: desmaritaleni@gmail.com

Diterima: 07.04.2023, Disetujui: 12.05.2023, Diterbitkan: 24.05.2023

### ABSTRACT

*This study designs and compares optimal machine learning models to predict the mechanical properties of aluminum, including Yield Strength (YS) and Tensile Strength (TS), based on the percentage composition of aluminum's chemical elements. The machine learning modeling in this study has nine input variables consisting of aluminum chemical elements such as Al, Mg, Zn, Ti, Cu, Mn, Cr, Fe, Si, and two output or target variables consisting of YS and TS. Additionally, Heatmap correlation is used to observe the correlation between chemical elements and the mechanical properties of aluminum. Three machine learning algorithms, namely Decision Tree (DT), Random Forest (RF), and Artificial Neural Network (ANN), are compared in this study. The comparison of these algorithms shows that Random Forest (RF) outperforms the other algorithms in predicting YS with MAE of 11.44, RMSE of 14.282, and R value of 0.93. On the other hand, ANN performs better in predicting TS with MAE of 19.593, RMSE of 22.005, and R value of 0.947.*

**Key words :** Model, Machine Learning, Aluminum, Tensile Strength, Algorithm

### ABSTRAK

Penelitian ini merancang dan membandingkan pemodelan *machine learning* yang optimal untuk memprediksi sifat mekanik aluminium diantaranya adalah *Yield Strength* (YS) dan *Tensile Strength* (TS), berdasarkan persentase komposisi unsur kimia aluminium. Pemodelan *machine learning* pada penelitian ini memiliki 9 variabel masukan yang terdiri dari unsur kimia aluminium seperti, Al, Mg, Zn, Ti, Cu, Mn, Cr, Fe, Si, dan 2 output atau target yang terdiri dari YS dan TS, selain itu untuk melihat korelasi antara unsur kimia dan sifat mekanik aluminium digunakan Heatmap correlation. Dalam penelitian ini dibandingkan 3 algoritma *machine learning* yang terdiri dari *Decision Tree* (DT), *Random Forest* (RF), dan *Artificial Neural Network* (ANN). Hasil perbandingan ketiga algoritma diperoleh bahwa, *Random Forest* (RF) memiliki kinerja lebih baik dalam memprediksi nilai YS dengan nilai MAE 11.44, RMSE 14.282, dan R 0.93, sedangkan ANN memiliki kinerja lebih baik dalam memprediksi nilai TS dengan nilai MAE 19.593, RMSE 22.005, dan R 0.947.

**Kata kunci :** Pemodelan, Machine learning, Aluminium, Tensile Strength, Algoritma

### I. Pendahuluan

Aluminium adalah jenis material non ferrous yang paling banyak digunakan dalam berbagai aplikasi industri modern seperti industri dirgantara, struktural dan otomotif. Menurut data (Usgs, 2022), badan survei geologis Amerika Serikat (AS) atau US *Geological Survey*, produksi aluminium di seluruh dunia pada tahun 2021 mencapai 68 juta metrik ton. Jumlah tersebut naik 4,45% dibandingkan produksi tahun sebelumnya yang berjumlah 65,1 juta metrik ton (Data aluminium 2022). Penggunaan aluminium yang luas di berbagai aspek industri modern mendorong *engineer* untuk lebih teliti dalam mengetahui sifat mekanik aluminium, hal ini

bertujuan untuk mencegah terjadinya kegagalan pada material, sifat mekanik suatu material memiliki peran yang penting dalam menentukan bahan untuk komponen industri modern demi mencegah terjadinya kegagalan terhadap komponen industri secara prematur (Branco, 2018). Kekuatan tarik dan kekuatan luluh merupakan dua diantara sifat mekanik material, sifat mekanik material dipengaruhi oleh beberapa faktor seperti struktur mikro, komposisi kimia dan berbagai perlakuan panas (*heat treatment*) (Antonio Augusto Morini, Manuel J. Ribeiro, 2019). Sifat mekanik material dapat ditingkatkan dengan cara menambahkan atau mengurangi unsur kimia tertentu sesuai kebutuhan, seperti penambahan

unsur nikel (Ni) untuk meningkatkan kekuatan rekat pengelasan pada aluminium (Branco, 2018). Perbedaan komposisi kimia aluminium mempengaruhi nilai kekuatan tarik aluminium, dengan kata lain apabila ditambahkan satu unsur kimia pada komposisi aluminium maka nilai kekuatan tarik pada aluminium tersebut akan ikut berubah sehingga perlu dilakukan pengujian ulang (George Krauss, 2015). Pengujian sifat mekanik aluminium seperti kekuatan tarik dan dan kekuatan luluh pada umumnya masih dilakukan secara manual menggunakan mesin uji tarik. Banyak model dan variasi unsur kimia aluminium yang bertujuan untuk membentuk sifat mekanik aluminium sesuai kebutuhan kerja mengakibatkan proses pengujian secara manual menjadi kurang efektif secara waktu dan biaya.

Perkembangan teknologi komputer yang begitu pesat di bidang ilmu material mendorong para ahli dan peneliti dalam mengembangkan pendekatan komputasi dalam menganalisa dan memecahkan berbagai permasalahan di bidang material (Hutchinson et al., 2011). Perkembangan teknologi berbasis komputasi di bidang material memungkinkan untuk dilakukannya pengujian material tanpa harus merusak spesimen. Beberapa tahun terakhir, kompleksitas masalah teknik telah mendorong peningkatan penerapan metode *machine learning* yang memanfaatkan algoritma matematika untuk belajar cepat dari pola yang diperkenalkan sebelumnya, dan teknik ini dapat berhasil memperoleh hubungan timbal balik yang kompleks antara beberapa parameter dan dengan cepat memprediksi output yang diinginkan (Ling Qiao, Zibo Wang, 2020). *Machine learning* adalah metode analisis statistik yang efisien untuk menangkap hubungan internal linear atau nonlinier dengan belajar dari data empiris (Ling Qiao, Zibo Wang, 2020), saat ini ada berbagai macam jenis algoritma *machine learning* yang digunakan untuk mengetahui sifat mekanik material dari komposisi paduannya dan perlakuan panas tanpa harus merusak spesimen. Penggunaan algoritma *machine learning* seperti Random Forest, Neural

Network, dan Decision Tree dalam memprediksi kekuatan tarik baja memberikan hasil prediksi yang baik (Amiri et al., 2020) selain itu, metode *machine learning* juga terbukti lebih praktis dalam memprediksi

kekuatan leleh material berdasarkan lembar data kelelahan (Kevin P. Murphy, 2012).

Ragam variasi dan model komposisi unsur kimia aluminium sangat mempengaruhi nilai kekuatan tarik dan kekuatan luluh aluminium, apabila ditambahkan sedikit unsur kimia pada komposisi aluminium maka kekuatan tarik dan kekuatan luluh aluminium tersebut akan ikut berubah. Selain komposisi kimia, perlakuan *heat treatment* juga mempengaruhi kekuatan tarik dan kekuatan luluh sehingga pengujian sifat mekanik aluminium secara manual membutuhkan waktu yang lama, biaya yang mahal, membutuhkan keahlian dan tidak ramah lingkungan karna menghasilkan gas emisi untuk mendaur ulang limbah aluminium hasil uji tarik. Mesin Uji Tarik Universal (UTM) digunakan untuk menguji aluminium untuk mengetahui sifat mekanik baja seperti kekuatan tarik dan kekuatan luluh, tetapi ketika material harus diuji pada suhu tinggi yang bervariasi dibutuhkan lebih banyak waktu untuk menghitung sifat mekanik material. Penelitian ini bertujuan untuk merancang sebuah model *machine learning* menggunakan beberapa algoritma seperti Random Forest (RF), Neural Network (NN), dan Decision Tree (DT) untuk memprediksi kekuatan tarik dan kekuatan luluh aluminium. Penelitian ini memberikan beberapa variasi parameter pada masing-masing algoritma *machine learning* dan membandingkan nilai performance seperti Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), dan correlation coefficient (R).

## II. Metode Penelitian

Data pada penelitian ini diambil pada periode September 2022 dari *website matmatch.com* yang merupakan sebuah website informasi material bahan terbesar didunia, website ini terdapat menyediakan 70.000 lebih data material dan di portal website ini, dimungkinkan untuk mengakses informasi disediakan oleh ribuan pemasok bahan dari berbagai jenis, termasuk paduan aluminium (Sandhya, N., Sowmya, V., Bandaru, C. R., & Babu, 2019). *Website* ini merupakan perpustakaan bahan online akses terbuka, yang terdiri dari ribuan entri. Untuk setiap record, dimungkinkan untuk memperoleh datasheet (secara umum, heterogen dan tidak lengkap) yang harus disaring, diatur dan diproses untuk mendapatkan korpus informasi yang akurat dan berguna (Merayo

et al., 2020). Data yang tersedia seperti sifat mekanik material, komposisi kimia, korosi, perlakuan panas, sifat termal dan masih banyak lagi, namun pada penelitian ini data yang diambil adalah data komposisi kimia aluminium, nilai uji tarik aluminium (TS) dan kekuatan luluh aluminium (YS). Data komposisi kimia aluminium yang diambil pada penelitian ini adalah presentase kadar Al, Mg, Zn, Ti, Cu, Mn, Cr, Fe, Si yang dapat dilihat

pada tabel 1. Pengambilan data aluminium ini menggunakan web scraper, web scrapper adalah sebuah *platform open source* yang memungkinkan *user* untuk mengambil data dari sebuah website dan dapat di download dalam bentuk file Csv, pada penelitian ini menggunakan dataset Ayang berjumlah 120 data aluminium dengan 9 data input dan 2 data output

Tabel 1. Statistik dataset aluminium

Keterangan	Tipe data	Min	Max	Mean
Mg (%)	Input	0	3.1	0.978
Zn (%)	Input	0.05	0.500	0.207
Ti (%)	Input	0	0.250	0.069
Cu (%)	Input	0	0.500	0.130
Mn (%)	Input	0.02	1	0.404
Cr (%)	Input	0	0.3	0.11
Fe (%)	Input	0.150	1	0.576
Si (%)	Input	0.15	1	0.436
Al (°C)	Input	94.650	99.58	97.089
UTS (Mpa)	Output	55	230	112.408
YS (Mpa)	Output	90	375	232.95

### 1. Perbandingan Algoritma *Machine Learning*

Perbandingan algoritma *machine learning* ini dibandingkan menggunakan bahasa pemrograman python yang dijalankan pada google colab. Penelitian ini menggunakan tiga algoritma *machine learning* yang berbeda untuk merancang pemodelan untuk memprediksi Sifat mekanik aluminium. Heatmap correlation digunakan untuk melihat seberapa besar hubungan unsur kimia dan sifat mekanik aluminium, dengan menggunakan heatmap *correlation* dapat memberikan informasi seberapa besar pengaruh unsur kimia terhadap YS dan TS (D Leni, F Earnestly, R Sumiati, A Adriansyah, 2023). Algoritma yang digunakan pada penelitian ini adalah Random Forest (RF), Artificial Neural Network (ANN), dan Decision Tree (DT). Pada penelitian ini digunakan 80% data sebagai data training dan 20% digunakan sebagai data testing kemudian performa model yang dilatih dievaluasi dengan menghitung *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan *Correlation Coefficient* (R), dari nilai aktual dan nilai prediksi dari data uji.

#### a. Random Forest (RF)

Random forest merupakan salah satu metode *machine learning* yang di perkenalkan oleh leo breiman dan deleculte. Random Forest (RF) (Weinbub et al., 2015) adalah metode pembelajaran ensemble untuk klasifikasi yang digunakan dengan membangun pohon keputusan dari set pelatihan dalam iterasi k. Dalam setiap iterasi, algoritma pelatihan pertamamata secara acak memilih satu set sampel dari set pelatihan. Untuk mereproduksi pohon keputusan dari subset ini, RF secara acak memilih subset fitur sebagai kandidat fitur untuk setiap node. Dengan demikian, setiap pohon keputusan dibangun melalui ansambel menggunakan himpunan bagian independen acak dari fitur dan sampel. Pada penelitian ini algoritma DT menggunakan 80% data sebagai data *training* dan 20% digunakan sebagai data testing kemudian performa model yang dilatih dievaluasi dengan menghitung *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan *Correlation Coefficient* (R), dari nilai aktual dan nilai prediksi dari data uji.

b. Artificial Neural Network (ANN)

Neural network merupakan salah satu metode *machine learning classifier* atau metode yang sangat *powerfull*. Neural network hampir selalu cocok dengan berbagai permasalahan machine learning sehingga neural network dapat menjadi hypotheses untuk berbagai persoalan rill sebab Neural network dibentuk berdasarkan cara kerja syaraf otak manusia. Pada penelitian ini algoritma DT menggunakan 80% data sebagai data training dan 20% digunakan sebagai data testing kemudian performa model yang dilatih dievaluasi dengan menghitung Root Mean Squared Error (RMSE), *Mean Absolute Error* (MAE), dan *Correlation Coefficient* (R), dari nilai aktual dan nilai prediksi dari data uji.

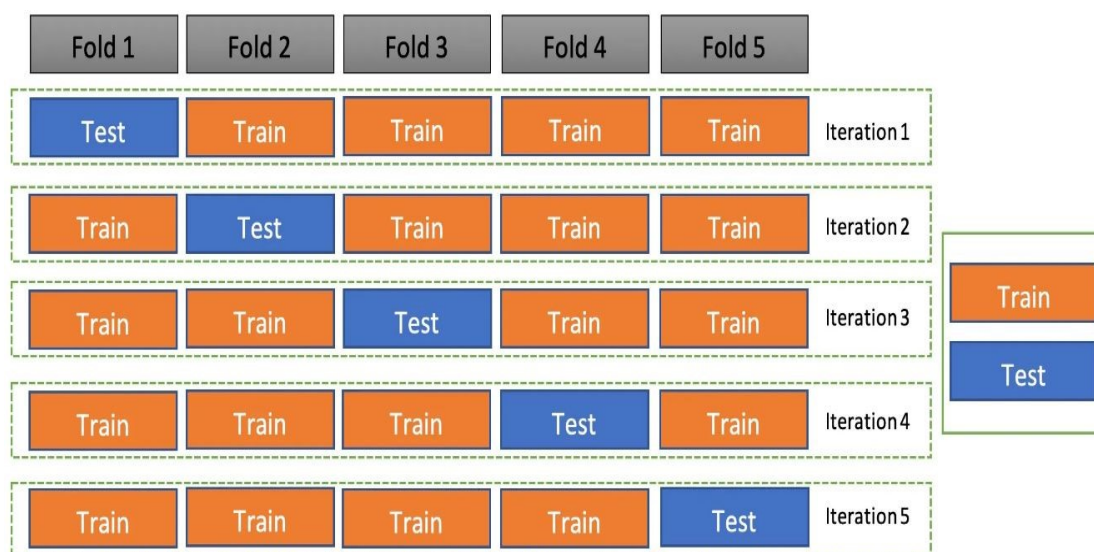
c. Decision Tree (DT)

*Decion tree reggression* merupakan salah satu predictive model yang digunakan pada *machine learning*, model ini adalah teknik terawasi yang melakukan teknik klasifikasi dan regresi. Pohon keputusan bekerja dengan baik untuk input kategorikal dan kontinu, variabel output. Ini mengembangkan pohon keputusan terkait dengan membagi dataset menjadi subset yang lebih kecil. Pada penelitian ini algoritma DT menggunakan 80% data sebagai data training dan 20% digunakan sebagai data testing kemudian performa model yang dilatih dievaluasi dengan menghitung *Root Mean*

*Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan *Correlation Coefficient* (R), dari nilai aktual dan nilai prediksi dari data uji.

2. Cross validation

Pemodelan dalam penelitian ini menggunakan pendekatan berdasarkan pencarian grid dan validasi silang k-fold, untuk mencegah terjadinya *overfitting*, oleh sebab itu proses pemilihan nilai parameter yang sesuai dari metode pembelajaran mesin memberikan dampak yang cukup besar pada akurasi. Selanjutnya, nilai optimal untuk parameter dapat bervariasi sesuai dengan masalahnya. Pencarian Grid adalah strategi untuk penyesuaian parameter otomatis dan optimal dari model. Teknik ini membangun mesh dari set nilai yang telah ditentukan untuk setiap parameter. Untuk setiap kemungkinan kombinasi parameter, model prediktif dilatih dengan beberapa data, menghasilkan satu set keluaran dan nilai parameter terbaik akan menghasilkan set keluaran terbaik (Agrawal et al., 2014). Pada langkah pelatihan, menggunakan validasi silang atau k-Fold, yang membagi kumpulan data menjadi k set. Model dilatih pada k-1 set dan divalidasi dengan bagian yang tersisa. Langkah-langkah pelatihan dan pengujian diulang k kali secara bergantian antara set pelatihan dan pengujian. Gambar 1 mengilustrasikan penerapan validasi silang k-Fold. Dalam penelitian ini menggunakan k = 5.



Gambar 1. Ilustrasi set pelatihan (orange) dan pengujian (biru) untuk k = 5

### 3. Evaluasi Model

Pemodelan dalam penelitian ini memiliki 9 variabel data input dan 2 data output sebagai target, setiap pemodelan yang dihasilkan dari setiap algoritma di evaluasi dengan menghitung nilai *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *Correlation Coefficient* (R), berdasarkan nilai aktual dan nilai prediksi dari data uji yang dapat dilihat pada perasamaan berikut:

$$MAE = \frac{1}{N} \sum |y_i - z_i| \quad (1)$$

Dimana  $i$  adalah indeks dari data pada sample,  $N$  adalah jumlah total sample,  $y_i$  adalah nilai aktual dari data ke- $i$ , sedangkan  $z_i$  adalah nilai prediksi dari model untuk data ke- $i$ .

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f(X_i) - Y_i)^2} \quad (2)$$

Dimana  $n$  adalah jumlah data yang digunakan untuk menguji model,  $f(X_i)$  adalah nilai yang diprediksi oleh model untuk data ke- $i$ ,  $Y_i$  adalah nilai sebenarnya untuk data ke- $i$ .

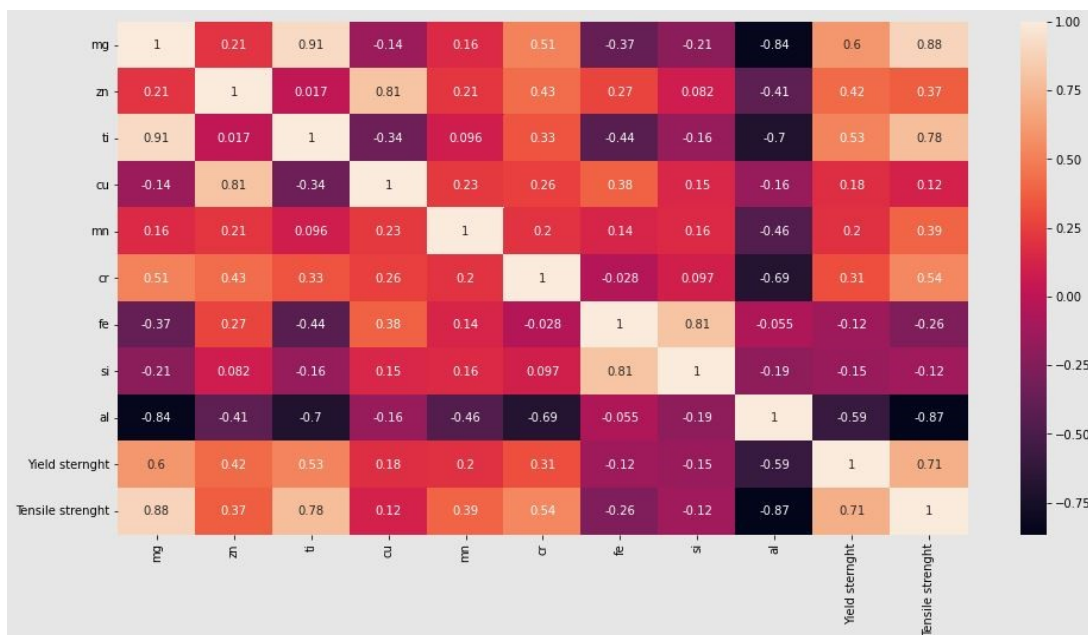
$$R = \frac{\sum_{i=1}^n (f(X_i) - \bar{f}(X)) (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (f(X_i) - \bar{f}(X))^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (3)$$

Dimana  $f(X_i)$  adalah nilai prediksi dari variabel dependen ( $Y$ ) berdasarkan variabel

independen ( $X$ ) pada observasi ke- $i$ ,  $\bar{f}(X)$  adalah rata-rata dari semua nilai prediksi  $f(X_i)$  pada seluruh observasi,  $Y_i$  adalah nilai observasi aktual dari variabel dependen pada observasi ke- $i$ ,  $\bar{Y}$  adalah rata-rata dari semua nilai observasi  $Y_i$  pada seluruh observasi, dan  $n$  adalah jumlah total observasi.

### III. Hasil dan Pembahasan

Pemodelan dibuat di google colab menggunakan bahasa pemrograman python, pemodelan dirancang dengan 9 variabel input yang terdiri dari unsur kimia paduan dan 2 variabel output yaitu YS dan TS. Hasil dari pemodelan ini dalam bentuk angka, yang dapat memprediksi nilai YS dan TS. Langkah pertama yang dilakukan adalah melihat korelasi antar variabel menggunakan *heatmap correlation* yang dapat dilihat pada Gambar 2. Hasil visualisasi dengan *heatmap correlation* dapat dilihat bahwa unsur Mg memiliki korelasi yang paling besar terhadap YS, kemudian diikuti oleh unsur Ti sebesar 0.53, unsur Zn 0.42, dan Cr sebesar 0.31. Kekuatan tarik (TS) sangat dipengaruhi oleh unsur Mg dengan korelasi terbesar yaitu 0.88, pada urutan kedua terdapat unsur Ti dengan korelasi 0.78 dan diikuti oleh unsur Cr sebesar 0.54 dan Mn sebesar 0.39.



Gambar 2. Heatmap correlation

### 1. Perbandingan algoritma machine learning

Pemodelan dirancang dengan tiga algoritma machine learning yaitu Random Forest (RF), Artificial Neural Network (ANN), dan *Decision Tree* (DT), setiap algoritma dicari parameter terbaik dengan cara memvariasikan setiap parameter pada algoritma yang dapat dilihat perlakuan

parameternya pada tabel 2. Pada Tabel 2 di kolom pertama dapat dilihat ada tiga algoritma yang digunakan yaitu DT, RF, dan ANN, kolom kedua nama parameter yang akan diberikan variasi berbeda, dan kolom yang ketiga adalah pengaturan parameter pada setiap algoritma.

Tabel 2. Perlakuan variasi parameter untuk mencari parameter terbaik.

Metode	Nama parameter	Pengaturan parameter
DT	Depth	1,2,3,4,5,6,7,8,9,10,15,20,25,30
	Number of tree	20,50, 80, 110, 140,150
RF	Maximal depth	5,10,15,25,30
	Training cycle	10.000, 20,000, 30.000, 40,000
ANN	Learning rate	0,001, sd 0,1
	Momentum	0.1 sd 1
	Neuron Hidden Layer 1	5, 10, 15, 20, 25, 30,40, 50
	Neuron Hidden Layer 2	5, 10, 15, 20, 25, 30,40, 50

Hasil pemodelan ketiga algoritma ini dengan berbagai variasi parameter terbaik sebagai berikut :

a. *Decision Tree* (DT)

Perlakuan variasi parameter pada pemodelan algoritma *Decision Trees* (DT) didapatkan parameter terbaik dengan jumlah 10 number of trees dengan menghasilkan nilai MAE 16,59, RMSE 20,267, dan R 0,854 pada pemodelan kekuatan luluh aluminium (YS) berdasarkan nilai aktual dan nilai prediksi dari data uji. Penambahan jumlah of trees diatas 10 hingga 20 memberikan nilai konstan terhadap nilai MAE, RMSE dan R sedangkan jumlah trees diatas 20 mengakibatkan kenaikan nilai terhadap nilai peforma, peforma terbaik adalah nilai terkecil dari MAE dan MRSE. Pemodelan DT ini menghasilkan nilai MAE 23,048, RMSE 29,021, dan R 0,854 pada pemodelan prediksi kekuatan tarik (TS) berdasarkan nilai aktual dan nilai prediksi dari data uji.

b. Randon forest (RF)

Pada pemodelan menggunakan algoritma random forest (RF) ada 2 parameter yang di variasikan diantaranya number of trees dan maximal depth yang dapat dilihat pada tabel 2. Peningkatan number of trees diatas 20 mengakibatkan kenaikan nilai RMSE dan penambahan jumlah maximal depth diatas 10 tidak memberikan dampak terhadap hasil permance. Pemodelan menggunakan algoritma RF didapatkan parameter terbaik number of trees 20 dan maximal depth 10 dengan nilai MAE 11,44, RMSE 14,282, dan R 0,93 pada pemodelan kekuatan luluh aluminium (YS). Pemodelan RF ini menghasilkan nilai MAE 21,669, RMSE 27,301, dan R 0,871 pada uji kekuatan tarik (TS) berdasarkan nilai aktual dan nilai prediksi dari data uji.

c. Artificial Neural Network (ANN)

Pemodelan menggunakan algoritma ANN ada 4 variasi parameter yang terdiri dari *training cycle*, *learning rate*, momentum dan size hidden layer. Peningkatan *training cycle* diatas 30.000 tidak memberikan dampak terhadap nilai parameter, RMSE konstan 19,111 begitu pula dengan nilai MAE dan R. Peningkatan nilai nilai *learning rate* mengakibatkan penurunan *performance*, dan nilai momentum diatas 0.9 meengakibatkan terjadinya error, variasi percobaan pada hidden layer didapatkan nilai terbaik pada 2 hidden layer dan 20 neuron. Pemodelan ANN didapatkan parameter terbaik dengan training cycle 30.000, *learning rate* 0,001 momentum 0,1, dan 2 hidden layer dan 20 neuron, yang menghasilkan nilai MAE 23,045, RMSE 28,448, dan R 0,865 pada pemodelan kekuatan luluh aluminium. Pemodelan ANN ini juga menghasilkan nilai MAE 19,593, RMSE 22,005, dan R 0,947 dari nilai aktual dan nilai prediksi dari data uji.

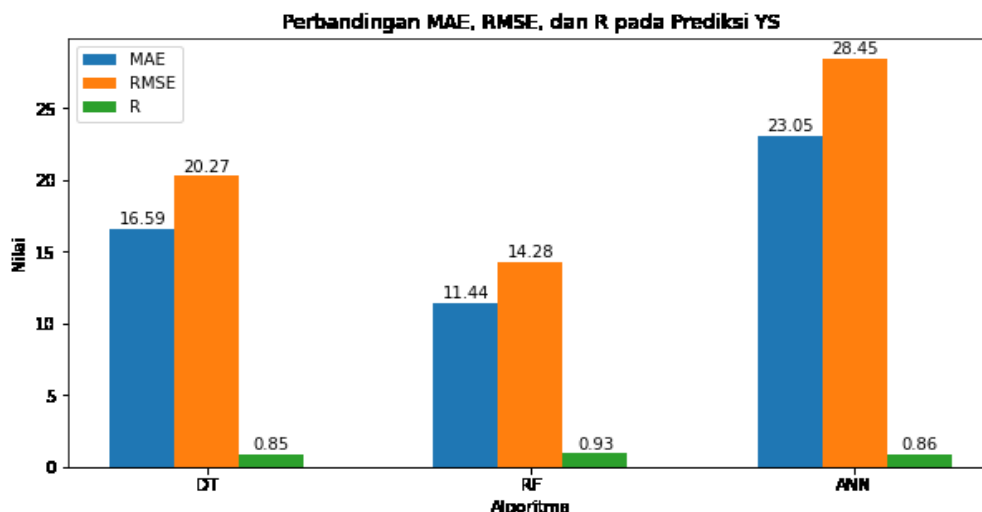
Berdasarkan hasil pengujian ketiga algoritma *machine learning* yang sudah dilakukan dapat dilihat bahwa, RF memiliki kinerja lebih baik dibandingkan DT dan ANN dalam memprediksi kekuatan luluh aluminium (YS), sedangkan untuk memprediksi kekuatan tarik aluminium (TS) ANN lebih baik dibandingkan DT dan RF. Hasil ini sejalan dengan hasil penelitian (Hasan, 2021), dimana Penelitian ini juga membandingkan hasil performa lima algoritma machine learning yang berbeda: K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), dan Gradient Boosting Machine (GBM). Analisis performa menunjukkan bahwa model machine learning dapat memprediksi gesekan dan keausan logam paduan Al dengan memuaskan dari data variabel bahan dan uji tribologi.

Lebih lanjut, hasil analisis menunjukkan bahwa metode Random Forest lebih unggul dalam memprediksi laju keausan dibandingkan dengan algoritma machine learning lainnya.

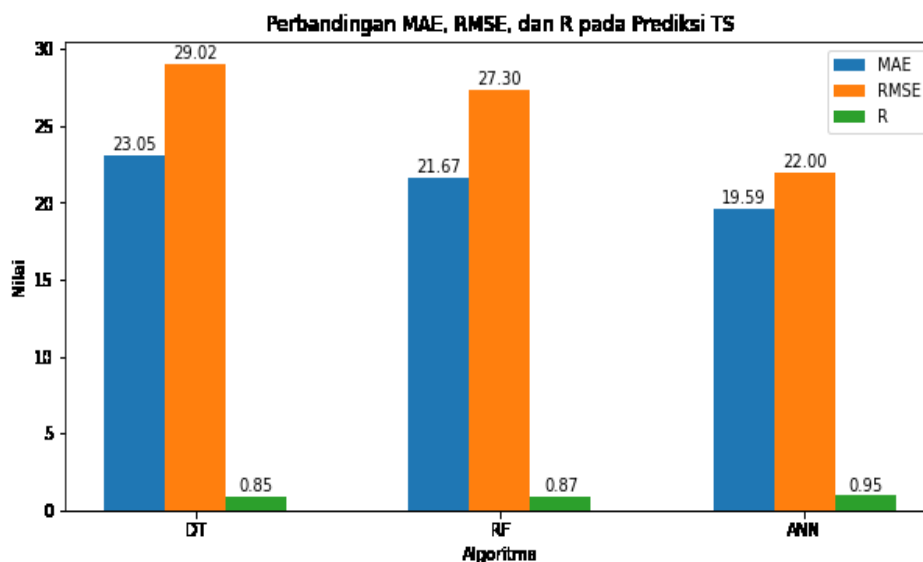
**2. Cross validation**

Cross validation adalah salah satu metode dalam validasi model yang melibatkan pembagian data menjadi beberapa subset atau fold. Pada setiap iterasi, satu fold dijadikan sebagai data validasi dan fold yang lain sebagai data latih (Prasetyo & Laksana, 2022). Pada pemodelan ini digunakan cross validation dengan

jumlah K 5, dimana data dibagi menjadi 5 subset atau fold yang berbeda dan dilakukan iterasi 5 kali dengan memilih setiap subset secara bergantian sebagai data uji dan subset lainnya sebagai data latih. Kemudian, untuk setiap iterasi, nilai performa model dihitung dengan menggunakan metrik evaluasi seperti MAE, RMSE, dan R-squared. Setelah dilakukan 5 iterasi, nilai performa model pada setiap fold diambil rata-ratanya untuk mendapatkan nilai performa model yang stabil dan konsisten. Dengan menggunakan metode k-fold cross validation ini, validitas model prediksi dapat diuji secara lebih akurat dan reliabel karena memperhitungkan variasi dari data yang berbeda-beda dalam setiap fold. Selain itu, dengan menggunakan k 5, jumlah data pelatihan yang digunakan untuk melatih model lebih banyak, sehingga hasil evaluasi model menjadi lebih akurat dan terpercaya (Intan et al., 2021). Hasil validasi model ketiga algoritma machine learning ini, tidak mengalami perbedaan dengan hasil pengujian sebelumnya. Hasil ini menunjukkan bahwa model yang digunakan memiliki performa yang stabil dan dapat diandalkan dalam melakukan prediksi pada dataset yang lebih besar dan beragam. Hasil perbandingan cross validation dengan jumlah k 5 untuk ketiga algoritma dapat dilihat pada gambar 3 untuk YS dan gambar 4 untuk TS.



Gambar 3. Perbandingan performa pemodelan prediksi YS



Gambar 4. Perbandingan performa pemodelan prediksi TS

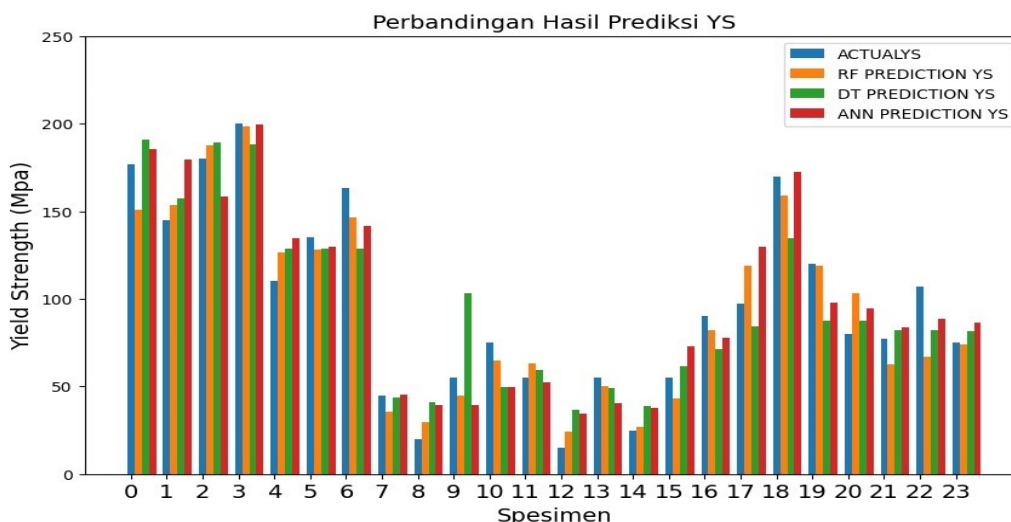
### 3. Evaluasi model

Evaluasi model bertujuan untuk mengetahui seberapa baik model yang dihasilkan dalam melakukan prediksi (Desmarita Leni, Yuda Perdana Kusuma, Ruzita Sumiati, Muchlisinalahuddin, 2022). Hasil pemodelan dengan perbandingan ketiga algoritma didapatkan bahwa algoritma Random Forest (RF) memiliki kinerja lebih baik dibandingkan algoritma Decision Trees (DT) dan Artificial Neural Network (ANN) dalam memprediksi kekuatan luluh aluminium (YS), sedangkan untuk memprediksi kekuatan tarik aluminium (TS) ANN lebih baik dibandingkan DT dan RF. Hasil ini tidak jauh berbeda dengan hasil

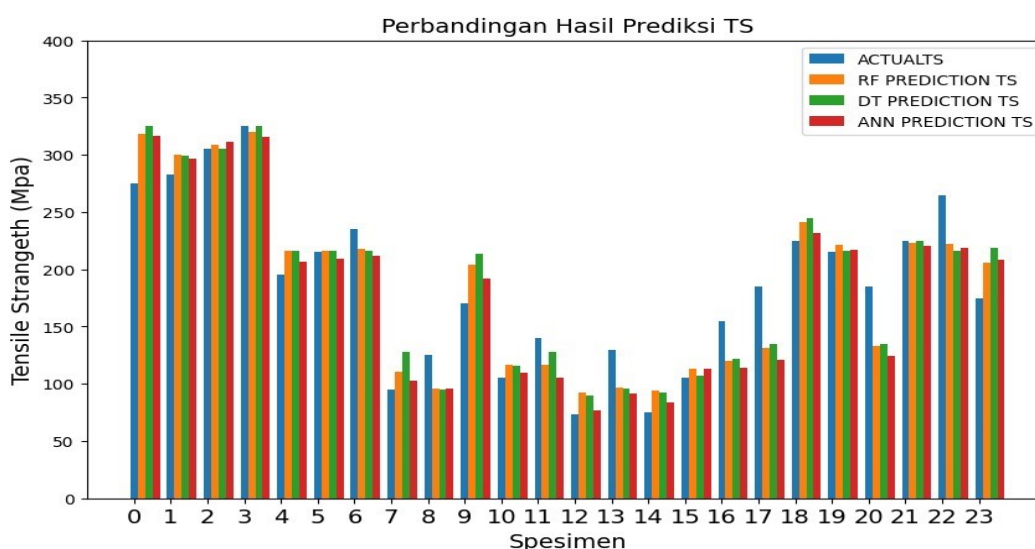
penelitian yang dilakukan oleh (D. Merayo, 2020), bahwa ANN dapat digunakan untuk memprediksi sifat mekanik logam, khususnya kekuatan luluh dan kekuatan tarik ultimat pada paduan aluminium. Pemodelan prediksi ini berdasarkan komposisi kimia paduan dan suhu temper yang diperoleh dari big data, hasil penelitian menunjukkan bahwa ANN mampu memberikan hasil prediksi kekuatan tarik sesuai dengan data experimental aluminium.

Hasil perbandingan prediksi ketiga algoritma machine learning dalam memprediksi sifat mekanik aluminium dapat dilihat pada Gambar 5 untuk YS dan Gambar 6 untuk TS.





Gambar 5. Grafik perbandingan prediksi kekuatan luluh(YS)



Gambar 6. Grafik perbandingan prediksi kekuatan tarik (TS)

#### IV. Kesimpulan

Berdasarkan hasil penelitian tentang pemodelan prediksi sifat mekanik aluminium menggunakan metode machine learning dapat disimpulkan bahwa, heatmap correlation mampu memberikan informasi unsur kimia yang paling berkorelasi dengan sifat mekanik aluminium seperti YS dan TS. Hasil perbandingan ketiga algoritma machine learning yang terdiri dari Decision Tree (DT), Random Forest (RF), dan Artificial Neural Network (ANN), diperoleh hasil pemodelan yang berbeda yaitu RF memiliki kinerja lebih baik dalam memprediksi YS, sedangkan ANN memiliki kinerja lebih baik dalam memprediksi TS. Hasil ini menunjukkan bahwa, pemilihan algoritma yang tepat dapat meningkatkan

akurasi prediksi sifat mekanik aluminium. Penelitian ini dapat memberikan pandangan yang lebih mendalam tentang penggunaan *machine learning* dalam pemodelan prediksi sifat mekanik logam, khususnya aluminium, yang dapat berguna dalam bidang industri dan rekayasa material.

#### Daftar Pustaka

- Agrawal, A., Deshpande, P. D., Cecen, A., Basavarsu, G. P., Choudhary, A. N., & Kalidindi, S. R. (2014). Exploration of data science techniques to predict fatigue strength of steel from composition and processing parameters. *Integrating Materials and Manufacturing Innovation*, 3(1), 90–108. <https://doi.org/10.1186/2193-9772-3-8>

- Amiri, N., Farrahi, G. H., Kashyzadeh, K. R., & Chizari, M. (2020). Applications of ultrasonic testing and machine learning methods to predict the static & fatigue behavior of spot-welded joints. *Journal of Manufacturing Processes*, 52, 26–34. <https://doi.org/10.1016/j.jmapro.2020.01.047>
- Antonio Augusto Morini, Manuel J. Ribeiro, D. H. (2019). Early-stage materials selection based on embodied energy and carbon footprint. *Materials & Design*, 178, 107861. <https://www.sciencedirect.com/science/article/pii/S0264127519302990>
- Branco, R. B. (2018). *Mechanical Behaviour of Aluminium Alloys*. MDPI Applied Sciences.
- D. Merayo, A. R.-P. and A. M. C. (2020). Prediction of Physical and Mechanical Properties for Metallic Materials Selection Using Big Data and Artificial Neural Networks. *IEEE Access*, 8, 13444–13456. <https://doi.org/10.1109/ACCESS.2020.2965769>.
- D Leni, F Earnestly, R Sumiati, A Adriansyah, Y. K. (2023). Evaluasi sifat mekanik baja paduan rendah berdasarkan komposisi kimia dan suhu perlakuan panas menggunakan teknik exploratory data analysis ( EDA ). *Dinamika Teknik Mesin*, 13(1), 74–83.
- Desmarita Leni, Yuda Perdana kusuma, Ruzita Sumiati, Muchlisinalahuddin, A. (2022). *Perbandingan Alogaritma Machine Learning Untuk Prediksi Sifat Mekanik Pada Baja Paduan Rendah*. 5(2), 167–174. <https://jurnal.umsu.ac.id/index.php/RMME/article/view/11407>
- George Krauss. (2015). *Steels: processing, structure, and performance* (Second edi). ASM International. [https://books.google.co.id/books?hl=id&lr=&id=ETJ7CgAAQBAJ&oi=fnd&pg=PR11&dq=Krauss,+G.+\(2015\).+Steels:+Processing,+Structure,+and+Performance%3B+ASM+International.+Russell,+OH,+USA:+Asm+International&ots=Nyj46WRGs0&sig=Htf3twiJgA2VdS2oWAGzCKog8dI&redir\\_esc=y#v=onepage&q&f=false](https://books.google.co.id/books?hl=id&lr=&id=ETJ7CgAAQBAJ&oi=fnd&pg=PR11&dq=Krauss,+G.+(2015).+Steels:+Processing,+Structure,+and+Performance%3B+ASM+International.+Russell,+OH,+USA:+Asm+International&ots=Nyj46WRGs0&sig=Htf3twiJgA2VdS2oWAGzCKog8dI&redir_esc=y#v=onepage&q&f=false)
- Hutchinson, B., Hagström, J., Karlsson, O., Lindell, D., Tornberg, M., Lindberg, F., & Thuvander, M. (2011). Microstructures and hardness of as-quenched martensites (0.1-0.5%C). *Acta Materialia*, 59(14), 5845–5858. <https://doi.org/10.1016/j.actamat.2011.05.061>
- Intan, I., Ghani, S. A. D., Nurdin, & Koswara, A. T. C. (2021). Performance Analysis of Weather Forecasting using Machine Learning Algorithms. *Jurnal Pekommas*, 6(2), 1–8. <https://doi.org/10.30818/jpkm.2021.2060221>
- Kevin P. Murphy. (2012). *Machine learning: a probabilistic perspective*. The MIT Press. [https://books.google.co.id/books?hl=id&lr=&id=RC43AgAAQBAJ&oi=fnd&pg=PR7&dq=Murphy,+K.+P.+\(2012\).+Machine+learning:+a+probabilistic+perspective.+MIT+press.&ots=umnxcDOv5b&sig=\\_dgX1nN92fUKC0vjj2dUrdRort8&redir\\_esc=y#v=onepage&q=Murphy%2C+K.+P.+\(2012\).+Machine+learning%3A+a+probabilistic+perspective.+MIT+press.&f=false](https://books.google.co.id/books?hl=id&lr=&id=RC43AgAAQBAJ&oi=fnd&pg=PR7&dq=Murphy,+K.+P.+(2012).+Machine+learning:+a+probabilistic+perspective.+MIT+press.&ots=umnxcDOv5b&sig=_dgX1nN92fUKC0vjj2dUrdRort8&redir_esc=y#v=onepage&q=Murphy%2C+K.+P.+(2012).+Machine+learning%3A+a+probabilistic+perspective.+MIT+press.&f=false)
- Ling Qiao, Zibo Wang, J. Z. (2020). Application of improved GRNN model to predict interlamellar spacing and mechanical properties of hypereutectoid steel. *Materials Science and Engineering, A*, 792, 139845. <https://doi.org/https://doi.org/10.1016/j.msea.2020.139845>.
- Merayo, D., Rodríguez-Prieto, A., & Camacho, A. M. (2020). Prediction of mechanical properties by artificial neural networks to characterize the plastic behavior of aluminum alloys. *Materials*, 13(22), 1–22. <https://doi.org/10.3390/ma13225227>
- Prasetyo, A. B., & Laksana, T. G. (2022). Optimasi Algoritma K-Nearest Neighbors dengan Teknik Cross Validation Dengan Streamlit (Studi Data: Penyakit Diabetes). *Journal of Applied Informatics and Computing (JAIC)*, 6(2), 194. <http://jurnal.polibatam.ac.id/index.php/JAIC>
- Sandhya, N., Sowmya, V., Bandaru, C. R., & Babu, G. R. (2019). Prediction of mechanical properties of steel using data science techniques. *Int. J. Recent Technol. Eng*, 235–241.
- Usgs. (2022). *effect-aluminium-and-sodium-impurities-vitro-toxicity-and-pro-inflammatory-potential*. <https://www.usgs.gov/publications/effect-aluminium-and-sodium-impurities-vitro-toxicity-and-pro-inflammatory-potential%0A>
- Weinbub, J., Wastl, M., Rupp, K., Rudolf, F., & Selberherr, S. (2015). ViennaMaterials - A dedicated material library for computational science and engineering. *Applied Mathematics and Computation*, 267, 282–293. <https://doi.org/10.1016/j.amc.2015.03.094>