

PREDIKSI PENJUALAN SUKU CADANG MOTOR DENGAN PENERAPAN RANDOM FOREST DI PT TERUS JAYA SENTOSA MOTOR

Sandi Reza Febrian, Asril Adi Sunarto, Agung Pambudi

Teknik Informatika, Universitas Muhammadiyah Sukabumi

Jl. R. Syamsudin, S.H. No. 50, Cikole, Kec. Cikole, Kota Sukabumi, Jawa Barat 43113.

sandireza006@ummi.ac.id

ABSTRAK

Suku cadang kendaraan memiliki peran vital dalam menjaga agar kendaraan tetap berfungsi optimal. PT Terus Jaya Sentosa Motor adalah perusahaan yang berfokus pada penjualan suku cadang motor Honda. Suku cadang ini memiliki berbagai kode barang yang berbeda karena banyaknya jenis dan variasi komponen motor Honda. Untuk meningkatkan penjualan, memahami suku cadang yang paling diminati oleh konsumen menjadi sangat penting. Algoritma *Random Forest* dapat digunakan untuk memprediksi penjualan suku cadang tersebut. *Random Forest* adalah teknik data mining yang membangun beberapa pohon keputusan secara acak dan menggabungkan hasil prediksi dari setiap pohon untuk mendapatkan hasil akhir yang lebih akurat. Penelitian sebelumnya membuktikan bahwa algoritma *Random Forest* sangat efektif dalam memprediksi harga jual rumah dan suku cadang mobil terlaris dengan tingkat akurasi yang tinggi. Dalam penelitian ini, dataset sebanyak 3528 data digunakan dengan pendekatan metodologi CRISP-DM. Hasil uji model menunjukkan nilai *Root Mean Squared Error* (RMSE) sebesar 23.38, *Mean Absolute Error* (MAE) sebesar 13.97, dan *Mean Squared Error* (MSE) sebesar 546.49, walaupun nilai MSE cukup besar, model ini tetap dianggap baik karena nilai RMSE dan MAE yang relatif kecil. Model prediksi ini diimplementasikan dalam sebuah situs web untuk mempermudah pengguna dalam memprediksi suku cadang motor di PT Terus Jaya Sentosa Motor.

Kata kunci : Suku Cadang, Random Forest, Prediksi Penjualan, Web, CRISP-DM.

1. PENDAHULUAN

Industri otomotif merupakan salah satu sektor ekonomi yang memiliki dampak luas dan signifikan di seluruh dunia. Kendaraan memainkan peran penting dalam memenuhi kebutuhan mobilitas manusia dan barang, menjadi bagian integral dari kehidupan sehari-hari sejak ditemukannya kendaraan bermesin seperti mobil dan sepeda motor. Penemuan ini memicu evolusi dalam desain teknologi dan peran sosial-ekonomi. Dalam beberapa dekade terakhir, industri otomotif mengalami transformasi signifikan yang dipicu oleh inovasi teknologi yang terus berkembang, adaptasi terhadap perubahan tren konsumen, dan respons terhadap tantangan global. Upaya terus-menerus untuk meningkatkan kinerja dan keamanan kendaraan telah mendorong industri ini untuk mengadopsi berbagai teknologi canggih, termasuk kecerdasan buatan (AI), ke dalam kendaraan.

Selain menjadi pusat inovasi dan perkembangan teknologi, industri otomotif juga berperan penting dalam perekonomian global dengan menciptakan banyak lapangan kerja dan mendukung pertumbuhan sektor terkait seperti manufaktur, logistik, dan jasa. Salah satu aspek penting dari industri otomotif adalah suku cadang kendaraan.

Suku cadang kendaraan berperan vital dalam menjaga kinerja kendaraan agar tetap optimal. Banyak perusahaan bergerak di bidang penjualan suku cadang kendaraan, salah satunya adalah PT Terus Jaya Sentosa Motor yang fokus pada suku cadang motor Honda. Suku cadang motor Honda memiliki beragam kode barang karena banyaknya tipe motor Honda yang beredar dan berbagai jenis suku cadang yang

dibutuhkan. Pengetahuan tentang suku cadang dengan permintaan tertinggi sangat penting untuk bisnis ini karena mempengaruhi tingkat penjualan. Stok barang yang memadai menjadi kunci sukses dalam meningkatkan penjualan, namun pada saat ini manajemen stok kurang optimal sehingga diperlukannya prediksi penjualan agar dapat memperkirakan stok barang yang dibutuhkan di masa mendatang.

Pada penelitian sebelumnya [1], Prediksi harga rumah dengan menggunakan metode Random Forest Regression menunjukkan bahwa model yang paling optimal terdiri dari 320 pohon keputusan. Dengan hasil akurasi 94.23% dan *Mean Absolute Percent Error* (MAPE) 26.48%.

Dan pada penelitian lain [2], Hasil prediksi menunjukkan bahwa akurasi penjualan suku cadang terlaris mencapai 80,00%. Perbandingan dilakukan antara prediksi penjualan suku cadang terlaris untuk merek Isuzu dan Mitsubishi.

Dengan demikian, penggunaan algoritma Random Forest untuk memprediksi penjualan suku cadang motor di PT Terus Jaya Sentosa Motor dapat dianggap sebagai metode yang efektif dan sesuai.

2. TINJAUAN PUSTAKA

2.1. Suku Cadang Motor

Suku cadang motor merupakan bagian-bagian atau komponen yang digunakan untuk memperbaiki atau mengganti bagian dari sepeda motor yang rusak. Suku cadang ini penting untuk menjaga kinerja, keamanan dan kenyamanan sepeda motor.

Penggunaan suku cadang yang sesuai dan berkualitas dapat memperpanjang umur sepeda motor.

2.2. Prediksi

Prediksi adalah perkiraan mengenai peristiwa yang mungkin terjadi di masa depan. Dalam bidang ilmu sosial, ketidakpastian sering kali sulit untuk diprediksi dengan tepat. Prediksi bertujuan untuk mengurangi pengaruh ketidakpastian dan meminimalkan kesalahan dalam peramalan [3].

Prediksi adalah upaya memperkirakan nilai variabel di masa depan. Ada tiga jenis prediksi: jangka pendek, menengah, dan panjang. Prediksi jangka pendek fokus pada pola data dalam periode singkat, sementara prediksi jangka menengah dan panjang digunakan untuk perencanaan strategis, seperti ekspansi dan kebutuhan jangka panjang [4].

2.3. Random Forest

Random Forest bekerja dengan membagi *dataset* menjadi sejumlah *subset* dan menerapkan algoritma pohon keputusan pada masing-masing *subset*. Struktur *Random Forest* terdiri dari *root node*, *internal node*, dan *leaf node*, yang masing-masing mewakili simpul dalam pohon keputusan. *Root node* adalah simpul paling atas, sering disebut sebagai akar, sementara *internal node* adalah simpul yang bercabang dan memiliki setidaknya dua keluaran. *Leaf node*, atau terminal, adalah simpul terakhir dalam pohon keputusan yang hanya memiliki satu keluaran dan tidak memiliki anak [5].

Adapun rumus dalam proses *random forest* :

a. Mean Squared Error (MSE)

MSE adalah metrik yang digunakan untuk mengukur keakuratan model prediksi dengan menghitung rata-rata dari kuadrat selisih antara nilai prediksi dan nilai aktual. Rumus MSE adalah sebagai berikut:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (1)$$

Keterangan :

- Y_i nilai sebenarnya.
- $\hat{Y}_i$ nilai yang diprediksi model.
- n jumlah total observasi.

MSE menilai kesalahan prediksi dengan memberikan bobot yang lebih besar pada kesalahan yang lebih besar, karena perhitungan berdasarkan kuadrat perbedaan. Nilai MSE yang lebih kecil menunjukkan bahwa model lebih baik dalam memprediksi nilai aktual [6].

b. Mean Absolute Error (MAE)

MAE adalah metrik yang digunakan untuk menilai akurasi suatu model dalam memprediksi nilai yang diamati dengan menghitung rata-rata dari nilai absolut perbedaan antara prediksi dan nilai sebenarnya. Rumus MAE adalah sebagai berikut:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (2)$$

Keterangan :

- Y_i nilai sebenarnya.

- $\hat{Y}_i$ nilai yang diprediksi model.

- n jumlah total observasi.

Nilai MAE yang lebih rendah mengindikasikan bahwa model lebih akurat dalam memprediksi nilai sebenarnya [6].

c. Root Mean Squared Error (RMSE)

RMSE adalah ukuran yang digunakan untuk mengevaluasi seberapa akurat sebuah model statistik dalam memprediksi nilai yang diamati. RMSE dihitung dengan cara mengambil akar kuadrat dari rata-rata perbedaan kuadrat antara nilai yang diprediksi dan nilai sebenarnya. Rumus RMSE adalah sebagai berikut:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (3)$$

Keterangan :

- Y_i nilai sebenarnya.
- $\hat{Y}_i$ nilai yang diprediksi model.
- n jumlah total observasi.

Keterangan :

- Y_i nilai sebenarnya.
- $\hat{Y}_i$ nilai yang diprediksi model.
- n jumlah total observasi.

RMSE memberikan gambaran tentang kesalahan prediksi dalam satuan yang sama dengan variabel yang diprediksi. Semakin kecil nilai RMSE, semakin baik model tersebut dalam memprediksi nilai sebenarnya [7].

2.4. CRISP-DM

Cross-Industry Standard Process for Data Mining (CRISP-DM) adalah kerangka kerja yang sering digunakan dalam data mining dan analisis data. Kerangka ini terdiri dari langkah-langkah terstruktur, mulai dari pemahaman bisnis hingga evaluasi hasil, untuk membantu para profesional mengembangkan solusi data mining yang efektif. Langkah-langkah dalam CRISP-DM meliputi:

a. Business Understanding

Tahap awal dalam proses *data mining* dan analisis data, yang bertujuan untuk memahami tujuan bisnis dan mengidentifikasi masalah yang harus diselesaikan menggunakan data. Pada tahap ini, berfokus pada memahami kebutuhan bisnis, merumuskan masalah dalam konteks data, serta menetapkan tujuan proyek yang akan mengarahkan seluruh proses data mining.

b. Data Understanding

Langkah ini adalah proses analisis data yang bertujuan untuk mendapatkan pemahaman mendalam tentang data yang akan digunakan dalam proyek. Pada tahap ini, data dianalisis secara menyeluruh untuk mengidentifikasi karakteristik, pola, dan struktur yang relevan. Selain itu, kualitas dan kegunaan data juga dievaluasi. Pemahaman data ini menjadi landasan penting dalam pengambilan keputusan yang akurat dan berbasis analisis yang informatif.

c. Data Preparation

Data preparation adalah tahap penting dalam analisis data, di mana data mentah dibersihkan dan disiapkan untuk analisis. Ini melibatkan pemilihan fitur, penghapusan data yang tidak relevan, dan transformasi data, dengan tujuan memastikan data berkualitas tinggi dan siap digunakan untuk menghasilkan hasil yang akurat.

d. Data Modelling

Langkah ini melibatkan pengembangan model Random Forest menggunakan data yang telah dipersiapkan. Pada tahap ini, hyperparameter seperti jumlah pohon (*n_estimators*), kedalaman maksimum pohon (*max_depth*), dan jumlah fitur yang dipertimbangkan saat membagi simpul (*max_features*) diatur. Kemudian, model dilatih dengan data yang telah disiapkan.

e. Evaluation

Evaluasi adalah tahap yang sangat penting untuk menilai kualitas dan kinerja model yang telah dibangun. Evaluasi kinerja *Random Forest* menggunakan metrik evaluasi RMSE, MSE, MAE dan R-Squared.

f. Deployment

Proses deployment mencakup integrasi model prediksi ke dalam aplikasi web dengan menggunakan *Flask*. Ini bertujuan untuk mempermudah akses dan penggunaan model tersebut melalui antarmuka situs web [8].

2.5. Bahasa Pemrograman Python

Python adalah bahasa pemrograman yang interaktif dan berorientasi objek, dirancang untuk menawarkan struktur data tingkat tinggi. Bahasa ini dikenal karena menggabungkan kemampuan tinggi dengan sintaks yang jelas, menjadikannya alat yang efisien dan ramah pengguna untuk pengembangan perangkat lunak. Dengan Python, pemula dapat dengan mudah mempelajari dan mengembangkan program tanpa harus menghadapi kompleksitas sintaks dari bahasa pemrograman yang lebih rumit [9].

2.6. Google Colaboratory

Google Colaboratory, atau Google Colab, adalah alat penelitian gratis yang memanfaatkan layanan cloud. Alat ini dirancang untuk mempermudah penelitian dengan menyediakan lingkungan pengembangan yang mirip dengan Jupyter Notebook, namun dengan kemampuan akses online yang ditingkatkan. Google Colab menawarkan solusi yang efektif dan ekonomis untuk penelitian dengan memanfaatkan keuntungan dari teknologi cloud [10].

3. METODE PENELITIAN

3.1. Business Understanding

Pada tahap ini, pemahaman bisnis dilakukan untuk mengidentifikasi tujuan penelitian yang ingin dicapai. Dalam konteks prediksi penjualan suku cadang motor, salah satu masalah yang dihadapi adalah pengelolaan stok barang yang tidak terkontrol,

yang dapat mempengaruhi hasil penjualan. Jika persediaan stok barang terlalu rendah, perusahaan mungkin kesulitan memenuhi permintaan pelanggan, yang dapat menyebabkan kehilangan penjualan dan kekecewaan pelanggan. Sebaliknya, jika persediaan dikelola berdasarkan prediksi penjualan yang terlalu tinggi, perusahaan dapat menghadapi penumpukan stok yang tidak terjual, yang berpotensi mengakibatkan biaya penyimpanan yang tinggi dan kerugian finansial.

3.2. Data Understanding

Data yang digunakan dalam penelitian ini diperoleh dari PT. Terus Jaya Sentosa Motor untuk periode 2021 hingga 2023. Data tersebut mencakup beberapa atribut, termasuk tahun dan bulan penjualan, kode barang, serta jumlah suku cadang yang terjual. Dengan menganalisis variabel-variabel ini, penelitian dapat dilakukan untuk memperkirakan jumlah barang yang akan terjual di tahun berikutnya dan mengidentifikasi pola penjualan yang dapat mempengaruhi kinerja bisnis.

3.3. Data Preparation

Tahapan selanjutnya adalah mempersiapkan dataset penjualan suku cadang motor dari PT. Terus Jaya Sentosa Motor. Beberapa langkah yang dilakukan dalam mempersiapkan data meliputi:

- Mengatasi nilai yang hilang: Langkah ini melibatkan pemeriksaan setiap variabel dalam data penjualan suku cadang untuk memastikan tidak ada nilai yang kosong. Jika ditemukan nilai yang hilang, dapat diatasi dengan menghapus baris yang mengandung nilai kosong atau mengisi nilai yang hilang dengan data yang sesuai.
- Pembagian data, pada tahap ini dataset penjualan suku cadang motor dibagi menjadi dua *subset*, yaitu data pelatihan (*training*) dan data pengujian (*test*).

3.4. Modelling

Validasi model dilakukan dengan menggunakan teknik validasi silang atau dengan membagi data menjadi set pelatihan dan pengujian. Kinerja model dievaluasi dengan metrik seperti akurasi, presisi, recall, dan F1-score untuk memastikan bahwa model dapat memprediksi penjualan dengan tepat.

3.5. Evaluation

Setelah membangun model Random Forest untuk memprediksi penjualan suku cadang motor di PT Terus Jaya Sentosa Motor, langkah evaluasi menjadi krusial untuk memastikan kualitas dan keandalannya. Evaluasi bertujuan untuk menilai sejauh mana model mampu memprediksi penjualan dengan akurat. Metrik evaluasi yang digunakan mencakup RMSE, MSE, MAE, dan R-Squared.

3.6. Deployment

Pada proses ini melibatkan integrasi model prediksi kedalam aplikasi web untuk mempermudah pengguna dalam mengakses model yang telah dibuat.

4. HASIL DAN PEMBAHASAN

4.1. Business Understanding

Penelitian ini bertujuan untuk menggunakan metode *Random Forest* dalam memprediksi penjualan suku cadang motor. Dengan prediksi yang lebih akurat, PT. Terus Jaya Sentosa Motor berharap dapat mengoptimalkan pengelolaan persediaan, meningkatkan efisiensi operasional, mengurangi biaya terkait kelebihan stok, dan memastikan ketersediaan suku cadang motor Honda guna meningkatkan kepuasan pelanggan.

4.2. Data Understanding

Langkah awal adalah mengumpulkan data historis penjualan suku cadang motor Honda. *Dataset* yang digunakan dalam penelitian ini berasal dari PT Terus Jaya Sentosa Motor dan diperoleh dari sistem penjualan suku cadang motor.

a. Memuat dataset

```
data_sukucadang =
pd.read_excel('/content/RandomForestSparepart.xlsx')
data_sukucadang
```

Gambar 1. Mengimpor data dari dokumen Excel

Membuat sebuah objek yang akan diproses oleh *library* untuk membaca dokumen lalu menampilkannya.

b. Menampilkan informasi dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3528 entries, 0 to 3527
Data columns (total 5 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Tahun            3528 non-null   int64
1   Bulan            3528 non-null   object
2   Kode_Barang      3528 non-null   object
3   Harga            3528 non-null   int64
4   Jumlah_Terjual   3528 non-null   int64
dtypes: int64(3), object(2)
memory usage: 137.9+ KB
```

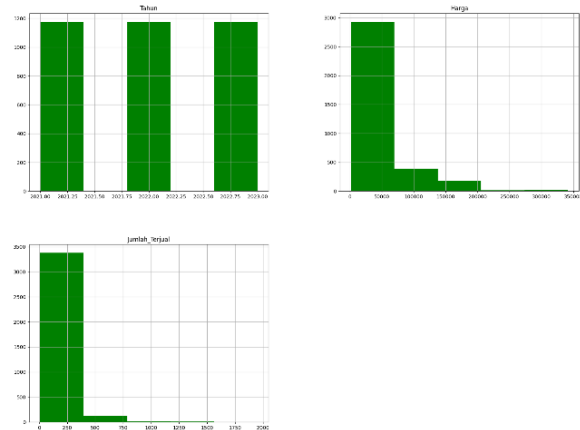
Gambar 2. Informasi ringkas

Menampilkan ringkasan informasi terkait *dataframe* yang berisikan jumlah data 3528, tipe data tahun, harga dan jumlah terjual yaitu *integer* sedangkan bulan dan kode_barang bertipe data *object*.

c. Persebaran data numerik

```
data_sperpat.hist(bins=5,figsize=(20,
15), color='green')
plt.show()
```

Gambar 3. Persebaran data numerik



Gambar 4. Diagram persebaran data numerik

Mengecek persebaran data numerik pada kolom Tahun, Harga dan Jumlah Terjual.

d. Identifikasi kode_barang

```
kode_barang=set(data_sukucadang['Kode_Barang'])
print(f'terdapat kode barang sebanyak {len(kode_barang)}')
print(kode_barang)
```

```
terdapat kode barang sebanyak 98
{'50639KCZ000', '082342MBK0LZ0', '082322MAU1JN3', ...}
```

Gambar 5. Mengidentifikasi kode_barang

Mengidentifikasi *kode_barang* bertujuan untuk mengetahui jumlah *kode_barang* unik. Hal ini dilakukan agar dapat memahami serta menghapus *kode_barang* yang duplikasi.

4.3. Data Preparation

a. Cek missing values

```
data_sukucadang.isna().sum()
```

```
Tahun            0
Bulan            0
Kode_Barang      0
Harga            0
Jumlah_Terjual   0
dtype: int64
```

Gambar 6. Cek missing values

Memeriksa missing values bertujuan mengidentifikasi data berlebihan yang dapat mempengaruhi analisis dan model prediksi.

b. Cek duplikasi

```
len(data_sukucadang[data_sukucadang.duplicated()])
#hasil 0
```

Gambar 7. Cek duplikasi

Metode `.duplicated()` menghasilkan boolean yang menunjukkan apakah baris di *DataFrame* duplikat, sedangkan `len()` menghitung jumlah baris duplikat dalam *DataFrame*.

c. Konversi data serta membuat JSON

```
data_sukucadang['Bulan'] =
data_sukucadang['Bulan'].astype(str)
data_sukucadang['Kode_Barang'] =
data_sukucadang['Kode_Barang'].astype(str)
kode_barang_list =
list(set(data_sukucadang['Kode_Barang']
))
kode_barang_json =
json.dumps(kode_barang_list)
with open('kode_barang.json', 'w') as
f: f.write(kode_barang_json)
```

Gambar 8. Konversi data dan membuat JSON

Membuat JSON bertujuan untuk penggunaan data yang dibutuhkan dalam *deployment* ke web. Dan konversi bulan serta kode_barang menjadi tipe data *string*.

d. Encoding

Pada proses ini untuk mengubah nilai-nilai kategorikal yang terdapat dalam kolom bulan dan kode_barang menjadi representasi numerik, dapat menggunakan LabelEncoder dari *library* scikit-learn. Proses ini memungkinkan untuk mengonversi data kategorikal yang berupa teks menjadi angka, sehingga lebih mudah digunakan dalam berbagai algoritma machine learning. Setelah konversi dilakukan, objek LabelEncoder yang telah digunakan akan disimpan dalam file menggunakan modul pickle. Penyimpanan ini bertujuan agar encoder yang sama dapat digunakan kembali pada data baru di masa depan, memastikan konsistensi dalam penanganan nilai kategorikal yang sama.

```
encode_bulan = LabelEncoder()
encode_kode_barang = LabelEncoder()
data_sukucadang['Bulan'] =
encode_bulan.fit_transform(data_sukucadang['Bulan'])
data_sukucadang['Kode_Barang'] =
encode_kode_barang.fit_transform(data_sukucadang['Kode_Barang'])
with open('encode_bulan.pkl', 'wb')
as file: pickle.dump(encode_bulan,
file)
with open('encode_kode_barang.pkl',
'wb') as file:
pickle.dump(encode_kode_barang,
file)
data_sukucadang
```

Gambar 9. Encoding

4.4. Modeling

a. Memisahkan label dan fitur

```
X =
data_sukucadang.drop('Jumlah_Terjual',
axis=1)
y =
data_sukucadang['Jumlah_Terjual']
X_train, X_test, y_train, y_test =
train_test_split(X, y,
test_size=0.2, random_state=72)
```

Gambar 10. Memisahkan label dan fitur

Mepersiapkan data untuk model pembelajaran mesin dengan memisahkan data menjadi fitur (x) dan target (y), lalu membaginya menjadi set pelatihan dan pengujian.

b. Pemodelan

Penggunaan *Random Forest* memainkan peran penting dalam membangun model prediktif yang kuat. Dalam konteks ini, akan mengatur dan melatih model *Random Forest Regressor* dengan parameter khusus untuk mengoptimalkan performa. Selain itu, proses ini mencakup ekstraksi salah satu pohon dari hutan acak yang dihasilkan, yang selanjutnya akan divisualisasikan untuk memberikan wawasan lebih mendalam tentang bagaimana model melakukan pemodelan dan keputusan.

```
RFR =
RandomForestRegressor(n_estimators=1
4, max_depth=50, random_state=65,
n_jobs=-1,
criterion='squared_error')
RFR.fit(X_train, y_train)
y_pred = RFR.predict(X_test)
```

Gambar 11. Pemodelan random forest

c. Menghitung Root Mean Square Error (RMSE)

Yang merupakan salah satu metrik untuk mengukur seberapa baik prediksi yang dihasilkan oleh model dibandingkan dengan nilai aktualnya. RMSE memberi tahu kita seberapa besar rata-rata kesalahan prediksi, dengan memberikan penalti lebih besar pada kesalahan yang lebih besar.

```
def rmse_metric(actual, predicted):
actual = np.array(actual)
predicted = np.array(predicted)
sum_error = 0.0
for i in range(len(actual)):
prediction_error = predicted[i] -
actual[i]
sum_error += (prediction_error ** 2)
mean_error = sum_error /
float(len(actual))
return np.sqrt(mean_error)
```

Gambar 12. Hitung RMSE

d. Menghitung MAE

Menghitung Mean Absolute Error (MAE), yang merupakan salah satu metrik evaluasi untuk mengukur seberapa akurat prediksi yang dibuat oleh model jika dibandingkan dengan nilai aktualnya. MAE menunjukkan rata-rata kesalahan absolut antara prediksi dan nilai sebenarnya, tanpa memperhatikan arah kesalahan (positif atau negatif).

```
def mae_metric(actual: np.ndarray,
               predicted: np.ndarray) -> float:
    actual = np.array(actual)
    predicted = np.array(predicted)
    sum_error = np.sum(np.abs(predicted
- actual))
    return sum_error /
float(len(actual))
```

Gambar 13. Hitung MAE

4.5. Evaluation

a. Menghitung MSE, RMSE dan MAE dengan data testing.

```
mse = mean_squared_error(y_test,
y_pred)
mae = mean_absolute_error(y_test,
y_pred)
rmse = rmse_metric(y_test, y_pred)
print("Mean Squared Error:", mse)
print("Mean Absolute Error:", mae)
print("Root Mean Squared Error:",
rmse)
```

```
Mean Squared Error: 546.4947832861191
Root Mean Squared Error: 23.3772278559925
Mean Absolute Error: 13.965722379603399
```

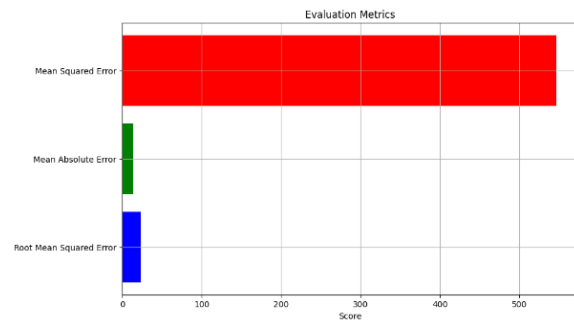
Gambar 14. Menghitung evaluasi metrik

Menghitung MSE, yaitu rata-rata kuadrat selisih antara nilai sebenarnya (y_{test}) dan prediksi (y_{pred}). Menghitung MAE, yang merupakan rata-rata selisih absolut antara nilai sebenarnya (y_{test}) dan prediksi (y_{pred}). Sedangkan RMSE, yaitu akar kuadrat dari rata-rata selisih kuadrat antara nilai aktual (y_{test}) dan prediksi (y_{pred}).

b. Menampilkan grafik evaluasi metrik

```
metrics = ['Root Mean Squared
Error', 'Mean Absolute Error',
'Mean Squared Error']
values = [rmse, mae, mse]
plt.figure(figsize=(10, 6))
plt.barh(metrics, values,
color=['blue', 'green', 'red',
'purple'])
plt.xlabel('Score')
plt.title('Evaluation Metrics')
plt.grid(True)
plt.show()
```

Gambar 15. Menampilkan evaluasi metrik



Gambar 16. Diagram evaluasi

Kode ini digunakan untuk mengevaluasi model dengan menghasilkan grafik batang yang membandingkan tiga metrik kesalahan: MAE, MSE, dan RMSE. Visualisasi ini mempermudah pemahaman kinerja model random forest dalam membantu mengidentifikasi area yang perlu diperbaiki untuk meningkatkan akurasi model.

4.6. Deployment

Tahap yang terakhir ini mengimplementasikan model *random forest* yang telah dilatih ke dalam sebuah *website*.

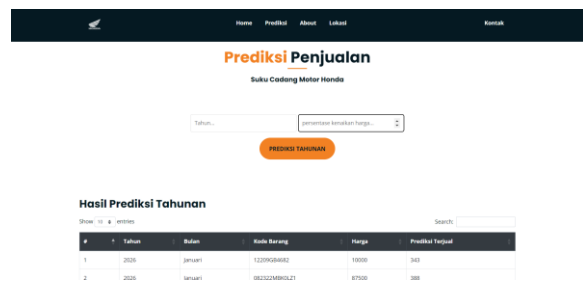
a. Menu awal



Gambar 17. Menu awal

Pada menu ini menampilkan halaman utama dari *website* prediksi penjualan suku cadang motor.

b. Menu prediksi



Gambar 18. Menu prediksi

Pada menu ini menyediakan *input* tahun dan persentase kenaikan harga untuk prediksi.

c. Menu hasil prediksi

#	Tahun	Bulan	Kode Barang	Harga	Prediksi Tahunan
1	2026	Januari	1220958482	10000	345
2	2026	Januari	082327884231	87000	388
3	2026	Januari	082049892923	24000	53
4	2026	Januari	50524307000	2000	43
5	2026	Januari	61103888800	3000	37
6	2026	Januari	8010361440026	36000	69
7	2026	Januari	22132949400	15000	115
8	2026	Januari	11103888800	31000	91
9	2026	Januari	08234950000	33000	89
10	2026	Januari	443038140026	40000	84

Gambar 19. Menu hasil prediksi

Menyajikan hasil dari prediksi model *random forest* sesuai dengan *input* yang telah diberikan.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengimplementasikan model prediksi penjualan suku cadang motor menggunakan *Random Forest*, yang menunjukkan hasil evaluasi dengan RMSE sebesar 23.37722, MAE sebesar 13.96572 serta nilai MSE sebesar 546.49478, hasil ini menandakan nilai RMSE menunjukkan rata-rata perbedaan prediksi dari nilai sebenarnya sekitar 23.38 unit, sedangkan rata-rata prediksi yang meleset sekitar 13.97 unit dari nilai sebenarnya. Meskipun dengan nilai MSE yang besar model ini dapat dikatakan cukup baik karena nilai RMSE dan MAE yang kecil. Model ini cukup efektif dalam memprediksi penjualan suku cadang, sehingga dapat membantu PT. Terus Jaya Sentosa Motor dalam mengelola stok dengan lebih optimal, serta mengurangi risiko kelebihan atau kekurangan persediaan. Untuk pengembangan lebih lanjut ataupun untuk penelitian selanjutnya yaitu, gunakan data aktual dari sistem penjualan perusahaan secara *real-time*, melakukan analisis lebih lanjut mengenai faktor eksternal yang mempengaruhi penjualan, memperbaharui *dataset* secara berkala serta menggunakan algoritma lain sebagai pembandingan untuk hasil yang lebih baik.

DAFTAR PUSTAKA

- [1] Putri Choirunisa, "Implementasi Artificial Intelligence Untuk Memprediksi Harga Penjualan Rumah Menggunakan Metode Random Forest Dan Flask," *UII.ac.id*, pp. 1–101, 2020.
- [2] S. Abdy, E. R. Br Gultom, S. Ramadhany, and A. Afifudin, "Prediksi Penjualan Sparepart Mobil Terlaris Menggunakan Metode K-Nearest

Neighbor," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 6, p. 2003, 2022, doi: 10.30865/jurikom.v9i6.5189.

- [3] G. N. Ayuni and D. Fitriana, "Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ," *J. Telemat.*, vol. 14, no. 2, pp. 79–86, 2019, [Online]. Available: <https://journal.ithb.ac.id/telematika/article/view/321>
- [4] R. D. Laksmana, E. Santoso, and B. Rahayudi, "Prediksi Penjualan Roti Menggunakan Metode Exponential Smoothing (Studi Kasus : Harum Bakery)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 5, pp. 4933–4941, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/5375/2525>
- [5] G. A. Sandag, "Prediksi Rating Aplikasi App Store Menggunakan Algoritma Random Forest," *CogITO Smart J.*, vol. 6, no. 2, pp. 167–178, 2020, doi: 10.31154/cogito.v6i2.270.167-178.
- [6] U. Azmi, Z. N. Hadi, and S. Soraya, "ARDL METHOD: Forecasting Data Curah Hujan Harian NTB," *J. Varian*, vol. 3, no. 2, pp. 73–82, 2020, doi: 10.30812/varian.v3i2.627.
- [7] H. W. Herwanto, T. Widiyaningtyas, and P. Indriana, "Penerapan Algoritme Linear Regression untuk Prediksi Hasil Panen Tanaman Padi," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 8, no. 4, p. 364, 2019, doi: 10.22146/jnteti.v8i4.537.
- [8] Y. Christian and K. O. Y. R. Qi, "Penerapan K-Means pada Segmentasi Pasar untuk Riset Pemasaran pada Startup Early Stage dengan Menggunakan CRISP-DM," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4, p. 966, 2022, doi: 10.30865/jurikom.v9i4.4486.
- [9] Dina Aqmila, "Perancangan Media Pembelajaran Bahasa Pemograman Python Menggunakan Aplikasi SCRATCH Untuk Siswa Sekolah Menengah Pertama (SMP)," *Skripsi*, pp. 1–51, 2022.
- [10] A. Trisnawati, D. Surani, B. Sri Kurniawan, and A. Fidriyanto, "Analysis of Material Understanding Usings Google Colaboratory in Class X Informatics Subject at SMAN 5 Serang," *Cakrawala Pedagog. (Jurnal Pendidik. dan Penelitian)*, vol. 8, no. 1, pp. 38–45, 2024.