

DAFTAR PUSTAKA

- Amri, Z., Kusrini, K., & Kusnawi, K. (2023). Prediksi Tingkat Kelulusan Mahasiswa menggunakan Algoritma Naïve Bayes, Decision Tree, ANN, KNN, dan SVM. *Edumatic: Jurnal Pendidikan Informatika*, 7(2), 187–196.
<https://doi.org/10.29408/edumatic.v7i2.18620>
- Budiyantara, A., & A, I. (2018). Prediksi Mahasiswa Lulus Tepat Waktu. *Infotech: Journal of Technology Information*, 5(2), 7–13.
<https://doi.org/10.37365/it.v5i2.39>
- Darmawan, A., Yudhisari, I., Anwari, A., & Makruf, M. (2023). Pola Prediksi Kelulusan Siswa Madrasah Aliyah Swasta dengan Support Vector Machine dan Random Forest. *Jurnal Minfo Polgan*, 12(1), 387–400.
<https://doi.org/10.33395/jmp.v12i1.12388>
- Latifah, S. L. S. N. H. (2020). *Prediksi Prestasi Akademik Mahasiswa Menggunakan Random Forest dan C.45*. VIII(1), 47–52. www.bsi.ac.id
- Linawati, S., Nurdiani, S., Handayani, K., & Latifah, L. (2020). Prediksi Prestasi Akademik Mahasiswa Menggunakan Algoritma Random Forest Dan C4.5. *Jurnal Khatulistiwa Informatika*, 8(1), 47–52.
<https://doi.org/10.31294/jki.v8i1.7827>
- Permatasari, R. P. (2021). *Implementasi algoritma decision tree untuk prediksi kelulusan mahasiswa tepat waktu laporan skripsi*.
- Plaosan, van S. (2019). Algoritma Random Forest.
Http://Learningbox.Coffeecup.Com/05_2_Randomforest.Html, 18(1), 10–14.
- Prediksi kelulusan pelajar menggunakan decision tree. (2023). 90–94.
- Satrio Junaidi, Valicia Anggela, R., & Kariman, D. (2024). Klasifikasi Metode Data Mining untuk Prediksi Kelulusan Tepat Waktu Mahasiswa dengan Algoritma Naïve Bayes, Random Forest, Support Vector Machine (SVM) dan Artificial Neural Nerwork (ANN). *Journal of Applied Computer Science and Technology*, 5(1), 109–119. <https://doi.org/10.52158/jacost.v5i1.489>
- Wahyudi, A. (2023). Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Decision Tree Dan Naïve Bayes. *Jurnal Permata Indonesia*, 14(2),

132–138. <https://doi.org/10.59737/jpi.v14i2.276>
Zeniarja, J., Salam, A., & Ma'ruf, F. A. (2022). Seleksi Fitur dan Perbandingan Algoritma Klasifikasi untuk Prediksi Kelulusan Mahasiswa. *Jurnal Rekayasa Elektrika*, 18(2), 102–108. <https://doi.org/10.17529/jre.v18i2.24047>



DAFTAR PUSTAKA

- Amri, Z., Kusrini, K., & Kusnawi, K. (2023). Prediksi Tingkat Kelulusan Mahasiswa menggunakan Algoritma Naïve Bayes, Decision Tree, ANN, KNN, dan SVM. *Edumatic: Jurnal Pendidikan Informatika*, 7(2), 187–196.
<https://doi.org/10.29408/edumatic.v7i2.18620>
- Budiyantara, A., & A, I. (2018). Prediksi Mahasiswa Lulus Tepat Waktu. *Infotech: Journal of Technology Information*, 5(2), 7–13.
<https://doi.org/10.37365/it.v5i2.39>
- Darmawan, A., Yudhisari, I., Anwari, A., & Makruf, M. (2023). Pola Prediksi Kelulusan Siswa Madrasah Aliyah Swasta dengan Support Vector Machine dan Random Forest. *Jurnal Minfo Polgan*, 12(1), 387–400.
<https://doi.org/10.33395/jmp.v12i1.12388>
- Latifah, S. L. S. N. H. (2020). *Prediksi Prestasi Akademik Mahasiswa Menggunakan Random Forest dan C.45*. VIII(1), 47–52. www.bsi.ac.id
- Linawati, S., Nurdiani, S., Handayani, K., & Latifah, L. (2020). Prediksi Prestasi Akademik Mahasiswa Menggunakan Algoritma Random Forest Dan C4.5. *Jurnal Khatulistiwa Informatika*, 8(1), 47–52.
<https://doi.org/10.31294/jki.v8i1.7827>
- Permatasari, R. P. (2021). *Implementasi algoritma decision tree untuk prediksi kelulusan mahasiswa tepat waktu laporan skripsi*.
- Plaosan, van S. (2019). Algoritma Random Forest.
Http://Learningbox.Coffeecup.Com/05_2_Randomforest.Html, 18(1), 10–14.
- Prediksi kelulusan pelajar menggunakan decision tree. (2023). 90–94.
- Satrio Junaidi, Valicia Anggela, R., & Kariman, D. (2024). Klasifikasi Metode Data Mining untuk Prediksi Kelulusan Tepat Waktu Mahasiswa dengan Algoritma Naïve Bayes, Random Forest, Support Vector Machine (SVM) dan Artificial Neural Nerwork (ANN). *Journal of Applied Computer Science and Technology*, 5(1), 109–119. <https://doi.org/10.52158/jacost.v5i1.489>
- Wahyudi, A. (2023). Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Decision Tree Dan Naïve Bayes. *Jurnal Permata Indonesia*, 14(2),

132–138. <https://doi.org/10.59737/jpi.v14i2.276>
Zeniarja, J., Salam, A., & Ma'ruf, F. A. (2022). Seleksi Fitur dan Perbandingan
Algoritma Klasifikasi untuk Prediksi Kelulusan Mahasiswa. *Jurnal Rekayasa
Elektrika*, 18(2), 102–108. <https://doi.org/10.17529/jre.v18i2.24047>



LAMPIRAN

1) Code Penelitian

```
2) # Langkah 1: Instalasi library yang diperlukan tanpa
   menampilkan pesan instalasi
3) !pip install PyPDF2 > /dev/null 2>&1
4) !pip install pandas > /dev/null 2>&1
5) !pip install scikit-learn > /dev/null 2>&1
6) !pip install matplotlib > /dev/null 2>&1
7) !pip install seaborn > /dev/null 2>&1
8)
9) # Langkah 2: Import library yang diperlukan
10)    import PyPDF2
11)    import pandas as pd
12)    import numpy as np
13)    from sklearn.model_selection import
        train_test_split
14)    from sklearn.ensemble import
        RandomForestClassifier # Menggunakan Random Forest
        Classifier
15)    from sklearn.metrics import f1_score,
        accuracy_score, recall_score, precision_score
16)    import matplotlib.pyplot as plt
17)    import seaborn as sns
18)
19) # Langkah 3: Fungsi untuk membaca data dari file
   PDF
20) def read_pdf(file_path):
21)     try:
22)         with open(file_path, 'rb') as file:
23)             reader = PyPDF2.PdfReader(file)
24)             text = ''
25)             for page in reader.pages:
26)                 text += page.extract_text()
27)             return text
28)     except Exception as e:
29)         print(f"Terjadi kesalahan saat membaca file
   PDF: {e}")
30)         return None
31)
32) # Langkah 4: Preprocessing data
33) def preprocess_data(text):
34)     try:
35)         # Contoh preprocessing sederhana:
            memisahkan teks menjadi baris-baris
```

```

36)             lines = text.strip().split('\n')
37)             # Anda perlu menyesuaikan bagian ini sesuai
            dengan struktur data PDF Anda
38)             # Misalnya, mengubah teks menjadi DataFrame
            pandas
39)             data = pd.DataFrame(lines,
            columns=['text'])
40)             # Anda mungkin perlu melakukan lebih banyak
            preprocessing, seperti encoding label, menangani missing
            values, dll.
41)             return data
42)         except Exception as e:
43)             print(f"Terjadi kesalahan saat memproses
            data: {e}")
44)             return None
45)
46)     # Langkah 5: Membaca dan memproses data dari file
        PDF
47)     file_path = 'data.pdf' # Ganti dengan path file
        PDF Anda
48)     text = read_pdf(file_path)
49)     if text:
50)         data = preprocess_data(text)
51)         if data is not None:
52)             print(data.head())
53)
54)     # Langkah 6: Menyiapkan data untuk model (contoh
        sederhana)
55)     # Anda perlu menyesuaikan bagian ini sesuai dengan
        struktur data Anda
56)     # Misalnya, memisahkan fitur dan label
57)     # X = data[['fitur1', 'fitur2', ...]]
58)     # y = data['label']
59)     # Contoh sederhana dengan data dummy
60)     np.random.seed(0)
61)     X = np.random.rand(100, 5) # 100 sampel, 5 fitur
62)     y = np.random.randint(0, 2, 100) # Label biner
63)
64)     # Langkah 7: Membagi data menjadi set pelatihan dan
        pengujian
65)     X_train, X_test, y_train, y_test =
            train_test_split(X, y, test_size=0.2, random_state=42)
66)
67)     # Langkah 8: Membangun model Random Forest

```

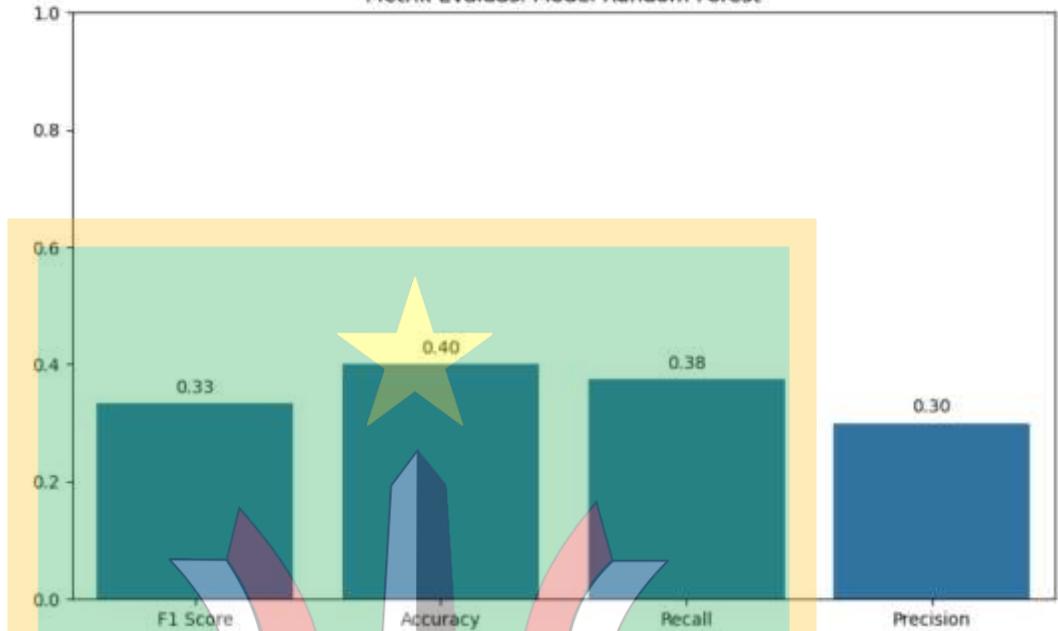
```

68)     model = RandomForestClassifier(random_state=42)    #
   Menggunakan Random Forest Classifier
69)     model.fit(X_train, y_train)
70)
71)     # Langkah 9: Melakukan prediksi
72)     y_pred = model.predict(X_test)
73)
74)     # Langkah 10: Menghitung metrik evaluasi
75)     f1 = f1_score(y_test, y_pred)
76)     accuracy = accuracy_score(y_test, y_pred)
77)     recall = recall_score(y_test, y_pred)
78)     precision = precision_score(y_test, y_pred)
79)
80)     # Langkah 11: Menampilkan metrik evaluasi
81)     metrics = {
82)         'F1 Score': f1,
83)         'Accuracy': accuracy,
84)         'Recall': recall,
85)         'Precision': precision
86)     }
87)     print(metrics)
88)
89)     # Langkah 12: Menampilkan hasil metrik dalam bentuk
   diagram batang
90)     plt.figure(figsize=(10, 6))
91)     ax = sns.barplot(x=list(metrics.keys()),
   y=list(metrics.values()))
92)     plt.title('Metrik Evaluasi Model Random Forest')
93)     plt.ylim(0, 1)
94)
95)     # Menambahkan nilai numerik di atas setiap batang
96)     for p in ax.patches:
97)         ax.annotate(f'{p.get_height():.2f}', 
   (p.get_x() + p.get_width() / 2.,
   p.get_height()), 
   ha='center', va='center',
98)             xytext=(0, 9),
99)             textcoords='offset points')
100)
101)
102)
103)     plt.show()
104)

```

```
{'F1 Score': 0.3333333333333333, 'Accuracy': 0.4, 'Recall': 0.375, 'Precision': 0.3}
```

Metrik Evaluasi Model Random Forest



```
import subprocess
import sys
import os
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import f1_score, accuracy_score, recall_score, precision_score,
confusion_matrix
import matplotlib.pyplot as plt
import seaborn as sns
```

```
# Function to install packages silently
def install(package):
    subprocess.check_call(
        [sys.executable, "-m", "pip", "install", package, "--quiet"],
        stdout=subprocess.DEVNULL,
        stderr=subprocess.DEVNULL
```

```
)
```

```
# Install required packages silently
```

```
install('PyPDF2')
```

```
install('pandas')
```

```
install('scikit-learn')
```

```
install('matplotlib')
```

```
install('seaborn')
```

```
# Import PyPDF2 after installation
```

```
import PyPDF2
```

```
# Langkah 3: Fungsi untuk membaca data dari file PDF
```

```
def read_pdf(file_path):
```

```
    try:
```

```
        with open(file_path, 'rb') as file:
```

```
            reader = PyPDF2.PdfReader(file)
```

```
            text = "
```

```
            for page in reader.pages:
```

```
                text += page.extract_text()
```

```
            return text
```

```
    except Exception as e:
```

```
        print(f"Terjadi kesalahan saat membaca file PDF: {e}")
```

```
    return None
```

```
# Langkah 4: Preprocessing data
```

```
def preprocess_data(text):
```

```
    try:
```

```
        # Contoh preprocessing sederhana: memisahkan teks menjadi baris-baris
```

```
        lines = text.strip().split("\n")
```

```
        # Anda perlu menyesuaikan bagian ini sesuai dengan struktur data PDF Anda
```

```
        # Misalnya, mengubah teks menjadi DataFrame pandas
```

```
        data = pd.DataFrame(lines, columns=['text'])
```

```

# Anda mungkin perlu melakukan lebih banyak preprocessing, seperti encoding
label, menangani missing values, dll.

return data

except Exception as e:
    print(f"Terjadi kesalahan saat memproses data: {e}")

return None

```

Langkah 5: Membaca dan memproses data dari file PDF

file_path = 'data.pdf' # Ganti dengan path file PDF Anda

text = read_pdf(file_path)

if text:

data = preprocess_data(text)

if data is not None:

print(data.head())

Langkah 6: Menyiapkan data untuk model (contoh sederhana)

Anda perlu menyesuaikan bagian ini sesuai dengan struktur data Anda

Misalnya, memisahkan fitur dan label

X = data[['fitur1', 'fitur2', ...]]

y = data['label']

Contoh sederhana dengan data dummy

np.random.seed(0)

X = np.random.rand(100, 5) # 100 sampel, 5 fitur

y = np.random.randint(0, 2, 100) # Label biner

Langkah 7: Membagi data menjadi set pelatihan dan pengujian

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

Langkah 8: Membangun model Decision Tree

model = DecisionTreeClassifier(random_state=42)

model.fit(X_train, y_train)

Langkah 9: Melakukan prediksi

y_pred = model.predict(X_test)

```

# Langkah 10: Menghitung metrik evaluasi
f1 = f1_score(y_test, y_pred)
accuracy = accuracy_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)
precision = precision_score(y_test, y_pred)

# Langkah 11: Menampilkan metrik evaluasi
metrics = {
    'F1 Score': f1,
    'Accuracy': accuracy,
    'Recall': recall,
    'Precision': precision
}
print(metrics)

# Langkah 12: Menampilkan hasil metrik dalam bentuk diagram batang
plt.figure(figsize=(10, 6))
ax = sns.barplot(x=list(metrics.keys()), y=list(metrics.values()))
plt.title('Metrik Evaluasi Model Decision Tree')
plt.ylim(0, 1)

# Menambahkan nilai numerik di atas setiap batang
for p in ax.patches:
    ax.annotate(f'{p.get_height():.2f}', 
                (p.get_x() + p.get_width() / 2., p.get_height()),
                ha='center', va='center',
                xytext=(0, 9),
                textcoords='offset points')

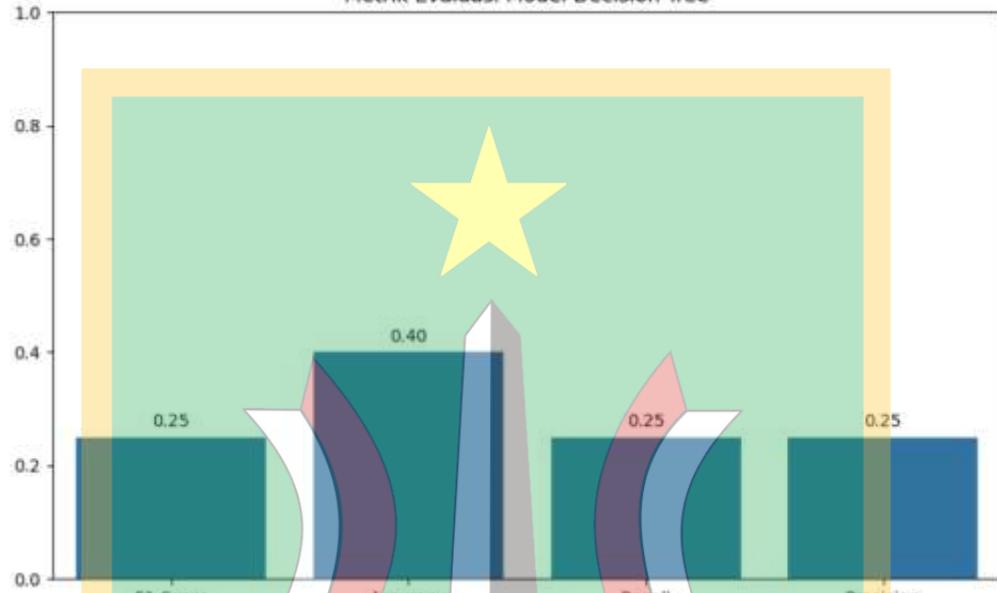
plt.show()

# Langkah 13: Menampilkan matriks konfusi
cm = confusion_matrix(y_test, y_pred)

```

```
print("Matriks Konfusi:")
print(cm)
```

```
('F1 Score': 0.25, 'Accuracy': 0.4, 'Recall': 0.25, 'Precision': 0.25)
Metrik Evaluasi Model Decision Tree
```



2. Bukti submit



Manuscript URL:
<https://ojs.unimal.ac.id/index.php/jreece/author/submission/21022>
Username: marlan255

If you have any questions, please contact me. Thank you for considering this journal as a venue for your work.

Arnawan Hasibuan



PRIMARY SOURCES

-
- 1 Dedy Hartama, Nanda Amalya.
"Perbandingan Algoritma Decision Tree, ID3,
dan Random Forest dalam Klasifikasi Faktor-
Faktor yang Mempengaruhi Karier Mahasiswa
Ilmu Komputer", Jurnal Indonesia :
Manajemen Informatika dan Komunikasi,
2025
Publication
- 2 etheses.uin-malang.ac.id
Internet Source
- 3 Submitted to University of Sydney
Student Paper
- 4 Submitted to Ajou University Graduate School
Student Paper
- 5 Submitted to University of North Carolina,
Greensboro
Student Paper
- 6 journal.isas.or.id
Internet Source
- 7 repository.pnj.ac.id
Internet Source
- 8 Royan Habibie Sukarna, Holilah Holilah, Fitri
Damyati, Mohamad Hilman. "Analisis Prediksi
Kelulusan Mahasiswa Universitas Sultan
Ageng Tirtayasa Menggunakan Algoritma

Machine Learning dan Feature Selection",
Jurnal Ilmiah Sains dan Teknologi, 2024

Publication

9	docplayer.info Internet Source	1 %
10	Safitri Linawati, Siti Nurdiani, Kartika Handayani, Latifah Latifah. "PREDIKSI PRESTASI AKADEMIK MAHASISWA MENGGUNAKAN ALGORITMA RANDOM FOREST DAN C4.5", Jurnal Khatulistiwa Informatika, 2020 Publication	1 %
11	www.jurnal.polgan.ac.id Internet Source	1 %
12	journal.binadarma.ac.id Internet Source	1 %
13	ejurnal.seminar-id.com Internet Source	1 %
14	jutif.if.unsoed.ac.id Internet Source	<1 %
15	www.slideshare.net Internet Source	<1 %
16	Nurdiyanto Yusuf. "PREDIKSI PRODUKSI DAGING SAPI DI INDONESIA MENGGUNAKAN RANDOM FOREST REGRESSION: ANALISIS DATA 2018-2025", Jurnal Ilmiah Teknik, 2024 Publication	<1 %
17	etd.umy.ac.id Internet Source	<1 %
18	Submitted to Sriwijaya University Student Paper	<1 %

19	e-journal.hamzanwadi.ac.id	<1 %
	Internet Source	

20	Dikan Ismafillah, Tatang Rohana, Yana Cahyana. "Analisis algoritma pohon keputusan untuk memprediksi penyakit diabetes menggunakan oversampling smote", INFOTECH : Jurnal Informatika & Teknologi, 2023	<1 %
	Publication	

21	ejournal.bsi.ac.id	<1 %
	Internet Source	

22	jurnal.polibatam.ac.id	<1 %
	Internet Source	

23	access-ai.com	<1 %
	Internet Source	

24	Submitted to Institut Teknologi Kalimantan	<1 %
	Student Paper	

25	Muhammad Galih Ramaputra, Hendri Purnomo. "Analisis Sentimen Opini Masyarakat Terhadap Penggunaan ChatGPT di Bidang Pendidikan Berbasis Twitter", Jurnal Pepadun, 2024	<1 %
	Publication	

26	repository.usbypkp.ac.id	<1 %
	Internet Source	

27	syariah.uin-malang.ac.id	<1 %
	Internet Source	

28	repository.unsri.ac.id	<1 %
	Internet Source	

29	Imelda Ruwae Lutunani, Adi Nugroho. "Analisis Prediksi Mahasiswa Terhadap	<1 %

Kelulusan Tepat Waktu Menggunakan Metode Data Mining Decision Tree (Studi Kasus: FTI UKSW)", Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi), 2023

Publication

- 30 j-innovative.org  <1 %
Internet Source
- 31 Submitted to Universitas Brawijaya <1 %
Student Paper
- 32 Submitted to publikasiubl <1 %
Student Paper
- 33 Enggar Novianto, Arief Hermawan, Donny Avianto. "KLASIFIKASI ALGORITMA K-NEAREST NEIGHBOR, NAIVE BAYES, DECISION TREE UNTUK PREDIKSI STATUS KELULUSAN MAHASISWA S1", Rabit : Jurnal Teknologi dan Sistem Informasi Univrab, 2023

Publication
- 34 Submitted to Konsorsium PTS Indonesia - Small Campus II <1 %
Student Paper
- 35 Mohamad Fajarianditya Nugroho, Setyoningsih Wibowo. "Fitur Seleksi Forward Selection Untuk Menetukan Atribut Yang Berpengaruh Pada Klasifikasi Kelulusan Mahasiswa Fakultas Ilmu Komputer UNAKI Semarang Menggunakan Algoritma Naive Bayes", 'Universitas PGRI Semarang', 2017
Internet Source
- 36 ejournal.unikadelasalle.ac.id <1 %
Internet Source
- 37 digilib.unila.ac.id <1 %
Internet Source

38	kc.umn.ac.id Internet Source	<1 %
39	repository.ub.ac.id Internet Source	<1 %
40	www.semanticscholar.org Internet Source	<1 %
41	www.tatestreetart.com Internet Source	<1 %
42	Submitted to Politeknik Negeri Bandung Student Paper	<1 %
43	Submitted to Southville International School and Colleges Student Paper	<1 %
44	Submitted to Universitas Negeri Jakarta Student Paper	<1 %
45	blog.binadarma.ac.id Internet Source	<1 %
46	uici.ac.id Internet Source	<1 %
47	www.prosiding.uim.ac.id Internet Source	<1 %
48	Bui Thanh Hung, M. Sekar, Ayhan Esi, R. Senthil Kumar. "Applications of Mathematics in Science and Technology - International Conference on Mathematical Applications in Science and Technology", CRC Press, 2025 Publication	<1 %
49	Submitted to Universitas Diponegoro Student Paper	<1 %

- 50 Zaenul Amri, Kusrini Kusrini, Kusnawi Kusnawi. "Prediksi Tingkat Kelulusan Mahasiswa menggunakan Algoritma Naïve Bayes, Decision Tree, ANN, KNN, dan SVM", Edumatic: Jurnal Pendidikan Informatika, 2023
Publication <1 %
- 51 digilib.esaunggul.ac.id Internet Source <1 %
- 52 Arief Maehendrayuga, Arief Setyanto, Kusnawi. "Analisa Prediksi Turnover Karyawan menggunakan Machine Learning", bit-Tech, 2024
Publication <1 %
- 53 Dinda Safitri, Shofa Shofiah Hilabi, Fitria Nurapriani. "ANALISIS PENGGUNAAN ALGORITMA KLASIFIKASI DALAM PREDIKSI KELULUSAN MENGGUNAKAN ORANGE DATA MINING", Rabbit : Jurnal Teknologi dan Sistem Informasi Univrab, 2023
Publication <1 %
- 54 Hanna Willa Dhany, Aminuddin Indra Permana, Fahmi Izhari, Andhika Pratama Ginting, Zahari Gilang Pratama. "Aplikasi Prediksi Serangan Jantung Untuk Warga Kelurahan Pelawi Utara", Jurnal Minfo Polgan, 2024
Publication <1 %
- 55 core.ac.uk Internet Source <1 %
- 56 id.123dok.com Internet Source <1 %

- 57 Putri Nabila Amir, Muhamad Fatchan, Edora
Edora. "Prediksi Kelulusan Siswa Dengan
Pendekatan Algortma C5.0 Pada SMAN 2
Cikarang Selatan", Jurnal Ilmiah Intech :
Information Technology Journal of UMUS,
2023

Publication

<1 %

