

LAMPIRAN

Lampiran 1. Riwayat Hidup

RIWAYAT HIDUP



Nyoman Dinda Indira Sudiasta Putri lahir di Anturan pada 5 Februari 2003. Penulis lahir dari pasangan suami istri Bapak Ketut Sudiasta dan Ibu Gusti Putu Suryani. Penulis berkebangsaan Indonesia dan beragama Hindu. Kini penulis beralamat di Banjar Dinas Munduk Desa Anturan, Kabupaten Buleleng, Provinsi Bali. Penulis menyelesaikan pendidikan dasar di SDN 125

Banyuasri dan lulus tahun 2015. Kemudian penulis melanjutkan pendidikan di SMP N 2 Singaraja dan lulus tahun 2018. Pada tahun 2021, penulis lulus dari SMA N 1 Singaraja jurusan Ilmu Pengetahuan Alam dan melanjutkan ke studi (S1) di Universitas Pendidikan Ganesha dengan Program Studi Sistem Informasi, Jurusan Teknik Informatika.



Lampiran 2. Surat Ketersediaan Pelabelan Data

	KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI UNIVERSITAS PENDIDIKAN GANESHA FAKULTAS TEKNIK DAN KEJURUAN Jalan Udayana Nomor 11 Singaraja Bali Laman: http://ftk.unediksha.ac.id												
<table border="0" style="width: 100%;"> <tr> <td style="width: 60%;"> Nomor : 1292/UN48.11.1/KM/2025 Perihal : Surat Permohonan Pengambilan Data </td> <td style="width: 40%; text-align: right; vertical-align: top;"> Singaraja, 2 Juni 2025 </td> </tr> </table>		Nomor : 1292/UN48.11.1/KM/2025 Perihal : Surat Permohonan Pengambilan Data	Singaraja, 2 Juni 2025										
Nomor : 1292/UN48.11.1/KM/2025 Perihal : Surat Permohonan Pengambilan Data	Singaraja, 2 Juni 2025												
Yth. Kepala SMK Negeri 1 Gerokgak c.q. Guru Bahasa Indonesia di tempat													
Dengan hormat, sehubungan dengan proses penyelesaian Tugas Akhir/Skripsi, maka melalui surat ini kami mohon Bapak/Ibu berkenan memberikan data yang terkait dengan data yang dibutuhkan. Adapun mahasiswa yang akan melakukan pengambilan data seperti tersebut di bawah ini: <table border="0" style="width: 100%;"> <tr> <td style="width: 30%;">Nama</td> <td>: Nyoman Dinda Indira Sudiasta Putri</td> </tr> <tr> <td>NIM</td> <td>: 2115091088</td> </tr> <tr> <td>Program Studi</td> <td>: Sistem Informasi</td> </tr> <tr> <td>Jurusan</td> <td>: Teknik Informatika</td> </tr> <tr> <td>Data yang dibutuhkan</td> <td>: Mengisi label Sentimen (Negatif dan Positif) pada komentar yang ada di formulir Yang telah diberikan</td> </tr> <tr> <td>Judul Skripsi</td> <td>: Komparasi Algoritma Naïve Bayes dan K-Nearest Neighbor Dalam Klasifikasi Sentimen Game Free Fire</td> </tr> </table>		Nama	: Nyoman Dinda Indira Sudiasta Putri	NIM	: 2115091088	Program Studi	: Sistem Informasi	Jurusan	: Teknik Informatika	Data yang dibutuhkan	: Mengisi label Sentimen (Negatif dan Positif) pada komentar yang ada di formulir Yang telah diberikan	Judul Skripsi	: Komparasi Algoritma Naïve Bayes dan K-Nearest Neighbor Dalam Klasifikasi Sentimen Game Free Fire
Nama	: Nyoman Dinda Indira Sudiasta Putri												
NIM	: 2115091088												
Program Studi	: Sistem Informasi												
Jurusan	: Teknik Informatika												
Data yang dibutuhkan	: Mengisi label Sentimen (Negatif dan Positif) pada komentar yang ada di formulir Yang telah diberikan												
Judul Skripsi	: Komparasi Algoritma Naïve Bayes dan K-Nearest Neighbor Dalam Klasifikasi Sentimen Game Free Fire												
Demikian kami sampaikan, atas perhatian dan kerjasamanya, diucapkan terima kasih.													
 Dekan Bidang Akademik, Made Winda Antara Kesiman NIP 198211120081210019													



**KEMENTERIAN PENDIDIKAN TINGGI,
SAINS, DAN TEKNOLOGI
UNIVERSITAS PENDIDIKAN GANESHA
FAKULTAS TEKNIK DAN KEJURUAN
Jalan Udayana Nomor 11 Singaraja Bali
Laman: <http://uk.undiksha.ac.id>**

Nomor : 1291/UN48.11.1/KM/2025

Singaraja, 2 Juni 2025

Perihal : Surat Permohonan Pengambilan Data

Yth. Kepala SMP Negeri 1 Gerokgak
c.q. Guru Bahasa Indonesia
di tempat

Dengan hormat, sehubungan dengan proses penyelesaian Tugas Akhir/Skripsi, maka melalui surat ini kami mohon Bapak/Ibu berkenan memberikan data yang terkait dengan data yang dibutuhkan. Adapun mahasiswa yang akan melakukan pengambilan data seperti tersebut di bawah ini:

Nama : Nyoman Dinda Indira Sudiasta Putri
NIM : 2115091088
Program Studi : Sistem Informasi
Jurusan : Teknik Informatika
Data yang dibutuhkan : Mengisi label Sentimen (Negatif dan Positif) pada komentar yang ada di formulir Yang telah diberikan
Judul Skripsi : Komparasi Algoritma Naive Bayes dan K-Nearest Neighbor Dalam Klasifikasi Sentimen Game Free Fire

Demikian kami sampaikan, atas perhatian dan kerjasamanya, diucapkan terima kasih.

a. Dekan
Wakil Dekan Bidang Akademik,

FTK
Made Winda Antara Kesiman
NIP 1982 1112008121001

Lampiran 3. Sertifikat Pendidik

No.0001331



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
REPUBLIK INDONESIA

SERTIFIKAT PENDIDIK

Nomor: 0711032115620519

Berdasarkan Surat Keputusan Menteri Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia Nomor 235/KPT/I/2019 tanggal 04 April 2019 tentang Ijin Pembukaan Program Studi Pendidikan Profesi Guru,
Rektor Universitas PGRI Kanjuruhan Malang menyatakan bahwa:

PUTU LIANITA RAHAYU
Nomor Induk Mahasiswa: 216101140390

lahir di BANJAR pada tanggal dua puluh satu bulan Agustus tahun seribu sembilan ratus delapan puluh sembilan telah memenuhi semua syarat penyelesaian Pendidikan Profesi Guru dan LULUS Uji Kompetensi Mahasiswa Pendidikan Profesi Guru.
Kepadaanya diberikan sebutan profesi GURU (Gr.) Bahasa Indonesia sesuai hak dan kewajiban yang melekat pada sebutan profesi tersebut.



0711032115620519-0711032115620519-1041007326





Malang, 28 Desember 2021
Rektor,

Dr. Pieter Sahertian, M.Si
NIP 298601102

No.0000737



KEMENTERIAN RISET, TEKNOLOGI DAN PENDIDIKAN TINGGI
REPUBLIK INDONESIA

SERTIFIKAT PENDIDIK

Nomor: 0010481915620783

Berdasarkan Surat Keputusan Menteri Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia Nomor 746/KPT/I/2018 tanggal 05 September 2018 tentang Ijin Pembukaan Program Studi Pendidikan Profesi Guru,
Rektor Universitas Pendidikan Ganesha menyatakan bahwa:

KOMANG AYU AGUSTINA
Nomor Induk Mahasiswa: 1974805019

lahir di GeroKgak pada tanggal delapan belas bulan Agustus tahun seribu sembilan ratus sembilan puluh tiga telah memenuhi semua syarat penyelesaian Pendidikan Profesi Guru dan LULUS Uji Kompetensi Mahasiswa Pendidikan Profesi Guru.
Kepadaanya diberikan sebutan profesi GURU (Gr.) Bahasa Indonesia sesuai hak dan kewajiban yang melekat pada sebutan profesi tersebut.



0010481915620783-0010481915620783-103702739008





Singaraja, 25 November 2019
Rektor,

Prof. Dr. I Nyoman Jampel, M.Pd.
NIP 195910101986031003

Lampiran 4. Surat Keterangan Validasi

	KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET DAN TEKNOLOGI UNIVERSITAS PENDIDIKAN GANESHA FAKULTAS TEKNIK DAN KEJURUAN Jalan Udayana, No. 11, Singaraja Bali Kode Pos 81116 Telepon (0362) 21541 Fax. (0362) 27561 Laman: http://undiksha.ac.id
SURAT KETERANGAN VALIDASI LABELING DATASET	
Yang bertanda tangan di bawah ini: Nama : Putu Lianita Rahayu, S.Pd NIP : 198908212022212018 Menerangkan bahwa Mahasiswa Universitas Pendidikan Ganesha di bawah ini: Nama : Nyoman Dinda Indira Sudiasta Putri NIM : 2115091088 Prodi/Jurusan : Sistem Informasi/Fakultas Teknik Dan Kejuruan Memang benar bahwa dataset yang sudah dilabeli telah divalidasi pada tanggal 13 Juni 2025. Demikian surat keterangan ini dibuat dengan sebenarnya untuk dapat digunakan sebagaimana mestinya.	
Singaraja, 13 Juni 2025 Ahli pakar I,  <u>Putu Lianita Rahayu, S.Pd</u> NIP 198908212022212018	
Catatan : • UU ITE No. 11 Tahun 2008 Pasal 5 ayat 1 "Informasi Elektronik dan/atau Dokumen Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah" • Dokumen ini tertanda ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BnRE • Surat ini dapat dibuktikan keasliannya dengan menggunakan qr code yang telah tersedia	
 Balai Sertifikasi Elektronik	





**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET DAN TEKNOLOGI
UNIVERSITAS PENDIDIKAN GANESHA**

FAKULTAS TEKNIK DAN KEJURUAN
Jalan Udayana, No. 11, Singaraja Bali Kode Pos 81116
Telepon (0362) 21541 Fax. (0362) 27561
Laman: <https://undiksha.ac.id>

SURAT KETERANGAN VALIDASI

LABELING DATASET

Yang bertanda tangan di bawah ini:

Nama : Komang Ayu Agustina, S.Pd

NIP : 199308182022212017

Menerangkan bahwa Mahasiswa Universitas Pendidikan Ganesha di bawah ini:

Nama : Nyoman Dinda Indira Sudiasta Putri

NIM : 2115091088

Prodi/Jurusan : Sistem Informasi/Fakultas Teknik Dan Kejuruan

Memang benar bahwa dataset yang sudah dilabeli telah divalidasi pada tanggal 13 Juni 2025. Demikian surat keterangan ini dibuat dengan sebenarnya untuk dapat digunakan sebagaimana mestinya.

Singaraja, 13 Juni 2025

Ahli pakar II,

Komang Ayu Agustina, S.Pd

NIP 199308182022212017



**Balai
Sertifikasi
Elektronik**

Catatan :

- UU ITE No. 11 Tahun 2008 Pasal 5 ayat 1 "Informasi Elektronik dan/atau Dokumen Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini tertanda ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSEI
- Surat ini dapat dibuktikan keasliannya dengan menggunakan qr code yang telah tersedia

Lampiran 5. Dokumentasi bersama Validator



Lampiran 6. Proses Scraping Data Play Store

```
#Install pustaka
!pip install google-play-scraper

#Import
from google_play_scraper import reviews, Sort
import pandas as pd

#Ambil review Free Fire (paling relevan)
result, _ = reviews(
    'com.dts.freefireth',      #ID aplikasi
    lang='id',                 #Bahasa
    country='id',              #Negara
    sort=Sort.MOST_RELEVANT,   #Ambil review paling relevan
    count=6500                 #Jumlah review
)

#Ubah ke DataFrame
df = pd.DataFrame(result)

#Tampilkan 5 data awal
df[['userName', 'score', 'content', 'at']].head()

df[['userName', 'score', 'content',
'at']].to_csv('freefire_reviews_most_relevant.csv',
index=False)
```

Lampiran 7. Proses Scraping Data App Store

```
!pip install app_store_scraper

from app_store_scraper import AppStore
import pandas as pd

app = AppStore(
    country="id",                #Negara Indonesia
    app_name="free-fire",        #Nama aplikasi
    app_id="1300146617"          #App ID dari App Store
)

#Mengambil review (maksimal sesuai yang tersedia)
app.review(how_many=4000)

#Simpan ke DataFrame
df = pd.DataFrame(app.reviews)

#Menampilkan 10 data pertama
df.head(10)

#Simpan ke file CSV
df.to_csv("app_reviews.csv", index=False, encoding="utf-8")
```

Lampiran 8. Proses Preprocessing

```

!pip install sastrawi
!pip install nltk

import pandas as pd
import re
import nltk
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
from google.colab import files
from IPython.display import display
#Download data nltk
nltk.download('punkt')
nltk.download('stopwords')

#Upload file
uploaded = files.upload()

#Ganti nama file jika berbeda
df = pd.read_csv('data playstore (label).csv')
#Asumsikan kolom ulasan adalah 'content' dan label 'label'
df = df[['ulasan', 'label']]
print(" ♦ Data Asli (10 teratas):")
display(df.head(5))

#tahapan cleansing
def cleansing(text):
    text = re.sub(r'^a-zA-Z\s', '', text) #hapus karakter
non-huruf
    text = re.sub(r'\s+', ' ', text) #hapus spasi
berlebih
    return text.strip()

df['cleansed'] = df['ulasan'].astype(str).apply(cleansing)

print(" ♦ Setelah Cleansing:")
display(df[['ulasan', 'cleansed']].head(5))

```

```

#tahapan Transform Cases
df['lower'] = df['cleansed'].str.lower()

print(" ♦ Setelah Lowercase:")
display(df[['cleansed', 'lower']].head(5))

#Tokenizing tanpa nltk
df['tokenized'] = df['lower'].apply(lambda x: x.split())
print(" ♦ Setelah Tokenizing (pakai .split()):")
display(df[['lower', 'tokenized']].head(5))

#Filter Stopword
stop_words = set(stopwords.words('indonesian'))
df['no_stopwords'] = df['tokenized'].apply(lambda tokens:
[word for word in tokens if word not in stop_words])

print(" ♦ Setelah Stopword Removal:")
display(df[['tokenized', 'no_stopwords']].head(5))

#Stemming
factory = StemmerFactory()
stemmer = factory.create_stemmer()

df['stemmed'] = df['no_stopwords'].apply(lambda tokens:
[stemmer.stem(word) for word in tokens])

print(" ♦ Setelah Stemming:")
display(df[['no_stopwords', 'stemmed']].head(5))

df_final = df[['ulasan', 'label', 'cleansed', 'lower',
'tokenized', 'stemmed']]
df_final.to_csv("hasil_preprocessing.csv", index=False)

print("✅ Hasil akhir berhasil disimpan ke
'hasil_preprocessing.csv'")

```

Lampiran 9. Proses TF-IDF (Term Frequency–Inverse Document Frequency)

```
import pandas as pd
from google.colab import files
#Upload file CSV hasil preprocessing
uploaded = files.upload()
#Baca file
df = pd.read_csv("hasil preprocessing playstore.csv")
#Tampilkan 5 data teratas
df.head()

df['stemmed_joined'] = df['stemmed'].apply(lambda x: "
".join(eval(x)) if isinstance(x, str) else "")
# Lihat hasil penggabungan kata
df[['stemmed', 'stemmed_joined']].head()

from sklearn.feature_extraction.text import TfidfVectorizer
#Inisialisasi TF-IDF Vectorizer
tfidf = TfidfVectorizer()
#Transformasi ke matriks TF-IDF
tfidf_matrix = tfidf.fit_transform(df['stemmed_joined'])
#Konversi hasil ke DataFrame
tfidf_df = pd.DataFrame(tfidf_matrix.toarray(),
columns=tfidf.get_feature_names_out())
#Tampilkan hasil TF-IDF
tfidf_df.head(10)

#Gabungkan dengan kolom label
tfidf_labeled = pd.concat([df[['label']], tfidf_df], axis=1)
#Tampilkan hasil
tfidf_labeled.head()
#Simpan ke file CSV
tfidf_labeled.to_csv("hasil_tfidf.csv", index=False)
print("✅ TF-IDF berhasil disimpan ke 'hasil_tfidf.csv'")
```

Lampiran 10. Proses Metode KNN dan Confusion Matrix

```
#Install dan Import Library
!pip install seaborn matplotlib --quiet

import pandas as pd
from scipy.sparse import load_npz
from sklearn.model_selection import KFold, cross_val_score,
cross_val_predict
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import classification_report,
confusion_matrix, accuracy_score
import seaborn as sns
import matplotlib.pyplot as plt

#Upload File
from google.colab import files
uploaded = files.upload()

#Load CSV
df = pd.read_csv('TF-IDF data gabungan.csv')
X = df.drop(columns=['label']) #Fitur TF-IDF
y = df['label'] #Target / label
(positif/negatif, dst)

#KNN dan K-Fold Cross Validation
knn = KNeighborsClassifier(n_neighbors=2)
kf = KFold(n_splits=10, shuffle=True, random_state=42)

# Akurasi tiap fold
scores = cross_val_score(knn, X, y, cv=kf,
scoring='accuracy')
print("🇮🇩 Akurasi per Fold:", scores)
print("📊 Rata-rata Akurasi: {:.2f}%".format(scores.mean()
* 100))

#Prediksi dan Evaluasi
y_pred = cross_val_predict(knn, X, y, cv=kf)
print("\n📄 Classification Report:")
```

```
print(classification_report(y, y_pred))

#Confusion Matrix
labels = y.unique().tolist() # Bisa jadi ['positif',
'negatif']
cm = confusion_matrix(y, y_pred, labels=labels)

#Gambar Heatmap Confusion Matrix
plt.figure(figsize=(6, 5))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues',
            xticklabels=labels, yticklabels=labels)
plt.title('Confusion Matrix (KNN + K-Fold)')
plt.xlabel('Predicted Label')
plt.ylabel('True Label')
plt.tight_layout()
plt.show()
```



Lampiran 11. Proses Metode Naive Bayes dan Confusion Matrix

```
#Install dan Import Library
!pip install seaborn matplotlib --quiet

import pandas as pd
from sklearn.model_selection import KFold, cross_val_score,
cross_val_predict
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import classification_report,
confusion_matrix
import seaborn as sns
import matplotlib.pyplot as plt

#Upload CSV
from google.colab import files
uploaded = files.upload()

#Load Data
df = pd.read_csv('TF-IDF data gabungan.csv')
X = df.drop(columns=['label']) #Fitur TF-IDF
y = df['label'] #Label / target

#Naive Bayes & K-Fold Cross Validation
nb = MultinomialNB()
kf = KFold(n_splits=10, shuffle=True, random_state=42)

# Akurasi per fold
scores = cross_val_score(nb, X, y, cv=kf,
scoring='accuracy')
print("🇮🇩 Akurasi per Fold:", scores)
print("📊 Rata-rata Akurasi: {:.2f}%".format(scores.mean()
* 100))

#Prediksi untuk Evaluasi
y_pred = cross_val_predict(nb, X, y, cv=kf)
print("\n📋 Classification Report:")
print(classification_report(y, y_pred))
```

```
#Confusion Matrix
labels = y.unique().tolist()
cm = confusion_matrix(y, y_pred, labels=labels)

#Visualisasi Confusion Matrix
plt.figure(figsize=(6, 5))
sns.heatmap(cm, annot=True, fmt='d', cmap='Oranges',
            xticklabels=labels, yticklabels=labels)
plt.title('Confusion Matrix (Naive Bayes + K-Fold)')
plt.xlabel('Predicted Label')
plt.ylabel('True Label')
plt.tight_layout()
plt.show()
```

