

Klasifikasi Abstrak Skripsi Teknik Informatika Menggunakan Metode Naïve Bayes

Natasya Umi Fitria^{1,*}, Rully Mujiatuti²

^{1,2}Teknik Informatika, Teknik, Universitas Muhammadiyah Jakarta, Jalan Cempaka Putih Tengah XXVII, Cempaka Putih, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta, 10510

*E-mail : 20200410700034@student.umj.ac.id

ABSTRAK

Skripsi merupakan tugas akhir yang wajib dibuat oleh mahasiswa sebagai salah satu persyaratan untuk memperoleh gelar kesarjanaan dalam lingkungan akademis. Pada saat mengajukan judul skripsi biasanya mahasiswa membutuhkan topik dan biasanya mahasiswa bingung dalam mendapatkan topik yang akan dibuat dalam skripsi. Salah satu cara yang dilakukan mahasiswa dalam menemukan topik yaitu dengan mencari referensi dari skripsi yang sebelumnya. dengan mengelompokan skripsi berdasarkan topik akan mempermudah mahasiswa dalam mencari referensi. Untuk pengelompokan skripsi dilakukan berdasarkan abstrak dalam skripsi karena di dalam abstrak terdapat informasi secara singkat yang terdapat dalam skripsi mengenai isinya dan sebuah ringkasan yang ada dalam karya ilmiah yang singkat, jelas, akurat dan padat serta mudah dibaca.karena itu penelitian kali ini mengambil domain *Natural Language Processing (NLP)* untuk menciptakan suatu project agar memudahkan para mahasiswa dalam pengerjaan skripsi, dengan judul *Klasifikasi Abstrak Skripsi Teknik Informatika Menggunakan Metode Naïve Bayes*. Penelitian ini dilakukan untuk melanjutkan penelitian yang sudah ada akan tetapi belum sempurna. Karena pada penelitian sebelumnya hasil dari akurasi pengelompokan data menggunakan metode *SVM* yang sangat kecil yaitu masih dibawah 1%. Maka dari itu dilakukan pengelompokan data kembali dengan menggunakan metode *naïve bayes* dengan harapan bisa menghasilkan output data yang lebih sempurna dan akurat atau dapat memiliki hasil akurasi 57% atau 0.57.

Kata kunci: *Abstrak, naïve bayes, Neural Language Processing, Skripsi*

ABSTRACT

The thesis is a final project that must be made by students as one of the requirements to obtain a bachelor's degree in an academic environment. When submitting a thesis title, students usually need a topic and students are usually confused about getting the topic to be made in the thesis. One way that students do in finding topics is by looking for references from previous theses. Grouping theses by topic will make it easier for students to find references. The grouping of these is done based on the abstract in the thesis because in the abstract there is brief information contained in the thesis regarding its contents and a summary contained in scientific work that is short, clear, accurate, concise, and easy to read. Therefore this research takes the Natural domain Language Processing (NLP) to create a project to make it easier for students to work on their thesis, with the title Abstract Classification of Informatics Engineering Thesis Using the Naïve Bayes Method. This research was conducted to continue existing research but it was not perfect. Because in previous research the results of the accuracy of grouping data using the SVM method were very small, which was still below 1%. Therefore, we grouped the data again using the Naïve Bayes method with the hope of producing more perfect and accurate data output or having an accuracy of 57% or 0.57.

Keywords: *Abstract, Naïve bayes, Neural Language Processing, Thesis.*

1. PENDAHULUAN

Skripsi merupakan tugas akhir yang wajib dibuat oleh mahasiswa sebagai salah satu persyaratan untuk memperoleh gelar kesarjanaan dalam lingkungan akademis (R., 2014). Pada saat mengajukan judul skripsi biasanya mahasiswa membutuhkan topik dan biasanya mahasiswa bingung dalam mendapatkan topik yang akan dibuat dalam skripsi. Salah satu cara yang dilakukan mahasiswa dalam menemukan topik yaitu dengan mencari referensi dari skripsi yang sebelumnya. Tetapi dalam hal mencari referensi kebanyakan perpustakaan masih menggunakan list tertulis dalam bentuk daftar judul skripsi dan masih belum dikelompokkan berdasarkan topiknya. Maka dari itu, dengan mengelompokkan skripsi berdasarkan topik akan mempermudah mahasiswa dalam mencari referensi.

Untuk pengelompokan skripsi dilakukan berdasarkan abstrak dalam skripsi karena di dalam abstrak terdapat informasi secara singkat yang terdapat dalam skripsi mengenai isinya dan sebuah ringkasan yang ada dalam karya ilmiah yang singkat, jelas, akurat dan padat serta mudah dibaca. Pada umumnya abstrak skripsi berupa teks sehingga dalam pengelompokannya melibatkan data teks yang tidak terstruktur, maka dari itu solusi untuk menemukan pola yang diinginkan untuk dijadikan kunci dari pengelompokan dapat digunakanlah Teknik text mining. (Yudiarta, 2018)

Dalam pengelompokan data abstrak skripsi ini dengan tepat digunakannya metode *Naive Bayes*. Metode *Naive Bayes classifier* ini bekerja sangat baik dibandingkan dengan model classifier lainnya. *Naive Bayes* memiliki kelebihan yaitu sederhana, cepat dan memiliki akurasi yang tinggi. Hal ini dibuktikan oleh (Xhemali, Hinde Stone) dalam jurnalnya "*Naive Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages*" mengatakan bahwa "*Naive Bayes Classifier* memiliki tingkat akurasi yang lebih baik dibanding model classifier lainnya" (Hidayat, 2016). Kelebihan dari *Naive Bayes* juga bisa dipakai untuk data kuantitatif dan kualitatif, tidak memerlukan jumlah data yang banyak, tidak memerlukan

data training yang banyak, perhitungan yang cepat dan efisien, mudah dipahami dan bisa untuk klasifikasi biner ataupun multiclass. Selain itu, untuk memperbaiki dari penelitian yang sudah ada sebelumnya yang belum sempurna dengan menggunakan Metode *K-Means Clustering* dimana pada saat visualisasi cluster belum terkelompok dengan sempurna dan masih banyak bercampur satu sama lain dan dengan menggunakan 200 data yang di dapatkan cluster sebanyak 9 cluster dengan hasil evaluasi clusternya yaitu 0.0447. Oleh sebab itu pada penelitian ini, dilakukanlah pengelompokan skripsi dengan judul "Klasifikasi Abstrak Skripsi Teknik Informatika Menggunakan Metode *Naive Bayes*".

2. METODE PENELITIAN

Dalam penelitian ini menggunakan enam metode sebagai berikut :

1. *Problem Scoping*

Proses awal untuk melihat permasalahan yang akan diambil nantinya untuk melanjutkan proses penelitian berikutnya.

2. *Data Acquisition*

Data Acquisition ini termasuk proses pengumpulan data awal yang dimana akan digunakan untuk ke proses selanjutnya.

3. *Data Exploration*

Pada proses ini data yang sudah dikumpulkan akan diproses kembali untuk mendapatkan data bersih atau teks bersih.

4. *Modelling*

Pada proses ini dilakukannya *modelling* ke dalam model python yang telah ditentukan, proses *modelling* kali ini menggunakan *Naive Bayes*.

5. *Evaluation*

Setelah proses *modelling* selesai dilakukan, kemudian dilanjutkan proses evaluasi ini, apakah dari proses *modelling* menggunakan metode tersebut berhasil baik atau tidak akurasinya.

6. *Deployment*

Proses ini dilanjutkan untuk proses penyambungan ke situs web atau

aplikasi yang akan dibuat nantinya dalam suatu project.

3. HASIL DAN PEMBAHASAN

Tahapan penelitian yang akan dilakukan saat melakukan klasifikasi abstrak skripsi teknik informatika, namun yang penulis gunakan ini hanya lima tahapan sesuai pengerjaan penulis sebelumnya. Dibawah ini akan dijelaskan lebih lanjut dalam tahapan penelitian.

1. Problem Scoping

Permasalahan yang diangkat adalah pengelompokan data skripsi khususnya mahasiswa teknik informatika berdasarkan abstrak. Penelitian ini dilakukan untuk melanjutkan penelitian yang sebelumnya sudah pernah ada akan tetapi belum berhasil. Maka dari itu kami ingin mengambil topik ini untuk melanjutkan dan membuktikan apakah penggunaan metode yang kami ambil sudah akurat. Selain itu penelitian ini dapat membantu mahasiswa semester akhir yang sering kali bingung dengan referensi dalam mencari topik untuk menyusun skripsi yang terhitung jarang dalam perpustakaan kampus serta dosen pembimbing yang sibuk dan kurang responsif.

Dalam perpustakaan kampus pengelompokan skripsi masih manual dengan melihat dari judul skripsi, di mana terkadang judul belum tentu menggambarkan semua isi dari topik yang dibahas. Untuk mengatasi masalah tersebut kami akan melakukan pengelompokan data skripsi berdasarkan abstrak dengan menggunakan metode *naive bayes* dengan judul **“Klasifikasi Abstrak Skripsi Teknik Informatika Menggunakan Metode Naive Bayes”**. Disamping itu kami juga ingin menciptakan teknologi yang dapat memberikan kemudahan kepada mahasiswa ketika ingin mencari referensi topik skripsi.

2. Data Acquisition

Langkah selanjutnya yaitu *Data Acquisition*. Data yang digunakan adalah data dari penelitian sebelumnya yaitu data skripsi dari mahasiswa teknik informatika Universitas Muhammadiyah Jakarta. Data ini diberikan kepada kami untuk

meneruskan penelitian yang sebelumnya belum berhasil dengan menggunakan metode berbeda dari penelitian sebelumnya. Data yang digunakan adalah dalam bentuk file *csv* dengan jumlah 200. Untuk penelitian ini kami hanya menggunakan 105 data dengan perbandingan yang sama 35 data dari masing-masing *Artificial Intelligence* (AI), Rekayasa Perangkat Lunak (RPL), dan Jaringan (JK), seperti *Tabel 3.1*

Tabel 3.1 Data Acquisition (Pengumpulan data awal)

No	NAMA	JUDUL	ABSTRAK
1	Ahmad Sujai	Perancangan Sistem Keamanan Kunci Rumah Dengan...	Tindakan pencurian barang berharga dan dokumen...
2	Akhmad Afandi	Implementasi Fiber To The Kos Dengan Mikrotik ...	Provider internet paling stabil dan sedang tre...
3	Ichsan Wiratama	Sistem Keamanan Jaringan Komputer Dan Data Den...	Keamanan data dan informasi membuat teknologi ...
4	Rizky Wahyu Ismail	Metode Penetration Testing pada Keamanan Jaringan...	jaringan Wireless selama ini digunakan sebagai...
5	Bukhari Hasan	Implementasi Log Event Management Web Application...	Aplikasi berbasis web merupakan salah satu jenis...
...
101	Rizky Shaumita Hefani	Pembuatan Aplikasi Help Desk Pelaporan Keluhan...	Help Desk merupakan centre point dimana keluhan...

Pada pengumpulan data dalam tabel 3.1 diatas ini kita mengumpulkan 105 data untuk digunakan di tahap selanjutnya, namun pada tabel diatas dimasukkan contoh data berisi 101 data karena suatu hal kondisi yang tidak memungkinkan memasukkan semua data.

3. Data Exploration

Dalam pembuatan klasifikasi abstrak pada skripsi teknik informatika ini, terdapat 200 data, akan tetapi yang digunakan sebanyak 105 data dalam melakukan

pengujian dengan data. Berikut data awal yang dipakai pada *Tabel 3.2*

Tabel 3.2 Data Awal

No	NAMA	JUDUL	ABSTRAK
1	Ahmad Sujai	Perancangan Sistem Keamanan Kunci Rumah Dengan...	Tindakan pencurian barang berharga dan dokumen...
2	Akhmad Afandi	Implementasi Fiber To The Kos Dengan Mikrotik ...	Provider internet paling stabil dan sedang tre...
3	Ichsan Wiratama	Sistem Keamanan Jaringan Komputer Dan Data Den...	Keamanan data dan informasi membuat teknologi ...
4	Rizky Wahyu Ismail	Metode Penetration Testing pada Keamanan Jaringan...	jaringan Wireless selama ini digunakan sebagai...
5	Bukhari Hasan	Implementasi Log Event Management Web Application...	Aplikasi berbasis web merupakan salah satu jen...
...
101	Rizky Shaumita Hefani	Pembuatan Aplikasi Help Desk Pelaporan Keluhan...	Help Desk merupakan centre point dimana keluhan...

Untuk melakukan tahapan lanjut, dimasukkan terlebih dahulu file data yang berjumlah 105 tersebut agar data awal dapat terbaca, maka digunakannya fungsi berikut :

```
dfuji=pd.read_csv("/content/dataAbstrak.csv")
```

Gambar 3.1 Perintah tampilkan data awal

Fungsi dari *dfuji* ini untuk menguji data dengan cara memanggil data untuk pemrosesan lanjut, maka dengan *dfuji* ini data awal dimasukkan dan *pd.read* data akan ditampilkan untuk dibaca.

Maka upload dulu data awal yang akan digunakan, panggil data tersebut untuk dimunculkan pada *collabnya*.

```
import seaborn as sns
sns.countplot(x="label",
data=dfuji,)
```

Gambar 3.2 Perintah untuk membuat label pada data

Selanjutnya, labelkan data menjadi 3 data peminatan sesuai dengan peminatan program studi teknik informatika yang ada di Universitas Muhammadiyah Jakarta yaitu *Artificial Intelligence (AI)*, *Rekayasa Perangkat Lunak (RPL)*, dan *Jaringan (JK)*. Hasilnya seperti contoh pada *Tabel 3.3*

Tabel 3.3 Hasil Label Data

Unnamed : 0	No	NAMA	JUDUL	ABSTRAK
0	1	Ahmad Sujai	Perancangan Sistem Keamanan Kunci Rumah Dengan...	Tindakan pencurian barang berharga dan dokumen...
1	2	Akhmad Afandi	Implementasi Fiber To The Kos Dengan Mikrotik ...	Provider internet paling stabil dan sedang tre...
2	3	Ichsan Wiratama	Sistem Keamanan Jaringan Komputer Dan Data Den...	Keamanan data dan informasi membuat teknologi ...
3	4	Rizky Wahyu Ismail	Metode Penetration Testing pada Keamanan Jaringan...	jaringan Wireless selama ini digunakan sebagai...
4	5	Bukhari Hasan	Implementasi Log Event Management Web Application..	Aplikasi berbasis web merupakan salah satu jen...
...
100	101	Rizky Shaumita Hefani	Pembuatan Aplikasi Help Desk Pelaporan Keluhan...	Help Desk merupakan centre point dimana keluhan...
101	102	Evi Ratmini	Pembangunan Sistem Informasi Administrasi Laya...	Pelayanan yang baik kepada pelanggan merupakan...
102	103	Anita Sari	Applikasi Penjualan Dan Inventori Produk Pipa...	PT. Rifeng International Indonesia (RII) adala...
103	104	Herawati Widi Astutik	Sistem Informasi Purchase Order	pengelolaan menjadi suatu

			Commercial Div...	rangkaian pekerjaan y...
104	105	Indra Sasmita	Penerapan Metode Fifo Pada Sistem Pengelolaan ...	Seiring dengan perkembangan waktu, kebutuhan a...

Pada tabel 3.3 telah dilakukan pemisahan untuk diberikan label berdasarkan peminatan, *Artificial Intelligence* (AI), Rekayasa Perangkat Lunak (RPL), maupun Jaringan (JK).

```
dfuji['Teksbersih']=dfuji.ABSTRAK.
apply(lambda x: preprocessing(x))
dfuji.head()
```

Gambar 3.3 Perintah menambah data baru, data Teksbersih

Fungsi ini akan menghasilkan data baru yaitu data teksbersih, maka di gunakannya fungsi pengujian seperti diatas. Dan *dfuji.head()* digunakan untuk menampilkan baris data teratas pada dataset.

Setelah data dilabeli. Pada gambar 3.3 dilakukan perintah untuk menambah data baru, langkah selanjutnya mencari teks bersihnya. Untuk data teks bersihnya dilakukan setelah *preprocessing*, kemudian diuji menggunakan *dfuji['Teksbersih']* seperti code yang ada diatas yang di mana dibersihkan dari abstrak sebelumnya dengan hasil teks bersihnya seperti pada Tabel 3.4

Tabel 3.4 Hasil Teks Bersih

Unnamed: 0	No	NAMA	JUDUL	ABSTRAK	label	Teksbersih
0	1	Ahmad Sujai	Perancangan Sistem Keamanan Kunci Rumah Dengan ...	Tindakan pencurian barang berharga dan dokumen ...	JK	tindak curi barang harga dokumen salah satu bo...
1	2	Akhmad Afandi	Implementasi Fiber To The Kos Dengan Mikrotik ...	Provider internet paling stabil dan sedang tre...	JK	provider internet stabil tren jaman provider ...

2	3	Ichsan Wiratama	Sistem Keamanan Jaringan Komputer Dan Data Den...	Keamanan data dan informasi membuat teknologi ...	JK	keamanan data informasi teknologi informasi baharu...
3	4	Rizky Wahyu Ismail	Metode Penetration Testing pada Keamanan Jaringan...	jaringan Wireless selama ini digunakan sebagai...	JK	jaring wireless selama sedia internet manfaat ...
4	5	Bukhari Hasan	Implementasi Log Event Management Web Application...	Aplikasi berbasis web merupakan salah satu jenis...	JK	aplikasi basis web salah jenis aplikasi perlu in...
...
100	101	Rizky Shaumita Hefani	Pembuatan Aplikasi Help Desk Pelaporan Keluhan...	Help Desk merupakan centre point dimana keluhan...	RP L	help desk centre point dimana keluhan minta lapor ...
101	102	Evi Ratmini	Pembangunan Sistem Informasi Administrasi Layanan...	Pelayanan yang baik kepada pelanggan merupakan...	RP L	layan langgan aspek perhati bisnis layan langgan n...
102	103	Anita Sari	Aplikasi Penjualan Dan Inventori Produk Pipa Pada...	PT. Rifeng Internasional Indonesia (RII) adala...	RP L	pt rifeng internasional indonesia a rii usaha ...
103	104	Herawati Widi Astutik	Sistem Informasi Purchase Order Commercial Div...	pengelolaan menjadi suatu rangkaian pekerjaan ...	RP L	pengelolaan rangkaian kerja kelompok orang proses...
104	105	Indra Sasmita	Penerapan Metode Fifo Pada Sistem	Seiring dengan perkembangan waktu,	RP L	seiring perkembangan butuh sistem informas

			Pengelo laan ...	kebutuha n a...		i tingkat s...
--	--	--	---------------------	--------------------	--	-------------------

Pada *tabel 3.4* merupakan data hasil teks bersih, yang dimana disini disinkronkan dalam pembahasan yang lebih mudah atau simple.

4. Modelling

Pada tahapan saat *modelling* klasifikasi ini, dilakukan dengan menggunakan metode *Naïve Bayes*

4.1. Naïve Bayes

Modelling pada penelitian ini menggunakan model *Gaussian Naive Bayes* di mana yang pertama kali harus dilakukan adalah import modul yang diperlukan, kemudian membagi *data train* dan *data test* dari *x* dan *y*. Setelah itu dilakukan pembobotan dengan perhitungan *tf-idf* dari *data train* dan *data test*. Lalu *data train* dan *data test* ini diterapkan pada model *Gaussian Naive Bayes* dengan dipanggil librarynya. Seperti *code* dibawah ini :

```
import time
from sklearn.naive_bayes import GaussianNB
from sklearn.model_selection import train_test_split

text_classifier_nb = GaussianNB()
t0_nb = time.time()
text_classifier_nb.fit(x_train, y_train)
t1_nb = time.time()
```

Fungsi pada data diatas ini digunakan untuk pemanggilan file dalam bentuk hitungan waktu

Kemudian dilakukan prediksi data untuk melihat akurasi dari data yang sudah di train tadi. Seperti pada *Gambar 3.4*

```
predictions_nb = text_classifier_nb.predict(x_test)
t2_nb = time.time()
time_linier_train_nb = t1_nb-t0_nb
time_linier_predict_nb = t2_nb-t1_nb

print("NB Training time: %fs; Prediction time: %fs" % (time_linier_train_nb, time_linier_predict_nb))
```

```
NB Training time: 0.004412s;
Prediction time: 7.151770s
```

Gambar 3.4 Proses hitungan waktu

```
y_pred = text_classifier_nb.predict(x_test)
y_pred

array(['AI', 'JK', 'AI', 'RPL', 'RPL', 'RPL', 'AI', 'JK', 'JK', 'AI', 'AI', 'JK', 'AI', 'JK', 'RPL', 'RPL', 'AI', 'RPL', 'RPL', 'JK', 'JK'], dtype='<U3')
```

Gambar 3.5 Hasil Prediksi

Pada gambar 3.5 merupakan hasil prediksi pada metode *naïve bayes classifier*. Dan fungsi ini menentukan hasil nilai probabilitas dari *x_test*.

5. Evaluation

Tahap ini menghasilkan metrics untuk prediksi tingkat akhir, seberapa banyak nilai prediksi yang dikeluarkan dalam penelitian kali ini. Setelah didapatkannya jumlah nilai prediksinya akan kita hasilkan juga pada perhitungan akurasi akhir, berapa yang akan didapat pada penelitian dengan menggunakan metode *naïve bayes* ini.

Input :

```
from sklearn.metrics import confusion_matrix
confusion_matrix(y_test, y_pred)
```

Output :

```
array([[3, 0, 3],
       [2, 6, 1],
       [2, 1, 3]])
```

Gambar 3.6 Hasil Confusion Matrix Dilakukannya *confusion_matrix* untuk mempermudah membaca kekeliruan matrix data

Input :

```
from sklearn.metrics import classification_report
print(classification_report(y_test, y_pred))
```

Output :

```
precision recall f1score support
```

AI	0.43	0.50	0.46	6
JK	0.86	0.67	0.75	9
RPL	0.43	0.50	0.46	6
accuracy			0.57	21
macro avg	0.57	0.56	0.56	21
weighted avg	0.61	0.57	0.59	21

Gambar 3.7 Hasil Akurasi Pada Klasifikasi Fungsi ini menghitung nilai akurasi dari klasifikasi *naïve bayes*

Metric yang digunakan untuk evaluasi model aplikasi ini adalah *precision*, *recall*, dan *f1-score*. Pada gambar 4.6 diatas Hasil data yang sudah dilakukan evaluasi memiliki akurasi 57% atau 0.57 dari *data train* yang sudah diprediksi. Dengan hasil *recall* rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Untuk *Artificial Intelligence* (AI) 0.50, Jaringan (JK) 0.67 dan Rekayasa Perangkat Lunak (RPL) mendapat *recall* 0.50 yang sama dengan *Artificial Intelligence* (AI) dan akurasinya 0.57.

4. KESIMPULAN

Berdasarkan penelitian diatas, dapat disimpulkan bahwa klasifikasi abstrak skripsi ini membutuhkan banyak data untuk bisa melihat perbandingan yang sempurna, lalu dengan menggunakan metode *naïve bayes* mendapatkan hasil yang lebih dibanding dengan metode penelitian terdahulu yang belum selesai menggunakan metode SVM. Dan juga akurasinya cukup memuaskan dengan metode *naïve bayes*.

DAFTAR PUSTAKA

- Abdhu, Y. (2023, Mei 6). *Apa itu Skripsi? Pengertian, Jenis, Struktur dan Karakteristik*. Retrieved September 26, 2023, from <https://deepublishstore.com/blog/apa-itu-skripsi/>
- dqlab.id. (2022). Mengenal naive bayes. *Mengenal naive bayes sebagai salah satu algoritma data science*.

- Dr. Rika Ariyani, M. (2021, November 21). *Skripsi Adalah : Pengertian, Tujuan, Dan Cara Menulisnya*. Retrieved Maret 10, 2023, from RIKAARIYANI.COM: <https://www.rikaariyani.com/2021/11/skripsi-adalah.html>

- Hidayat, A. (2016, Juny 2). *Algoritma Naive Bayes*. Retrieved November 2022, from Algoritma Naive Bayes: <https://arfianhidayat.com/algoritma-naive-bayes>

- Laraswati, B. D. (2022, Agustus 22). *Algoritma Naive Bayes: Definisi dan Contoh Penerapannya*. Retrieved September 22, 2023, from <https://blog.algoritma.algoritma-naive-bayes/>

- Nanda, S. (2022, October 22). *Apa itu Skripsi & Bagaimana Cara Membuatnya? Simak Tahapan & Strukturnya*. Retrieved September 26, 2023, from brainacademy: <https://www.brainacademy.id/blog/apa-itu-skripsi>

- Oliver, A. (2022, Januari 25). *Mengenal Google Colab: Mulai dari Definisi, Cara Menggunakan, hingga Manfaatnya*. Retrieved November 20, 2022, from glints.com: <https://glints.com/id/lowongan/google-colab-adalah/#.Y-AqQXZBzI>

- Rosanti, R. (2014). Pengaruh musik mozart dalam mengurangi stres pada mahasiswa yang sedang skripsi. *Journal of Educational, Health and Community Psychology*, 71-78.

- Tineges, R. (2022, Mei 23). *Mengenal Naive Bayes Sebagai Salah Satu Algoritma Data Science*. Retrieved September 22, 2023, from DQLab: <https://dqlab.id/mengenal-naive-bayes-sebagai-salah-satu-algoritma->

data-
science#:~:text=Definisi%20Algorit
ma%20Naive%20Bayes,-
Naive%20bayes%20merupakan&tex
t=Naive%20bayes%20merupakan%
20metode%20pengklasifikasian%20
berdasarkan%20probabilitas%20se
derhana%20

Yudiarta, N. G. (2018). Penerapan Metode Clustering Text Mining Untuk Pengelompokan Berita Pada Unstructured Textual Data. *Majalah Ilmiah Teknologi Elektro*, Vol. 17. *Majalah Ilmiah Teknologi Elektro*, 334-339.