

## IMPLEMENTASI NAIVE BAYES CLASSIFIER DALAM MENENTUKAN KEAKTIFAN REMAJA

Miftah Hadi Surbakti<sup>1\*</sup>, Ilka Zufria<sup>1</sup>, Suhardi<sup>1</sup>  
Universitas Islam Negeri Sumatera Utara, Medan  
e-mail: <sup>1</sup>miftahhadi14@gmail.com

**Abstract:** *This research aims to implement the Naive Bayes Classifier method to determine the activity level of adolescents, a group that tends to be active in various activities. This method is chosen due to its ability to classify complex data. The data is obtained through surveys involving adolescents from different social environments, covering activities such as sports, academics, social engagements, and others. The implementation process includes data preprocessing, model training with training data, and model testing using test data. The results show that the Naive Bayes Classifier can accurately classify adolescents into active or inactive categories based on the given data. These findings can serve as a reference for educational institutions, youth organizations, and government bodies in developing programs to enhance adolescent activity levels and as a basis for further research in classifying adolescent data using more advanced methods.*

**Keywords:** *Data Mining, Youth Activeness, Mosque Youth, Naive Bayes Classifier, Medan City*

**Abstrak:** Penelitian ini bertujuan untuk mengimplementasikan metode Naive Bayes Classifier dalam menentukan keaktifan remaja, yang merupakan kelompok usia aktif dalam berbagai kegiatan. Metode ini dipilih karena kemampuannya dalam klasifikasi data kompleks. Data diperoleh melalui survei terhadap remaja di berbagai lingkungan sosial, mencakup kegiatan olahraga, akademik, sosial, dan lainnya. Proses implementasi meliputi preprocessing data, pelatihan model dengan data latih, dan pengujian model menggunakan data uji. Hasil penelitian menunjukkan bahwa Naive Bayes Classifier mampu secara akurat mengklasifikasikan keaktifan remaja menjadi kategori aktif atau tidak aktif berdasarkan data yang ada. Hasil ini dapat menjadi acuan bagi lembaga pendidikan, organisasi pemuda, dan pemerintah dalam menyusun program untuk meningkatkan keaktifan remaja serta menjadi dasar untuk penelitian lanjutan dalam pengklasifikasian data remaja menggunakan metode lebih canggih.

**Kata kunci:** Data Mining, Keaktifan Remaja, Remaja Masjid, Naive Bayes Classifier, Kota Medan

### PENDAHULUAN

Forum Komunikasi Remaja Masjid (FKRM) se-Perumnas Mandala adalah lembaga resmi Kecamatan Percut Sei Tuan yang berada di Perumnas Mandala. FKRM se-Perumnas Mandala tergabung dari 12 Remaja Masjid aktif di Perumnas Mandala dan sekitarnya. FKRM se-Perumnas Mandala dibentuk bertujuan untuk menumbuhkan nilai-nilai solidaritas dan kebersamaan antar remaja masjid

yang ada di perumnas mandala dan sekitarnya (Maisurah, 2023). FKRM se-Perumnas Mandala juga memiliki peranan untuk membentuk kader-kader calon pengurus.

Pengurus FKRM se-Perumnas Mandala saat ini sudah memiliki data kader FKRM se-Perumnas Mandala yang akan dijadikan calon pengurus. Namun pengurus FKRM se-Perumnas Mandala masih terkendala karena belum menemukan cara tepat untuk menentukan kriteria

remaja aktif yang akan dijadikan pengurus (Huda, 2022). Dikarenakan jumlah data kader yang banyak, tidak memiliki tingkat akurasi pada Dataset remaja aktif dan adanya potensi perubahan data dari oknum-oknum yang tidak bertanggung jawab, sehingga membuat pengurus FKRM se-Perumnas Mandala sulit serta kurang efektif untuk menentukan kriteria calon pengurus tersebut. Dari fenomena diatas diperlukan Data Mining metode Naïve Bayes Classifier karena dapat mengklasifikasi serta menentukan remaja aktif di FKRM se-Perumnas Mandala secara efisien dengan nilai akurasi yang kompetitif (Zakasih & Handoko, 2022).

Nungky Asmiati dan Fatmawati (2020) judul penelitian ini adalah “Penerapan Algoritma Naïve Bayes Classifier untuk Mengklasifikasi Pengaruh Negatif Game Online Bagi Remaja Milenia”. Dalam penelitian ini menggunakan metode Naïve Bayes Classifier untuk menentukan pengaruh negatif Game Online dengan akurasi mencapai 89%. Suhendro Busono (2020) judul penelitian ini adalah “Optimasi Naïve Bayes Classifier Algoritma Genetika Sebagai Seleksi Fitur Untuk Memprediksi Performa Siswa”. Dalam penelitian ini menggunakan metode Naïve Bayes Classifier untuk menentukan tingkat performa siswa dengan akurasi mencapai 89%. Dari 2 penelitian di atas yang membahas tentang dampak negatif Game Online dan prediksi performa siswa terdapat pembaruan pada penilitan penulis yaitu kedua peneliti di atas menggunakan RapidMiner sedang penulis menggunakan Python. Penulis juga melakukan pembaruan terhadap studi kasus yaitu penentuan remaja aktif.

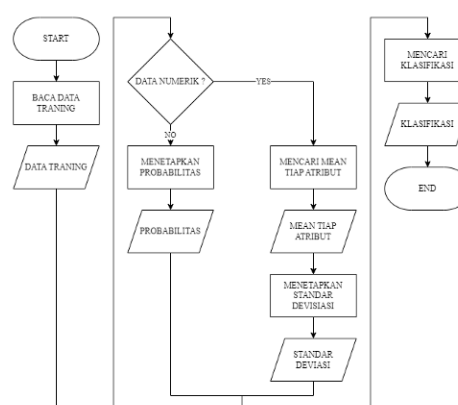
Dari latar belakang di atas peneliti mengambil judul “Implementasi Naïve Bayes Classifier Dalam Menentukan Remaja Aktif di Forum Komunikasi Remaja Masjid se-Perumnas Mandala” diharapkan dengan adanya penelitian ini dapat menjadi jawaban para pengurus Forum Komunikasi Remaja Masjid se-Perumnas Mandala.

## METODE

Teknik pengumpulan data yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Studi Pustaka  
Pada tahapan ini, penulis memperoleh berbagai data dari informasi penunjang lainnya yang berhubungan dan mendukung penelitian (Irawan et al., 2021) (Zulfa et al., 2021).
2. Observasi  
Penulis melakukan pengamatan secara langsung dalam kegiatan pendataan yang dilakukan oleh Forum Komunikasi Remaja Masjid se-Perumnas Mandala agar akurasi/ketepatan pada data yang akan diberikan oleh Forum Komunikasi Remaja Masjid se-Perumnas Mandala bisa menjadi lebih akurat (Suhardi et al., 2023).
3. Wawancara  
Pada tahapan ini, penulis mewawancarai Muhammad Rizky selaku ketua Forum Komunikasi Remaja Masjid se-Perumnas Mandala untuk mengetahui jumlah anggota kemudian melakukan pengumpulan data berdasarkan dari Forum Komunikasi Remaja Masjid se-Perumnas Mandala (Irwanto, 2021).

## Flowchart



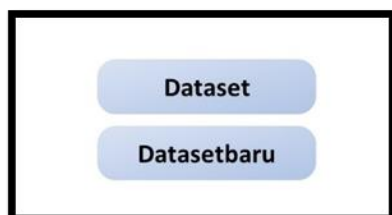
Gambar 1. Flowchart Metode Naïve Bayes Classifier

## HASIL DAN PEMBAHASAN

### Perancangan

#### Perancangan Database

Berikut adalah desain *database* yang digunakan dalam sistem data *mining* untuk menentukan keaktifan remaja menggunakan *naïve bayes classifier*, yaitu:



Gambar 2. Rancangan Database

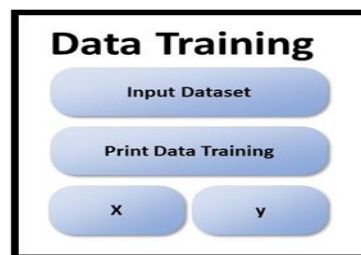
Folder dataset akan berisi file dataset berbasis CSV yang akan digunakan sebagai data *training* pada sistem penentuan keaktifan remaja forum komunikasi remaja masjid se-perumnas mandala. folder datasetbaru akan berisi file dataset baru berbasis csv yang akan digunakan sebagai data testing pada sistem penentuan keaktifan remaja forum komunikasi remaja masjid se-perumnas mandala.

#### Perancangan User Interface

Perancangan *user interface* *input/output* bertujuan untuk membentuk *user interface* sistem yang terintegrasi dengan perangkat lunak agar sistem lebih mudah dipahami. Adapun perancangan user interface terdiri dari halaman *login*, halaman data *training*, halaman data *testing*, halaman proses *naïve bayes classifier*, halaman klasifikasi, halaman laporan dan halaman *logout*. Berikut adalah perancangan *user interface* yang akan diimplementasikan di sistem berbasis *jupyter notebook*.

#### Rancangan Tampilan Halaman Data Training

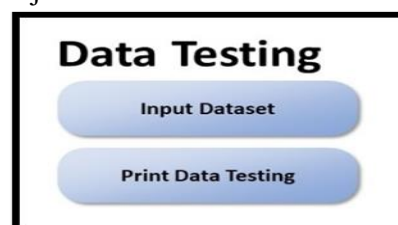
Halaman data *training* adalah menu untuk *user* menginput serta melihat data *training* pada sistem penentuan keaktifan remaja.



Gambar 3. Rancangan User Interface Data Training

#### Rancangan Tampilan Halaman Data Testing

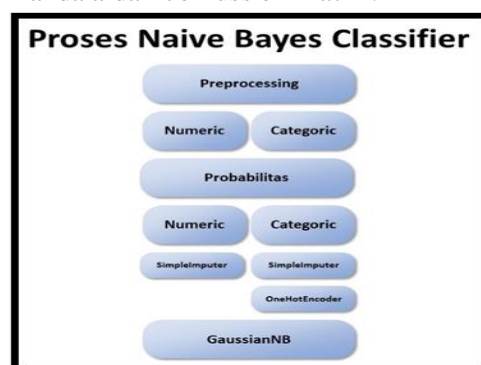
Halaman data *testing* adalah menu untuk *user* menginput serta melihat data *testing* pada sistem penentuan keaktifan remaja.



Gambar 4. Rancangan User Interface Data Testing

#### Naïve Bayes Classifier

Halaman proses *Naïve Bayes Classifier* adalah menu untuk sistem menjalankan *preprocessing*, menentukan probabilitas, klasifikasi keaktifan forum komunikasi remaja masjid se-perumnas mandala dan confusion matrix.



Gambar 5. Rancangan User Interface Proses Naïve Bayes Classifier

#### Rancangan Tampilan Halaman Laporan

Halaman ini merupakan halaman yang akan menampilkan *export* dataset

dan dataset baru yang ada pada *jupyter notebook* ke dalam bentuk *excel* dan disimpan ke dalam folder *python*.



Gambar 6. Rancangan *User Interface* Laporan

### Penerapan

#### Tampilan Halaman Data Training

Halaman data *training* akan dapat diakses, karena *user* telah berhasil memasukkan *password* dengan benar. *User* dapat meng-*input* dataset yang sebelumnya sudah dimasukkan di *database* yang akan digunakan sebagai data *training*. *User* juga dapat melihat *print* dari data *training*.

**Data Training**

```
In [2]: dt = pd.read_csv('Database/dataset/datatraining.csv')
dt
```

```
Out[2]:
```

	Nama	Usia	Kehadiran	Peran	Qiroah	Prestasi	Keaktifan
0	Nawal Ramadhinata	Mencukupi	Sering	Panitia	Mahir	3	1
1	Gilang Syahputra	Mencukupi	Jarang	Panitia	Bisa	1	1
2	Dinda Prasetya	Mencukupi	Sering	Panitia	Tidak Bisa	1	1
3	M. Ramdhani S.	Mencukupi	Sering	Panitia	Mahir	2	1
4	Muhammad Ali Imanan	Mencukupi	Jarang	Panitia	Mahir	2	1
...	...	...	...	...	...	...	...
65	M. Ziki	Mencukupi	Tidak Pernah	Tidak Panitia	Bisa	1	0
66	M. Izzah	Mencukupi	Sering	Panitia	Bisa	1	1
67	Faithi Nardiah	Mencukupi	Tidak Pernah	Panitia	Bisa	1	0
68	Mursyid Dalmardhe	Mencukupi	Jarang	Tidak Panitia	Bisa	1	1
69	Riski Ramadhani Nasution	Mencukupi	tidak pernah	tidak panitia	Bisa	1	0

70 rows x 7 columns

Gambar 7. Tampilan Data Training

#### Tampilan Halaman Data Testing

Halaman data *testing* akan bisa terakses, jika *user* telah menyelesaikan halaman data *training*. *User* dapat meng-*input* dataset baru yang sebelumnya sudah dimasukkan di *database* yang akan digunakan sebagai data *testing*. *User* juga dapat melihat *print* dari data *testing*.

#### Data Testing

```
In [4]: dt = pd.read_csv('Database/dataset/datatesting.csv')
dt
```

```
Out[4]:
```

	Nama	Usia	Kehadiran	Peran	Qiroah	Prestasi	Keaktifan
0	Adi Ihsan	Mencukupi	Tidak Pernah	Panitia	Bisa	1	0
1	Sabalia Sapah	Mencukupi	Jarang	Panitia	Bisa	1	1
2	Lico Gilva	Mencukupi	Jarang	tidak Panitia	Bisa	1	0
3	Dahlan Syahputra	Mencukupi	Sering	Panitia Inis	Bisa	1	1
4	Febrian Rizaldi	Mencukupi	Sering	Panitia Inis	Bisa	1	1
5	Muhammad Haiman	Mencukupi	tidak Pernah	tidak Panitia	Bisa	1	0
6	Pendaria Yasin	Mencukupi	Tidak Pernah	Tidak Panitia	Tidak Bisa	1	0
7	Solihudin N Ihsan	Mencukupi	Tidak Pernah	Tidak Panitia	Mahir	2	0
8	Isyah Aprilia	Mencukupi	Sering	Panitia	Bisa	2	1
9	Riky Syahputra	Mencukupi	Jarang	Tidak Panitia	Mahir	1	0
10	Naufal Nurzaki	Mencukupi	Jarang	Panitia	Mahir	1	1
11	Muhara Zulfar Ikhan	Mencukupi	Jarang	Panitia	Mahir	1	1
12	Anggi Al-hasan	Mencukupi	Tidak Pernah	Tidak Panitia	Tidak Bisa	1	0
13	Rafli Amir	Mencukupi	Jarang	Tidak Panitia	Bisa	1	0
14	Mta Harasta	Mencukupi	Sering	Panitia	Bisa	1	1
15	Alta Harastya	Mencukupi	Jarang	Tidak Panitia	Bisa	1	0
16	Yusuf Irfadi	Mencukupi	Sering	Panitia	Mahir	1	1
17	M. Izza	Mencukupi	tidak Pernah	tidak Panitia	Bisa	2	0
18	Arpan	Mencukupi	Tidak Pernah	Tidak Panitia	Bisa	1	0
19	Ricky Darus	Mencukupi	Tidak Pernah	Tidak Panitia	Bisa	3	0
20	Halm-Ayhan Stepha	Mencukupi	Sering	Panitia	Mahir	2	1
21	Riky Zulfahri	Mencukupi	Jarang	Panitia	Mahir	2	1
22	M. Iqbal Maslana Siregar	Mencukupi	Sering	Panitia Inis	Bisa	1	1
23	Witaya Saththa	Mencukupi	sering	Panitia Inis	Mahir	3	1
24	Dani	Mencukupi	Jarang	Tidak Panitia	Tidak Bisa	1	0
25	Safira Sapah	Mencukupi	Jarang	Tidak Panitia	Bisa	1	0
26	Almohamad	Mencukupi	Jarang	Panitia	Bisa	1	1
27	Muhammad Zainal	Mencukupi	Jarang	Tidak Panitia	Bisa	1	0
28	Azer Akbar	Mencukupi	Tidak Pernah	Tidak Panitia	Bisa	1	0
29	Roy Syahputra	Mencukupi	Jarang	Panitia	Bisa	1	1

Gambar 8. Tampilan Data Testing

#### Tampilan Halaman Proses Naïve Bayes Classifier

Halaman proses *naïve bayes classifier* akan bisa akses, jika *user* telah menyelesaikan halaman data *testing*. Sistem akan melakukan *preprocessing* serta membuat probabilitas dari data *training*, setelah probabilitas didapatkan maka sistem akan melakukan klasifikasi ke data *training* dan data *testing* dan melakukan perhitungan *accuration score*. *User* dapat melihat hasil klasifikasi dan perhitungan *accuration score* akan *diprint*.

**Proses Naive Bayes Classifier**

```
In [3]: preprocessor = ColumnTransformer([
('numeric', num_pipe(), ['prestasi']),
('categorical', cat_pipe(encoder='onehot'), ['usia', 'kehadiran', 'peran', 'qiroah']),
])
```

```
In [7]: from sklearn.naive_bayes import GaussianNB
pipeline = Pipeline([
('prep', preprocessor),
('algo', GaussianNB())
])
```

```
In [8]: pipeline.fit(x_train, y_train)
```

```
Out[8]:
```

```
Preprocessor: ColumnTransformer
- numeric: SimpleImputer
- categorical: SimpleImputer, OneHotEncoder
- GaussianNB
```

```
In [9]: dt["keaktifan"] = pipeline.predict(dt)
```

```
Out[9]:
```

	Nama	Usia	Kehadiran	Peran	Qiroah	Prestasi	Keaktifan
0	Nawal Ramadhinata	Mencukupi	Sering	Panitia	Mahir	3	1
1	Gilang Syahputra	Mencukupi	Jarang	Panitia	Bisa	1	0
2	Dinda Prasetya	Mencukupi	Sering	Panitia	Tidak Bisa	1	1
3	M. Ramdhani S.	Mencukupi	Sering	Panitia	Mahir	2	1
4	Muhammad Ali Imanan	Mencukupi	Jarang	Panitia	Mahir	2	0
...	...	...	...	...	...	...	...
65	M. Ziki	Mencukupi	Tidak Pernah	Tidak Panitia	Bisa	1	0
66	M. Izzah	Mencukupi	Sering	Panitia	Bisa	1	1
67	Faithi Nardiah	Mencukupi	Tidak Pernah	Panitia	Bisa	1	0
68	Mursyid Dalmardhe	Mencukupi	Jarang	Tidak Panitia	Bisa	1	0
69	Riski Ramadhani Nasution	Mencukupi	Tidak Pernah	Tidak Panitia	Bisa	1	0



- HUDA, M. R. (2022). strategi Pengurus Masjid dalam Memakmurkan Masjid Baiturrahman Desa Kuantan Sako Kecamatan Logas Tanah Darat Kabupaten Kuantan Singingi. Universitas Islam Negeri Sultan Syarif Kasim Riau.
- Irawan, T., Dahlan, T., & Fitriani, F. (2021). Analisis Penggunaan Media Video Animasi Terhadap Motivasi Belajar Siswa di Sekolah Dasar. *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, 7(01), 212–225.
- Irwanto, I. (2021). Perancangan Sistem Informasi Sekolah Kejuruan dengan Menggunakan Metode Waterfall (Studi Kasus SMK PGRI 1 Kota Serang-Banten). *Lectura : Jurnal Pendidikan*, 12(1), 86–107. <https://doi.org/10.31849/lectura.v12i1.6093>
- Maisurah, N. (2023). Peran Badan Kemakmuran Masjid Jamik Darussalam Kecamatan Syiah Kuala Kota Banda Aceh Dalam Meningkatkan Kesadaran Keberagamaan Masyarakat. UIN Ar-Raniry Fakultas Ushuluddin dan Filsafat.
- Suhardi, Lubis, A. H., Aprilia, A., & Ningrum, I. A. (2023). Penerapan Metode Simple Multi Attribute Rating Technique pada Pemilihan Cafe Terfavorit. *Sistem Pendukung Keputusan Dengan Aplikasi*, 2(1), 1–11. <https://doi.org/10.55537/spk.v2i1.114>
- Zakasih, M. I., & Handoko, W. T. (2022). Analisis Sentimen Pengguna Twitter Tentang Nft (Non Fungible Token) Dengan Metode Naive Bayes Classifier. *Jurnal Informatika Dan Rekayasa Elektronik*, 5(2), 221–229.
- Zulfa, A., Zufria, I., & Triase, T. (2021). Penerapan Metode Moora-Waspas Pada Sistem Penentuan Calon Penerima Subsidi Tunjangan Fungsional Guru Bpns Di Sma Tamansiswa Tapian Dolok. *JISTech (Journal of Islamic Science and Technology)*, 6(2).