P-ISSN: 1411-3635
E-ISSN: 2962-5645
https://doi.org/10.37817/Tekinfo.v26i1

PERBANDINGAN AKURASI METODE *NAÏVE BAYES*, *DECISION TREE* (C.45), DAN *RANDOM FOREST* DALAM MELAKUKAN PREDIKSI MASA TUNGGU KERJA ALUMNI BERDASARKAN DATA TRACER STUDY PADA FAKULTAS PSIKOLOGI UIN JAKARTA

[1]Sarip Hidayatuloh, [2]Dizar Wangsa Samudera

Sistem Informasi, Universitas Islam Negeri Syarif Hidayatullah Jakarta
E-mail : sarip_hidayatuloh@uinjkt.ac.id , dizar.wangsa21@mhs.uinjkt.ac.id

ABSTRAK

Perguruan tinggi menghadapi tantangan dalam menyiapkan lulusan yang sesuai dengan kebutuhan pasar kerja global. Tracer study menjadi alat penting untuk mengevaluasi masa tunggu kerja alumni, namun analisisnya masih sering dilakukan secara manual. Penelitian ini membandingkan akurasi tiga metode data mining—Naïve Bayes, Decision Tree (C4.5), dan Random Forest—dalam memprediksi masa tunggu kerja alumni Fakultas Psikologi UIN Jakarta berdasarkan data tahun 2020–2022. Dari 300 data mentah, sebanyak 170 data bersih dianalisis. Sistem informasi berbasis web dikembangkan menggunakan metode Rapid Application Development (RAD) dan pemodelan Unified Modeling Language (UML). Hasil evaluasi menunjukkan Naïve Bayes memiliki akurasi tertinggi (53,96%), diikuti Random Forest (50,79%) dan Decision Tree (46,03%). Meski akurasi masih rendah, sistem ini mempermudah analisis dan mengurangi ketergantungan pada metode manual. Penelitian ini menjadi dasar untuk pengembangan analisis tracer study yang lebih efektif dan berbasis teknologi.
Kata kunci: Data Mining, Rancang Bangun, Klasifikasi, RAD, Naïve Bayes

ABSTRACT

Universities today face a major challenge in preparing graduates who are relevant to the needs of the global job market. One way to measure this relevance is through tracer studies, which track alumni to evaluate their job waiting period. However, tracer study data analysis is still often done manually, which requires a more effective approach to obtain more accurate and informative results. This study aims to compare the accuracy of three data mining methods-Naïve Bayes, Decision Tree (C.45), and Random Forest in predicting the job waiting period of alumni of the Faculty of Psychology UIN Jakarta using alumni tracer study data in 2020-2022. From 300 raw data, 170 clean data were selected for analysis. A web-based information system was developed using the Rapid Application Development (RAD) method and Unified Modeling Language (UML) modeling. Three data mining algorithms were applied to classify the data and predict the waiting period of alumni employment. The evaluation results show Naïve Bayes has the highest accuracy of 53.96%, followed by Random Forest (50.79%) and Decision Tree (46.03%). Although the accuracy is still relatively low, this web-based system has succeeded in facilitating analysis and reducing dependence on manual analysis. In conclusion, despite limitations in accuracy, this system has the potential to assist universities in making strategic decisions related to the development of graduate quality and suitability for the labor market. This research provides a basis for further development of tracer study data analysis using data mining methods.
Keywords: Data Mining, Design, Classification, RAD, Naïve Bayes

https://doi.org/10.37817/Tekinfo.v26i1
P-ISSN: 1411-3635
E-ISSN: 2962-5645

1. Pendahuluan

Perguruan Tinggi (PT) adalah suatu lembaga pendidikan taraf akhir yang mencetak lulusan yang nantinya akan memasuki dunia kerja. Persaingan dunia global kerja yang semakin kompleks membawa setiap perguruan tinggi pada suatu konflik yang sama, yakni seberapa relevankah keluaran perguruan tinggi terhadap kebutuhan pengguna lulusan perguruan tinggi saat ini. Hal penting yang dihadapi institusi pendidikan tinggi di Indonesia saat ini yaitu persaingan global (Prabowo & Kodar, 2019). Setiap universitas menerapkan pendekatan berbeda untuk meningkatkan kualitas lulusannya agar lebih sesuai dengan tuntutan pasar tenaga kerja. Oleh karena itu, data tracer study dapat digunakan untuk menentukan apakah lulusan universitas telah mendapatkan pekerjaan (Rachmadiansyah et al., 2022).

Tracer study menjadi salah satu metode dan hal penting bagi perguruan tinggi untuk melacak jejak alumni setelah lulus dari pendidikannya sampai ke dunia kerja. Badan Akreditasi Nasional Perguruan Tinggi (BANPT) menjadikan tracer study sebagai salah satu penilaian dan juga dijadikan sebagai salah satu syarat kelengkapan akreditasi.

Fakultas Psikologi adalah salah satu fakultas yang ada di Universitas Islam Negeri Syarif Hidayatullah Jakarta. Sejauh ini fakultas ini sudah pernah dilakukan pelacakan terhadap status dari durasi masa tunggu kerja alumninya, namu yang sangat disayangkan adalah terdapat beberapa data yang tersebar dan tidak menyatu Kehadiran Pusat Karier UIN Syarif Hidayatullah Jakarta salah tujuannya untuk melakukan penggabungan data tracer study agar menjadi satu kesatuan. Dari pelacakan tersbut nantinya akan di lakukan proses pengelolaan data mining yang diharapkan dapat menghasilkan suatu pengetahuan (Syahputra et al., 2018).

Data mining merupakan suatu proses yang menerapkan keilmuan matematika, statistic, kecerdasan buatan, dan machine learning untuk mengindentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai sumber database. Salah satu metode analisis yang terdapat pada dalam data mining adalah klasifikasi. Salah satu teknik klasifikasi yang popular digunakan adalah Naïve Bayes Classification (NBC).

Pada penelitian ini, penulis bermaksud untuk membandingkan ketiga metode yaitu Naïve Bayes, Decision tree (C45), dan Random Forest untuk menyelesaikan permasalahan dalam tracer study yaitu untuk memprediksi masa tunggu kerja alumni berdasarkan data-data yang tersedia dengan harapan bisa mendapatkan akurasi yang tinggi. Setelah menemukan hasil uji akurasi pada ketiga algoritma tersebut peneliti juga membuat aplikasi data mining berbasis website berdasarkan hasil dari uji akurasi yang paling tinggi.

2. Landasan Teori

a.) Tracer Study

Tracer Study adalah sebuah metode penelitian yang digunakan oleh perguruan tinggi untuk melacak dan menganalisis perkembangan alumni mereka setelah lulus. Hal ini melibatkan pengumpulan data terkait dengan pekerjaan, pendidikan lanjutan, dan pencapaian lainnya dari lulusan perguruan tinggi tersebut. Tujuan dari Tracer Study adalah untuk mengetahui hasil pendidikan dalam bentuk transisi dari dunia pendidikan tinggi ke dunia usaha dan industri, serta untuk mengevaluasi karier dan kondisi para lulusan dari suatu perguruan tinggi. Dengan demikian, Tracer Study dapat membantu mengatasi permasalahan kesenjangan kesempatan kerja dan upaya

P-ISSN: 1411-3635
E-ISSN: 2962-5645
https://doi.org/10.37817/Tekinfo.v26i1

perbaikannya, serta memberikan informasi penting mengenai hubungan antara dunia pendidikan tinggi dengan dunia usaha dan industri (Purwantiningsih et al., 2021).

Tracer Study melibatkan pengembangan sistem informasi, seperti aplikasi web, untuk mengumpulkan data lulusan terkait pekerjaan, pendidikan lanjutan, dan pencapaian lainnya. Data ini digunakan untuk mengevaluasi kualitas pendidikan dan merumuskan strategi peningkatan peluang kerja bagi lulusan (RAZZAQ, 2022).

Dalam sintesis, implementasi Tracer Study adalah proses yang penting untuk meningkatkan kualitas pendidikan dan kesempatan kerja bagi lulusan. Sistem informasi yang dikembangkan harus memungkinkan pengumpulan data yang akurat dan efektif, serta memudahkan penggunaan oleh end user.

b.) Data Mining

Data Mining adalah proses mengekstraksi dan mengidentifikasi informasi relevan dan pengetahuan terkait dari berbagai database besar dengan menggunakan teknik matematika, statistik, kecerdasan buatan, dan pembelajaran mesin. Data mining, juga disebut sebagai penemuan pengetahuan dalam database (KDD), adalah praktik menganalisis data tersimpan dalam jumlah besar menggunakan pendekatan statistik dan/atau matematika untuk mengekstrak hubungan atau pola yang berguna (Papakyriakou & Barbounakis, 2022). Data mining saat ini dapat menangani kumpulan data yang berbeda, dimensi data yang sangat tinggi, dan volume data yang sangat besar. Berikut metode dalam pengelolaan data mining :

1. Predictive modeling adalah penerapan teknik penambangan data untuk menghasilkan perkiraan. Membangun nilai model prediksi dengan atribut tertentu adalah tujuannya. Jaringan saraf, mesin vektor pendukung, regresi linier, dan algoritma lainnya.
2. Association (Asoiasi) Salah satu teknik penambangan data yang melihat hubungan antar data. Menganalisis perilaku siswa yang datang terlambat merupakan salah satu contoh penerapannya. FPGrowth, A Priori, dan algoritma lainnya.
3. Clustering (Klastering) atau pengelompokan merupakan teknik untuk mengidentifikasikan data ke dalam suatu kelompok tertentu. Contoh algoritmanya K-Means, K Medoids, Self-Organitation Map (SOM), Fuzzy C-Means, dan lain-lain.
4. Clasification dalam data mining membagi data menjadi banyak kelompok atau kategori sesuai dengan properti yang terkait dengan data tersebut. Beberapa metode antara lain Decision Trees, Neural Networks, Support Vector Machines, kNN, Naïve Bayes, dan GA digunakan dalam proses kategorisasi ini.

c.) Evaluasi Model

Evaluasi model merupakan proses penting dalam penilaian kinerja suatu model atau algoritma dalam pemrosesan data. Evaluasi model dapat memberikan wawasan tentang seberapa baik atau buruk model tersebut dalam menyelesaikan tugas yang diberikan, seperti klasifikasi, regresi, atau tugas lainnya. Confusion Matrix adalah suatu alat evaluasi dalam bentuk matriks yang digunakan untuk mengevaluasi kinerja model klasifikasi. Menurut[7] Confusion Matrix memiliki empat jenis kejadian yang meliputi :

P-ISSN: 1411-3635
E-ISSN: 2962-5645

1. TP (True Positive) : data yang berhasil terklasifikasi sebagai positif dengan benar.
2. TN (True Negative) : data yang berhasil terklasifikasi sebagai negatif dengan benar.
3. FP (False Positive) : data yang keliru terklasifikasi sebagai positif.
4. FN (False Negative) : data yang keliru terklasifikasi sebagai negatif.

Dengan menggunakan nilai yang terdapat dalam Confusion Matrix, maka dapat menghitung berbagai matrix evaluasi berikut :

1. Akurasi (Accuracy), mengukur sejauh mana model dapat mengklasifikasikan data dengan benar, dapat dihitung menggunakan rumus sebagai berikut :

$$\textbf{accuracy} = \frac{(TP+TN)}{TP+TN+FP+FN}$$

2. Resisi (Precision), mengukur seberapa tepat model dalam memprediksi kelas positif, dapat dihitung menggunakan rumus sebagai berikut :

$$\textbf{presisi} = \frac{TP}{(TP+FP)}$$

3. Recall, mengukur seberapa baik model dalam mengidentifikasi data kelas positif berdasarkan semua data actual yang seharusnya positif, dapat dihitung menggunakan rumus sebagai berikut :

$$\textbf{recall} = \frac{TP}{(TP+FN)}$$

4. F1-Score, matrix yang memberikan keseimbangan antara presisi dan recall, dapat dihitung menggunakan rumus sebagai berikut :

$$\textbf{F1-Score} = 2 \times \frac{(Presisi \times Recall)}{(Presisi+R\ ecall)}$$

d.) Rancang Bangun Sistem Informasi

Rancang Bangun Sistem Informasi (RBSI) adalah proses perancangan dan pengembangan sistem informasi yang dirancang untuk memenuhi kebutuhan spesifik organisasi atau individu. Sistem informasi ini dapat berupa aplikasi web, desktop, atau mobile yang digunakan untuk mengumpulkan, memproses, menyimpan, menganalisis, dan menyebarkan informasi untuk mencapai tujuan yang spesifik (Raffi et al., 2023).

Dalam RBSI, beberapa langkah yang umum dilakukan meliputi:

1. Analisis : Penyusun mengumpulkan dan menganalisis kebutuhan sistem informasi yang akan dibangun. Analisis ini melibatkan identifikasi kebutuhan, tujuan, dan sumber daya yang tersedia.
2. Perancangan : Berdasarkan hasil, penyusun analisis membuat desain sistem informasi yang akan dibangun. Desain ini meliputi definisi sistem, struktur data, dan algoritma yang akan digunakan.
3. Pengembangan : Penyusun mengembangkan sistem informasi yang telah dirancang. Pengembangan ini meliputi pengkodean, pengujian, dan pengembangan fiturfitur yang diperlukan.
4. Pengujian : Sistem informasi yang telah dibangun diuji untuk memastikan bahwa sistem tersebut berfungsi dengan baik dan memenuhi kebutuhan yang telah ditentukan.
5. Pemeliharaan : Sistem informasi yang telah dibangun untuk memastikan bahwa sistem tersebut tetap berfungsi dengan baik dan memenuhi kebutuhan yang telah ditentukan.

3. Metodologi

P-ISSN: 1411-3635
E-ISSN: 2962-5645
https://doi.org/10.37817/Tekinfo.v26i1

3.1 Metode Pengumpulan dan Pengelolaan Data

Proses pengumpulan data tracer study yang dilaksanakan oleh Pusat Karier UIN Jakarta dirancang dengan mempertimbangkan tujuan pimpinan organisasi. Informasi yang diperoleh berasal dari hasil survei tracer study dalam format Comma Separated Values (.CSV). Survei ini menggunakan instrumen baru yang dikembangkan untuk menggali data tentang alumni Fakultas Psikologi UIN Jakarta yang lulus pada tahun 2020, 2021, dan 2022.

1. Jumlah Data :
   - Total data mentah yang dikumpulkan: 300.
   - Total data bersih setelah proses pengolahan: 170.
2. Pengelolaan Data :
   - Data yang telah dikumpulkan dari survei diolah dan dikategorisasi ke dalam beberapa kategori lain sesuai kebutuhan analisis.
   - Data ini diubah menjadi sampel atau dataset dari data master untuk digunakan sebagai data training.
3. Proses Pengelolaan Data :
   - Data dipecah menjadi dua bagian utama:
   - Data Training: Digunakan untuk melatih model.
   - Data Testing: Digunakan untuk menguji model yang telah dilatih.
4. Metode Analisis :
   - Metode yang digunakan adalah Naive Bayes, dengan data training dan testing yang dimanfaatkan untuk membuat tabel probabilitas.

Proses ini bertujuan menghasilkan informasi yang akurat dan relevan untuk mendukung keputusan organisasi terkait alumni.

3.2 Metode Pengembangan Sistem

Penelitian ini dilakukan dengan mengembangkan sistem dengan menggunakan metode pengembangan sistem yang dikenal dengan Rapid Application Development (RAD) dan pemodelan sistem dengan Unified Modeling Language (UML). RAD merupakan pendekatan berorientasi objek untuk pengembangan sistem yang di dalamnya melibatkan metodologi pengembangan dan berbagai perangkat lunak yang mendukungnya[8]. RAD bertujuan mempercepat siklus hidup pengembangan sistem informasi dengan meminimalkan waktu yang biasanya dibutuhkan dalam proses perancangan hingga implementasi pada metode tradisional.

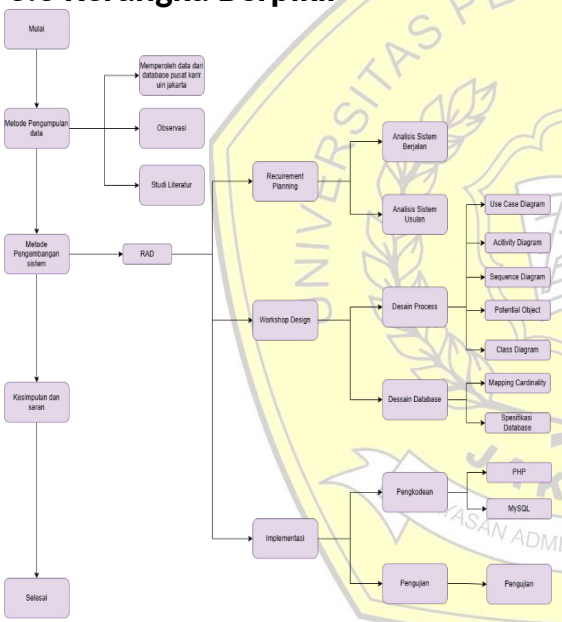
1. Fase Requirement Planning
   Pada tahap ini, peneliti melakukan identifikasi terhadap sistem yang sedang berjalan untuk mengevaluasi kekuatan, kelemahan, dan permasalahannya. Proses ini bertujuan untuk menghasilkan analisis baru sebagai dasar perancangan sistem informasi yang akan dikembangkan.
2. Fase Desain Proses
   Fase ini merupakan tahap perancangan yang didasarkan pada analisis dari tahap sebelumnya, di mana peneliti mulai merancang sistem data mining setelah mengidentifikasi seluruh masalah dan kebutuhan yang ada. Dengan menggunakan *tools* UML, sejumlah diagram dibuat untuk mendukung pengembangan sistem informasi data mining. Tahapan perancangan ini mencakup tiga aspek utama, yaitu *desain proses*, desain basis data, dan desain antarmuka. Pada desain proses, diagram yang dibuat meliputi *use case diagram*, *activity diagram*, *sequence diagram*, dan *class diagram*. Selanjutnya, desain basis data mencakup pemetaan kardinalitas dan spesifikasi basis data.

P-ISSN: 1411-3635
E-ISSN: 2962-5645

Terakhir, desain antarmuka berfokus pada pengembangan tampilan sistem yang akan digunakan oleh pengguna.

3. Fase Implementasi
   Tahap akhir dalam pengembangan sistem informasi data mining naïve bayes adalah implementasi, di mana hasil dari tahap perancangan sebelumnya diterapkan. Pada tahap ini, peneliti melaksanakan dua langkah utama, yaitu proses pengkodean dan pengujian menggunakan metode black box.

3.3 Kerangka Berpikir



Kerangka berpikir penelitian ini diawali dengan mengidentifikasi permasalahan dalam pengelolaan data trecer study yang masih dilakukan secara manual menggunakan ms excel. Langkah selanjutnya adalah pengumpulan data melalui database pusat karir, Observasi, dan studi literatur untuk memahami kebutuhan sistem. Berdasarkan analisis terhadap sistem yang ada dan usulan pengembangan, metode Rapid Application Development (RAD) diterapkan untuk merancang sistem. Pada tahap implementasi, sistem dibangun menggunakan PHP dan MySQL, lalu diuji dengan metode Black Box. Penelitian ini menghasilkan sistem informasi yang dirancang untuk meningkatkan efisiensi dalam pengelolaan data trecer study.

4. Hasil & Pembahasan

4.1 Pemodelan dan Evaluasi Algoritma

Dalam studi penelitian ini, peneliti dapat menggunakan analisis klasifikasi daripada analisis regresi. Hal ini dikarenakan analisis klasifikasi lebih tepat digunakan untuk menentukan kategori atau label yang memiliki bias yang kuat, seperti perpindahan kerja alumni dalam penelitian ini. Metode klasifikasi dapat mengelompokkan data sesuai dengan kategori yang telah ditentukan, seperti "antara tiga sampai dua belas bulan", "tiga sampai enam bulan", "enam sampai dua belas bulan", dan seterusnya, tergantung pada atributnya.

Selain itu, metode klasifikasi yang digunakan dalam penelitian ini - *Naïve Bayes*, *Decision Tree* (C.45), *dan KNearest Neighbors* - telah terbukti efektif dalam berbagai aplikasi dan menyoroti praktik kerja yang tepat untuk mengklasifikasikan atau melabeli data. Selain itu, dengan menggunakan analisis klasifikasi dalam penelitian ini memungkinkan kita untuk memahami faktor-faktor yang mempengaruhi jam kerja alumni berdasarkan banyak atribut seperti program studi, indeks prestasi, pengalaman kerja, dan sebagainya. Dengan demikian, analisis ini dapat memberikan wawasan lebih lanjut mengenai hubungan antara variabelvariabel yang disebutkan di atas dengan masa kerja alumni.

Setelah dilakukan tahapan – tahapan diatas selanjutnya peneliti akan mengevaluasi dari ketiga algortima tersebut dengan menggunakan kode yang akan di tampilkan dibawah ini:

1. Evaluasi Algortima Naïve Bayes

P-ISSN: 1411-3635
E-ISSN: 2962-5645
https://doi.org/10.37817/Tekinfo.v26i1

Setelah dilakukan pemodelan *Naïve Bayes* terhadap data tracer study alumni Fakultas Psikologi UIN Jakarta, dilakukan evaluasi untuk mengetahui performa dari model *Naïve Bayes*.Evaluasi dilakukan dengan menggunakan library *GaussianNB*, accuracy_score,confusion_matrix, classification_report dari sklearn.metrics. Hasil evaluasi model menunjukkan bahwa performa dari model Navie Bayes untuk memprediksi masa tunggu kerja alumni Fakultas Psikologi UIN Jakarta memiliki akurasi sebesar 53,96%.

```
[ ] from sklearn.naive_bayes import GaussianNB

    gnb = GaussianNB()
    prediks = gnb.fit(x_train, y_train).predict(x_test)
    prediks

array([3, 1, 3, 3, 4, 1, 3, 3, 3, 3, 1, 2, 3, 1, 1, 3, 1, 3, 3, 1, 1, 3,
       4, 3, 2, 3, 3, 1, 3, 1, 3, 3, 1, 1, 3, 2, 2, 1, 3, 3, 3, 2, 3, 1,
       3, 1, 4, 3, 3, 1, 3, 1, 3, 4, 1, 3, 1, 1, 1, 1, 3, 1, 3])

[ ] print("Akurasi :", accuracy_score(y_test, prediks))

Akurasi : 0.5396825396825397
```

2. Evaluasi Algoritma Random Forest

Setelah dilakukan pemodelan Random Forest terhadap data tracer study alumni Fakultas Psikologi UIN Jakarta, dilakukan evaluasi untuk mengetahui performa dari model Random Forest. Evaluasi dilakukan dengan menggunakan library RandomForestClassifier, accuracy_score,confusion_matrix, classification_report dari sklearn.metrics. Hasil evaluasi model menunjukkan bahwa performa dari model Random Forest untuk memprediksi masa tunggu kerja alumni Fakultas Psikologi UIN Jakarta memiliki akurasi sebesar 50,79%.

```
[ ] from sklearn.ensemble import RandomForestClassifier

[ ] RF = RandomForestClassifier()
    RF.fit(x_train, y_train)
    pred = RF.predict(x_test)
    pred

array([0, 0, 3, 0, 4, 3, 3, 0, 3, 2, 1, 3, 3, 0, 1, 0, 0, 2, 1, 1, 1, 2,
       3, 3, 3, 2, 3, 1, 3, 1, 3, 3, 0, 1, 3, 3, 3, 1, 3, 3, 1, 2, 3, 1,
       3, 5, 3, 3, 0, 2, 2, 1, 3, 3, 1, 3, 1, 0, 5, 3, 3, 3, 2])

[ ] print("Akurasi :", accuracy_score(y_test, pred))

Akurasi : 0.5079365079365079
```

3. Evaluasi Algoritma Decision Tree

Setelah dilakukan pemodelan Desicion Tree terhadap data tracer study alumni Fakultas Psikologi UIN Jakarta, dilakukan evaluasi untuk mengetahui performa dari model Decision Tree. Evaluasi dilakukan dengan menggunakan library Tree,accuracy_score,confusion_matrix, classification_report dari sklearn.metrics. Hasil evaluasi model menunjukkan bahwa performa dari model Random Forest untuk memprediksi masa tunggu kerja alumni Fakultas Psikologi UIN Jakarta memiliki akurasi sebesar 46,03%.

```
[ ] from sklearn import tree

    pohon = tree.DecisionTreeClassifier()
    pohon.fit(x_train, y_train)
    predik = pohon.predict(x_test)
    predik

array([0, 0, 2, 0, 3, 3, 3, 3, 5, 1, 3, 3, 0, 1, 0, 0, 2, 1, 1, 1, 5,
       3, 3, 3, 0, 2, 1, 3, 1, 2, 2, 0, 2, 2, 3, 3, 1, 3, 2, 1, 2, 3, 1,
       2, 5, 3, 2, 0, 2, 2, 1, 2, 3, 1, 3, 1, 0, 5, 3, 0, 3, 2])

[ ] print("Akurasi :", accuracy_score(y_test, predik))

Akurasi : 0.4603174603174603
```

4. Pengkategorian Hasil Pemdodelan dan Evaluasi

Setelah melakukan hasil uji pemodelan dan evaluasi diperoleh hasil bahwasannya akurasi naïve bayes menjadi algortima yang tertinggi dengan nilai akurasi 53,96%, lalu disusul dengan algoritma Random forest dengan hasil akurasi sebesar 50,79%, dan yang terakhir yaitu algoritma Decision Tree dengan hasil akurasi sebesar 46,03%. Setelah diperoleh hasil akurasi dari ketiga algoritma, langkah selanjutnya adalah membuat pengelompokan kriteria presentasi akurasi dari yang sangat baik hingga sangat buruk. Berikut adalah table presentasi kriteria hasil uji akurasi algortima dalam persentase:

Kategori	Rentang persentase Akurasi
Sangat baik	90% - 100%
Baik	75% - 89%

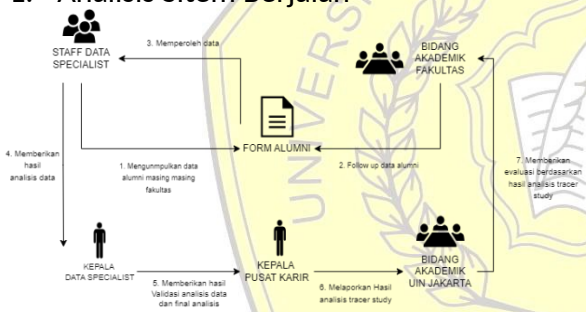
https://doi.org/10.37817/Tekinfo.v26i1
P-ISSN: 1411-3635
E-ISSN: 2962-5645

Cukup	60% - 74%
Buruk	50% - 59 %
Sangat Buruk	< 50%

Dari hasil di atas dapat disimpulkan bahawasannya hasil akurasi berdasarkan data yang dianalisis menggunakan algoritma Naïve bayes, Random Forest, dan Decision Tree mendapatkan kategori buruk untuk Naïve bayes dengan score akurasi 53,96%, lalu utnuk Random Forest mendapatkan kategori buruk dengan score akurasi 50,79% dan yang terakhir adalah algoritma Decision Tree dengan kategori sangat buruk dengan score akurasi46,03%.

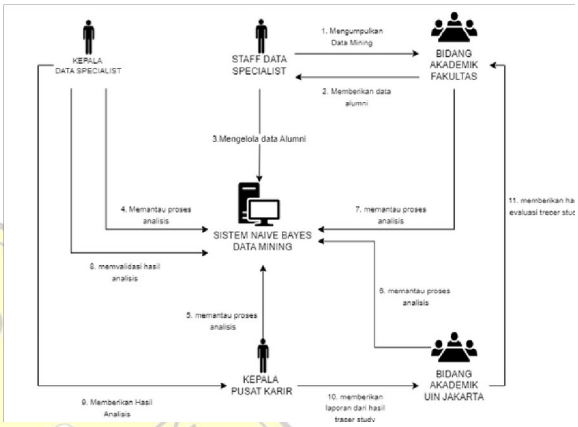
4.2 Fase Requirement Planing

1. Analisis Sitem Berjalan



Berdasarkan Gambar diatas, analisis sistem berjalan untuk proses Tracer Study di Pusat Karir UIN Jakarta dapat dijelaskan sebagai berikut. Proses diawali dengan staf data specialist yang mengumpulkan data untuk diolah lebih lanjut. Selanjutnya, bidang akademik fakultas melakukan tindak lanjut terhadap data alumni fakultas mereka dengan menghubungi alumni dari setiap jurusan yang ada di fakultas tersebut. Setelah itu, staf data specialist mengolah data yang telah dikumpulkan menjadi informasi yang lebih terstruktur. Hasil analisis data tersebut kemudian diserahkan kepada kepala divisi data specialist untuk dilakukan validasi dan finalisasi. Setelah data divalidasi, kepala divisi data specialist memberikan hasil akhir analisis data kepada kepala Pusat Karir. Kepala Pusat Karir kemudian melaporkan hasil analisis Tracer Study kepada bidang akademik universitas. Akhirnya, bidang akademik universitas memberikan evaluasi berdasarkan hasil analisis Tracer Study kepada fakultas masing-masing.

2. Analisis Sistem Usulan



Berdasarkan Gambar diatas, analisis sistem usulan untuk penerapan sistem Naïve Bayes Data Mining pada Pusat Karir UIN Jakarta dapat dijelaskan sebagai berikut. Proses dimulai dengan staf Data Specialist yang mengumpulkan data melalui permintaan kepada pihak akademik fakultas. Setelah itu, pihak akademik fakultas memberikan data alumni kepada staf Data Specialist. Data alumni yang telah diterima kemudian dikelola oleh staf Data Specialist untuk diolah lebih lanjut. Selanjutnya, kepala divisi Data Specialist memantau proses pengelolaan data yang dilakukan pada sistem Naïve Bayes Data Mining. Begitu pula, kepala Pusat Karir, bidang akademik universitas, dan bidang akademik fakultas juga dapat melakukan pemantauan terhadap proses pengelolaan data di sistem Naïve Bayes Data Mining. Setelah data dianalisis, kepala divisi Data Specialist memvalidasi hasil analisis yang dihasilkan oleh sistem Naïve Bayes Data Mining. Hasil analisis tersebut kemudian diserahkan kepada kepala Pusat Karir. Kepala Pusat Karir melaporkan hasil Tracer Study kepada bidang akademik universitas, yang selanjutnya memberikan evaluasi berdasarkan hasil Tracer Study yang telah

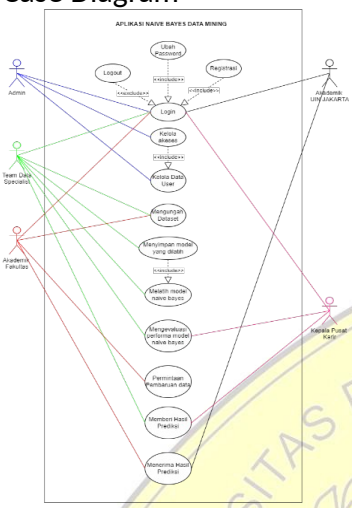
P-ISSN: 1411-3635
E-ISSN: 2962-5645
https://doi.org/10.37817/Tekinfo.v26i1

dilakukan, serta memberikan kebijakan pembelajaran kepada pihak fakultas.

4.3 Fase Workshop Design

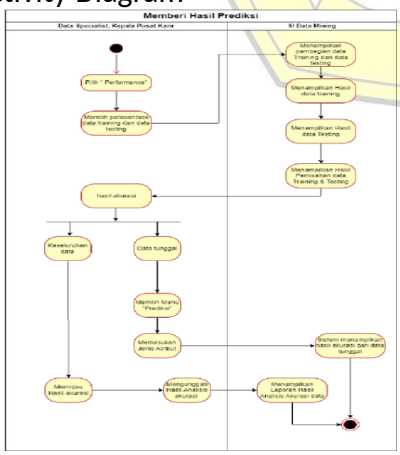
1. Desain Proses

a. Use Case Diagram



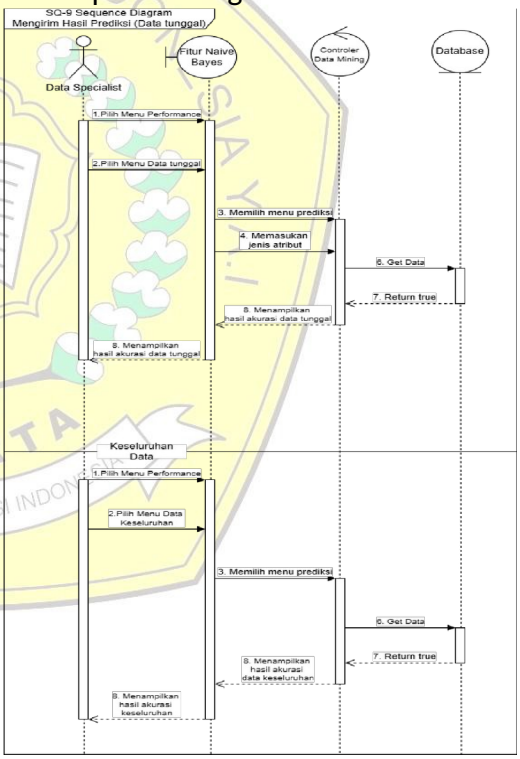
Pada *use case diagram* pengembang bisa melihat alur penggunaan fitur naïve bayes pada web pusat karir uinjakarta oleh setiap actor dan interaksi yang dilakukan dengan sistem. Pada gambar diatas menjelaskan bagaimana use case diagram beperan sebagai petunjuk dalam mengidentifikasi alur sistem, pada gambar tersebut juga akan berisi narasi *use case diagram*.

b. Activity Diagram



Use case ini mengilustrasikan langkah-langkah yang dilakukan oleh Data Specialist dan Kepala Pusat Karir dalam mengevaluasi hasil prediksi dari model yang telah dilatih, serta memberikan fasilitas untuk melakukan prediksi pada data tunggal. Proses dimulai dengan Data Specialist atau Kepala Pusat Karir memilih opsi "Performance" untuk menentukan metrik kinerja yang akan digunakan dalam evaluasi model Setelah itu, sistem menampilkan pembagian data pelatihan dan pengujian yang sudah ditetapkan sebelumnya, diikuti dengan tampilan hasil pelatihan model pada data pelatihan dan hasil pengujian model pada data pengujian. Sistem juga menyajikan hasil evaluasi secara keseluruhan, termasuk perbandingan kinerja model pada kedua set data tersebut, serta menampilkan akurasi model secara keseluruhan.

c. Sequence Diagram



Sequence diagram Mengirim Hasil Prediksi yang menggambarkan alur interaksi antara aktor dan berbagai komponen sistem selama proses Mengirim Hasil Predikisi dapat dilihat pada Gambar diatas.

d. Potential Object

Tabel dibawah merupakan daftar potential object dalam aplikasi informasi data science naive bayes. Objek-objek ini membantu mengidentifikasi kebutuhan dan fitur yang perlu diimplementasikan

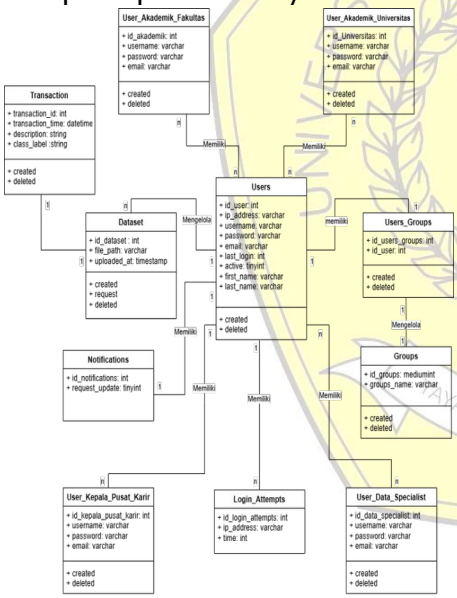
https://doi.org/10.37817/Tekinfo.v26i1
P-ISSN: 1411-3635
E-ISSN: 2962-5645

untuk mendukung fungsi operasional serta interaksi pengguna dengan aplikasi secara efektif

Users
Groups
User_Groups
Notification
Dataset
Login_attempts
*User_*Data *_Specialist*
*User_*Akademik_Fakultas
User_Akademik_Universitas
User_Kepala_Pusat_Karir

e. Class Diagram

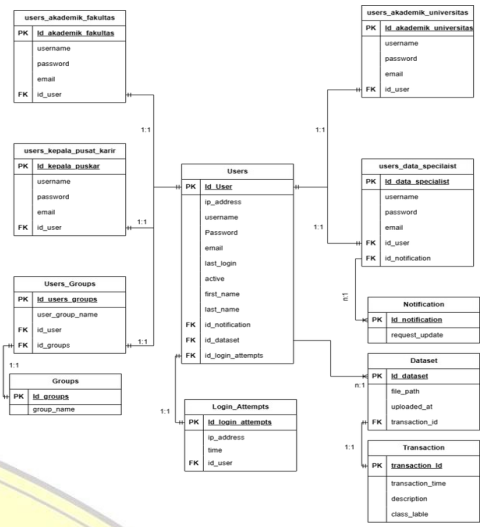
Pada Gambar 10 merupakan pemodelan *class diagram* dari sistem informasi data mining naïve bayes berdasarkan tahaptahap sebelumnya.



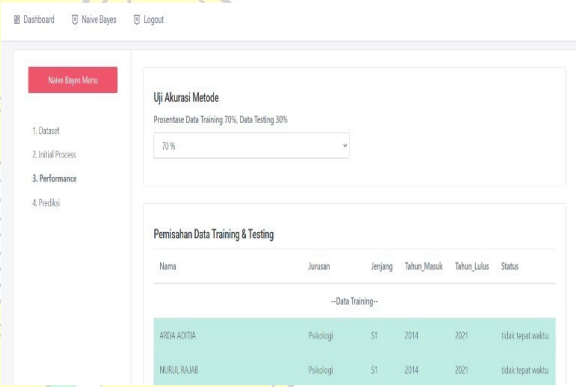
4.4 Desain Database

1. Mapping Cardinality

Setelah membuat *class diagram*, langkah berikutnya adalah membuat *mapping cardinality* yang menjelaskan *primary key* dan *foreign key* di setiap tabel. Berikut adalah *mapping cardinality* dari sistem informasi data mining naïve bayes yang dapat dilihat pada Gambar dibawah ini.



4.5 Desain Interface



Pada Gambar 12 sistem informasi data mining menggunakan metode Naive Bayes, halaman Performance menyajikan fitur "Uji Akurasi Metode" yang berfungsi untuk mengevaluasi keandalan model klasifikasi. Pada halaman ini, pengguna dapat melakukan pembagian dataset menjadi dua bagian: Data Training dan Data Testing dengan fleksibilitas pemilihan persentase pembagian mulai dari 10% hingga 90%. Sebagai contoh default yang ditampilkan, sistem menunjukkan rasio 70% untuk Data Training dan 30% untuk Data Testing, namun pengguna memiliki kebebasan untuk menyesuaikan rasio ini sesuai dengan kebutuhan analisis mereka melalui dropdown menu yang disediakan.

4.6 Fase Implementasi

Tahap terakhir dalam pengembangan sistem informasi data mining adalah tahap

P-ISSN: 1411-3635
E-ISSN: 2962-5645
https://doi.org/10.37817/Tekinfo.v26i1

implementasi, yang melibatkan penerapan hasil dari tahap perancangan sistem sebelumnya. Pada tahap ini, peneliti melaksanakan dua langkah utama, yaitu pengkodean dan pengujian. Pengujian sistem informasi data mining pada dilakukan menggunakan teknik black box testing untuk menguji aspek fungsional dari sistem. Tabel III menyajikan hasil pengujian fungsionalitas sistem pada proses pengelolaan data yang ada yaitu pengevaluasian performa model.

No.	Skenario Pengujian	Hasil yang Diharapkan
1	Memilih model yang telah dilatih	Model berhasil dipilih
2	Menjalankan evaluasi performa model	Sistem memberikan hasil evaluasi performa

No.	Skenario Pengujian	Hasil yang Diharapkan
1	Mengakses halaman permintaan pembaruan dataset	Menampilkan halaman untuk permintaan pembaruan dataset
2	Memasukkan data pembaruan dataset	Data pembaruan berhasil dimasukkan
3	Mengklik tombol permintaan pembaruan	Permintaan pembaruan berhasil diproses

5. Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dijelaskan pada bab sebelumnya, peneliti dapat menyimpulkan beberapa hal. Pertama, hasil penelitian mengenai perbandingan ketiga algoritma menunjukkan bahwa algoritma Naïve Bayes memiliki akurasi tertinggi dalam uji akurasi menggunakan dataset mahasiswa alumni psikologi, dengan nilai akurasi sebesar 53,96%. Berdasarkan hasil uji akurasi tersebut, algoritma Naïve Bayes dipilih sebagai algoritma utama dalam pembuatan aplikasi data mining. Kedua, sistem data mining Naïve Bayes yang dikembangkan mampu membantu dalam analisis data dan pengolahan informasi, mempermudah proses prediksi, mengurangi ketergantungan pada analisis manual, serta meningkatkan efisiensi dalam pengambilan keputusan berbasis data. Ketiga, sistem ini dapat memberikan hasil prediksi dengan cepat dan akurat sesuai dengan kebutuhan, sehingga meminimalisir kesalahan dalam pengolahan data dan memberikan wawasan yang lebih baik mengenai tren serta pola dalam dataset yang dianalisis. Keempat, sistem data mining Naïve Bayes dikembangkan berbasis web menggunakan bahasa pemrograman PHP, yang memungkinkan akses dari berbagai platform hanya dengan menggunakan browser. Sistem ini juga dilengkapi dengan fitur peringatan dan rekomendasi prediksi, yang membantu pengguna dalam mengambil keputusan berdasarkan analisis data dan prediksi yang dihasilkan, sehingga dapat menghindari kesalahan prediksi atau keputusan yang tidak tepat.

P-ISSN: 1411-3635
E-ISSN: 2962-5645

DAFTAR PUSTAKA

Papakyriakou, D., & Barbounakis, I. S. (2022). Data Mining Methods: A Review. *International Journal of Computer Applications*, *183*(48), 5–19. https://doi.org/10.5120/ijca2022921884

Prabowo, F. E., & Kodar, A. (2019). Analisis Prediksi Masa Studi Mahasiswa Menggunakan Algoritma Naïve Bayes. *Jurnal Ilmu Teknik Dan Komputer*, *3*(2), 155–161.

Purwantiningsih, A., Sardjiyo, S., Sudrajat, A., & Permana, S. A. (2021). Layanan Informasi Digital Sebagai Studi Penelusuran Alumni S1 Program Studi PPKn Universitas Terbuka. *Refleksi Edukatika: Jurnal Ilmiah Kependidikan*, *12*(1), 102–109. https://doi.org/10.24176/re.v12i1.6762

Rachmadiansyah, R., Rumlaklak, N. D., & Mauko, A. Y. (2022). Prediksi Masa Tunggu Kerja Alumni Menggunakan Naïve Bayes Classifier Pada Program Studi Ilmu Komputer Universitas Nusa Cendana. *Jurnal Komputer Dan Informatika*, *10*(2), 143–150. https://doi.org/10.35508/jicon.v10i2.7426

Raffi, M., Suharso, A., & Maulana, I. (2023). Analisis Sentimen Ulasan Aplikasi Binar Pada Google Play Store Menggunakan Algoritma Naïve Bayes. *Journal of Information Technology and Computer Science (INTECOMS)*, *6*(1), 450–462. https://doi.org/https://doi.org/10.31539/intecoms.v6i1.6117

RAZZAQ, M. F. (2022). *SISTEM INFORMASI LULUSAN DAN BUKU WISUDA DI UIN Mahmud Yunus Batusangkar MENGGUNAKAN METODE WATERFALL*.

Syahputra, I. K., Abdurrachman Bachtiar, F., & Wicaksono, S. A. (2018). Implementasi Data Mining untuk Prediksi Mahasiswa Pengambil Mata Kuliah dengan Algoritme Naive Bayes. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *2*(11), 5902–5910. http://j-ptiik.ub.ac.id