

*Research article/Review article/Book review*

## Model Klasifikasi Untuk Menentukan Kesiapan Kerja Mahasiswa Dan Kelulusan Tepat Waktu Dengan Metode Machine Learning

Fadya Rizka Yudana<sup>1\*</sup>, M. Suyanto<sup>2\*</sup>, Asro Nasiri<sup>3\*</sup>

<sup>1</sup> Universitas AMIKOM Yogyakarta, Indonesia

<sup>2</sup> Universitas AMIKOM Yogyakarta, Indonesia

<sup>3</sup> Universitas AMIKOM Yogyakarta, Indonesia

\* [vadya@amikom.ac.id](mailto:vadya@amikom.ac.id)

### Abstract

Graduate on time is one of the indicators of a student's success in obtaining their degrees. However, graduating on time also raises several issues, one of which is related to the employability of students after they graduate. Previous study using Three Different machine learning techniques. found that the accuracy level using these algorithms yielded satisfactory results. However, these research results were obtained using only one type of algorithm, so there is no comparison to determine whether the algorithm used is the most optimal. After conducting a classification model with KNIME using three machine learning algorithms, namely decision tree, naïve Bayes, and KNN, it was found that the Naive Bayes algorithm has the highest accuracy (0.537) compared to Decision Tree (0.509) and KNN (0.511).

Lulus tepat waktu adalah salah satu indikator keberhasilan seorang mahasiswa dalam meraih gelar mereka. Namun, lulus tepat waktu juga menimbulkan beberapa masalah, salah satunya adalah terkait dengan keterserapan kerja mahasiswa setelah mereka lulus. Studi sebelumnya menggunakan tiga teknik machine learning yang berbeda menemukan bahwa tingkat akurasi menggunakan algoritma ini menghasilkan hasil yang memuaskan. Namun, hasil penelitian ini diperoleh hanya dengan menggunakan satu jenis algoritma, sehingga tidak ada perbandingan untuk menentukan apakah algoritma yang digunakan adalah yang paling optimal. Setelah melakukan model klasifikasi dengan KNIME menggunakan tiga algoritma machine learning, yaitu decision tree, naïve Bayes, dan KNN, ditemukan bahwa algoritma Naive Bayes memiliki akurasi tertinggi (0.537) dibandingkan dengan Decision Tree (0.509) dan KNN (0.511).

**Keywords** Decision Tree, Employability, KNIME, KNN, Naive bayes

**Article history** Submitted: 01/01/2023; revised: 03/03/2023; accepted: 04/04/2023.



© 2023 by the author(s). Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY SA) license, <https://creativecommons.org/licenses/by-sa/4.0/>.

## DECLARATION FORM

*For manuscript with multi-authors, this form is to be filled in by the corresponding author.*

Name\* : Fadya Rizka Yudana  
Residence address : Jalan Tluki 4 no 128 Condong Catur, Provinsi DI Yogyakarta  
Affiliation\* : Mahasiswa  
Edu. Background : Masters at Universitas AMIKOM Yogyakarta  
Field of expertise : Public Relation,  
WhatsApp No.\* : 0818.0270.2158

Declares that

1. This work is original, free of plagiarism, and has not been published elsewhere, nor is it currently under consideration for publication elsewhere;
2. The potential ethical research oversight is clear under the ethical clearance guidance by The National Research and Innovation Agency (BRIN) of the Republic of Indonesia;
3. The author(s) have no conflict of interest, including personal and financial.
4. *For multiauthor-manuscript*, Author's Role/Participation:  
Fadya Rizka Yudana : Study design, data collection, manuscript preparation.  
M. Suyanto: final approval;  
Asro Nasiri: final approval;

*Yogyakarta, 15 September 2023*



*Fadya Rizka Yudana*

## 1. Pendahuluan

Lulus tepat waktu merupakan salah satu indikator keberhasilan mahasiswa dalam menyelesaikan studinya. Mahasiswa dikatakan lulus tepat waktu apabila mereka mampu menyelesaikan studinya di perguruan tinggi selama tiga setengah tahun atau setidaknya empat tahun untuk program sarjana (S1), sedangkan mahasiswa dikatakan tidak lulus tepat waktu apabila menyelesaikan studinya lebih dari empat tahun.

Namun lulus tepat waktu juga menimbulkan beberapa permasalahan, salah satunya terkait dengan kesiapan kerja (Employability) Mahasiswa setelah mereka lulus. Menurut Caballero, Walker dan Fuller (2011), kesiapan kerja adalah suatu sikap, keterampilan dan pengetahuan yang dimiliki seorang yang telah lulus dari institusi pendidikan tinggi sehingga memiliki kesiapan dalam bekerja atau siap untuk sukses dalam lingkungan kerja. Ketepatan Mahasiswa dalam menyelesaikan Studinya tentu saja juga harus ditunjang juga dengan kesiapan mereka ketika akan memasuki dunia kerja.

Dalam Penelitian yang lain, Setyaningsih, Dwi (2019) menyampaikan bahwa ada beberapa faktor yang mempengaruhi kesiapan kerja, yaitu faktor internal dan external. Faktor internal yang mempengaruhi employability yaitu: efikasi diri ((Keyakinan diri sendiri bahwa bisa berhasil dan sukses) dan kepercayaan diri. Sedangkan untuk faktor eksternalnya, yaitu: dukungan sosial, kepuasan kerja dan *career development training*.

Berdasarkan Penelitian yang dilakukan oleh Hartinah (2016), periode rata-rata waktu tunggu sarjana (S1) di Indonesia hingga mendapatkan pekerjaan pertama adalah 0 (nol) hingga 9 (sembilan) bulan. Hal ini mengindikasikan bahwa rata-rata waktu tunggu lulusan untuk mendapatkan pekerjaan pertama cukup bervariasi bahkan cenderung memiliki waktu tunggu yang lama dalam mendapat pekerjaan pertama.

Bahkan dalam penelitian yang dilakukan Kementerian Riset, Teknologi, dan Pendidikan Tinggi (2019) yang mencatat sekitar 8,8% dari total 7 juta pengangguran di Indonesia adalah sarjana. Pada tahun 2017, diketahui bahwa hanya ada 17,5% jumlah tenaga kerja lulusan perguruan tinggi. Angka presentase ini jauh lebih kecil jika dibandingkan dengan tenaga kerja lulusan SMA/SMK sebesar 82% dan tenaga kerja lulusan SD sebesar 60%. Hal ini menunjukkan bahwa penting untuk mempersiapkan Mahasiswa agar bisa menghadapi tantangan di dunia kerja.

Salah satu faktor penting dalam mempersiapkan kesiapan kerja mahasiswa adalah dengan membiasakan mahasiswa untuk bekerja sebelum mereka diwisuda agar memiliki pengalaman pekerjaan sehingga mahasiswa akan lebih mudah beradaptasi di dunia kerja, setelah mereka lulus nantinya. Menurut Jackson, Denise dan Tomlinson, Thomas (2021), Pengalaman Kerja, terutama yang terstruktur dan relevan, adalah jembatan penting antara studi formal yang saat ini ditempuh mereka dengan pekerjaan yang akan mereka lakukan di masa depan.

Universitas AMIKOM Yogyakarta sendiri sangat memperhatikan faktor kesiapan kerja Mahasiswa. Hal ini terbukti dalam rating evaluasi QS Star yang menunjukkan bahwa tingkat Kesiapan Kerja (Graduate Employability) Lulusan Universitas Amikom Yogyakarta mendapat predikat tertinggi (Bintang lima). Tingginya rating employability yang diraih Universitas Amikom Yogyakarta ini, menunjukkan tingginya kesempatan lulusan amikom dalam mendapatkan pekerjaan atau menjadi pengusaha setelah mereka diwisuda. Berdasarkan data yang didapat dari Buku wisuda

periode ke 69 Universitas Amikom Yogyakarta, terhitung dari wisuda periode 1 (21 Maret 1998) hingga periode 69 (25 Januari 2019), Amikom sudah meluluskan 23814 wisudawan, dan Jumlah ini akan terus bertambah di tiap periode wisuda.

Untuk mengetahui seberapa besar kiprah alumni di masyarakat, serta relevansinya dengan proses pendidikan yang telah dia lalui selama masa perkuliahan, BPC melakukan tracer studi terhadap para lulusan secara rutin, satu tahun setelah mereka diwisuda. Data hasil tracer studi ini nantinya akan dimanfaatkan untuk berbagai macam tujuan, seperti akreditasi kampus, evaluasi proses pendidikan, akuntabilitas bagi pemangku kebijakan, serta informasi relevansi kontribusi dunia pendidikan terhadap kehidupan mereka setelah lulus. Pada penelitian ini, penulis menggunakan data yang didapat Business Placement Center (BPC) Universitas Amikom Yogyakarta tentang tracer alumni yang nantinya akan diolah dengan menggunakan metode Machine Learning.

Salah satu metode dalam Machine Learning yang biasa digunakan dalam membandingkan dan menggolongkan data berdasarkan ciri tertentu adalah Klasifikasi. Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data dengan tujuan untuk memperkirakan kelas yang tidak diketahui dari suatu objek. Tujuan klasifikasi adalah memprediksi label kelas, yang merupakan pilihan dari daftar kemungkinan yang telah ditentukan.

Berdasarkan kajian pada tiga jurnal publikasi yang melakukan penelitian tentang Machine Learning terhadap kelulusan mahasiswa dengan berbagai metode klasifikasi, yaitu: SVM (Rooy Thaniket, 2020), Decision Tree J48 (Yusran Timur Samuel, 2019), dan ID3 (Ismail setiawan, 2019). Pada ketiga penelitian tersebut didapat bahwa tingkat akurasi dengan menggunakan algoritma tersebut memiliki hasil yang memuaskan. Meski demikian hasil penelitian tersebut didapat hanya dengan menggunakan satu jenis algoritma saja, sehingga tidak ada pembanding untuk mengetahui apakah algoritma yang digunakan sudah yang paling optimal.

Berdasarkan semua keterangan penelitian diatas, penulis akan mencoba membahas tentang Model Klasifikasi Untuk Menentukan Kesiapan Kerja Mahasiswa dan Kelulusan Tepat Waktu dengan mengimplementasikan beberapa teknik machine learning, yaitu algoritma Decision Tree, KNN dan Naive Bayes, serta membandingkan ketiganya untuk mendapatkan tingkat akurasi yang paling baik.

Dalam penelitian in. data yang digunakan diambil dari dari database Business Placement Center (BPC) Universitas Amikom Yogyakarta. Jumlah data sampel sebanyak 1644 data alumni untuk tahun penelusuran 2019 dan 2018, yang akan dibuat model klasifikasinya dengan metode Naive bayes, Decision Tree, dan KNN. Hasil penelitian akan didapat dengan membandingkan hasil akurasi yang terbaik dari ketiga algoritma diatas. Hasil dari Klasifikasi tersebut nantinya akan dilakukan Uji Korelasi, untuk mengetahui hubungannya dengan Status Kelulusan.

## **2. Hasil dan Pembahasan**

### ***2.1. Dasar Teori***

#### **2.1.1 Lulus Tepat Waktu dan Kesiapan Kerja (Employability)**

Ketepatan waktu lulus mahasiswa memiliki kriteria yang berbeda untuk setiap program yang tersedia pada jenjang perguruan tinggi. Dalam Permenristekdikti RI No. 44 Tahun 2015, yang diubah dengan Permenristekdikti No 50 Tahun 2018, dan saat ini telah ditetapkan dengan Permendikbud No 3 Tahun 2020 sebagai Standar Nasional Pendidikan Tinggi (SN Dikti). Masa studi untuk program strata 1 (S1) dengan beban belajar mahasiswa paling sedikit 144 satuan kredit semester (SKS), Idealnya 4 tahun akademik dan paling lama masa adalah 7 tahun akademik dengan hak pemberhentian sementara (cuti) kuliah seorang mahasiswa maksimal 2 semester. Dalam Penelitiannya, Srinadi (2020) menyampaikan bahwa Prosentase mahasiswa lulus tepat waktu menjadi salah satu kriteria dalam penilaian akreditasi Program Studi, sebagai salah satu indikator apakah standar Proses Pembelajaran yang ditetapkan dalam implementasinya telah terpenuhi.

Kesiapan Kerja (Employability) merupakan kemampuan yang mencakup pengetahuan, sikap, dan kognisi untuk menyesuaikan diri dalam meningkatkan hubungan antara diri dan pekerjaan (Fugate, Kinicki dan Ashforth, 2004). Dalam beberapa penelitian sebelumnya, terdapat beberapa faktor yang mempengaruhi employability seseorang, salah satunya adalah faktor pengalaman kerja. Jackson, Denice dan Tomlisson (2021), menyampaikan bahwa Pengalaman Kerja, terutama yang terstruktur dan relevan, adalah jembatan penting antara studi formal yang saat ini ditempuh mereka dengan pekerjaan yang akan mereka lakukan di masa depan. Hal tersebut meningkatkan kepercayaan diri mahasiswa dalam menghadapi dunia kerja, meningkatkan persepsi kesiapan kerja, dan mempersiapkan sumber daya yang terkait dengan karir mereka agar sesuai dengan tuntutan pasar tenaga kerja yang terus berubah seiring perkembangan zaman.

### 2.1.2 Machine Learning dan Klasifikasi

\_Machine Learning (ML) adalah disiplin ilmu yang menugaskan komputer untuk belajar dan bertindak seperti manusia, dan meningkatkan kemampuan belajarnya dari waktu ke waktu secara otomatis, dengan cara menyuplai data dan informasi sebagai bentuk pengalaman dan interaksi dunia nyata. Itulah kenapa, salah satu ciri khas dari ML adalah adanya proses pelatihan, pembelajaran, atau training. Oleh karena itu, ML membutuhkan data untuk dipelajari yang disebut sebagai data training.

Klasifikasi adalah sebuah proses menemukan definisi kesamaan karakteristik dalam suatu kelompok atau kelas (class). klasifikasi bertujuan untuk memprediksi label kelas, yang merupakan pilihan dari daftar kemungkinan yang telah ditentukan. Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data dengan tujuan untuk memperkirakan kelas yang tidak diketahui dari suatu objek. Klasifikasi data sendiri terdiri dari 2 langkah proses. Pertama adalah learning (fase training), dimana algoritma klasifikasi dibuat untuk menganalisa data training lalu direpresentasikan dalam bentuk aturan klasifikasi. Proses kedua adalah klasifikasi dimana data tes digunakan untuk memperkirakan akurasi dari aturan klasifikasi

Klasifikasi dapat pula diartikan sebagai proses untuk mengelompokkan sekumpulan data ke dalam kelas-kelas tertentu yang sudah diberikan label berdasarkan kesamaan sifat dan pola yang terdapat dalam data-data tersebut. Secara umum, proses klasifikasi dimulai dengan diberikannya sejumlah data yang menjadi acuan untuk membuat aturan klasifikasi data.

Data-data ini nantinya akan disebut dengan data training sets. Dari training sets tersebut kemudian akan dibuat menjadi model untuk mengklasifikasikan data. Model tersebut kemudian

digunakan sebagai acuan untuk mengklasifikasikan data-data yang belum diketahui kelasnya yang biasa disebut dengan test sets (Rifqi, Maharani, & Shaufiah, 2011).

Terdapat beberapa algoritma yang dapat digunakan untuk melakukan klasifikasi, Beberapa di Algoritma yang digunakan dalam penelitian ini adalah :

### 1. Algoritma Naïve Bayes

Naïve Bayes merupakan sebuah metode klasifikasi probabilistik sederhana untuk menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas.

Dalam “*An Essay Towards the Solving a Problem in the Doctrines of Chance*”, Thomas Bayes menyajikan tentang “Proposition 9”, yang akhirnya dikenal dengan “Teorema Bayes”. Pendekatan penalaran statistik yang dilakukan oleh Bayes pada saat itu berbeda dibandingkan dengan matematis tradisional lainnya yang hanya fokus pada pada tingkah laku sampel dari populasi yang diketahui saja. Namun Bayes mengemukakan ide yang lebih lanjut untuk menentukan property dari populasi berdasarkan sampel tersebut. Prediksi Bayes yang berdasarkan pada teorema Bayes memiliki rumus umum seperti pada persamaan berikut:

$$P(H|E) = \frac{P(E \vee H) \cdot P(H)}{P(E)}$$

Keterangan

- $P(H|E)$  = Probabilitas akhir bersyarat (conditional probability) suatu hipotesis H terjadi jika diberikan bukti (evidence) E terjadi
- $P(E|H)$  = Probabilitas sebuah bukti E terjadi akan mempengaruhi hipotesis H.
- $P(H)$  = Probabilitas awal (prior) hipotesis H terjadi tanpa memandang bukti apapun.
- $P(E)$  = Probabilitas awal (prior) bukti E terjadi tanpa memandang bukti yang lain

Ide dasar dari aturan Bayes adalah bahwa hasil dari hipotesis atau peristiwa (H) dapat diperkirakan berdasarkan pada beberapa bukti (E) yang diamati. Ada beberapa hal penting dari aturan Bayes tersebut, yaitu probabilitas awal/prior H atau  $P(H)$  adalah probabilitas dari suatu hipotesis sebelum bukti diamati. Serta probabilitas akhir H atau  $P(H|E)$  adalah probabilitas dari suatu hipotesis setelah bukti diamati.

### 2. Algoritma Decision Tree

Decision tree merupakan suatu metode klasifikasi yang menggunakan struktur pohon, dimana setiap node merepresentasikan atribut dan cabangnya merepresentasikan nilai dari atribut, sedangkan daunnya digunakan untuk merepresentasikan kelas. Level node teratas dari sebuah Decision Tree disebut dengan node akar (root) yang biasanya berupa atribut yang paling memiliki pengaruh terbesar pada suatu kelas tertentu.

Decision Tress atau Pohon Keputusan adalah jenis algoritma untuk pembelajaran mesin pemodelan prediktif. Konsep dari pohon keputusan adalah mengubah data menjadi decision tree dan aturan-aturan keputusan. Prosesnya adalah dengan mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi rule, dan menyederhanakan rule tersebut. Manfaat utama dari menggunakan pohon keputusan adalah kemampuan untuk menyederhanakan proses

pengambilan keputusan yang kompleks menjadi lebih simpel sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan.

### 3. Algoritma KNN

Algoritma K-nearest Neighbors (KNN) merupakan metode yang digunakan untuk mengklasifikasikan suatu data berdasarkan dataset pelatihan, dengan mempertimbangkan k tetangga terdekatnya. Nilai k dalam konteks ini merujuk pada jumlah tetangga terdekat. Prinsip utama dari metode ini melibatkan dua tahapan utama. Tahap pertama adalah perhitungan validitas untuk semua data dalam dataset pelatihan. Tahap kedua melibatkan perhitungan Weight Voting pada semua data uji dengan menggunakan validitas data (Hamid Parvin, dkk 2010).

#### 2.1.3. KNIME

Knime adalah platform analisis data berperingkat tinggi dengan penerapan yang luas dan banyak integrasi dengan produk lain, seperti dengan basis data, bahasa, kerangka Machine Learning, dan kerangka kerja pembelajaran dalam. Platform ini dimaksudkan sebagai media kolaborasi dan penelitian. Platform juga dapat berperan sebagai platform integrasi untuk berbagai proyek analisis data lainnya

KNIME Analytics memungkinkan pengguna melakukan visualisasi aliran data masuk (pipelines), menjalankan tahapan analisis tertentu atau semua tahapan, dan melihat hasil dari pemodelan secara interaktif. KNIME ditulis dengan bahasa Java (dengan platform eclipse) dan dilengkapi dengan mekanisme plugin untuk kemudahan pengembangan dan penambahan fungsionalitas. Versi utamanya sudah meliputi ratusan modul integrasi data (file I/O, dukungan terhadap DBMS umum), transformasi data (filter, converter, combiner), dan metoda umum yang digunakan untuk analisis dan visualisasi data.

Pada KNIME Analytics Platform, setiap tugas direpresentasikan oleh node. Setiap node memiliki warna dan input-output port. Masukan dari setiap node merupakan data untuk diproses, sedangkan keluarannya adalah dataset. Setiap node memiliki konfigurasi spesifik sesuai kebutuhan pengguna. Selain itu, node memiliki indikator status yang terletak tepat di bawah node. Jika indikator berwarna merah, berarti node belum dikonfigurasi. Indikator yang berwarna kuning bermakna node telah dikonfigurasi tetapi belum dijalankan. Indikator berwarna hijau berarti node telah dieksekusi. Jika terdapat tanda silang berwarna merah, hal tersebut menunjukkan bahwa terdapat error pada node

## **2.2 Model Data**

Dalam penelitian ini, data yang digunakan diambil dari database Business Placement Center (BPC) Universitas Amikom Yogyakarta. Jumlah data sampel sebanyak 1644 data alumni untuk tahun penelusuran 2019 dan 2018, yang akan dibuat model klasifikasinya dengan metode Naïve Bayes, Decision Tree, dan SVM. Hasil penelitian akan didapat dengan membandingkan hasil akurasi yang terbaik dari ketiga algoritma diatas. Hasil dari Klasifikasi tersebut nantinya akan dilakukan Uji Korelasi, untuk mengetahui hubungannya dengan Status Kelulusan.

Sifat dari penelitian ini adalah deskriptif, penelitian menggambarkan obyek tertentu dan menjelaskan hal-hal terkait yang menggambarkan karakteristik populasi tertentu dalam bidang

tertentu secara faktual dan cermat. Pendekatan yang dilakukan pada penelitian ini adalah pendekatan kuantitatif

Berikut ini adalah rincian dari Tabel dari variabel penelitian yang digunakan dalam penelitian berikut ini:

**Table 1.** General information of IJITECH paper submission

|                          |   |
|--------------------------|---|
| <b>Subject</b>           | Social sciences and cultural studies, and any relevant fields   |
| <b>NIM</b>               | 4000 to 6000 words  |
| <b>Gender</b>            | Laki-Laki, Perempuan  |
| <b>Program Studi</b>     | S1 Sistem Informasi, S1 Teknik Informatika, S1 Teknologi Informasi, S1 Teknik Komputer, D3 Manajemen Informatika, D3 Teknik Informatika, S1 Arsitektur, S1 Perencanaan Wilayah dan Kota, S1 Geografi, S1 Ilmu Komunikasi, S1 Ekonomi, S1 Akuntansi, S1 Ilmu Pemerintahan, S1 Hubungan Internasional, S1 Kewirausahaan , S2 MTL. |
| <b>Tahun Masuk</b>       | Tahun Saat Mahasiswa Masuk.   |
| <b>Tahun Keluar</b>      | Tahun saat Mahasiswa KEluar   |
| <b>Lama Lulus</b>        | Tepat Waktu, Tidak tepat Waktu  |
| <b>Mulai Kerja</b>       | Sebelum Lulus, Setelah Lulus  |
| <b>Status Mulaikerja</b> | Sebelum Lulus, Setelah Lulus  |

### ***2.3 Pembahasan dan Hasil***

Data yang digunakan dalam penelitian ini diolah menggunakan platform KNIME sebagai alat analisis data. Data penelitian tersebut disimpan pada file excel dan akan dibaca oleh KNIME menggunakan node excel reader. Jumlah data sampel sebanyak 19947 data alumni yang terdiri dari 918 data Tracer Studi untuk tahun penelusuran 2018 dan 1029 data Tracer Studi untuk tahun penelusuran 2019.

Semua data akan dilakukan praproses terlebih dahulu dalam Proses Transformasi data. Untuk menghindari data-data kosong, digunakan node Row Filter untuk melakukan imputasi pada missing value yakni menghapus nilai yang kosong. Selanjutnya data yang berasal dari dua file excel yang berbeda itu disatukan dengan menggunakan node concatenate rows.

Tahap selanjutnya adalah menambahkan beberapa Kolom tambahan, agar data yang ada bisa dibaca oleh Model klasifikasi yang sudah ditentukan. Kolom pertama yang ditambahkan adalah kolom Lama Lulus, yang didapat dengan memasukkan node Math Formula, yang menambahkan kolom baru yang berisi hasil dari pengurangan kolom Tahun Lulus yang dikurangi Tahun masuk.

Kolom selanjutnya yang dibuat adalah kolom Status Kelulusan, yang didapat dengan memasukkan node Rule Engine, dengan ketentuan Jika Lama Lulus  $\leq 4.0$  maka Valuenya adalah "Tepat Waktu", jika tidak valuenya adalah "Tidak Tepat Waktu", Teknik yang sama digunakan untuk menambahkan Kolom Status Mulai Kerja.

Agar data lebih sederhana, saat dilakukan proses klasifikasi di Tahap selanjutnya, Maka digunakan Node Column Filter, yang akan menghilangkan kolom-kolom yang tidak dibutuhkan dalam proses selanjutnya. Hasil dari Proses Transformasi data ini akan menyisakan 4 kolom, yaitu

Status Mulai Kerja, Var Karier, Var Kompetensi, dan Var Softskill. Workflow pada KNIME untuk tahapan Transformasi data dapat dilihat pada Gambar 1.

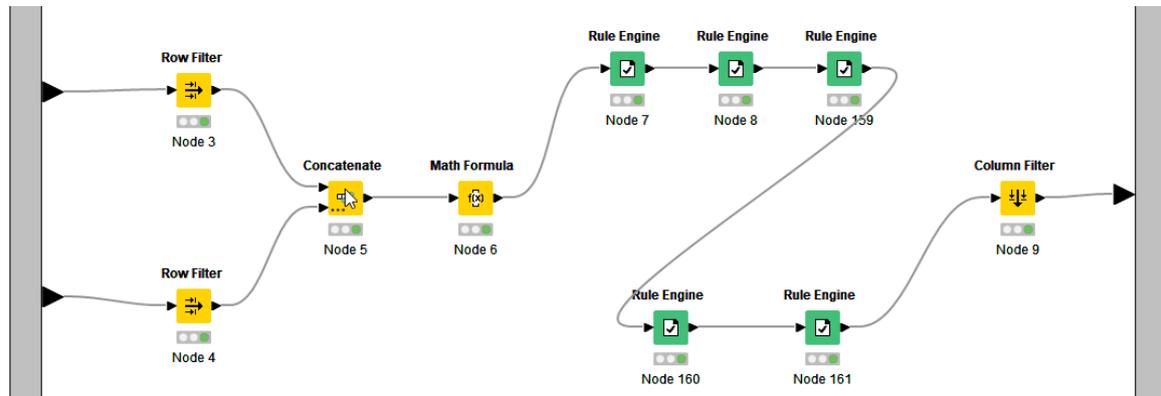


Figure 1. Workflow Transformasi Data di KNIME

Setelah tahapan praproses selesai, tahapan selanjutnya adalah pembagian data. Pada KNIME, kita dapat menggunakan node partitioning dengan menggunakan 5-fold cross validation. Hasil dari pembagian data ini, data akan terbagi menjadi data latih sebanyak 80% dan data uji sebanyak 20%. Hasil dari tahap ini adalah Data latih sebanyak 1557 data dan data uji sebanyak 390 data dengan perbandingan data setiap kelasnya sama.

Tahap selanjutnya adalah pembentukan model klasifikasi untuk tiga algoritma machine learning yakni Decision Tree, Naïve Bayes, dan KNN di KNIME, Berikut ini adalah Workflow yang digunakan untuk algoritma decision tree, Naïve Bayes, dan KNN dapat dilihat pada Gambar 5 sebagai berikut:

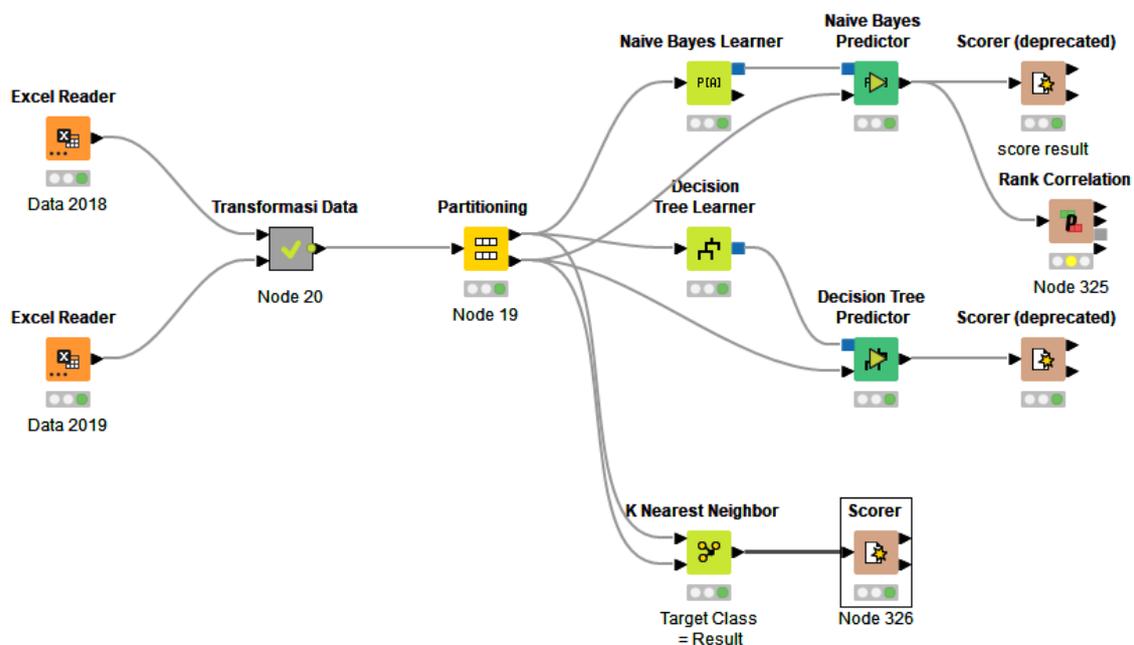


Figure 2. Workflow Transformasi Data Model Klasifikasi di KNIME

Pembentukan model decision tree dapat diperoleh dengan menggunakan node Decision tree Learner. Model yang dihasilkan dapat ditampilkan menggunakan node Decision tree View. Kemudian, model yang dihasilkan akan diuji menggunakan data uji. Node Decision tree Detector pada KNIME digunakan untuk melakukan proses pengujian. Hasil pengujian dapat dilihat menggunakan node Scorer.

Pembentukan model klasifikasi pada algoritma naïve Bayes pada KNIME dapat menggunakan node Naïve Bayes Learner. Model yang dihasilkan akan diujikan menggunakan node Naïve Bayes Predictor dan hasilnya pengujiannya akan ditampilkan pada node Scorer. Sedangkan untuk Pembentukan model klasifikasi pada algoritma k Nearest Neighbor pada KNIME dapat menggunakan node k Nearest Neighbor, dimana input data latih dan data uji ada dalam nodes yang sama.

Hasil evaluasi model klasifikasi dari ketiga algoritma machine learning yaitu decision tree, naïve Bayes dan KNN dapat dilihat pada Tabel 3.

**Tabel 2.** Tabel Perbandingan decision tree, naïve Bayes dan KNN

| Nilai Evaluasi | Presisi | Recall | Akurasi |
|----------------|---------|--------|---------|
| Decision Tree  | 0,4875  | 0,49   | 0,509   |
| Naïve Bayes    | 0,509   | 0,504  | 0,537   |
| KNN            | 0,499   | 0,499  | 0,511   |

Nilai akurasi pada algoritma decision tree yakni 0.518 yang artinya model ini dapat memprediksi status mulai kerja mahasiswa (baik sebelum atau setelah lulus) dengan benar sekitar 51.8% dari total keseluruhan data. Kelas Alumni yang bekerja sebelum lulus yang benar terklasifikasi sebanyak 13 data dan dari Kelas Alumni yang bekerja setelah lulus sebanyak 189 data dengan total keseluruhan data uji adalah 390 data. Nilai recall pada algoritma decision tree yakni 0.482 yang artinya decision tree mampu memprediksi benar kelas Alumni sebesar 48.2% dari total keseluruhan data. Sementara itu, untuk tingkat precision pada algoritma decision tree adalah sebesar 0.447. Pada algoritma ini nilai recall memiliki nilai yang rendah artinya masih terdapat banyak kesalahan pada kelas yang ada (false negatif). Confussion matrix pada algoritma decision tree dapat dilihat pada Tabel 4.

**Tabel 3.** Tabel Perbandingan Confussion matrix pada decision tree

| Row ID        | Sebelu... | Setelah... |
|---------------|-----------|------------|
| Sebelum Lulus | 71        | 196        |
| Setelah Lulus | 91        | 227        |

Nilai akurasi pada algoritma Naive Bayes yakni 0.537 yang artinya model ini dapat memprediksi status mulai kerja mahasiswa (baik sebelum atau setelah lulus) dengan benar sekitar 53.7% dari total keseluruhan data. Kelas Alumni yang bekerja sebelum lulus yang benar terklasifikasi sebanyak 34 data dan dari Kelas Alumni yang bekerja setelah lulus sebanyak 280 data dengan total keseluruhan data uji adalah 585 data. Nilai recall pada algoritma decision tree yakni 0.504 yang artinya decision tree mampu memprediksi benar kelas Alumni sebesar 50.4% dari total keseluruhan data. Sementara itu, untuk tingkat precision pada algoritma decision tree adalah sebesar 0.509. Pada algoritma ini nilai recall memiliki nilai yang rendah artinya masih terdapat banyak

kesalahan pada kelas yang ada (false negatif). Confussion matrix pada algoritma decision tree dapat dilihat pada Tabel 4.

**Tabel 4.** Tabel Perbandingan Confussion matrix pada Naïve Bayes

| Row ID        | Sebelu... | Setelah... |
|---------------|-----------|------------|
| Sebelum Lulus | 34        | 233        |
| Setelah Lulus | 38        | 280        |

Nilai akurasi pada algoritma KNN yakni 0.511 yang artinya model ini dapat memprediksi status mulai kerja mahasiswa (baik sebelum atau setelah lulus) dengan benar sekitar 51.1% dari total keseluruhan data. Kelas Alumni yang bekerja sebelum lulus yang benar terklasifikasi sebanyak 95 data dan dari Kelas Alumni yang bekerja setelah lulus sebanyak 204 data dengan total keseluruhan data uji adalah 585 data. Nilai recall pada algoritma decision tree yakni 0.499 yang artinya decision tree mampu memprediksi benar kelas Alumni sebesar 49.9 % dari total keseluruhan data. Sementara itu, untuk tingkat precision pada algoritma decision tree adalah sebesar 0.49. Pada algoritma ini nilai recall memiliki nilai yang rendah artinya masih terdapat banyak kesalahan pada kelas yang ada (false negatif). Confussion matrix pada algoritma decision tree dapat dilihat pada Tabel 4.

**Tabel 5.** Tabel Perbandingan Confussion matrix pada KNN

| Row ID        | Sebelu... | Setelah... |
|---------------|-----------|------------|
| Sebelum Lulus | 95        | 172        |
| Setelah Lulus | 114       | 204        |

Dari hasil tersebut, dapat dilihat bahwa algoritma Naive Bayes memiliki akurasi tertinggi (0.537) dibandingkan dengan Decision Tree (0.509) dan KNN (0.511). Namun, dalam hal presisi dan recall, ketiga algoritma tersebut memiliki nilai yang hampir sama. Perlu diperhatikan bahwa jumlah data uji untuk Decision Tree lebih sedikit dibandingkan dengan Naive Bayes dan KNN, sehingga perbandingan mungkin tidak sepenuhnya adil.

Selain itu, dalam konteks ini, Anda mungkin juga ingin mempertimbangkan interpretasi model dan kecepatan pelatihan/prediksi, karena beberapa model mungkin lebih mudah diinterpretasikan atau lebih cepat dalam pelatihan/prediksi dibandingkan dengan yang lain.

### 3. Kesimpulan

Pada penelitian ini, telah diimplementasikan tiga algoritma machine learning yakni decision tree, naïve Bayes dan KNN untuk membandingkan akurasi antar machine learning. Setelah dilakukan model Klasifikasi dengan KNIME., didapat hasil bahwa algoritma Naive Bayes memiliki akurasi tertinggi (0.537) dibandingkan dengan Decision Tree (0.509) dan KNN (0.511). Hasil yang mirip untuk Presisi, didapat bahwa algoritma Naive Bayes memiliki akurasi tertinggi (0,509) dibandingkan dengan Decision Tree (0,4875) dan KNN (0,499). Sedangkan untuk Algoritma KNN, didapat bahwa Nilai Recallnya cenderung Lebih tinggi dari Decision Tree dan KNN, yang memiliki Nilai Recal yang sama, Yaitu 0.499.

## Daftar Pustaka

- Caballero, C.L., Walker, A., Fuller-Tyszkiewicz, M., (2011). The Work Readiness Scale (WRS): Developing a measure to assess work readiness in college graduates. *JTLGE* 2, 41–54.  
<https://doi.org/10.21153/jtlge2011vol2no1art552>
- Cristianini, N., Shawe-Taylor, J., (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, 1st ed. Cambridge University Press.  
<https://doi.org/10.1017/CBO9780511801389>
- Darujati, C., Gumelar, A.B., n.d. (2012). Pemanfaatan Teknik Supervised Untuk Klasifikasi Teks Bahasa Indonesia. *JURNAL LINK* Vol 16 No. 1
- Fugate, M., Kinicki, A.J., Ashforth, B.E., (2004). Employability: A psycho-social construct, its dimensions, and applications. *Journal of Vocational Behavior* 65, 14–38.  
<https://doi.org/10.1016/j.jvb.2003.10.005>
- Hartinah, N.Y.Siti., (2016). Analisis Uji Ketahanan Hidup Data Waktu Tunggu Sarjana Dengan Metode Kaplanmeier Berbantuan Software Medcalc. Skripsi. Universitas Negeri Semarang, Semarang.
- Istiawan, I., Purbianto, E., (2019). Pemanfaatan Algoritma ID3 sebagai Prediktor Kelulusan Mahasiswa AMIK Harapan Bangsa Surakarta. *J. n.a Infotama* 15.  
<https://doi.org/10.37676/jmi.v15i1.745>
- Jackson, D., Tomlinson, M., (2021). The relative importance of work experience, extra-curricular and university-based activities on student employability. *Higher Education Research & Development* 1–17. <https://doi.org/10.1080/07294360.2021.1901663>
- Jung, D., Kwon, K., Kim, H., (2006). Bayesian Network based Human Pose Estimation, in: *AIA'06: Proceedings of the 24th IASTED International Conference on Artificial Intelligence and Applications*. pp. 268–272.
- Rifqi, N., Maharani, W., (2011). Analisis dan Implementasi Klasifikasi Data Mining Menggunakan Jaringan Syaraf Tiruan dan Evolution. *Konferensi Nasional Sistem dan Informatika 2011*; Bali, November 12, 2011
- Samuel, Y.T., Jonathan, B., Naibaho, J., (2019). Memprediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Decision Tree J48 Di Universitas Advent Indonesia. *TeIKa* 9, 43–52.  
<https://doi.org/10.36342/teika.v9i01.790>
- Srinadi, I.G.A.M., Nilakusmawati, D.P.E., (2020). Analisis Waktu Kelulusan Mahasiswa Fmipa Universitas Udayana Dan Faktor-Faktor Yang Memengaruhinya. *MTK* 9, 205.  
<https://doi.org/10.24843/MTK.2020.v09.i03.p300>
- Syani, R.S., Tentama, F., Diponegoro, A.M., (2021). Pengaruh Pengalaman Kerja, Kepuasan Kerja Dan Dukungan Sosial Terhadap Employability Pada Guru Smk Muhammadiyah Dalam Krisis Covid 19. *PSYCHE: Jurnal Psikologi Universitas Muhammadiyah Lampung* 3.
- Thaniket, R., Luthfi, E.T., Kusriani, (2020). Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Algoritma Support Vector Machine. *JURNAL FATEKSA: Jurnal Teknologi dan Rekayasa*, 5, 10.
- Vermeulen, L., Schmidt, H.G., (2008). Learning environment, learning process, academic outcomes and career success of university graduates. *Studies in Higher Education* 33, 431–451.  
<https://doi.org/10.1080/03075070802211810>
- Wiley, M., (2014). A review of the literature on current practice in the development of employability skills. *European Journal of Engineering Education* 31(6):637-650 *SCONUL*.  
<https://doi.org/10.1080/03043790600911704>