

IJCIT

(Indonesian Journal on Computer and Information Technology)

Journal Homepage: <http://ejournal.bsi.ac.id/ejurnal/index.php/ijcit>

Prediksi Lama Masa Tunggu Alumni USM dalam Mendapatkan Pekerjaan dengan Algoritma *KNN*

Saifur Rohman Cholil¹, Vensy Vydia², Susanto³

^{1,2} Sistem Informasi, Universitas Semarang, Semarang, Indonesia

³ Teknik Informatika, Universitas Semarang, Semarang, Indonesia

e-mail: cholil@usm.ac.id¹, vensy@usm.ac.id², susanto@usm.ac.id³

ABSTRAK

Pencarian informasi alumni perguruan tinggi dilakukan supaya dapat menentukan seberapa lama masa yang dibutuhkan agar mendapatkan pekerjaan. Dalam tahap peningkatan mutu sebuah perguruan tinggi, alumni mengambil peran penting, sebab kualitas pembelajaran pada perguruan tinggi dapat dikatakan sudah cukup baik apabila alumninya cepat terserap didalam dunia kerja. Pencarian informasi (tracer study) diperoleh melalui kuesioner yang dibagikan oleh pihak USM *Career and graduated Class* (UCAC) kepada para alumni. Informasi tersebut terdiri dari data-data para alumni pada tahun 2019-2020 dan kemudian digunakan untuk mendapatkan dan memperkirakan lama waktu tunggu yang dibutuhkan para lulusan USM hingga memperoleh pekerjaan setelah dinyatakan lulus *studi* sarjana. Algoritma KNN digunakan dalam penelitian ini, hal tersebut dikarenakan metode KNN mampu memprediksi masa tunggu alumni dibandingkan dengan metode lainnya. Adapun kompetensi yang berdampak terhadap masa tunggu alumni dalam mendapatkan pekerjaan dapat diperoleh melalui hasil analisis penelitian ini, seperti keterampilan dalam menggunakan komputer, manajemen pengolahan waktu, kemampuan dalam menganalisis, serta disiplin ilmu. Hasil implementasi algoritma KNN dengan mencoba nilai K dari 1-100 yang memiliki akurasi tertinggi mencapai 98,84%.

Kata Kunci: alumni, data mining, KNN, masa tunggu

ABSTRACTS

A search for information on university alumni was carried out in order to determine how long it would take to get a job. In the stage of progress in the quality of a university, alumni play an important role, because the quality of learning at a university can be said to be quite good if the alumni are quickly absorbed into the world of work. The search (tracer study) was obtained through a questionnaire distributed by the USM Career and Graduate Class (UCAC) to alumni. This information consists of data from alumni in 2019-2020 and is then used to obtain examples of the length of time to employment needed for USM graduates to get a job after graduating from undergraduate studies. The KNN algorithm was used in this research, this is because the KNN method is able to predict the time to employment for alumni compared to other methods. Competencies that have an impact on the time to employment for alumni to get a job can be obtained through the results of this research analysis, such as skills in using computers, time management, analytical skills, and scientific discipline. The results of implementing the KNN algorithm by trying K values from 1-100 have the highest accuracy reaching 98.84%.

Keywords: alumni, data mining, KNN, time to employment



1. PENDAHULUAN

Perguruan tinggi merupakan lembaga pendidikan tingkat akhir dalam pendidikan formal yang menghasilkan lulusan dan kemudian memasuki dunia kerja. Lingkungan kerja global yang semakin kompleks dan kompetitif menempatkan semua universitas pada konflik yang sama, yaitu apakah *output* pendidikan tinggi sesuai dengan kebutuhan lulusan universitas saat ini atau tidak. Dalam skala global, tracer study telah diakui sebagai alat penting untuk mengukur efektivitas sistem pendidikan tinggi dalam menghasilkan lulusan yang kompeten dan relevan dengan kebutuhan pasar kerja. Dengan memanfaatkan data alumni, tracer study memungkinkan evaluasi terhadap kesenjangan keterampilan, penyesuaian kurikulum, dan peningkatan daya saing universitas di tingkat internasional. Pendidikan awal perlu memberikan keterampilan yang dibutuhkan untuk setiap mahasiswa agar dapat mengatasi tantangan yang ada, baik dalam mendapatkan pekerjaan maupun menciptakan lapangan pekerjaan sendiri (Albina & Sumagaysay, 2020). Setelah dinyatakan lulus program *studi* sarjana, maka memperoleh pekerjaan adalah salah satu tujuan utama para *fresh graduate* (Kiri & Atti, 2021). Dibutuhkan informasi mengenai alumni dengan cara *tracer study* yang sudah mendapatkan pekerjaan dengan menganalisis pada saat studi (Akbar & Mukhtar, 2020).

Tracer study adalah sebuah program pada perguruan tinggi yang digunakan untuk mengevaluasi kelemahan yang ada pada sistem pembelajaran dan pendidikan untuk pedoman agar lebih baik kedepannya, Sebagai umpan balik dimasa yang akan datang bagi para alumni, maka pelaksanaan *tracer study* ini sangat diperlukan. Selain dijadikan sebagai evaluasi, *tracer study* dapat memperlihatkan gambaran suatu perguruan tinggi dalam mencetak lulusan yang berkualitas (Noveandini & Wulandari, 2019). *Tracer study* harus dilakukan supaya keberadaan mahasiswa yang telah lulus dilokasi kerja dapat diketahui. Tujuan dari pelacakan ini agar dapat mengetahui seberapa tingkat kesuksesan lembaga pendidikan dalam menjalankan program yang direncanakan dengan melihat bagaimana kondisi para lulusan saat di lapangan kerja (Alfarisi et al., 2022).

Universitas Semarang (USM) merupakan sebuah perguruan tinggi swasta, yang dibangun pada 23 Juni 1987 dengan nama awalnya adalah Politeknik Semarang dan kemudian pada tahun

1994 resmi berganti nama menjadi Universitas Semarang. Visi Universitas Semarang adalah "Menjadi perguruan tinggi yang menghasilkan individu yang profesional, beradab, dan berkeindonesiaan, memiliki wawasan dalam teknologi informasi dan pembangunan berkelanjutan, serta mampu bersaing baik di tingkat nasional maupun global."

Strategi guna mewujudkan visi dari USM dalam meningkatkan mutu, yang mana salah satu dari kriteria keberhasilan bagi perguruan tinggi dalam aspek Pendidikan yaitu masuknya alumni kedalam dunia kerja. Standar bagi suatu perguruan tinggi yang ditetapkan oleh Standar penjamin Mutu *Internal* (SPMI) yaitu kelulusan mahasiswa (Widaningsih, 2019). *Tracer study* dilakukan setiap tahunnya oleh USM untuk pengembangan kurikulum, menanggapi kebutuhan data untuk akreditasi, serta perbaikan proses pembelajaran. Apabila alumni dapat dengan cepat dalam mendapatkan pekerjaan, maka semakin baik mutu mahasiswanya. Data tersebut didapatkan melalui kuisioner yang disebarkan kepada para alumni.

Proses ini memanfaatkan teknik data mining yang diimplementasikan dalam berbagai langkah proses pengolahan data, hal yang sangat penting dalam pemilihan metode atau algoritma yang akan digunakan (Xu et al., 2023). Klasifikasi merupakan metode data mining untuk memprediksi jenis atau kategori berdasarkan atribut dalam data. Apabila ingin menangani data yang besar dan banyak, maka algoritma KNN ialah teknik data mining yang paling cocok sebab keefisienan dan keefektifannya (Shokrzade et al., 2021). Namun, jika dibandingkan dengan algoritma lain seperti Naïve Bayes, algoritma KNN memiliki keunggulan dalam kesederhanaan prinsip kerjanya tetapi membutuhkan lebih banyak sumber daya komputasi untuk dataset yang besar. Sementara itu, algoritma seperti C4.5 lebih cocok untuk data dengan kompleksitas tinggi dan memberikan interpretasi yang lebih mudah melalui pohon keputusan (Wang et al., 2022). Karena ada kumpulan data besar yang dibuat dan diperluas setiap hari di berbagai server *online* dan *offline*, maka metode yang efisien untuk data tersebut merupakan metode KNN. Melalui metode KNN, akan dilakukan klasifikasi terhadap kelompok sampel data yang dianggap paling relevan jika dibandingkan dengan objek berdasarkan jarak paling dekat. Algoritma ini memiliki keunggulan pada prinsip yang sederhana, namun juga disertai dengan beberapa kelemahan, seperti lamanya waktu

eksekusi dan kebutuhan ruang yang berlebihan, serta kesulitan dalam menentukan nilai kluster. (Wang et al., 2022). Algoritma KNN dipilih guna mengolah data prediksi masa tunggu alumni pada penelitian ini. Diharapkan nantinya penelitian ini mampu menghasilkan keputusan yang akurat dan nilai akurasi yang meningkat, sehingga mampu memudahkan pihak Universitas Semarang dalam keperluan mengembangkan kurikulum serta keperluan dalam data akreditasi.

Saat ini, banyak sistem komputer telah dikembangkan dengan metode dan perangkat lunak yang berbeda. Dari sejumlah referensi berupa penelitian-penelitian sebelumnya yang berhubungan dengan objek penelitian. Penggunaan referensi digunakan untuk memberikan batasan pada sistem akan semakin dikembangkan di kemudian hari. Penelitian yang dilakukan oleh (Yahdin et al., 2021) menggunakan algoritma *K-Nearest Neighbor* dengan 372 data lulusan alumni tahun 2014-2016 dan menghasilkan nilai akurasi sebesar 74,77%.

Penelitian lain yang dilaksanakan dari (Maricar & Dian Pramana, 2019) yang membandingkan Di antara metode *Naïve Bayes* dan *K-Nearest Neighbor* guna memperkirakan waktu tunggu para alumni untuk mendapatkan pekerjaan. Jumlah data uji yang dimanfaatkan mencapai 344, sementara data latih terdiri dari 1335 entri. Hasil evaluasi menunjukkan tingkat akurasi metode *Naïve Bayes* sekitar 83,83%, sedangkan metode *K-Nearest Neighbor* mencapai 82,34%.

Riset terdahulu yang dilaksanakan peneliti (Adnyana, 2020) memperkirakan masa tunggu alumni untuk mendapatkan pekerjaan menggunakan implementasi dengan parameter bernilai 10 pada aplikasi WEKA dan algoritma *Naïve Bayes* dengan tingkat akurasi sebesar 48,629%. Penelitian yang lain tentang masa tunggu alumni mendapatkan pekerjaan dilakukan oleh (Aulia et al., 2023) untuk memprediksi masa tunggu bagi alumni dalam memperoleh pekerjaan dengan mengambil data pelatihan dan pengujian dari 193 alumni penelitian *tracer* Alumni STMIK *Royal* yang lulus dan sudah bekerja pada tahun 2021 menjadi subjek penelitian. Faktor-faktor seperti jenis kelamin, jurusan, tahun masuk, tahun lulus, IPK, dan lama pekerjaan digunakan dalam perhitungan waktu tunggu kerja alumni. Pemodelan dengan menggunakan algoritma *Naïve Bayes* menghasilkan tingkat akurasi sekitar 94%.

Penelitian selanjutnya juga dilakukan oleh (Widhiantoyo, 2021) menggunakan algoritma *Naïve Bayes* dengan fitur *Backward Elimination* dalam memprediksi waktu tunggu alumni mendapatkan pekerjaan dan menghasilkan nilai akurasi sebesar 68.52%. Penelitian lain menggunakan algoritma *Naïve Bayes Classifier* dilakukan oleh (Rachmadiansyah et al., 2022) dengan menggunakan data yang diperoleh melalui hasil kuesioner *tracer study*, variabel-variabel yang diobservasi mencakup Masa Studi, IPK, keaktifan organisasi, dan penguasaan kompetensi. Masa tunggu kerja dibagi menjadi empat kelas, yaitu ≤ 10 bulan, 11 bulan - 2 tahun 1 bulan, 2 tahun 2 bulan - 3 tahun 4 bulan, dan > 3 tahun 4 bulan. Hasil penelitian ini menunjukkan prediksi masa tunggu kerja, yang direpresentasikan dalam bentuk *confusion matrix* dengan tingkat akurasi sekitar 81,82%. Adapun penelitian menggunakan algoritma lain seperti penelitian yang telah dilakukan oleh (Noviyanto & Mukti, 2022) dengan mengimplementasikan algoritma C.45 dengan data dari alumni 5 tahun terakhir, dalam kurung waktu 2015-2019 dan mendapatkan data *training* sebanyak 270 serta data testing sebanyak 462. Hasil akurasi yang termasuk pada *excellent classification* yaitu 82,86%.

Studi yang dilakukan oleh (Rezkika et al., 2021) juga mengadopsi pendekatan algoritma C4.5 yang didasarkan pada Forward Selection dalam mengklasifikasikan masa tunggu alumni untuk memperoleh pekerjaan. Data terdiri dari para alumni pada tahun 2016-2020, yang terdapat 270 data *training* dan 462 data testing. Berdasarkan hasil uji dari data lulusan alumni tahun 2016-2020 menunjukkan bahwa nilai akurasi tinggi sebesar 80,37%.

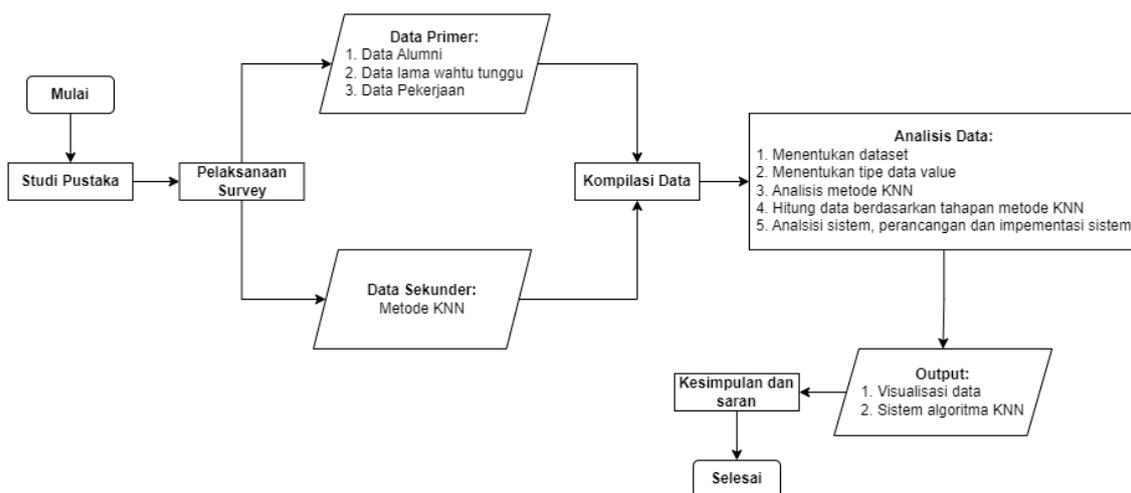
Berdasarkan beberapa penelitian yang terdahulu, belum ada penelitian yang menggunakan algoritma *K-Nearest Neighbor* (KNN) digunakan untuk melakukan prediksi. masa tunggu alumni dalam mendapatkan pekerjaan. Sehingga menjadikan metode ini sebagai jalan alternatif untuk menyelesaikan permasalahan yang ada pada Universitas Semarang (USM). Diharapkan dengan penelitian ini dapat membantu dalam kebutuhan akreditasi dan pengembangan kurikulum.

2. METODE PENELITIAN

Diagram proses disajikan dengan merinci alur tahapan penelitian, mencakup langkah-langkah seperti algoritma, jalur, pemodelan,

pengumpulan data, serta desain yang digunakan dalam perancangan sistem (Bianto et al., 2020). Gambaran keseluruhan proses penelitian dapat ditemukan dalam Alur Pembuatan Penelitian, sebagaimana ditampilkan dalam Gambar 1. Gambar ini merupakan penjelasan mengenai tahapan dari bagan alir penelitian. Diawali dari studi pustaka tentang metode KNN, pelaksanaan survei kompilasi dan analisis data sampai dengan

output dari penelitian. Metode yang diterapkan dalam mengumpulkan data adalah survei bersama alumni-alumni dengan rentang waktu kelulusan 2 tahun yang lalu (2019-2020) dari Universitas Semarang, sehingga menghasilkan informasi berupa data alumni, data pekerjaan yang mereka dapatkan, dan data lama waktu tunggu alumni dalam mendapatkan pekerjaan tersebut.



Gambar 1. Alur Tahapan Penelitian

Tahapan penelitian diawali dengan studi kasus, yaitu pemeriksaan lebih lanjut secara terperinci dan mendalam mengenai alumni Universitas Semarang. Penelitian ini mencakup pengumpulan data yang relevan untuk mendapatkan gambaran tentang profil alumni, termasuk data pekerjaan, lama waktu tunggu untuk mendapatkan pekerjaan, dan informasi umum lainnya. Survei dilakukan sebagai langkah berikutnya, dengan tujuan untuk memperoleh data yang dibutuhkan dalam penelitian ini. Data yang dikumpulkan meliputi lama waktu tunggu, tempat bekerja, dan informasi umum alumni yang diperlukan untuk mendukung proses analisis lebih lanjut.

Setelah survei selesai, data yang terkumpul dikompilasi untuk membentuk dataset yang akan dianalisis. Proses analisis data melibatkan langkah-langkah penting, seperti menentukan tipe nilai data, penerapan metode KNN, perhitungan berdasarkan algoritma KNN, serta analisis, perancangan, dan implementasi sistem yang relevan. Dataset alumni yang digunakan dalam penelitian ini disajikan pada tabel 1, mencakup berbagai atribut yang relevan untuk mendukung penelitian.

Tabel 1. Skala Penilaian *Survey*

No.	Dataset	Keterangan
1.	NIM	Atribut
2.	Nama	Atribut
3.	Program Studi	Atribut
4.	Tahun Lulus	Atribut
5.	Jenis Kelamin	Atribut
6.	Semester Masuk	Atribut
7.	Semester Lulus	Atribut
8.	Tanggal Lulus	Atribut
9.	IPK	Atribut
10.	Masa Studi	Atribut
11.	Lama Waktu Mendapatkan Pekerjaan	Label

Tahapan selanjutnya adalah mendapatkan output berupa visualisasi data dan sistem berbasis algoritma KNN. Output ini diharapkan dapat memberikan gambaran yang jelas dan mendalam terkait hasil analisis. Penelitian diakhiri dengan tahap kesimpulan dan saran, yang dirancang untuk memberikan rekomendasi strategis guna pengembangan kurikulum dan kebutuhan akreditasi di Universitas Semarang.

Data yang diperoleh kemudian akan diolah dengan mengimplementasikan metode KNN dengan mempertimbangkan kinerja para karyawan yang sebelumnya memiliki jam operasional kerja normal pada tahun 2019, yaitu sebelum pemerintah mencanangkan PPKM dikarenakan penyebaran COVID-19.

Pada tahap *data preparation*, dataset yang akan diolah akan mengalami proses pengolahan data *cleaning* tujuannya agar dapat menghapus data yang tidak diperlukan pada proses pengolahan data, sehingga menghasilkan data yang bersih dan terhindar dari *missing value* (Rahmadayanti et al., 2023). Data yang digunakan yaitu 6709. Selanjutnya mengolah data guna memperoleh dataset untuk digunakan di tahap pemodelan. Pada tahap ini mencakup pengambilan kelas dan atribut-atribut data yang kemudian dilakukan proses pembersihan (*cleaning*) serta regulasi data, pada proses ini mencoba untuk menghilangkan *missing value*. Adapun totalnya terdapat *atribut* sebanyak 11 data, dan label yang digunakan adalah lama waktu mendapatkan pekerjaan.

Metode KNN merupakan suatu algoritma yang diimplementasikan sebagai teknik klasifikasi pada data latih (*training dataset*), kemudian diekstraksi dari nilai K tetangga yang paling dekat (*nearest neighbor*). Konsep paling dasar dalam algoritma KNN ini diimplementasikan dengan menghitung jarak *Euclidean*, yang merupakan rumus mencari jarak diantara dua titik dalam sebuah lingkup 2D. Kelas terdekat dengan jumlah terbanyak akan menjadi sumber data evaluasi tersebut berada.

Kelebihan dari metode KNN adalah sederhana, mudah dalam mengimplementasikannya, memiliki efisiensi yang tinggi serta, serta ketahanan dan ketangguhannya terhadap data training (Feng et al., 2021). Selain itu, metode KNN juga memiliki kekurangan seperti diperlukannya perhitungan untuk menentukan nilai dari parameter K dan tidak mampu menangani nilai yang hilang secara implisit. Sebagai perbandingan, algoritma Naïve Bayes sering digunakan dalam penelitian serupa karena kemampuan prediktifnya yang tinggi dengan data berlabel diskret. Sementara itu, algoritma C4.5 unggul dalam memberikan interpretasi hasil melalui pohon keputusan, namun memiliki keterbatasan dalam menangani data yang terlalu banyak atribut. Dalam penelitian ini, KNN dipilih karena keunggulannya dalam menangani data besar dan kesederhanaannya dalam implementasi,

meskipun memiliki kelemahan pada waktu komputasi yang tinggi untuk dataset yang sangat besar. Dengan demikian, pemilihan KNN dalam penelitian ini lebih relevan untuk dataset alumni Universitas Semarang yang berukuran besar.

Prosedur pelaksanaan metode KNN dimulai dengan menentukan jumlah tetangga terdekat atau parameter K, yang merupakan elemen kunci dalam algoritma ini. Parameter K menentukan jumlah data tetangga yang akan digunakan sebagai dasar klasifikasi. Langkah berikutnya adalah melakukan perhitungan jarak Euclidean pada data training untuk mengukur kedekatan antara data yang diuji (*testing data*) dengan data pelatihan (*training data*). Perhitungan ini menggunakan rumus pada persamaan (1).

$$D_{xy} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Dimana D_{xy} adalah kedekatan jarak, x merupakan data *training*, y merupakan data *testing*, n merupakan jumlah atribut individu antara 1 s.d. n , dan i adalah ttribut individu antara 1 sampai dengan n

3. HASIL DAN PEMBAHASAN

3.1. Implementasi Algoritma K-NN

Data penelitian ini diperoleh melalui kegiatan pengumpulan data alumni Universitas Semarang (USM) tahun kelulusan 2019-2020. Data tersebut dikumpulkan menggunakan kuesioner yang disebar oleh UCAC Universitas Semarang. Dari total 6709 data alumni yang terkumpul, data ini kemudian dibagi menjadi dua bagian: 3743 data digunakan sebagai data pelatihan (*training data*) dan 2966 data sebagai data pengujian (*testing data*). Setiap entri data memiliki 12 atribut, yaitu NIM, Nama, Program Studi, Tahun Lulus, Jenis Kelamin, Semester Masuk, Semester Lulus, Tanggal Lulus, IPK, Masa Studi, dan Lama Waktu Tunggu Mendapatkan Pekerjaan. Atribut-atribut ini digunakan untuk melatih dan menguji model klasifikasi yang dikembangkan dalam penelitian ini.

Model klasifikasi yang dihasilkan mampu mengelompokkan data ke dalam tiga kategori berdasarkan lama waktu tunggu mendapatkan pekerjaan: " ≤ 10 bulan," " ≤ 15 bulan," dan " > 15 bulan." Proses klasifikasi bertujuan untuk mengidentifikasi pola dalam data alumni terkait dengan faktor-faktor yang memengaruhi waktu tunggu pekerjaan. Klasifikasi tersebut dirangkum dalam Tabel 2, sementara Tabel 3 menyajikan

rencian data pelatihan, dan Tabel 4 menyajikan data pengujian. Model ini dapat membantu Universitas Semarang dalam menganalisis efektivitas program pendidikan serta memberikan wawasan untuk pengembangan kurikulum yang lebih relevan dengan kebutuhan pasar kerja.

Tabel 2. Klasifikasi Waktu Tunggu

Parameter	Nilai
Cepat	<=10 bulan
Sedang	<=15 bulan
Lambat	>15 bulan

Tabel 3. Data Training

No	NIM	Nama	Prodi	Thn Lulus	Jenis Kelamin	Smt Masuk	Smt Lulus	Tgl Lulus	IPK	Masa Studi	Waktu Tunggu	Hasil
1	A111100097	FIRMANDA DESIANDO ARSITO	S1 ILMU HUKUM	2019	Laki-laki	GASAL TH 2010/2011	GENAP TH 2018/2019	2019-08-30	3.26	9	0	CEPAT
2	A111140044	WAHYU PUJI PURNI	S1 ILMU HUKUM	2019	Perempuan	GASAL TH 2014/2015	GENAP TH 2018/2019	2019-08-30	3.68	5	0	CEPAT
3	A111140058	ANANDA BIMA SAPUTRA	S1 ILMU HUKUM	2019	Laki-laki	GASAL TH 2014/2015	GENAP TH 2018/2019	2019-08-30	3.21	5	0	CEPAT
4	A111140120	TAUFIK AGUNG KURNIAWAN	S1 ILMU HUKUM	2019	Laki-laki	GASAL TH 2014/2015	GASAL TH 2018/2019	2019-03-21	3.01	4,5	0	CEPAT
5	A111150002	ADI PRASETYO	S1 ILMU HUKUM	2019	Laki-laki	GASAL TH 2015/2016	GENAP TH 2018/2019	2019-08-30	3.34	4	0	CEPAT
...
3743	B3123718067	RIRIN DWI APSARI, S.K.M.	S2 MAGISTER MANAJEMEN	2020	Perempuan	GENAP TH 2018/2019	GENAP TH 2019/2020	2020-08-27	3,68	1,5	0	CEPAT

Tabel 4. Data Testing

No	NIM	Nama	Prodi	Thn Lulus	Jenis Kelamin	Smt Masuk	Smt Lulus	Tgl Lulus	IPK	Masa Studi	Waktu Tunggu
1	A111100054	ASFITA MARINA PALUPI	S1 ILMU HUKUM	2019	Perempuan	GASAL TH 2010/2011	GENAP TH 2018/2019	2019-08-30	3.28	9	25
2	A111110020	SAUT BOANG MANALU	S1 ILMU HUKUM	2019	Laki-laki	GASAL TH 2011/2012	GENAP TH 2018/2019	2019-08-30	3.09	8	20
3	A111130019	RAFIQ IMAN RAHMAWAN	S1 ILMU HUKUM	2019	Laki-laki	GASAL TH 2013/2014	GASAL TH 2018/2019	2019-03-21	3.44	5,5	10
4	A111130078	WEMONA AMELINDA MELITA DIVA	S1 ILMU HUKUM	2019	Perempuan	GASAL TH 2013/2014	GENAP TH 2018/2019	2019-08-30	3.33	6	10
5	A111130091	RAGA LAPLATA CHAERUBA	S1 ILMU HUKUM	2019	Laki-laki	GASAL TH 2013/2014	GASAL TH 2018/2019	2019-03-21	2.87	5,5	10
...
2966	B3123517020	NADIA RIZKI RACHMAWATI, S.S.T.	S2 MAGISTER MANAJEMEN	2020	Perempuan	GENAP TH 2017/2018	GASAL TH 2019/2020	2020-08-27	3,68	2	0

3.2. Implementasi Algoritma K-NN

Data alumni mahasiswa Universitas Semarang tahun 2019/2020 sebanyak 3743 dipilih sebagai data *training* dan sisanya 2966 data dijadikan untuk data pengujian. dilakukan perhitungan jarak *Euclidean* dengan menggunakan rumus persamaan 1. Semua proses penjumlahan menggunakan aplikasi *orange* serta proses akurasi perhitungan menggunakan *rapidminer*. Aplikasi *orange* digunakan untuk melakukan prediksi dari data *testing*. Visualisasi tampilan penggunaan aplikasi *orange* disajikan pada gambar 2. Gambar ini menunjukkan visualisasi prediksi dari data *training*. Prosesnya dengan cara mengambil data *training* yang diimport dari Ms. Excel disambungkan ke widget algoritma KNN dan mengambil data *testing* untuk dilakukan prediksi.

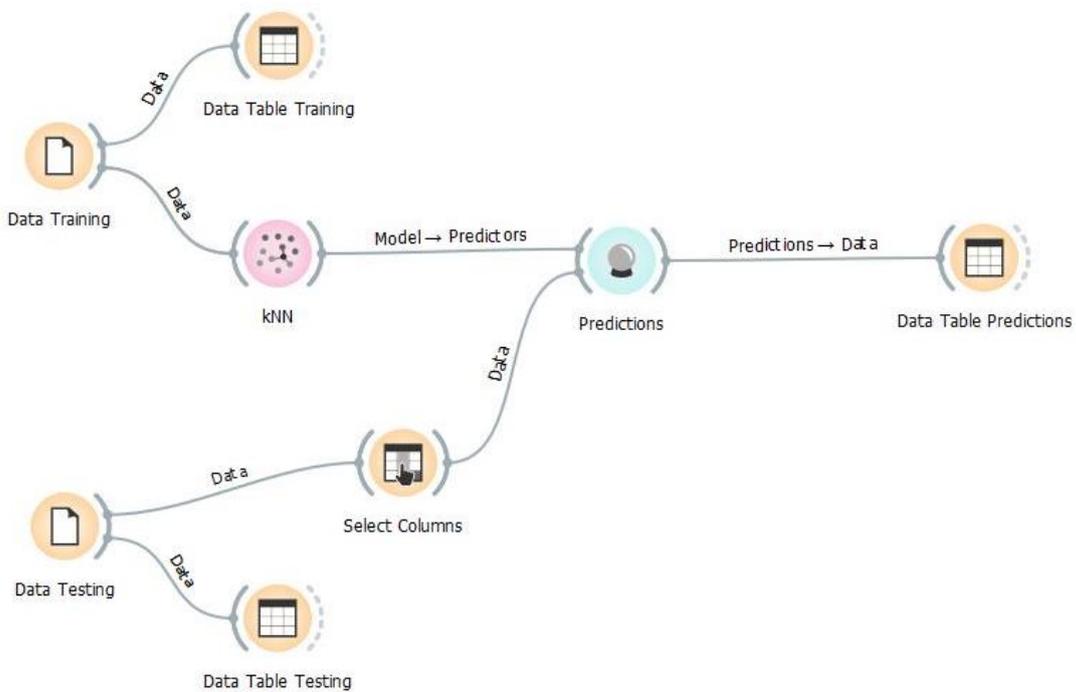
3.3. Perhitungan Akurasi

Data dalam penelitian ini diolah menggunakan *RapidMiner*, suatu perangkat lunak sumber terbuka yang digunakan untuk analisis data *mining*, *text mining*, dan analisis prediksi. (Yuliarina & Hendry, 2022), dengan data *testing* dan data *training* alumni mahasiswa tahun 2019-2020. Visualisasi akurasi menggunakan *rapidminer* disajikan pada gambar 3.

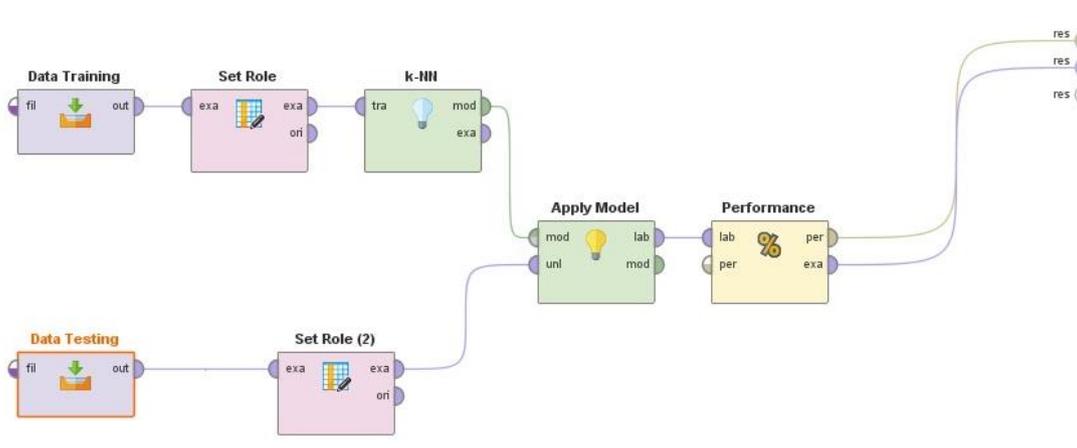
Perhitungan akurasi dengan nilai K = 1 sampai dengan K = 100 dapat dijadikan sebagai solusi atas problem yang ada, yaitu untuk mengetahui berapa lama waktu alumni Universitas Semarang dalam mendapatkan pekerjaan. Hasil dari perhitungan akurasi dengan nilai K = 1 sampai dengan K = 100 diperoleh hasil rata-rata akurasi sebesar 98,84% yang disajikan pada Tabel 5.

Hasil akurasi sebesar 98,84% yang diperoleh dari implementasi algoritma KNN dalam penelitian ini memiliki signifikansi yang penting dalam konteks tracer study di Universitas Semarang (USM). Tingkat akurasi yang sangat tinggi ini menunjukkan bahwa model mampu secara konsisten dan tepat mengklasifikasikan lama waktu tunggu alumni untuk mendapatkan pekerjaan berdasarkan atribut-atribut yang tersedia. Dalam konteks tracer study, akurasi ini sangat penting karena memberikan kepercayaan tinggi pada hasil analisis, yang pada akhirnya mendukung pengambilan keputusan strategis. Misalnya, hasil ini dapat digunakan oleh USM untuk mengevaluasi efektivitas program

akademik, merancang intervensi yang lebih tepat sasaran untuk meningkatkan daya saing lulusan, dan memberikan masukan yang valid untuk pengembangan kurikulum. Selain itu, akurasi tinggi ini juga menunjukkan bahwa data yang digunakan sudah cukup representatif dan relevan untuk menggambarkan situasi alumni USM, sehingga hasil penelitian dapat menjadi dasar yang kuat untuk mengukur kinerja universitas dalam mencetak lulusan yang siap bersaing di dunia kerja. Dengan demikian, nilai akurasi ini bukan hanya angka statistik, tetapi juga alat strategis untuk mendukung perbaikan mutu pendidikan di USM.



Gambar 2. Visualisasi menggunakan Orange



Gambar 3. Visualisasi Menggunakan Rapidminer

Tabel 5. Hasil Akurasi

Nilai K	Hasil Akurasi (%)
1	98,42
2	98,45
3	98,85
4	98,82
5	98,85
6	98,85
7	98,85
...	...
...	...
...	...
...	...
93	98,85
94	98,85
95	98,85
96	98,85
97	98,85
98	98,85
99	98,85
100	98,85
Rata-rata	98,84

4. KESIMPULAN

Sesuai hasil *studi* yang sudah dilaksanakan, dapat ditarik kesimpulan bahwa berapa lama waktu tunggu mahasiswa dalam mendapatkan pekerjaan menggunakan algoritma *K-Nearest Neighbor*. Hasil perhitungan memakai perangkat lunak *RapidMiner* dengan terdapat data *training* sebanyak 3743 dan data *testing* sebanyak 2966. Hasil dari perhitungan menunjukkan bahwa rata-rata akurasi yang diperoleh dengan nilai $K = 1$ sampai dengan $K = 100$ yaitu sebesar 98,84%. Penelitian ini dapat membantu memudahkan dalam keperluan data akreditasi, pengembangan kurikulum dan mengevaluasi kekurangan yang terjadi pada proses pendidikan.

Ada poin-poin saran yang dapat dipertimbangkan untuk malakukan penelitian berikutnya yaitu: 1). Perlu adanya sebuah sistem yang dimanfaatkan untuk mengklasifikasikan data *tracer study* alumni mahasiswa Universitas Semarang dalam mendapatkan pekerja; 2). Perlu adanya perbandingan antara algoritma *K-Nearest Neighbor* dengan algoritma data mining lainnya dalam proses *tracer study*

5. REFERENSI

Adnyana, I. M. B. (2020). Implementasi Naïve Bayes Untuk Memprediksi Waktu Tunggu Alumni Dalam Memperoleh Pekerjaan. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, 131–134. <https://prosiding.seminar-id.com/index.php/sainteks>

Akbar, R., & Mukhtar, M. (2020). Perancangan E-Tracer Study berbasis Sistem Cerdas. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 9(1), 8–12. <https://doi.org/10.32736/sisfokom.v9i1.631>

Albina, A. C., & Sumagaysay, L. P. (2020). Employability tracer study of Information Technology Education graduates from a state university in the Philippines. *Social Sciences and Humanities Open*, 2(1), 100055. <https://doi.org/10.1016/j.ssaho.2020.10055>

Alfarisi, W., Fauziyah, & Gunawan Sudarsono, B. (2022). Rancang Bangun Aplikasi Tracer Study Berbasis Web Studi Kasus Pada Fakultas Ilmu Komputer University Bung Karno. *Jurnal Ilmiah Sistem Informasi*, 2(1), 47–58. <http://simasi.lppmbinabangsa.id/index.php/home>

Aulia, T. D., Siagian, Y., & Putri, P. (2023). Penerapan Algoritma Naïve Bayes Untuk Prediksi Waktu Tunggu Alumni Mendapatkan Pekerjaan Pada Lembaga Pusat Layanan Karir Stmik Royal. *J-Com (Journal of Computer)*, 3(2), 85–92.

Bianto, M. A., Kusriani, K., & Sudarmawan, S. (2020). Perancangan Sistem Klasifikasi Penyakit Jantung Menggunakan Naïve Bayes. *Creative Information Technology Journal*, 6(1), 75. <https://doi.org/10.24076/citec.2019v6i1.231>

Feng, K., González, A., & Casero, M. (2021). A kNN algorithm for locating and quantifying stiffness loss in a bridge from the forced vibration due to a truck crossing at low speed. *Mechanical Systems and Signal Processing*, 154, 107599. <https://doi.org/10.1016/j.ymssp.2020.107599>

Kiri, T. P., & Atti, A. (2021). Pengaruh Ipk Dan Masa Studi Terhadap Waktu Tunggu Mendapatkan Pekerjaan. *Jurnal Diferensial*, 3(1). <https://doi.org/10.35508/jd.v3i1.3998>

Maricar, M. A., & Dian Pramana. (2019). Perbandingan Akurasi Naïve Bayes dan K-Nearest Neighbor pada Klasifikasi untuk Meramalkan Status Pekerjaan Alumni ITB STIKOM Bali. *Jurnal Sistem Dan Informatika (JSI)*, 14(1), 16–22. <https://doi.org/10.30864/jsi.v14i1.233>

- Noveandini, R., & Wulandari, M. S. (2019). Analisis Clustering K-Means Pada Pengelompokan Hasil Tracer Study Sebagai Media Informasi Dalam Pengembangan Kurikulum Program Studi. *Seminar Nasional Teknologi Informasi & Komunikasi (SeNTIK)*, 3(1), 133–141.
- Noviyanto, H., & Mukti, B. (2022). Prediksi Kesiapan Kerja Mahasiswa menggunakan Algoritme K-Means dan C4.5. *JOSTECH: Journal of Science and Technology*, 2(2), 179–188. <https://doi.org/10.15548/jostech.v2i2.4422>
- 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>
- Rahmadayanti, F., Anggraini, I., & Susanti, T. (2023). Pengklasterisian Data Penyakit Hipertensi dengan Menggunakan Metode K-Means. *Journal of Information System Research (JOSH)*, 4(2), 737–741. <https://doi.org/10.47065/josh.v4i2.2905>
- Rezkika, F., Sari, B. N., & Irawan, A. S. Y. (2021). Klasifikasi Masa Tunggu Alumni Untuk Mendapatkan Pekerjaan Berdasarkan Kompetensi Menggunakan Algoritma C4.5 (Studi Kasus : Fasilkom Unsika). *Progresif: Jurnal Ilmiah Komputer*, 17(2), 95. <https://doi.org/10.35889/progresif.v17i2.652>
- Shokrzade, A., Ramezani, M., Akhlaghian Tab, F., & Abdulla Mohammad, M. (2021). A novel extreme learning machine based kNN classification method for dealing with big data. *Expert Systems with Applications*, 183(December 2019), 1–18. <https://doi.org/10.1016/j.eswa.2021.115293>
- Wang, H., Xu, P., & Zhao, J. (2022). Improved KNN algorithms of spherical regions based on clustering and region division. *Alexandria Engineering Journal*, 61(5), 3571–3585. <https://doi.org/10.1016/j.aej.2021.09.004>
- Widaningsih, S. (2019). Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, Knn Dan Svm. *Jurnal Tekno Insentif*, 13(1), 16–25. <https://doi.org/10.36787/jti.v13i1.78>
- Widhiantoyo, A. (2021). Penerapan Algoritma Naïve Bayes Dengan Backward Elimination Untuk Prediksi Waktu Tunggu Alumni Mendapatkan Pekerjaan. *JIKO (Jurnal Informatika Dan Komputer)*, 4(3), 145–151. <https://doi.org/10.33387/jiko.v4i3.3272>
- Xu, Q., Ning, L., Yuan, T., & Wu, H. (2023). Application of data mining combined with power data in assessment and prevention of regional atmospheric pollution. *Energy Reports*, 9, 3397–3405. <https://doi.org/10.1016/j.egyr.2023.02.016>
- Yahdin, S., Desiani, A., Gofar, N., & Agustin, K. (2021). Application of the Relief-f Algorithm for Feature Selection in the Prediction of the Relevance Education Background with the Graduate Employment of the Universitas Sriwijaya. *Computer Engineering and Applications Journal*, 10(2), 71–80. <https://doi.org/10.18495/comengapp.v10i2.369>
- Yuliarina, A. N., & Hendry, H. (2022). Comparison of Prediction Analysis of Gofood Service Users Using the Knn & Naive Bayes Algorithm With Rapidminer Software. *Jurnal Teknik Informatika (Jutif)*, 3(4), 847–856. <https://doi.org/10.20884/1.jutif.2022.3.4.294>