TUGAS 2

RESUME JURNAL & STATE OF THE ART
PENGANTAR DATA SAINS

Dosen Pengampu: Prima Dina Atika, S.Kom.,M. Kom.

Disusun oleh Kelompok II:

1.	Angsal Pangestu	202110715143
2.	Krisna Novianto	202110715003
3.	Raihan Malik	202110715155
4.	Situmorang Ondo Palito Tua	202110715145

Kelas : F3A4

Program Studi Informatika

Fakultas Ilmu Komputer

Universitas Bhayangkara Jakarta Raya

2022

RESUME JURNAL

✔ Resume Jurnal 1

Jurnal yang berjudul Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Classifier Dan C4.5 Untuk Memprediksi Kelulusan Mahasiswa ditulis oleh Endang Etriyant, Dedy Syamsuar, dan Yesi Novaria Kunang pada tahun 2020 dengan metode penelitian Algoritma Naive Bayes Classifier dan Algoritme C4.5. Dalam penelitian dalam jurnal variabel yang diukur sebanyak 227 data set mahasiswa tahun angkatan 2013-2014 dengan jumlah 11 atribut diantaranya adalah NIM, Nama, Jenis Kelamin, Status Sekolah, Asal Sekolah, ip-s1, ip-s2, ip-s3, ip-s4, dan status kelulusan.

Penelitian ini mengangkat sebuah permasalahan yaitu mengukur tingkat akurasi dengan membandingkan kinerja antara Algoritma Naive Bayes Classifier dan Algoritma C4.5. Hal ini dilakukan agar memudahkan peneliti selanjutnya dalam memilih algoritma yang paling baik dalam melakukan penelitian. Jurnal tersebut masih mempunyai kekurangan yaitu sedikitnya jumlah data yang digunakan dalam penelitian. Alangkah baiknya jumlah data ditambah lebih banyak supaya meningkatkan nilai akurasi dari kedua algoritma yang dibandingkan. Sedangkan kelebihan dari penelitian dalam jurnal tersebut ialah metode Algoritma C4.5 memiliki nilai akurasi yang lebih tinggi dibandingkan dengan nilai akurasi metode Naive Bayes Classifier.

Dari hasil penelitian yang didapatkan bahwa metode Naive Bayes Classifier menunjukan nilai akurasi sebesar 78,46% sedangkan metode Algoritma C4.5 diperoleh nilai akurasi sebesar 79,0% yang diperoleh dari data set 227 mahasiswa tahun angkatan 2013-2014. Dengan demikian diperoleh kesimpulan bahwa implementasi data mining dalam memprediksi kelulusan mahasiswa menggunakan metode Algoritma C4.5 mempunyai kinerja yang lebih baik dibandingkan metode Naive bayes Classifier.

✔ Resume Jurnal 2

Jurnal yang berjudul Prediksi Kelulusan Mahasiswa Stikom Tunas Bangsa Prodi Sistem Informasi Dengan Menggunakan Algoritma C4.5 ditulis oleh Lydia Yohana Lumban Gaol, M. Safii, dan Dedi Suhendro pada tahun 2021 dengan metode penelitian Data Mining yaitu Algoritma C4.5. Dalam penelitian dalam jurnal variabel yang diukur adalah data alumni dan data mahasiswa tingkat akhir prodi Sistem Informasi STIKOM Tunas Bangsa.

Penelitian ini mengangkat sebuah permasalahan yaitu pengukuran terhadap pohon keputusan yang terbentuk untuk memperoleh nilai precision, recall dan akurasinya. Jurnal tersebut masih mempunyai kekurangan diantaranya; tidak disebutkan berapa banyak jumlah data yang akan diteliti, dan tidak ada yang namanya evaluasi dalam melakukan perhitungan terhadap kinerja dari pohon keputusan. Sedangkan kelebihan dari penelitian dalam jurnal tersebut ialah metode Algoritma C4.5 bisa digunakan dalam membentuk sebuah pohon keputusan.

Dari hasil penelitian yang didapatkan diperoleh nilai akurasi sebesar 90.00% dengan precision sebesar 91.38% dan recall sebesar 98.15%. Dengan demikian diperoleh kesimpulan bahwa algoritma C4.5 dinilai baik dalam menentukan prediksi kelulusan mahasiswa Stikom Tunas Bangsa Prodi Sistem Informasi.

✔ **Resume Jurnal 3**

Jurnal yang berjudul Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4.5, Naïve Bayes, Knn, Dan Svm ditulis oleh Sri Widaningsih pada tahun 2019 dengan metode penelitian Proses knowledge discovery in database (KDD). Dalam penelitian dalam jurnal variabel yang diukur sebanyak 162 data mahasiswa dengan adanya variabel predictor yang terdiri dari Jennis kelamin dan Indeks prestasi dari semester 3 hingga 6.

Penelitian ini mengangkat sebuah permasalahan yaitu mencari solusi dalam memprediksi waktu dan nilai kelulusan mahasiswa Teknik Informatika dengan algoritma C4.5, Naïve Bayes, kNN, dan SVM. Jurnal tersebut masih mempunyai kekurangan diantaranya; Algoritma kNN dengan k = 3 memiliki nilai yang sangat kecil dibandingkan dengan metode algoritma yang lain. Sedangkan kelebihan dari penelitian dalam jurnal tersebut ialah metode Algoritma Naïve Bayes dianggap memiliki nilai yang paling baik untuk semua kategori performansi dibandingkan dengan metode algoritma lainnya.

Dari hasil penelitian yang didapatkan adalah; Algoritma Naïve Bayes memiliki nilai akurasi sebesar 76,79%, algoritma SVM memiliki nilai akurasi sebesar 74,04%, algoritma kNN dengan k = 3 memiliki nilai akurasi sebesar 68,05% sedangkan algoritma C4.5 memiliki nilai akurasi sebesar 75,96%. Dengan demikian diperoleh kesimpulan dari beberapa metode Algoritma yang digunakan bahwa Algoritma C4.5 dinilai paling baik dalam memprediksi kelulusan mahasiswa dan untuk bisa mendapatkan nilai yang tinggi dapat dilakukan dengan cara memperbaiki pelaksanaan pembelajaraan untuk menghasilkan lulusan yang berkualitas.

✔ **Resume Jurnal 4**

Jurnal yang berjudul Pemodelan Prediksi Predikat Kelulusan Mahasiswa Menggunakan Decision Tree C4.5 ditulis oleh Abdul Rohman dan Sri Mujiyono pada tahun 2021 dengan metode penelitian Algoritma decision tree C4.5. Dalam penelitian dalam jurnal variabel yang diukur sebanyak 35 data mahasiswa yang terdiri dari 20 mahasiswa lulusan S1 Teknik Informatika, 4 lulusan S1 Pendidikan Guru PAUD, dan 11 lulusan S1 Pendidikan Guru SD.

Penelitian ini mengangkat sebuah permasalahan yaitu memprediksi kelulusan mahasiswa berdasarkan predikat kelulusan mahasiswa. Jurnal tersebut masih mempunyai kekurangan yaitu jumlah data yang digunakan sedikit. Sedangkan kelebihan dari penelitian dalam jurnal tersebut dijelaskan bahwa Decision tree C4.5 dapat membantu menerapkan pola atau model yang menghasilkan pohon keputusan yang digunakan

Dari hasil penelitian ini mendapat model pohon keputusan dengan variabel atau atribut indeks prestasi dari semester 1,2,3, dan yang berpengaruh terhadap predikat kelulusan mahasiswa dengan nilai akurasi sebesar 71,67%.

✔ **Resume Jurnal 5**

Jurnal yang berjudul Prediksi Ketepatan Waktu Lulus Mahasiswa Menggunakan Algoritma C4.5 Pada Fakultas Dakwah Dan Ilmu Komunikasi UIN Raden Intan Lampung ditulis oleh Fiqih Satria, Zamhariri , dan M. Apun Syaripudin pada tahun 2020 dengan metode penelitian Algoritma C4.5. Dalam penelitian dalam jurnal variabel yang diukur sebanyak 280 data mahasiswa angkatan 2014 Fakultas Dakwah Dan Ilmu Komunikasi UIN Raden Intan Lampung.

Penelitian ini mengangkat sebuah permasalahan yaitu melakukan analisis ketepatan waktu lulus mahasiswa dengan Algoritma C4.5. Pada penelitian masih memiliki kekurangan yaitu dataset memiliki jumlah data False Negative (FN) dan False Positive (FP) yang tidak mendekati hitungan nilai akurasi. Sedangkan kelebihan dari penelitian dalam jurnal tersebut ialah dijelaskan bahwa metode Algoritma C.45 dinilai baik dalam mengklasifikasi dan melakukan prediksi terhadap mahasiswa yang lulus tepat waktu. Dari hasil penelitian yang didapatkan software WEKA dengan menggunakan algoritma C4.5 menjadikan jenis kelamin root pada pohon keputusan, kemudian dilanjutkan dengan atribut jenis sekolah, jalur masuk dan peringkat. Dengan demikian diperoleh kesimpulan dari pengukuran kinerja algoritma yang telah dilakukan, dapat disimpulkan dalam kasus ini algoritma C4.5 memiliki nilai akurasi, presisi dan recall masing-masing 58,2%, 86,8% dan 60,4%.

✔ **Resume Jurnal 6**

Jurnal yang berjudul Prediksi Kelulusan Mahasiswa Fakultas Teknologi Informasi Umby Menggunakan Metode Decision Tree Penerapan Algoritma C4.5 ditulis oleh Vidya Anggraini Nurislamiaty dan Anief Fauzan Rozi pada tahun 2021 dengan metode penelitian Algoritma Decision Tree C4.5. Dalam penelitian dalam jurnal variabel yang diukur sebanyak 7.955 data mahasiswa Fakultas Teknologi Informasi Universitas Mercu Buana Yogyakarta Angkatan 2012-2015.
Penelitian ini mengangkat sebuah permasalahan yaitu masa studi mahasiswa lebih dari waktu yang di tetapkan oleh perguruan tinggi yang menyebabkan kerugian fakultas dan membuat akreditasi fakultas menjadi menurun. Pada penelitian masih memiliki kekurangan yaitu data latih masih sedikit dan diperlukan jumlah data yang lebih banyak agar validasi data jauh lebih baik dengan cara menggunakan metode lain. Sedangkan kelebihannya adalah metode Algoritma Decision Tree C4.5 dijelaskan dapat membentuk pohon keputusan yang mengubah fakta menjadi aturan dengan bahasa alami.

Dari hasil penelitian yang didapatkan dalam melakukan analisis dan penerapan metode decision tree pada data mahasiswa FTI UMBY Angkatan 2012-2015 dapat diperoleh kesimpulan bahwa :

1. Dari hasil uji coba menggunakan 263 data sampel didapatkan akurasi menunjukkan kinerja algoritma mengklasifikasikan data sebanyak 215 data dengan presentase sebesar (82,8897%) serta data tidak berhasil di klasifikasikan dengan benar (error) sebanyak 48 data dengan presentase sebesar (17,1103%).

2. Pada hasil penerapan Algoritma C4.5 pada 263 data sampel Mahasiswa FTI UMBY Angkatan 2012-2015 dengan atribut sks tempuh kamulatif semester 7, prodi, ips semester 5-ips semester 7 menghasilkan suatu rule dimana mahasiswa yang lulus tepat waktu ialah mahasiswa pada saat semester 7 memiliki IPS dengan indeks A (3,51- 4,00) dan indeks B (3,00-3,50) serta telah menempuh SKS 101-150 SKS pada semester 7.

✔ **Resume Jurnal 7**

Jurnal yang berjudul Metode Klasifikasi Data Mining Algoritma C4.5 Dan Part Untuk Prediksi Waktu Kelulusan Mahasiswa Di Universitas Darwan Ali ditulis oleh Selviana Yunita dan Nurahman pada tahun 2022 dengan metode penelitian Algoritma C4.5, Algoritma PART. Dalam penelitian dalam jurnal variabel yang diukur sebanyak 707 mahasiswa/i, tahun 2015 sebanyak 290 mahasiswa/i. Variabel yang diukur adalah jurusan program studi, IPK dan jenis kelamin mahasiswa.

Penelitian ini mengangkat sebuah permasalahan yaitu Universitas Darwan Ali sampai saat ini belum memanfaatkan data kondisi akademik mahasiswa secara maksimal. Jurnal tersebut masih mempunyai kekurangan yaitu tingkat performa yang dilakukan dalam penelitian pada Algoritma PART nilai akurasinya masih rendah saat dibandingkan dengan Algoritma C4.5. Sedangkan kelebihan dari jurnal tersebut dijelaskan bahwa metode klasifikasi bisa membantu mengembangkan suatu sistem cerdas serta variabel yang sesuai untuk mendukung sistem tersebut dan bisa menentukan algoritma mana dengan akurasi yang lebih tinggi.

Dari hasil penelitian yang didapatkan berdasarkan 1012 jumlah data yang berisi data Prodi, Indeks Prestasi dari semester 1 hingga 4, jenis kelamin serta status kelulusan mahasiswa, diperoleh beberapa rule yang dapat dijadikan panduan oleh Universitas untuk mengetahui ketepatan waktu lulus mahasiswa berdasarkan Indeks Prestasi 4 semester pertama serta jenis kelamin mahasiswa. Algoritma yang direkomendasikan untuk digunakan dalam proses klasifikasi adalah algoritma C4.5 karena memiliki akurasi yang lebih baik dibandingkan algoritma PART, yaitu sebesar 83.004 %. Dengan demikian diperoleh kesimpulan bahwa metode Algoritma C4.5 lebih baik daripada metode Algoritma PART.

TABEL RELATED RESEACH / STATE of THE ART
JUDUL: PENERAPAN C4.5 DALAM MEMPREDIKSI KELULUSAN MAHASISWA

No	Tahun	Peneliti	Judul	Metode	Yang Diukur	Masalah	Kelebihan	Kekurangan	Hasil
1	2020	Endang Etriyant, Dedy Syamsuar, Yesi Novaria Kunang	Implementasi Data Mining Menggunakan Algoritma Naive Bayes Classifier Dan C4.5 Untuk Memprediksi Kelulusan Mahasiswa	Algoritma Naive Bayes Classifier dan Algoritme C4.5	227 data set mahasiswa tahun angkatan 2013-2014 dengan 11 atribut yaitu NIM, Nama, Jenis Kelamin, Status Sekolah, Asal Sekolah, IP-S1, IP-S2, IP-S3, IP-S4, Status Kelulusan.	Mengukur tingkat akurasi dengan membandingkan kinerja antara Algoritma Naive Bayes Classifier dan Algoritma C4.5.	Algoritma C4.5 memiliki nilai akurasi yang lebih tinggi dibandingkan dengan nilai akurasi metode Naive Bayes Classifier	Sedikitnya jumlah data yang digunakan dalam penelitian.	Data set dengan metode Naive Bayes Classifier menunjukan nilai Accuracy 78,46% dan metode Algoritma C4.5 diperoleh nilai Accuracy yang lebih besar yaitu 79,0%.
2	2021	Lydia Yohana Lumban Gaol, M. Safii, Dedi Suhendro	Prediksi Kelulusan Mahasiswa Stikom Tunas Bangsa Prodi Sistem Informasi Dengan Menggunakan Algoritma C4.5	Data Mining, Algoritma C4.5	Data Alumni dan data mahasiswa tingkat akhir prodi Sistem Informasi STIKOM Tunas Bangsa	Pengukuran terhadap pohon keputusan yang terbentuk untuk memperoleh nilai precision, recall dan akurasinya.	Algoritma C4.5 dapat digunakan dalam membentuk sebuah pohon keputusan.	Jumlah data tidak disebutkan, Tidak ada evaluasi dalam melakukan perhitungan terhadap kinerja dari pohon keputusan.	Diperoleh tingkat akurasi sebesar 90.00% dengan precision sebesar 91.38% dan recall sebesar 98.15%.

No	Tahun	Peneliti	Judul	Metode	Yang Diukur	Masalah	Kelebihan	Kekurangan	Hasil
3	2019	Sri Widaningsih	Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4.5, Naïve Bayes, Knn, Dan Svm	Proses knowledge discovery in database (KDD).	162 data mahasiswa dengan variabel-variabel prediktor terdiri dari jenis kelamin dan indeks prestasi dari semester 3 hingga 6.	Mencari solusi dalam memprediksi waktu dan nilai kelulusan mahasiswa Teknik Informatika dengan algoritma C4.5, Naïve Bayes, kNN, dan SVM.	Algoritma Naïve Bayes memiliki nilai yang paling baik untuk semua kategori performansi dibandingkan dengan algoritma lainnya.	Algoritma kNN dengan k = 3 memiliki nilai yang sangat kecil.	Algoritma Naïve Bayes memiliki akurasi 76,79%, algoritma SVM memiliki akurasi sebesar 74,04%, algoritma kNN dengan k = 3 memiliki akurasi 68,05% sedangkan algoritma C4.5 memiliki akurasi sebesar 75,96%
4	2021	Abdul Rohman, Sri Mujiyono	Pemodelan Prediksi Predikat Kelulusan Mahasiswa Menggunakan Decision Tree C4.5.	Algoritma decision tree C4.5.	35 data mahasiswa yang terdiri dari 20 lulusan S1 Teknik Informatika, 4 lulusan S1 Pendidikan Guru PAUD, dan 11 lulusan S1 Pendidikan Guru SD.	Memprediksi kelulusan mahasiswa berdasarkan predikat kelulusan mahasiswa.	Decision tree C4.5 membantu menerapkan pola atau model yang menghasilkan pohon keputusan yang digunakan.	Jumlah data yang digunakan sedikit.	Decision Tree C4.5 memprediksi kelulusan mahasiswa dengan nilai akurasi 71,67%.

No	Tahun	Peneliti	Judul	Metode	Yang Diukur	Masalah	Kelebihan	Kekurangan	Hasil
5	2020	Fiqih Satria, Zamhariri , M. Apun Syaripudin	Prediksi Ketepatan Waktu Lulus Mahasiswa Menggunakan Algoritma C4.5 Pada Fakultas Dakwah Dan Ilmu Komunikasi UIN Raden Intan Lampung.	Algoritma C4.5	280 data mahasiswa angkatan 2014 Fakultas Dakwah Dan Ilmu Komunikasi UIN Raden Intan Lampung	Melakukan analisis ketepatan waktu lulus mahasiswa dengan Algoritma C4.5	Algoritma C.45 dinilai baik dalam mengklasifikasi dan melakukan prediksi terhadap mahasiswa yang lulus tepat waktu.	Dataset memiliki jumlah data False Negative (FN) dan False Positive (FP) yang tidak mendekati nilai akurasi.	Nilai akurasi sebesar 71 % melalui perhitungan F-Measure.
6	2021	Vidya Anggraini Nurislamiaty, Anief Fauzan Rozi	Prediksi Kelulusan Mahasiswa Fakultas Teknologi Informasi Umby Menggunakan Metode Decision Tree Penerapan Algoritma C4.5	Algoritma Decision Tree C4.5.	7.955 data mahasiswa Fakultas Teknologi Informasi Universitas Mercu Buana Yogyakarta Angkatan 2012-2015.	Masa studi mahasiswa lebih dari waktu yang di tetapkan oleh kampus, Akreditasi fakultas menurun.	Algoritma Decision Tree C4.5 dapat membentuk pohon keputusan yang mengubah fakta menjadi aturan dengan bahasa alami.	Data latih masih sedikit.	Algoritma C.45 nilai akurasi sebesar 82.8897% .

No	Tahun	Peneliti	Judul	Metode	Yang Diukur	Masalah	Kelebihan	Kekurangan	Hasil
7	2022	Selviana Yunita, Nurahman	Metode Klasifikasi Data Mining Algoritma C4.5 Dan Part Untuk Prediksi Waktu Kelulusan Mahasiswa Di Universitas Darwan Ali	Algoritma C4.5, Algoritma PART	Data mahasiswa tahun 2011 sebanyak 707 orang, tahun 2015 sebanyak 290 orang. Variabel yang diukur adalah jurusan program studi, IPK dan jenis kelamin mahasiswa.	Universitas Darwan Ali sampai saat ini belum memanfaatkan data kondisi akademik mahasiswa secara maksimal.	Dapat membantu mengembangkan suatu sistem cerdas serta variabel yang sesuai untuk mendukung sistem tersebut dan bisa menentukan algoritma mana dengan akurasi yang lebih tinggi .	Tingkat performa yang dilakukan dalam penelitian pada Algoritma PART nilai akurasi masih rendah.	Nilai akurasi pada Algoritma C4.5 adalah 83,004%. dan nilai akurasi pada Algoritma PART adalah 82.8063 %

DAFTAR PUSTAKA

Etriyanti, E., Syamsuar, D., & Kunang, Y. N. (2020, Februari). Implementasi Data Mining Menggunakan Algoritme Naive Bayes Classifier dan C4.5 untuk Memprediksi Kelulusan Mahasiswa. *Telematika, XIII*(1), 56-67.

Lumban Gaol, L. Y., Safii, M., & Suhendro, D. (2021, Juni). Prediksi Kelulusan Mahasiswa Stikom Tunas Bangsa Prodi Sistem Informasi Dengan Menggunakan Algoritma C4.5. *BRAHMANA, II*(2), 97-106.

Nurislamiaty, V. A., & Rozi, A. F. (2021). Prediksi Kelulusan Mahasiswa Fakultas Teknologi Informasi Umby Menggunakan Metode Decision Tree Penerapan Algoritma C4.5. -, 1-8.

Rohman, A., & Mujiyono, S. (2021, Desember). Permodelan Prediksi Predikat Kelulusan Mahasiswa Menggunakan Decision Tree C4.5. *Jurnal Prodi Teknik Informatika UNW, III*(2), 1-5.

Satria, F., Zamhariri, & Syaripudin, M. A. (2020, April). Prediksi Ketepatan Waktu Lulus Mahasiswa Menggunakan Algoritma C4.5 Pada Fakultas Dakwah Dan Ilmu Komunikasi. *Jurnal Ilmiah MATRIK, XXII*(1), 28-25.

Widaningsih, S. (2019, April). PERBANDINGAN METODE DATA MINING UNTUK PREDIKSI NILAI DAN WAKTU KELULUSAN MAHASISWA PRODI TEKNIK INFORMATIKA. *Jurnal Tekno Insentif, XIII*(1), 16-25.

Yunita, S., & Nurahman. (2022). METODE KLASIFIKASI DATA MINING ALGORITMA C4.5 DAN PART UNTUK PREDIKSI WAKTU KELULUSAN MAHASISWA DI UNIVERSITAS DARWAN ALI.