

Perbandingan Algoritma ELM Dan Backpropagation Terhadap Prestasi Akademik Mahasiswa

Heny Pratiwi, Kusno Harianto

STMIK Widya Cipta Dharma

Jl. Prof. M. Yamin No.25 Samarinda Kalimantan Timur, (0541)736071

henypratiwi@wicida.ac.id, dzakiraharianto97@gmail.com

Abstract

Extreme Learning Machine and Backpropagation Algorithms are used in this study to find out which algorithm is most suitable for knowing student academic achievement. The data about students are explored to get a pattern so that the characteristics of new students can be known every year. The evaluation process of this study uses confusion matrix for the introduction of correctly recognized data and unknown data. Comparison of this algorithm uses student data at the beginning of the lecture as early detection of students who have problems with academics to be anticipated. The variables used are the value of the entrance examination for new students, the first grade IP value, Gender, and Working Status, while the output variable is the quality value as a classification of academic performance. The results of this study state that the Extreme Learning Machine algorithm has a 14.84% error rate lower than Backpropagation 28.20%. From the model testing stage, the most accurate result is the Extreme Learning Machine algorithm because it has the highest accuracy and the lowest error rate.

Keywords: Algorithm, ELM, Backpropagation, Achievement

Abstrak

Algoritma Extreme Learning Machine dan Backpropagasi digunakan pada penelitian ini untuk mengetahui algoritma mana yang paling sesuai dalam mengetahui prestasi akademik mahasiswa. Data-data tentang mahasiswa ini digali untuk mendapatkan pola sehingga dapat diketahui karakteristik mahasiswa baru setiap tahunnya. Proses evaluasi dari penelitian ini menggunakan confusion matrix untuk pengenalan data yang dikenali dengan benar dan data yang tidak dikenali. Perbandingan algoritma ini menggunakan data mahasiswa di awal perkuliahan sebagai deteksi dini terhadap mahasiswa yang memiliki masalah dengan akademik untuk segera diantisipasi. Adapun variabel yang digunakan adalah nilai ujian masuk mahasiswa baru, nilai IP Semester 1, Jenis Kelamin, dan Status Bekerja, sedangkan variabel keluaran adalah nilai mutu sebagai klasifikasi kinerja akademik. Hasil penelitian ini menyatakan bahwa algoritma Extreme Learning Machine memiliki 14,84% tingkat kesalahan lebih rendah dibandingkan dengan Backpropagasi 28,20%. Dari tahapan pengujian model didapatkan hasil yang paling akurat adalah algoritma Extreme Learning Machine karena mempunyai tingkat akurasi paling tinggi dan tingkat kesalahan paling rendah.

Kata kunci: Algoritma, ELM, Backpropagasi, Prestasi

1. PENDAHULUAN

Algoritma Extreme Learning Machine memiliki kelebihan dalam bidang kecepatan pembelajaran dan tingkat akurasi, sedangkan Backpropagasi memiliki keunggulan dalam memecahkan suatu permasalahan dibandingkan dengan jaringan lapisan tunggal dengan pelatihan terbimbing sehingga pada setiap masing-masing pola masukan terdapat pasangan keluaran.

Kedua algoritma ini akan diujicoba dan dibandingkan untuk menghasilkan informasi prestasi akademik mahasiswa. Berlimpahnya data mahasiswa di Perguruan Tinggi terutama mahasiswa baru membuka peluang pihak akademik untuk dapat mengukur prestasi akademik mahasiswa secara lebih tepat. Data-data tersebut apabila digali dengan tepat akan membuat pihak Perguruan Tinggi mendapatkan pengetahuan atau pola-pola prestasi akademik mahasiswa baru sehingga dapat mengatasi masalah-masalah yang biasanya muncul dalam semester awal mahasiswa baru seperti kurang bisa bersosialisasi dan beradaptasi terhadap proses perkuliahan, mendapatkan nilai rendah di suatu mata kuliah sehingga menyebabkan ketidaklulusan, atau faktor ekonomi.

Sebagai contoh apabila pihak akademik mengetahui ada mahasiswa di semester awal perkuliahan ada kemungkinan mahasiswa tidak lulus kuliah atau memiliki indeks prestasi kurang dari atau sama dengan dua akan bisa diantisipasi sejak awal dan dapat menentukan kebijakan serta langkah yang tepat untuk menangani prestasi akademik mahasiswa tersebut sejak awal semester. Tujuan dari penelitian ini adalah mengevaluasi kinerja algoritma Extreme Learning Machine dan Backpropagasi terhadap prestasi akademik mahasiswa di semester mendatang. Adapun penelitian ini bermanfaat karena saat ini bagian akademik selalu menemukan adanya mahasiswa baru yang tidak dapat melanjutkan studi ke semester lanjut sehingga terjadi penurunan mahasiswa aktif kuliah secara signifikan. Hal ini terjadi hampir di semua perguruan tinggi yang ada. Ruang lingkup pada penelitian ini adalah melihat keakuratan dua algoritma dan melakukan perbandingan yaitu ELM dan Backpropagasi dalam melakukan evaluasi data-data mahasiswa baru untuk meningkatkan prestasi akademik mahasiswa.

Extreme Learning Machine merupakan algoritma pembelajaran yang muncul untuk jaringan *neural single feed layer* tersembunyi, dimana parameter simpul tersembunyi dihasilkan secara acak dan bobot output dapat dihitung secara analitik. [9]. ELM juga memberikan solusi pembelajaran terpadu yang efisien untuk aplikasi pembelajaran fitur, pengelompokan, regresi, dan klasifikasi. Teori ELM menunjukkan bahwa neuron tersembunyi penting tetapi tidak perlu disetel secara iteratif. Faktanya semua parameter dari *node* tersembunyi dapat mandiri dari sample pelatihan dan dihasilkan secara acak sesuai dengan setiap distribusi probabilitas kontinu. [5]

ELM memiliki fitur yang menarik dan signifikan, berbeda dengan algoritma pembelajaran berbasis gradien yang populer untuk jaringan saraf *feed-forward*. Fitur yang dimaksud adalah :

- a. Kecepatan belajar ELM sangat cepat. Dalam simulasi yang dilaporkan dalam literatur, fase pembelajaran ELM dapat diselesaikan dalam hitungan detik untuk banyak aplikasi. Sebelumnya, tampaknya ada penghalang kecepatan virtual yang sebagian besar algoritma pembelajaran klasik tidak dapat menembusnya. Dan bukan hal yang tidak biasa lagi kalau pelatihan jaringan saraf *feed-forward* yang

- menggunakan algoritma pembelajaran klasik memerlukan waktu yang cukup lama bahkan untuk aplikasi yang sederhana.
- b. ELM memiliki kinerja generalisasi yang lebih baik dibandingkan pembelajaran berbasis gradien, seperti backpropagasi dalam kebanyakan kasus.
 - c. Algoritma pembelajaran klasik berbasis gradien dan beberapa algoritma pembelajaran lainnya menghadapi beberapa masalah seperti minima lokal, tingkat pembelajaran yang tidak tepat, dan lain-lain. Untuk menghindari masalah ini, beberapa metode seperti peluruhan bobot dan metode pemberhentian lebih awal sering digunakan pada algoritma klasik ini. Algoritma pembelajaran ELM terlihat jauh lebih sederhana dari algoritma pembelajaran jaringan saraf *feed-forward* kebanyakan.
 - d. Tidak seperti algoritma pembelajaran berbasis gradien yang hanya bekerja untuk fungsi aktivasi terdiferensiasi, algoritma extreme learning machine dapat digunakan untuk melatih SLFNs dengan banyak fungsi aktivasi yang tidak terdiferensiasi.

Backpropagasi merupakan metode umum untuk pelatihan jaringan saraf. Metode ini beroperasi pada dalam dua fase untuk menghitung gradien kerugian pada ketentuan bobot jaringan. Pada fase pertama, fitur masukan disebarkan ke depan melalui jaringan untuk menghitung keluaran fungsi, fase kedua merupakan turunan dari kehilangan pelatihan sehubungan dengan bobot yang disebarkan kembali dari lapisan keluaran menuju lapisan masukan. Turunan ini digunakan untuk memperbaharui bobot. [4] Algoritma backpropagasi digunakan untuk mengembangkan model jaringan syaraf tiruan. [7] Tipikal topologi dari jaringan syaraf tiruan backpropagasi melibatkan tiga lapisan yaitu lapisan masukan (tempat data diperkenalkan kepada jaringan), lapisan tersembunyi (tempat data diproses), dan lapisan keluaran (tempat hasil dari masukan yang diberikan diproduksi). [2] Algoritma backpropagasi biasa digunakan untuk pelatihan bahkan dengan data yang kompleks. [3] Metode pelatihan backpropagasi melibatkan *feedforward* dari pola pelatihan masukan, perhitungan, dan backpropagasi dari eror serta penyesuaian bobot dalam sinapsis. [6]

Saat ini banyak Perguruan Tinggi mengukur prestasi mahasiswa dengan indeks prestasi 0,00 sampai 4,00. Indeks prestasi akademik diperoleh dari rata-rata poin matakuliah yang diambil. Indeks prestasi diukur pada tiap semester dan kumulatif. Indeks prestasi semester (IPS) merupakan indeks prestasi yang dihitung pada tiap semester berdasarkan pada nilai akhir dari matakuliah pada semester tersebut. Indeks prestasi kumulatif (IPK) merupakan indeks prestasi yang dihitung dari keseluruhan nilai matakuliah.

Pertumbuhan luar biasa yang terdapat pada data elektronik Perguruan Tinggi menciptakan kebutuhan untuk memiliki beberapa informasi bermakna yang diambil dari volume data yang cukup besar ini. Kemajuan di bidang penambangan data (*data mining*) ini memungkinkan untuk menambang data yang dapat digunakan untuk meningkatkan kualitas

pendidikan. [1]

Data Mining merujuk pada keseluruhan proses yang terdiri dari pengumpulan dan analisis data, pengembangan model pembelajaran induktif dan adopsi praktek pengambilan keputusan serta konsekuensi dari tindakan berdasarkan pengetahuan yang diperoleh. Kegiatan ini dapat dibagi menjadi dua bagian utama penelitian, sesuai dengan tujuan utama dari analisisnya yaitu interpretasi dan prediksi. Ada tujuh kegiatan dasar penggalian data yaitu karakterisasi dan diskriminasi, klasifikasi, regresi, *time series*, asosiasi, *clustering*, serta dekripsi, dan visualisasi. [8]

Jaringan saraf tiruan telah secara luas diterapkan untuk klasifikasi pola dan masalah regresi. Alasan utama keberhasilan dari jaringan saraf tiruan adalah kemampuannya dalam memperoleh fungsi model perkiraan non-linier yang menggambarkan hubungan antara variabel dependen dan independen dengan menggunakan input sampel yang diberikan. Meskipun jaringan saraf tiruan memiliki banyak kelebihan seperti kemampuan pendekatan yang lebih baik serta struktur jaringan yang sederhana, namun, masih memiliki kelemahan seperti minima lokal, tingkat pembelajaran yang kurang tepat, serta pemilihan jumlah hidden neuron yang kurang tepat.

Kegiatan data mining dibagi menjadi dua bagian utama penelitian sesuai dengan tujuan utama dari analisisnya yaitu interpretasi dan prediksi.

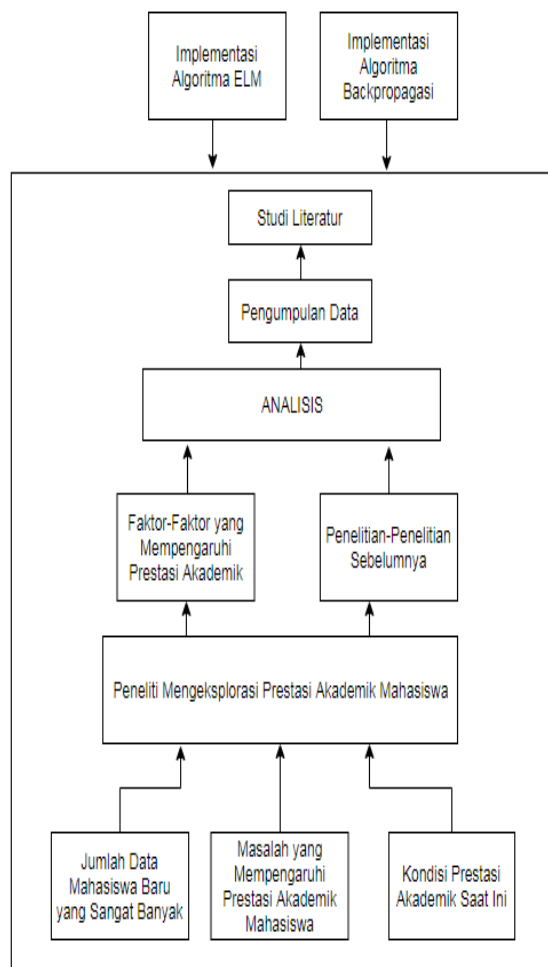
- a. Interpretasi. Tujuan interpretasi adalah untuk mengidentifikasi pola-pola teratur dalam data dan untuk mengekspresikannya melalui aturan dan kriteria yang mudah dipahami oleh para ahli dalam domain aplikasi. Contohnya adalah, untuk sebuah perusahaan di industri ritel mungkin akan menguntungkan bila bisa mengelompokkan para pelanggan yang setia berdasarkan profil pembelanjaan mereka. Segmen yang dihasilkan dengan cara ini terbukti berguna dalam mengidentifikasi relung pasar baru dan mengarahkan promosi pemasaran perusahaan di masa depan.
- b. Prediksi. Tujuan dari prediksi adalah untuk mengantisipasi nilai dimana variabel acak akan mengasumsikan masa depan atau memperkirakan kemungkinan kejadian di masa depan. Contohnya adalah sebuah perusahaan retail dapat memprediksikan penjualan produk tertentu untuk minggu-minggu berikutnya hal ini dikarenakan sebagian besar teknik data mining berasal dari nilai sekumpulan variabel yang berkaitan dengan entity di dalam database. Model data mining dapat menunjukkan bahwa kemungkinan seorang pelanggan akan melakukan pembelian lagi di masa depan tergantung pada fitur-fitur seperti usia, pekerjaan, dan sebagainya.

Data mining memiliki kemampuan cepat dalam melakukan analisis dan sangat fokus pada variabel-variabel penting. *Patterns* (pola) dan *rules* (aturan, kaidah) yang dapat digunakan dalam membuat keputusan dan *forecast* (meramalkan) dampak dari keputusan tersebut. *Intelligent* data mining meliputi informasi dalam data *warehouse* dimana *query* dan laporan biasa tidak bisa mengungkapkan informasi secara efektif. *Data mining tools* mampu menemukan pola dalam data dan memberi dugaan berupa *rules*.

2. METODOLOGI PENELITIAN

2.1. Kerangka Berpikir

Algoritma ELM dan backpropagasi digunakan untuk melihat metode mana yang paling akurat dalam mengukur prestasi akademik mahasiswa apalagi prestasi dapat menjadi acuan dari sebuah kinerja perguruan tinggi. Adapun kerangka berpikir dari penelitian ini yaitu berawal dari prestasi akademik mahasiswa baru saat ini apalagi dengan jumlah mahasiswa baru yang cukup banyak. Kemudian adanya masalah-masalah yang mempengaruhi prestasi akademik mahasiswa, dan kondisi prestasi akademik mahasiswa saat ini. Dari hal-hal tersebut kemudian peneliti mengeksplorasi prestasi akademik mahasiswa dengan cara melihat faktor-faktor yang mempengaruhi prestasi akademik dan melihat penelitian-penelitian sebelumnya. Semua hal tersebut akan dianalisis serta dilakukan pengumpulan data dan melaksanakan studi literatur. Seperti yang dapat dilihat pada gambar 1.



Gambar 1. Kerangka Berpikir

2.2. Prosedur Pengumpulan Data

Data pada penelitian ini menggunakan data primer yaitu mahasiswa baru kemudian data sekunder yang diambil dengan mempelajari studi pustaka berupa *data mining*, *ebook*, dan jurnal tentang prestasi akademik mahasiswa. Data yang digunakan adalah data mahasiswa baru angkatan 2014 sampai dengan 2016 program studi Manajemen Informatika, Sistem Informasi, dan Teknik Informatika. Di tiap kategori terdiri dari empat variabel masukan yaitu nilai tes masuk, nilai indeks prestasi semester 1, jenis kelamin, dan status bekerja. Variabel keluaran adalah nilai mutu berupa A B C D E. Pemilihan variabel ini didasarkan pada acuan penelitian yang sudah dilakukan sebelumnya, sedangkan lainnya merupakan variabel yang dicoba digunakan dalam penelitian ini. Pemilihan variabel ini didasarkan pula pada ketersediaan data di Perguruan Tinggi.

2.3. Desain Penelitian

Pada penelitian ini digunakan model standarisasi data mining yaitu CRISP-DM (*Cross Industry Standart Process for Data Mining*) dengan langkah-langkah sebagai berikut:

2.3.1. Pemahaman Data (*Data Understanding*)

Data yang digunakan dalam penelitian ini adalah sumber data primer. Data diperoleh dari Sistem Informasi Akademik berupa data nilai tes masuk, nilai Indeks Prestasi semester 1, Jenis Kelamin, dan Status Bekerja.

2.3.2. Pengolahan Data (*Data Preparation*)

Tahapan pengolahan data adalah sebagai berikut:

- a. Tahap Pertama, penentuan data yang akan diolah. Dari data yang diperoleh, tidak semua data akan diolah karena penelitian yang dilakukan memiliki batasan-batasan data yang akan digunakan.
- b. Tahap Kedua, penanganan data *missing value*. *Missing Value* adalah data yang tidak lengkap karena atribut tidak tercatat maupun atribut memang tidak dimiliki. Penanganan *missing value* dilakukan dengan penghapusan record yang kosong.
- c. Tahap Ketiga, menentukan atribut yang akan digunakan dari tahap pertama. Atribut yang digunakan adalah nilai tes masuk, nilai Indeks Prestasi semester 1, Jenis Kelamin, dan Status Bekerja.
- d. Tahap Keempat, melakukan konversi data. Data dengan atribut yang telah dipilih kemudian dikonversikan untuk memudahkan proses data mining pada sebagian atribut, karena data akan diproses oleh MATLAB.

2.3.3. Permodelan (*Modelling*)

Metode yang digunakan pada penelitian ini adalah Extreme Learning Machine dan Backpropagasi, untuk melakukan pengukuran akurasi akan digunakan MATLAB.

2.3.4. Evaluasi (*Evaluation*)

Pada tahapan ini akan dilakukan validasi serta pengukuran keakuratan hasil yang dicapai oleh model yaitu Confusion Matrix.

2.4. Tahapan Penelitian

Tahapan dalam penelitian perbandingan algoritma ELM dan Backpropagasi terhadap Prestasi Akademik Mahasiswa adalah pengumpulan data, implementasi metode Extreme Learning Machine dan Backpropagasi untuk prediksi melalui pembagian data, training dan testing serta analisis hasil prediksi.

Pada tahap awal penelitian dimulai dengan menentukan latar belakang dan tujuan penelitian serta mendefinisikan ruang lingkup. Studi literatur dilakukan untuk memperdalam pemahaman mengenai cara kerja metode serta tahapan-tahapan apa saja yang diperlukan untuk prediksi. Selain itu, studi literatur juga dilakukan untuk mengetahui prestasi akademik mahasiswa baru supaya pengumpulan data memiliki fokus yang lebih baik.

Tahap ke dua dari penelitian ini adalah pengumpulan data. Tahap ke tiga adalah implementasi metode Extreme Learning Machine dan Backpropagasi melalui pembagian data, *training* dan *testing*. Tahap ke empat adalah analisis hasil prediksi, kemudian tahap terakhir adalah menarik kesimpulan perbandingan algoritma mana yang paling akurat dalam memberikan hasil.

2.5. Metode Analisis Kinerja

Pada metode analisis kinerja Extreme Learning Machine digunakan Confusion Matrix. Metode klasifikasi akan dilakukan evaluasi terutama pada bagian akurasi dari hasil klasifikasi. Akurasi sebuah klasifikasi berpengaruh terhadap performa dari suatu klasifikasi. Untuk melakukan analisa dapat digunakan confusion matrix yaitu sebuah matrik dari prediksi yang akan dibandingkan dengan kelas yang asli dari data inputan. Sebagai contoh, sebuah *test* (i,j) dari sebuah confusion matrix adalah persentase dari waktu dari sebuah *classifier* yang melakukan identifikasi input I sebagai pattern dari kelas j. Setiap kolom dari matriks berkorespondensi kepada classifier output dan setiap baris pada input. Akurasi sebuah klasifikasi dimana $i=j$ menerangkan akurasi dari klasifikasi pada setiap kelas.

Untuk menunjukkan kecepatan proses pembelajaran dan keakuratan Extreme Learning Machine maka kinerja Extreme Learning Machine akan dibandingkan dengan kinerja jaringan syaraf tiruan yaitu Backpropagasi pada data yang sama.

3. HASIL DAN PEMBAHASAN

3.1. Pemahaman Data

Dengan memanfaatkan sumber data mahasiswa dari angkatan 2014 sampai dengan 2016 yang ada. Maka dapat dilakukan pembahasan data penelitian. Komponen *data mining* yang digunakan dalam penelitian ini adalah sebagai berikut:

a. Basis Data

Berhubungan dengan informasi mahasiswa dan akademik.

b. *Extract, Transform, Load*

Data dari basis data akan di ekstrak dan dipindah ke pusat penyimpanan data.

c. *Data Warehouse* Sistem Informasi Akademik

Disini disimpan data yang berasal dari basis data operasional, data ini yang kemudian digunakan untuk proses penggalian data.

d. Penggalian Data (*Data Mining*)

Membangun dua model penggalian data yaitu algoritma ELM dan Backpropagasi.

e. Aplikasi MATLAB.

Keakuratan kedua metode akan diuji menggunakan MATLAB.

3.2. Pengolahan Data Dengan Membangun Struktur

Membangun struktur *data mining* termasuk menentukan variabel masukan dan keluaran yang digunakan pada penggalian data.

Tabel 1. Atribut Penggalian Data

| No | Variabel | Deskripsi |
|----|----------------------------------|---|
| 1 | Nilai Tes Masuk | Nilai yang diambil dari hasil ujian mahasiswa baru yaitu Matematika, Bahasa Inggris, dan Pengetahuan Umum Komputer. Nilai Ujian Masuk didapat dari jumlah seluruh nilai ujian dibagi 3. |
| 2 | Nilai Indeks Prestasi Semester 1 | Nilai Indeks Prestasi yang Didapat di Semester 1 |
| 3 | Jenis Kelamin | Laki-Laki dan Perempuan |
| 4 | Status Bekerja | Terdiri dari Sudah atau Belum Bekerja |
| 5 | Nilai Mutu | Nilai Akhir Mata Kuliah terdiri dari A, B, C, D, E dan Lulus atau Mengulang |

3.3. Permodelan

Setelah struktur data mining selesai dibuat, maka tahap selanjutnya adalah membangun model. Ada dua model yang dibangun dengan algoritma *neural network* yaitu Backpropagasi dan Extreme Learning Machine. Pada tiap tahapan pembuatan model, ditentukan atribut kolom yang menjadi key, input dan output (*predicted variable*). Pada tabel 2 dan tabel 3 dapat dilihat dokumentasi model dari penggalian data.

Tabel 2. Penggalian Data dengan Algoritma ELM

| Variabel | Deskripsi |
|------------|--------------------------|
| Nama Model | Extreme Learning Machine |
| Deskripsi | Jaringan Syaraf Tiruan |
| Algoritma | Extreme Learning Machine |

| Variabel | Deskripsi |
|---|-----------|
| Input | 4 |
| Output | 1 |
| Hidden | 4 |
| Informasi Keluaran untuk "Students Grades" | |
| Tipe Isi | Diskrit |
| Berhasil | 85.16% |
| Gagal | 14.48% |

Tabel 3. Penggalan Data dengan Algoritma Backpropagasi

| Variabel | Deskripsi |
|---|------------------------|
| Nama Model | Backpropagasi |
| Deskripsi | Jaringan Syaraf Tiruan |
| Algoritma | Backpropagasi |
| Input | 4 |
| Output | 1 |
| Hidden | 4 |
| Informasi Keluaran untuk "Students Grades" | |
| Tipe Isi | Diskrit |
| Berhasil | 71.80% |
| Gagal | 28.20% |

Selanjutnya adalah pengujian akurasi model.

3.3.1. Validasi Model

Pada penelitian ini akan menguji algoritma mana yang memiliki kemampuan paling baik dalam mengetahui *grade* prestasi akademik mahasiswa.

3.3.2. Target Populasi dan Model Probabilitas

Pada pengujian ini menggunakan nilai akademik "A" sebagai parameter untuk diujicoba. Semakin tinggi persentase populasi target pada model maka semakin baik model tersebut dalam melakukan prediksi. Dalam menentukan model yang terbaik, dapat dibandingkan dari tingkat akurasi antar model, dilihat dari persentase populasi target pada tabel 4.

Tabel 4. Target Populasi dan Model Probabilitas

| Persentase Populasi = 62.5% | | | |
|-----------------------------|-------|-----------------|---------------------|
| Model | Nilai | Populasi Target | Probilitas Prediksi |
| ELM | 154 | 200 | 71% |
| Backpropagasi | 171 | 200 | 85.50% |

Dari tabel 4 dapat dilihat bahwa dengan target populasi yang sama yaitu 200, ketepatan prediksi yang lebih besar adalah algoritma Extreme Learning Machine sebesar 85.50%

3.3.3. Klasifikasi Matriks

Klasifikasi ini membandingkan nilai data *testing* dengan nilai yang diprediksi. Kemudian dibuat rangkuman untuk menampilkan jumlah atau persentase prediksi yang akurat atau tidak akurat pada setiap model. Klasifikasi Matriks dapat dilihat pada tabel 5.

Tabel 5. Klasifikasi Matriks

| 1. Results as Percentages for Model 'ELM' | | | | | |
|---|----|----|----|---|---|
| Actual | A | B | C | D | E |
| Correct | 70 | 78 | 18 | 1 | 4 |
| Misclassified | 5 | 18 | 6 | 0 | 0 |
| 2. Results as Percentages for Model 'Backpropagasi' | | | | | |
| Actual | A | B | C | D | E |
| Correct | 66 | 62 | 21 | 1 | 4 |
| Misclassified | 5 | 28 | 8 | 6 | 0 |

Pada tahapan pengujian model ini juga terlihat bahwa model yang paling akurat adalah Extreme Learning Machine dengan tingkat kesalahan paling rendah.

3.3.4. Pengujian Akurasi dari Data Perkuliahan Sebelumnya

Pada pengujian ini menggunakan data transaksi perkuliahan mahasiswa yang sudah berlalu untuk memprediksi *grade* setiap matakuliah yang diambil dengan tiga model yang sudah ada. Hasil prediksi *grade* dibandingkan dengan *grade* aktual, karena ini data transaksi perkuliahan yang sudah berlalu maka sudah pasti ada *grade* matakuliah sebelumnya. Perbandingan ini dilakukan dengan cara mencocokkan *grade* hasil prediksi dengan *grade* aktual, kemudian dikalkulasikan jumlah yang benar untuk mendapatkan persentase akurasi dibandingkan dengan jumlah total sampel data yang digunakan.

Jumlah data *testing* pada setiap tiap periode adalah jumlah keseluruhan matakuliah yang diambil oleh mahasiswa pada periode berjalan. Tabel 6 berikut ini merupakan hasil dari perbandingan data prediksi dan aktual, pada setiap model menunjukkan jumlah akurasi (jumlah prediksi benar) dan kemudian dibandingkan dalam bentuk persentase.

Tabel 6. Perbandingan *Grade* Prediksi dan Aktual

| Nama | Akurasi | Nilai A | Nilai B | Nilai C | Nilai D | Nilai E |
|------------------|---------|---------|---------|---------|---------|---------|
| Keberhasilan ELM | 93.33 % | 81.25 % | 75.00 % | 100% | 100% | 93.33 % |
| Kegagalan ELM | 6.67 % | 18.75 % | 25.00 % | 0.00 % | 0.00 % | 6.67 % |
| Keberhasilan ELM | 92.96 % | 68.89 % | 72.41 % | 14.29 % | 100 % | 92.96 % |
| Kegagalan ELM | 7.04 % | 31.11 % | 27.59 % | 85.71 % | 0.00 % | 7.04 % |

Maka pada tahapan pengujian ini, bila membandingkan nilai persentase akurasi dari setiap periode, Extreme Learning Machine mempunyai persentase akurasi lebih besar dibandingkan Backpropagation. Maka, bisa disimpulkan pada tahapan pengujian ini, Extreme Learning Machine adalah model yang lebih akurat dalam memprediksi grade matakuliah mahasiswa.

3.4. Evaluasi

3.4.1. Confusion Matrix

Pada tahap ini dilakukan pengujian terhadap sistem yang telah dibangun. Pengujian sistem dilakukan dengan menghitung tingkat keakuratan sistem dalam mengenali atribut berdasarkan data yang telah di buat. Langkah yang harus ditempuh yakni menggunakan confusion matrix. Untuk melihat arsitektur jaringan yang paling efektif, dapat dilihat dari confusion matrix yang ada pada metode Ekstreme Learning Machine dan metode Backpropagasi. Berikut *confusion matrix* yang ada pada sistem.

a. Variasi 1 Extreme Learning Machine

Hasil pengenalan data:

- a) Data dapat dikenali dengan benar = 171
- b) Data dikenali dengan salah = 29

Tabel 7. Confusion Matrix untuk ELM

| | A | B | C | D | E |
|---|----|----|----|---|---|
| A | 70 | 2 | 3 | 0 | 0 |
| B | 9 | 78 | 6 | 3 | 0 |
| C | 1 | 4 | 18 | 1 | 0 |
| D | 0 | 0 | 0 | 1 | 0 |
| E | 0 | 0 | 0 | 0 | 4 |

Tabel 7 ini memberikan penjelasan tingkat akurasi dari pengenalan data untuk setiap kelas yang disiapkan. Tabel yang berwarna biru menjelaskan jumlah data dikenali dengan benar (*True*). Sedangkan tabel yang berwarna putih menjelaskan jumlah data yang dikenali dengan salah atau citra yang tidak dikenal (*False*).

b. Variasi 2 Backpropagasi

Hasil pengenalan data:

- c) Data dapat dikenali dengan benar = 154
- d) Data dikenali dengan salah = 46

Tabel 8. Confusion Matrix untuk Backpropagasi

| | A | B | C | D | E |
|---|----|----|----|---|---|
| A | 70 | 2 | 3 | 0 | 0 |
| B | 9 | 78 | 6 | 3 | 0 |
| C | 1 | 4 | 18 | 1 | 0 |
| D | 0 | 0 | 0 | 1 | 0 |
| E | 0 | 0 | 0 | 0 | 4 |

Tabel 8 ini memberikan penjelasan tingkat akurasi dari pengenalan data untuk setiap kelas yang disiapkan. Tabel yang berwarna biru menjelaskan jumlah data dikenali dengan benar (*True*). Sedangkan tabel yang berwarna putih menjelaskan jumlah data yang dikenali dengan salah atau citra yang tidak dikenal (*False*).

3.4.2. Penerapan Hasil Penggalan Data

Tahapan ini merupakan penerapan hasil penggalan data yang dikembangkan dengan menggunakan MATLAB untuk dapat memprediksi prestasi akademik mahasiswa sejak awal untuk dapat diakses oleh Perguruan Tinggi.

4. SIMPULAN

Penelitian ini membandingkan dua algoritma yang biasa digunakan untuk penelitian dalam bidang data mining dan jaringan syaraf tiruan yaitu Extreme Learning Machine dan Backpropagasi. Dengan menggunakan data-data akademik mahasiswa baru yang banyak tersedia maka dilakukan perbandingan antar dua metode untuk mengetahui algoritma mana yang paling sesuai dalam mengetahui prestasi akademik mahasiswa. Variabel yang digunakan adalah nilai ujian masuk mahasiswa baru, nilai IP Semester 1, Jenis Kelamin, dan Status Bekerja, sedangkan variabel keluaran adalah nilai mutu sebagai klasifikasi kinerja akademik.

Pengujian akurasi model algoritma menggunakan validasi model, probabilitas prediksi, klasifikasi matriks, dan pengujian akurasi dari data perkuliahan sebelumnya. Sedangkan proses evaluasi menggunakan Confusion Matrix, setelah itu dapat dilakukan penerapan hasil data mining. Hasil penelitian ini menyatakan bahwa algoritma Extreme Learning Machine memiliki 14,84% tingkat kesalahan lebih rendah dibandingkan dengan Backpropagasi 28,20%. Dari tahapan pengujian model didapatkan hasil yang paling akurat adalah algoritma Extreme Learning Machine karena mempunyai tingkat akurasi paling tinggi dan tingkat kesalahan paling rendah.

DAFTAR PUSTAKA

- [1] ASIF, R., MERCERON, A., ALI, S. A., & HAIDER, N. G. 2017. Analyzing undergraduate students' performance using educational data mining. *Computers & Education*, 113, 177-194.
- [2] D. HUANG AND Z. WU, 2017. "Forecasting Outpatient Visits Using Empirical Mode Decomposition Coupled With Backpropagation Artificial Neural Networks Optimized by Particle Swarm Optimization," *PLoS One*, vol. 12, no. 2, pp. 1-18.
- [3] EHRET, D. HOCHSTUHL, D. GIANOLA, AND G. THALLER, 2015. "Application of neural networks with back-propagation to genome-enabled prediction of complex traits in HolsteinFriesian and German Fleckvieh cattle," *Genetics Selection Evolution*, vol. 47, no. 1, p. 22.

- [4] HERNÁNDEZ-LOBATO, José Miguel; ADAMS, 2015. Ryan. Probabilistic backpropagation for scalable learning of bayesian neural networks. In: International Conference on Machine Learning. p. 1861-1869.
- [5] HUANG, Guang-Bin, et al. 2015. Local receptive fields based extreme learning machine. IEEE, 10.2: 18-29.
- [6] J. TARIGAN, NADIA, R. DIEDAN, AND Y. SURYANA, 2017. "Plate Recognition Using Backpropagation Neural Network and Genetic Algorithm," Procedia Computer Science, vol. 116, pp. 365–372.
- [7] P. ANTWI ET AL., 2017. "Estimation of Biogas and Methane Yields in an UASB Treating Potato Starch Processing Wastewater With Backpropagation Artificial Neural Network," Journal Bioresour Technology, vol. 228, pp. 106–115.
- [8] Pangaribuan, J. J. (2016). Mendiagnosis Penyakit Diabetes Melitus Dengan Menggunakan Metode Extreme Learning Machine. Journal Information System Development (ISD), 1(2).
- [9] TANG, Jiexiong; DENG, Chenwei; HUANG, Guang-Bin. 2015. Extreme learning machine for multilayer perceptron. IEEE transactions on neural networks and learning systems, 27.4: 809-821.