



BAB I

PENDAHULUAN

BAB I

PENDAHULUAN

I.1. Latar Belakang

Sebagai penyelenggara proses pendidikan yang tertinggi, perguruan tinggi memiliki peranan yang penting sebagai agen perubahan masyarakat, dalam mengembangkan potensi dan ketrampilan yang dimilikinya, sehingga menghasilkan sumber daya manusia yang bermanfaat bagi bangsa dan negara (Sedyati, 2022). Mahasiswa merupakan salah satu komponen penting di dalam perguruan tinggi, dimana rasio antara jumlah mahasiswa baru yang diterima dengan mahasiswa yang lulus pada satu tahun ajaran haruslah seimbang, sehingga perguruan tinggi harus terus mengevaluasi kinerja akademiknya, sehingga tidak terjadi jumlah mahasiswa yang diterima jauh lebih tinggi daripada mahasiswa yang lulus (Gunawan et al., 2021). STMIK Time, sebagai salah satu perguruan tinggi yang berfokus pada program studi bidang komputer, mengalami masalah dimana rasio ini sering tidak berimbang, dengan ditemukannya mahasiswa yang lulus tidak tepat pada waktunya. Hal ini tentunya berpengaruh pada evaluasi kinerja akademik program studi di STMIK Time, yang pada akhirnya juga mempengaruhi tingkat akreditasi perguruan tinggi tersebut. Untuk memecahkan masalah ini, penulis memberikan alternatif solusi berupa sebuah model yang dapat memprediksi tingkat kelulusan mahasiswa pada program studi yang ditawarkan oleh STMIK Time, dengan menggunakan sistem kecerdasan buatan berbasis *machine learning*.

Di dalam *machine learning*, terdapat pendekatan *supervised learning*, dimana pendekatan ini membutuhkan dataset yang sudah memiliki target atau label untuk menghasilkan keluaran berupa hasil prediksi atau klasifikasi, sebagai output dari solusi pemecahan masalah (Kusuma et al., 2022). Salah satu algoritma *machine learning* yang populer digunakan dalam memecahkan masalah prediksi adalah Backpropagation, yang memproses data melalui *input layer*, *hidden layer* dan *output layer* (Franciska et al., 2023). Algoritma ini memproses data inputan menjadi hasil prediksi melalui tahapan *feed-forward* dan *backward propagation*, sehingga mampu menyesuaikan nilai bobot dan bias berdasarkan kesalahan output prediksinya (Hayadi et al., 2021). Dengan kemampuannya untuk menyesuaikan nilai bobot dan bias melalui pembelajaran yang dilakukan secara berulang-ulang melalui kedua tahapan tersebut, algoritma Backpropagation mampu menghasilkan pola sistem yang tahan terhadap kerusakan serta konsisten bekerja dengan baik (Hartono et al., 2020).

Penulis memilih algoritma ini dalam merancang model prediksi kelulusan mahasiswa STMIK Time, berdasarkan referensi penelitian yang sudah mengimplementasikan algoritma ini ke dalam kasus yang berhubungan dengan prediksi kelulusan mahasiswa, seperti penelitian yang dilakukan oleh (Yalidhan, 2018) dengan hasil akurasi sebesar 98.97 %, penelitian yang dilakukan oleh (Riyanto et al., 2019), dengan hasil akurasi sebesar 84.68%, penelitian yang dilakukan oleh (Rismayati et al., 2022), dengan hasil akurasi sebesar 91.87%. Ketiga penelitian ini penulis gunakan sebagai *benchmark*, dimana hasil akurasi

yang diharapkan dari model yang dihasilkan setidaknya mendekati atau melebihi akurasi dari penelitian-penelitian tersebut.

Algoritma Backpropagation menggunakan *learning rate* sebagai acuan dalam kecepatan pembelajaran, yang sangat penting pengaruhnya dalam mendekati nilai *minimum error* (Suhendra & Saputra, 2020). Penentuan *learning rate* secara otomatis, atau lebih dikenal dengan istilah *learning rate scheduler* merupakan topik penelitian yang menarik perhatian peneliti, khususnya pada bidang jaringan syaraf tiruan, untuk mempelajari bagaimana mengoptimalkan metode pembelajaran algoritma *machine learning* (Donini et al., 2020). Ada beberapa metode *learning rate scheduler* di dalam algoritma Backpropagation, dan pada penelitian ini, penulis memilih untuk membandingkan tiga metode, yaitu *constant learning rate (fixed learning rate)*, *step decay learning rate*, dan *cosine annealing learning rate*.

Constant learning rate merupakan metode dimana nilai *learning rate* tidak berubah selama proses pelatihan berlangsung, dan untuk menentukan nilai terbaik biasanya dilakukan dengan cara *trial and error* (Zhang et al., 2020). Kelebihan metode ini adalah cara kerjanya yang sederhana, bersifat stabil dan konsisten, serta lebih tahan terhadap terjadinya *overfitting* (Wu et al., 2019). Beberapa penelitian sudah menggunakan metode *fixed learning rate* pada penelitian mereka, seperti pada penelitian yang dilakukan oleh (Ridwan et al., 2020), dengan hasil akurasi 98.27%, penelitian yang dilakukan oleh (Yuhelmi et al., 2022) dengan hasil akurasi 89.63%, dan penelitian yang dilakukan oleh (Firzatullah, 2021) dengan hasil akurasi 85%. Ketiga penelitian ini menggunakan nilai *fixed learning rate* dengan

rentang 0.1 sampai 1, sehingga penulis menggunakan nilai ini sebagai *benchmark* dalam penelitian ini.

Step decay merupakan salah satu *learning rate scheduler* yang populer di gunakan di dalam algoritma Backpropagation, karena kemampuannya dalam menggeneralisasi data dan memiliki kualitas konvergen yang sangat baik (X. Wang et al., 2021). Beberapa penelitian sudah menggunakan metode *step decay* pada penelitian mereka, seperti pada penelitian yang dilakukan oleh (Ge et al., 2019), dengan rata-rata akurasi 99.4733%, penelitian yang dilakukan oleh (Martínez-Fernandez et al., 2023) dengan hasil akurasi 93%, dan penelitian yang dilakukan oleh (Nakamura et al., 2021) dengan rata-rata akurasi 98.585%. Ketiga penelitian menggunakan *decay rate* sebesar 0.1, sehingga nilai ini penulis gunakan sebagai *benchmark* di dalam penelitian ini.

Cosine annealing merupakan salah satu *learning rate scheduler* yang populer di gunakan di dalam algoritma Backpropagation, karena kemampuannya dalam menggeneralisasi data dan memiliki kualitas konvergen yang sangat baik (X. Wang et al., 2021). Beberapa penelitian sudah menggunakan metode *cosine annealing* pada penelitian mereka, seperti pada penelitian yang dilakukan oleh (Li et al., 2021), dengan hasil akurasi tertinggi sebesar 92.85%, penelitian yang dilakukan oleh (Pradiptatmaka et al., 2021) dengan hasil akurasi 82.8%, dan penelitian yang dilakukan oleh (Z. Wang et al., 2022) dengan hasil akurasi 92.6%. Ketiga penelitian ini menggunakan nilai *max* dan *min learning rate* yang berada pada rentang nilai 1 sampai 0.1, sehingga penulis memilih nilai ini sebagai *benchmark* di dalam penelitian ini.

Berdasarkan penjelasan dan referensi dari ketiga metode *learning rate scheduler* di atas, penulis tertarik untuk membandingkan ketiga metode tersebut ke dalam model prediksi tingkat kelulusan mahasiswa STMIK Time, dan menganalisis nilai *learning rate* paling optimal menurut masing-masing metode berdasarkan nilai performa akurasi, presisi dan *recall* yang dihasilkan masing-masing model.

Berdasarkan latar belakang tersebut, penelitian ini mencoba melakukan kombinasi variasi nilai *learning rate* dan jumlah *neuron* pada algoritma Backpropagation, dimana untuk masing-masing model yang dihasilkan diukur performanya dan dibandingkan sehingga diperoleh model terbaik dalam memecahkan masalah prediksi tingkat kelulusan mahasiswa STMIK Time, dan menuangkan hasil penelitian tersebut dalam bentuk sebuah tesis yang “**Analisis Pengaruh Learning Rate Pada Model Backpropagation Dalam Prediksi Kelulusan Mahasiswa Stmik Time**”.

I.2. Rumusan Masalah

Berdasarkan latar belakang masalah sebelumnya, ada beberapa rumusan masalah yang menjadi fokus dalam penelitian ini, yaitu:

1. Bagaimana variasi metode *learning rate* mempengaruhi pemilihan nilai *learning rate* paling optimal pada algoritma Backpropagation dalam memprediksi tingkat kelulusan mahasiswa STMIK Time?
2. Bagaimana pengaruh pemilihan metode *learning rate scheduler* terhadap kecepatan proses *training* maupun *testing* pada model prediksi tingkat

kelulusan mahasiswa STMIK Time yang menggunakan algoritma Backpropagation?

3. Bagaimana pengaruh nilai *learning rate* terhadap performa model yang diukur melalui nilai akurasi, presisi dan *recall* pada masalah prediksi tingkat kelulusan mahasiswa STMIK Time?
4. Apakah model yang dihasilkan dari penelitian ini sudah dapat membantu STMIK Time dalam memecahkan masalah prediksi tingkat kelulusan mahasiswa?

I.3. Tujuan

Berdasarkan pertanyaan penelitian di atas, tujuan yang ingin dicapai pada penelitian yang akan dilakukan ini adalah sebagai berikut:

1. Menganalisis perbandingan nilai *learning rate* yang dihasilkan oleh metode *constant learning rate*, *step decay*, dan *cosine annealing* pada model prediksi tingkat kelulusan mahasiswa STMIK Time yang menggunakan algoritma Backpropagation.
2. Menganalisis performa model, khususnya pada *training time* dan *testing time*, setelah menggunakan variasi ketiga metode *learning rate scheduler* yang berbeda, dalam masalah prediksi tingkat kelulusan mahasiswa STMIK Time.
3. Menghasilkan sebuah model terbaik yang menggunakan nilai *learning rate* terbaik dari hasil analisis perbandingan ketiga metode *learning rate scheduler* pada algoritma Backpropagation, yang menghasilkan output berupa prediksi tingkat kelulusan mahasiswa STMIK Time.

4. Memberikan alternatif solusi berupa sebuah model prediksi, yang dapat membantu pihak STMIK Time dalam memprediksi mahasiswa yang lulus tidak tepat pada waktunya.

I.4. Manfaat

Manfaat yang diharapkan dari hasil penelitian yang akan dilakukan adalah sebagai berikut:

1. Memberikan gambaran mengenai perbandingan nilai *learning rate* yang dihasilkan oleh ketiga metode *learning rate scheduler* yang diimplementasikan pada algoritma Backpropagation dalam kasus prediksi tingkat kelulusan mahasiswa STMIK Time.
2. Memberikan gambaran berupa pengaruh pemilihan metode *learning rate scheduler*, khususnya pada *training time* dan *testing time* pada algoritma Backpropagation yang diimplementasikan ke dalam model prediksi tingkat kelulusan mahasiswa STMIK Time.
3. Sebagai bahan referensi bagi peneliti lain mengenai hasil nilai *learning rate* yang dihasilkan oleh masing-masing *learning rate scheduler* dan pengaruhnya terhadap performa model prediksi tingkat kelulusan mahasiswa STMIK Time.
4. Membantu pihak STMIK Time dalam memecahkan masalah prediksi mahasiswa yang lulus tidak tepat pada waktunya dalam bentuk sebuah model yang dapat diimplementasikan menjadi sebuah aplikasi.

I.5. Batasan Masalah

Agar penelitian ini tidak melenceng dari pokok permasalahan, penulis memberikan beberapa batasan masalah sebagai berikut:

1. Dataset yang digunakan adalah data primer yang diperoleh langsung dari STMIK Time, berupa data nilai mahasiswa dan jumlah ujian perbaikan yang diambil oleh masing-masing mahasiswa, pada periode tahun ajaran 2019.
2. Pada penelitian ini dibandingkan tiga metode *learning rate scheduling*, yaitu *Constant Learning Rate*, *Step Decay Learning Rate*, dan *Cosine Annealing*.
3. Evaluasi model dilakukan dengan menggunakan tabel *confusion matrix*, yaitu dengan membandingkan nilai *true positive*, *true negative*, *false positive*, dan *false negative* yang dihasilkan masing-masing model.
4. Penentuan model terbaik diukur dengan menggunakan nilai akurasi, presisi, dan *recall* dari hasil evaluasi masing-masing model.

I.6. Sistematika Penulisan

Sistematika proposal tesis yang akan dilakukan dalam penelitian ini terdiri dari 3 bab, yaitu:

BAB I PENDAHULUAN

Pendahuluan akan membahas tentang latar belakang dari masalah yang akan di angkat beserta rumusan masalah, tujuan penelitian, dan batasan masalah serta sistematika penulisan yang akan diuraikan dalam penyusunan laporan penelitian ini.

BAB II TINJAUAN PUSTAKA

Tinjauan pustaka menjelaskan tentang teori-teori yang akan diuraikan mengenai Indeks Prestasi Mahasiswa, Machine Learning, Backpropagation, dan perangkat analisis sistem beserta metode analisis yang digunakan.

BAB III METODOLOGI PENELITIAN

Metodologi penelitian menjelaskan rangkaian tahapan dalam pembuatan sistem yang terdiri dari pengumpulan data, analisis data, perancangan sampai implementasi beserta analisis yang akan dilakukan.

BAB IV HASIL

Hasil membahas mengenai hasil prediksi tingkat kelulusan mahasiswa dari masing-masing model, perbandingan performa masing-masing model berdasarkan nilai akurasi, presisi dan *recall*, serta nilai *learning rate* yang dihasilkan oleh masing-masing metode *learning rate scheduler*.

BAB V PEMBAHASAN

Pembahasan berisi analisis yang dilakukan terhadap hasil penelitian sebelumnya, berupa model dengan performa terbaik dan terburuk, serta pengaruh pemilihan metode *learning rate scheduler* terhadap kecepatan *training* dan *testing* masing-masing model.

BAB VI KESIMPULAN DAN SARAN

Kesimpulan membahas intisari dari temuan yang diperoleh dari hasil penelitian, sedangkan saran berisi saran-saran yang penulis berikan untuk pengembangan penelitian berikutnya yang berhubungan dengan topik penelitian ini.