

BAB I

PENDAHULUAN

I.1. Latar Belakang

Pendidikan merupakan suatu hak yang wajib dimiliki oleh peserta didik, dan memiliki tujuan untuk membangun karakter serta pengetahuan bagi pesertanya (Putri & Adirakasiwi, 2021). Sesuai Edaran Mendikbud nomor 4 tahun 2020 (Mendikbud RI, 2020), pembelajaran jarak jauh harus memberikan pengalaman belajar yang bermakna, tanpa terbebani tuntutan capaian kurikulum untuk kenaikan kelas maupun kelulusan, dan difokuskan pada pendidikan kecakapan hidup antara lain mengenai pandemi covid-19 (Jusuf et al., 2020). Sejak COVID-19 masuk di Indonesia dan membawa dampak bagi kehidupan pada segi ekonomi, sosial maupun pendidikan, pemerintah provinsi mengeluarkan kebijakan cara belajar secara daring, dimana pelaksanaan aktivitas belajar menggunakan sistem pembelajaran *online* yang menyebabkan guru dan siswa tidak dapat berinteraksi langsung secara fisik (Riyani et al., 2021). sistem pembelajaran daring yang dianjurkan oleh pemerintah, dimana di dalam prosesnya menyebabkan terjadinya perubahan besar, seperti guru harus bekerja extra untuk memberikan materi bagi para siswa, dan menurunnya minat belajar siswa karena sistem menyebabkan siswa cenderung kurang memperhatikan pembelajaran (Elva et al., 2022). Sekolah menengah Kejuruan Tunas Pelita Binjai merupakan sekolah yang juga ikut menjalankan sistem pembelajaran daring ini, sehingga tidak luput dari masalah yang sama. Oleh karena itu, dibutuhkan sebuah sistem yang dapat memberikan

informasi mengenai tingkat minat belajar siswa pada sistem pembelajaran daring, yang dapat secara otomatis memberikan *output* berdasarkan data proses pembelajaran sebelumnya. Penelitian ini mencoba memecahkan masalah tersebut, dengan cara menambang data proses pembelajaran daring yang dilakukan SMK Tunas Pelita Binjai dalam dua semester terakhir, dan memberikan alternatif solusi berupa sebuah model yang efektif dalam mengklasifikasikan tingkat minat belajar siswa pada sistem pembelajaran daring.

Untuk mengumpulkan data dalam jumlah besar, *data mining* sangat berguna dalam hal mengekstraksi informasi di dalam data tersebut sehingga menghasilkan informasi pola data yang dapat dimanfaatkan dalam proses klasifikasi (Alkhairi & Situmorang, 2022). Pemanfaatan *data mining* dalam klasifikasi minat belajar siswa sudah dilakukan oleh beberapa peneliti, seperti penelitian yang dilakukan oleh (Madya & Abdurahman, 2021) yang menggunakan nilai ujian *online* sebagai acuan, penelitian yang dilakukan oleh (R.H. Zer et al., 2022) yang menggunakan cara dosen mengajar sebagai acuan, penelitian yang dilakukan oleh (Setiana et al., 2021) yang menggunakan tingkat kehadiran siswa sebagai acuan, dan penelitian yang dilakukan oleh (Abdelkader et al., 2022) yang menggunakan kesempatan bertanya pada saat pembelajaran sebagai salah satu acuan. Berdasarkan referensi penelitian-penelitian ini, pada penelitian ini digunakan kategori kehadiran, aktivitas siswa bertanya dan evaluasi nilai ujian sebagai acuan pada proses klasifikasi.

Linear regression adalah salah satu algoritma *data mining* yang menggunakan persamaan linier sederhana dalam memodelkan hubungan antara variabel independen (fitur) dan dependen (target), melalui hasil perhitungan

estimasi kuadrat terkecil (Jasni et al., 2022). Algoritma ini memiliki kelebihan dalam hal algoritmanya yang sederhana dan mudah untuk dipahami (Yulianto et al., 2020), memiliki koefisien regresi yang berfungsi untuk memudahkan interpretasi hubungan antara fitur dan target (Muhammad Ahmad Iqbal, 2021), dan merupakan algoritma yang memiliki waktu kompilasi yang cepat dan efisien untuk kasus data yang tidak memiliki banyak fitur (Jiao et al., 2020). Beberapa penelitian telah mengimplementasikan algoritma ini ke dalam model *data mining*, seperti pada analisis minat belajar siswa pada metode *blended learning* dengan hasil nilai RMSE sebesar 0.00 (Ariyanto et al., 2020), prediksi performa belajar siswa berdasarkan tingkat motivasi dengan hasil nilai RMSE sebesar 0.5 (Bramhe et al., 2023), prediksi performa akademik siswa berdasarkan perilaku belajarnya dengan hasil nilai RMSE sebesar 0.48 (Dabhade et al., 2021), penerapan *linear regression* pada *learning analytic* untuk memprediksi performa mahasiswa jurusan *computer science* dengan rata-rata nilai RMSE sebesar 0.1758 (Alamgir et al., 2023), dan prediksi minat siswa terhadap sistem pembelajaran *online* dengan hasil nilai RMSE sebesar 0.883 (Ho et al., 2021). Dari penelitian-penelitian ini, terlihat bahwa *linear regression* merupakan salah satu alternatif algoritma yang cocok digunakan dalam memprediksi minat belajar siswa, dan untuk evaluasi performanya dapat digunakan nilai RMSE.

Selain *linear regression*, algoritma *data mining* yang sering digunakan dalam penelitian masalah regresi adalah *support vector machine* (SVM). Algoritma ini memiliki beberapa kelebihan seperti kemampuannya dalam mengolah data non-linear (Ghosh et al., 2019), memiliki tingkat ketahanan yang cukup baik terhadap

data *outlier* (Al-Mejibli et al., 2020), memiliki parameter yang dapat digunakan untuk mengontrol terjadinya *overfitting* (Pimentel et al., 2021), serta dapat digunakan untuk memecahkan masalah prediksi untuk model regresi (Agustina et al., 2018). Beberapa penelitian telah menerapkan algoritma ini dalam penelitian seputar masalah prediksi model regresi, seperti pada prediksi performa siswa dalam pembelajaran *online* dengan hasil nilai RMSE sebesar 0.2557 (Samsudin et al., 2022), prediksi kecelakaan lalu lintas dengan hasil nilai RMSE sebesar 0.001 (Gatera et al., 2023), prediksi umlah partikel halus di udara dengan hasil RMSE sebesar 23.94 (Letchumanan & B, 2022), prediksi kualitas air sungai dengan hasil RMSE sebesar 0.1292 (Najwa Mohd Rizal et al., 2022), dan prediksi laju inflasi di Indonesia dengan hasil RMSE sebesar 0.002 (Rais, 2022). Dari referensi penelitian-penelitian ini, terlihat bahwa SVM merupakan salah satu alternatif algoritma yang cocok digunakan dalam memecahkan masalah prediksi regresi, sehingga penulis pilih untuk dibandingkan dengan algoritma *linear regression* sebelumnya.

Topik dalam penelitian ini adalah menganalisa performa model yang menggunakan algoritma *linear regression* dan SVM dalam memprediksi minat belajar siswa SMK Tunas Pelita Binjai terhadap media pembelajaran daring. Penelitian ini membandingkan model tersebut dengan menggunakan nilai MSE, RMSE, MAE, dan R^2 , untuk menentukan alternatif model terbaik. Dalam penelitian ini dibangun dua model yang mengimplementasikan algoritma *linear regression*, yaitu model yang menggunakan regulasi *lasso* dan *ridge*, serta dua model yang mengimplementasikan algoritma SVM, yaitu model yang menggunakan fungsi kernel Linear dan Polynomial. Alternatif model yang dihasilkan pada penelitian ini

diharapkan dapat membantu pihak SMK Tunas Pelita Binjai dalam membangun aplikasi yang dapat memprediksi efektivitas pembelajaran daring yang dilakukan serta mengevaluasinya.

Berdasarkan deskripsi latar belakang yang telah diuraikan, penulis menuangkan hasil penelitian tersebut ke dalam sebuah tesis yang berjudul “**Analisa Perbandingan Regresi *Linear* dan *Support Vector Machine* pada Minat Belajar Siswa**”

I.2. Rumusan Masalah

Berdasarkan uraian latar belakang permasalahan, rumusan masalah pada penelitian ini adalah Bagaimana perbandingan akurasi dari metode Regresi *Linear* dan *Support Vector Machine* pada minat belajar siswa.

I.3. Tujuan

Adapun tujuan penelitian ini adalah memperoleh hasil akurasi dengan metode Regresi *Linear* dan *Support Vector Machine* pada minat belajar siswa.

I.4. Manfaat

Manfaat yang diharapkan dari hasil penelitian ini adalah sebagai berikut:

1. Dapat mengetahui lebih dalam tentang tahapan analisis pada data kehadiran dan data penilaian siswa.

2. Dapat mengetahui penilaian akurasi algoritma *Support Vector Machine* dan Regresi *Linear* pada proses analisis data kehadiran dan data penilaian siswa untuk dijadikan data minat belajar siswa.
3. Hasil penelitian dapat menjadi rujukan para pembaca dalam memahami algoritma *Support Vector Machine* dan algoritma Regresi *Linear* terutama dalam bagian akurasi.
4. Hasil Analisis dari Algoritma *Support Vector Machine* dan algoritma Regresi *Linear* dapat membantu sekolah untuk melihat hasil rating yang di analisis berdasarkan data grafik minat belajar siswa sehingga proses perbaikan model belajar dapat dilakukan untuk masa yang datang.

I.5. Batasan Masalah

Dalam menganalisis data minat belajar siswa menggunakan algoritma *Support Vector Machine* dan algoritma Regresi *Linear*, juga visualisasi dengan menggunakan Bahasa program *Python* dilakukan beberapa batasan sebagai berikut:

1. Data set yang digunakan adalah data belajar siswa setiap semester pada saat masa pandemi dimulai hingga semester saat ini, yang diperoleh dari Sekolah Menengah Kejuruan Tunas Pelita Binjai
2. Model yang mengimplementasikan algoritma *linear regression* dibangun dengan variasi regulasi *lasso* dan *ridge*.
3. Model yang mengimplementasikan algoritma SVM dibangun dengan variasi fungsi kernel *linear* dan *polynomial*.

4. Output dari penelitian adalah model terbaik untuk memecahkan permasalahan penelitian, yang dievaluasi berdasarkan nilai MSE, RMSE, MAE, dan R^2 .

I.6. Sistematika Penulisan

Sistematika proposal tesis yang akan dilakukan dalam penelitian ini terdiri dari 6 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 minat belajar siswa, beserta perangkat analisa sistem beserta metode analisa 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 analisa yang akan dilakukan.

BAB IV HASIL

Hasil penelitian yaitu berupa pemaparan hasil pelaksanaan metode/teknik penelitian dan menyajikan data yang mendukung hasil

tersebut. Penyajian data dan penjelasannya dilakukan secara berurutan dan lengkap menggunakan teks dan ilustrasi lainnya (misalnya, tabel dan gambar) sebagai pendukung hasil.

BAB V PEMBAHASAN

Pembahasan penelitian yaitu menerjemahkan makna dari hasil yang diperoleh untuk menjawab pertanyaan atau masalah penelitian. Fungsi lainnya adalah untuk menjelaskan pemahaman baru yang didapatkan dari hasil penelitian, yang diharapkan berguna dalam pengembangan keilmuan. Dalam penelitian tingkat lanjut, hasil pembahasan diharapkan menunjukkan kontribusi penulis terhadap pengembangan keilmuan. Dalam penelitian ini, hasil pembahasan bukan pengembangan keilmuan secara substansial, tetapi cukup terhadap pemahaman personal dalam implementasi konsep atau teori.

BAB VI KESIMPULAN DAN SARAN

Kesimpulan dan saran berisi hasil penelitian data minat belajar siswa dengan menggunakan algoritma Support Vector Machine dan algoritma Regresi Linear, Visualisasi dan simulasi interaktif dengan menggunakan python. Bagian kesimpulan berisi jawaban atas rumusan masalah pada penelitian dan pada bagian saran berupa pengembangan penelitian dan kendala yang dihadapi dalam menyusun penelitian sehingga dapat memberikan masukan dan pengetahuan baru yang bermanfaat bagi penelitian selanjutnya.

