

BAB II

TINJAUAN PUSTAKA

II.1. Minat Belajar Siswa

Kata minat secara etimologi berasal dari bahasa Inggris yaitu “*interest*” yang berarti kesukaan, perhatian (kecenderungan hati pada sesuatu), keinginan. Jadi dalam proses belajar siswa harus mempunyai minat atau kesukaan untuk mengikuti kegiatan belajar yang berlangsung, karena dengan adanya minat mendorong siswa untuk menunjukkan perhatian, aktivitasnya dan partisipasinya dalam mengikuti belajar yang berlangsung. Minat adalah suatu rasa lebih suka dan rasa keterikatan pada suatu hal atau aktivitas, tanpa ada yang menyuruh (Tanjung et al., 2021).

Sedangkan belajar merupakan suatu proses perubahan dalam tingkah laku menuju perubahan tingkah laku yang baik, dimana perubahan tersebut terjadi melalui latihan atau pengalaman (Riyani et al., 2021), berdasarkan dari beberapa pengertian di atas, dapat disimpulkan pengertian dari minat belajar adalah dorongan yang muncul dari dalam diri seseorang untuk mempelajari sesuatu hal dengan kemauannya sendiri tanpa ada yang menyuruh untuk mendapatkan kepuasan tersendiri. Minat belajar dapat dipengaruhi oleh beberapa faktor yaitu:

1. Faktor internal:

- a. Rasa senang terhadap mata pelajaran,
- b. Rasa ketertarikan siswa terhadap mata pelajaran,
- c. Rasa siswa yang tinggi terhadap mata pelajaran.

2. Faktor eksternal:

- a. Keluarga, pendampingan orang tua pada saat siswa belajar di rumah,

- b. Lingkungan masyarakat dan teman bergaul,
- c. Sarana dan prasarana, fasilitas yang tersedia dilingkungan sekolah maupun di rumah sangat mendukung minat belajar siswa

II.2. Pembelajaran Daring

Pembelajaran daring merupakan kepanjangan dari pembelajaran dalam jaringan (daring) dengan pola pembelajarannya melalui bantuan jaringan internet sehingga akan terjadi interaksi kegiatan belajar mengajar antara siswa dan guru, Menurut (Friantini & Winata, 2021) “Pembelajaran daring merupakan sebuah inovasi pendidikan yang melibatkan unsur teknologi informasi dalam pembelajaran. Adapun kelebihan proses belajar tidak lagi dibatasi oleh ruang dan waktu”.

Pembelajaran daring adalah pembelajaran yang dilakukan menggunakan internet sebagai tempat menyalurkan ilmu pengetahuan. Pembelajaran daring adalah sistem belajar yang terbuka dan tersebar dengan menggunakan perangkat pedagogi (alat bantu pendidikan), yang dimungkinkan melalui internet dan teknologi berbasis jaringan untuk memfasilitasi pembentukan proses belajar dan pengetahuan melalui aksi dan interaksi yang berarti. Sarana yang dapat digunakan sebagai media pembelajaran online antara lain schoology, google classroom, google form, whatsapp grup dan media lainnya. Adapun kelebihan penggunaan media pembelajaran online adalah pembelajaran bersifat mandiri dan interaktivitas yang tinggi, mampu meningkatkan tingkat ingatan, memberikan lebih banyak pengalaman belajar, dengan teks, audio, video dan animasi yang semuanya digunakan untuk menyampaikan informasi, dan juga memberikan kemudahan menyampaikan, meng-update isi, mengunduh, para siswa juga bisa mengirim email kepada siswa lain, mengirim komentar pada forum diskusi, memakai ruang chat, hingga link video conference untuk berkomunikasi langsung.

II.3. Support Vector Machine (SVM)

Support Vector Machines (SVM) yaitu sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi – fungsi linier dalam sebuah fitur yang berdimensi tinggi dan dilatih dengan menggunakan algoritma pembelajaran yang didasarkan pada teori optimasi. SVM pertama kali diperkenalkan pada tahun 1992 oleh Vapnik sebagai rangkaian dari beberapa konsep–konsep unggulan dalam bidang *pattern recognition*.

Konsep SVM mencari hyperplane terbaik yang berfungsi sebagai pemisah dua buah kelas pada *input space*. *Pattern* yang merupakan anggota dari dua buah kelas: +1 dan -1 dan berbagi alternative garis pemisah (*discrimination boundaries*). *Margin* adalah jarak antara hyperplane tersebut dengan pattern terdekat dari masing-masing kelas. *Pattern* yang paling dekat ini disebut sebagai *support vector*. Usaha untuk mencari lokasi hyperplane ini merupakan inti dari proses pembelajaran pada SVM. Secara intuitif, model SVM merupakan representasi dari data sebagai titik dalam ruang, dipetakan sehingga kategori contoh terpisah dibagi oleh celah jelas yang selebar mungkin. Data baru kemudian dipetakan ke dalam ruang yang sama dan diperkirakan termasuk kategori berdasarkan sisi mana dari celah data tersebut berada. (Adijaya, 2020)

SVM bekerja atas prinsip *Structural Risk Minimization (SRM)* dengan tujuan menemukan *hyperlane* terbaik yang memisahkan dua buah kelas pada input *space*. Prinsip dasar SVM adalah klasifikasikan linear. Pada SVM pemisahan yang baik dicapai oleh *hyperplane* yang memiliki jarak terbesar ke titik data *training* terdekat dari setiap kelas (margin fungsional disebut), karena pada umumnya semakin besar margin semakin rendah *error* generalisasi dari pemilah. Misalkan terdapat *m data training* $\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ adalah sampel data dan $y_1 \in \{1, -1\}$ adalah target atau kelas dari sampel data. Misalkan juga bahwa data untuk kedua kelas terpisah maka ingin dicari fungsi pemisah (*hyperplane*)

$$f(x) = xw + b \quad (II.2)$$

di mana w adalah parameter bobot dan b adalah parameter bias.

Pattern $\vec{x}_i$ yang termasuk kelas +1 (sampel positif) dapat dirumuskan sebagai pattern yang memenuhi pertidaksamaan

$$x_i w + b > 0 \text{ untuk } y_i = 1 \quad (\text{II.3})$$

Pattern $\vec{x}_i$ yang termasuk kelas -1 (sampel negatif) dapat dirumuskan sebagai *pattern* yang memenuhi pertidaksamaan

$$x_i w + b < 0 \text{ untuk } y_i = -1 \quad (\text{II.4})$$

Masalah mencari parameter w dan b yang optimal agar diperoleh *hyperplane* yang optimal merupakan *quadratic programming*.

$$\min_{w,b} \frac{1}{2} w^T w + C \sum_{i=1}^m \mu_i \quad (\text{II.5})$$

Masalah ini dapat dipecahkan dengan berbagai teknik komputasi, diantaranya dengan *Lagrange Multiplier*.

$$\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j \vec{x}_i \vec{x}_j \quad (\text{II.6})$$

α_i adalah *Lagrange multipliers*, yang bernilai nol atau positif ($\alpha_i \geq 0$). Dari hasil perhitungan ini diperoleh α_i yang kebanyakan bernilai positif. Data yang berkorelasi dengan α_i yang positif inilah yang disebut sebagai *Support Vector Machine*.

SVM pada prinsipnya adalah klasifikasi linear. Tetapi SVM memiliki keunggulan dalam klasifikasi untuk problem non-linier. Pada prosesnya data diproyeksikan ke ruang vektor baru, sering disebut *feature space*, berdimensi lebih tinggi sedemikian hingga data itu dapat terpisah secara linier. Selanjutnya, di ruang baru, SVM mencari *hyperplane* optimal yaitu bekerja sebagai klasifier linear.

Dapat diasumsikan bahwa kedua belah kelas dapat terpisah secara sempurna oleh *hyperplane* (*linear separable*). Akan tetapi, pada umumnya dua belah kelas pada *input space*

tidak dapat terpisah secara sempurna (*non linear separable*). Untuk mengatasi masalah ini, SVM dirumuskan ulang dengan memperkenalkan metode *Margin Soft*. Cortes et al. (1995) menyarankan ide margin maksimal dimodifikasi yang memungkinkan untuk contoh *mislabeled*. Jika terdapat hyperplane yang dapat memecah contoh "ya" dan "tidak", metode *Margin Soft* akan memilih hyperplane yang membagi contoh-contoh sebersih mungkin. Metode ini memperkenalkan variabel slack ξ_i ($\xi_i \geq 0$), yang mengukur tingkat kesalahan klasifikasi.

$$y_i(\vec{w} \cdot \vec{x}_i + b) \geq 1 - \xi_i \quad (i = 1, 2, \dots, l) \quad (\text{II.7})$$

Sedangkan objective function (2.5) yang dioptimalkan menjadi.

$$\text{minimize } \frac{1}{2} \|\vec{w}\|^2 + C \sum_{i=1}^l \xi_i \quad (\text{II.8})$$

C merupakan parameter yang mengontrol *trade-off* antara margin dan error klasifikasi ξ . Semakin besar nilai C , berarti nilai akhir terhadap kesalahan menjadi semakin besar, sehingga proses training menjadi lebih ketat. Parameter C ditentukan dengan mencoba beberapa nilai dan dievaluasi efeknya terhadap akurasi yang dicapai oleh SVM. (Norwawi, 2020)

II.4. Regresi Linier

Model regresi digunakan untuk membuat taksiran mengenai variabel dependen disebut persamaan regresi estimasi, yaitu suatu formula matematis yang menunjukkan hubungan keterkaitan antara satu atau beberapa variabel yang nilainya sudah diketahui dengan suatu variabel yang nilainya belum diketahui. (Ningsih & Dukalang, 2019). Metode analisis regresi sederhana adalah metode statistik yang berfungsi untuk menguji sejauh mana hubungan sebab akibat antara variabel kualitas Pelayanan (X) terhadap variabel Kepuasan Pelanggan (Y) (Reza Nurul Ichsan, Ahmad Karim, 2021).

Tujuan dari metode ini adalah untuk memprediksi nilai Y untuk nilai X yang diberikan. Model regresi linier sederhana adalah model regresi yang paling sederhana yang hanya memiliki satu variabel bebas X. Analisis regresi memiliki beberapa kegunaan, salah satunya untuk melakukan prediksi terhadap variabel terikat Y (Novianty et al., 2021)

Analisis yang memiliki variabel bebas lebih dari satu disebut analisis regresi berganda. Analisis regresi linier berganda memberikan kemudahan bagi pengguna untuk memasukkan lebih dari satu variabel bebas hingga k dimana banyaknya k kurang dari jumlah observasi (n). Sehingga model regresi linier berganda untuk populasi dapat ditunjukkan sebagai berikut:

$$Y = a + b_x \quad (II.9)$$

Dimana:

Y = Variabel Response atau Variabel Akibat (Dependent)

X = Variabel Predictor atau Variabel Faktor Penyebab (Independent)

a = konstanta

b = koefisien regresi (kemiringan); besaran Response yang ditimbulkan oleh Predictor.

II.5. Penelitian Terkait

Penelitian yang dilakukan oleh (Madya & Abdurahman, 2021) tentang analisis minat belajar siswa terhadap sistem pembelajaran *blended learning* menyatakan bahwa nilai ujian dapat digunakan sebagai salah satu acuan dalam menentukan seberapa tinggi minat belajar siswa terhadap sistem pembelajaran tersebut.

Penelitian yang dilakukan oleh (R.H. Zer et al., 2022) tentang analisis tingkat kepuasan mahasiswa dalam sebuah sistem pembelajaran menyatakan bahwa cara dosen mengajar merupakan faktor yang penting dalam penilaian tingkat kepuasan tersebut.

Penelitian yang dilakukan oleh (Setiana et al., 2021) tentang tingkat ketertarikan siswa terhadap media pembelajaran *online* pada masa pandemi Covid-19 menyatakan bahwa tingkat kehadiran siswa, dalam hal ini seberapa aktif siswa tersebut mengikuti proses pembelajaran secara daring, merupakan faktor penting dalam penilaian tingkat ketertarikan tersebut.

Penelitian yang dilakukan oleh (Abdelkader et al., 2022) tentang tingkat kepuasan siswa terhadap sistem pembelajaran secara *online* menyatakan bahwa kesempatan bertanya yang diberikan pada saat proses pembelajaran merupakan salah satu faktor penting dalam menilai tingkat kepuasan tersebut.

Penelitian yang dilakukan oleh (Ariyanto et al., 2020) membandingkan algoritma *linear regression* dengan *neural network* pada prediksi minat belajar siswa terhadap sistem pembelajaran *blended learning*, dengan hasil bahwa *linear regression* menghasilkan model dengan performa terbaik dengan capaian nilai RMSE sebesar 0.00.

Penelitian yang dilakukan oleh (Bramhe et al., 2023) tentang prediksi performa belajar siswa berdasarkan hasil survei opini menyatakan bahwa algoritma *linear regression* mampu menghasilkan model dengan performa yang cukup baik dengan nilai RMSE 0.5.

Penelitian yang dilakukan oleh (Dabhade et al., 2021) tentang prediksi performa belajar siswa berdasarkan perilaku belajarnya menyatakan bahwa algoritma *linear regression* menghasilkan model dengan performa prediksi yang baik dengan nilai RMSE sebesar 0.48.

Penelitian yang dilakukan oleh (Alamgir et al., 2023) tentang prediksi performa mahasiswa jurusan *Computer Science* berdasarkan metode *Educational Data Mining* menyatakan bahwa algoritma *linear regression* menghasilkan tingkat prediksi yang tinggi, dengan nilai rata-rata RMSE sebesar 0.1758.

Penelitian yang dilakukan oleh (Ho et al., 2021) tentang prediksi tingkat kepuasan siswa terhadap media pembelajaran *online* pada masa pandemi Covid-19 menyatakan

bahwa algoritma *linear regression* mampu menghasilkan model prediksi dengan tingkat kesalahan yang rendah, dengan capaian nilai RMSE 0.883.

Penelitian yang dilakukan oleh (Samsudin et al., 2022) tentang prediksi performa belajar siswa pada sistem pembelajaran *online* menyatakan bahwa algoritma SVM menghasilkan model dengan performa prediksi yang baik dengan nilai RMSE sebesar 0.2557.

Penelitian yang dilakukan oleh (Gatera et al., 2023) tentang prediksi kecelakaan lalu-lintas, membandingkan algoritma SVM dengan *Random Forest*, dengan hasil bahwa algoritma SVM lebih unggul dalam prediksi data latih, yang mencapai nilai RMSE hingga 0.001.

Penelitian yang dilakukan oleh (Letchumanan & B, 2022) tentang prediksi jumlah partikel halus PM2.5 di udara, membandingkan algoritma *linear regression* dan SVM, dengan hasil bahwa algoritma SVM mencapai nilai RMSE sebesar 23.94, lebih rendah dibandingkan algoritma *linear regression* yang mencapai nilai 23.42.

Penelitian yang dilakukan oleh (Najwa Mohd Rizal et al., 2022) tentang prediksi tingkat kualitas air sungai membandingkan algoritma SVM dengan *neural network*, dengan hasil bahwa algoritma SVM mencapai nilai RMSE sebesar 0.1292, di bawah pencapaian *neural network* dengan nilai RMSE 0.0104.

Penelitian yang dilakukan oleh (Rais, 2022) tentang prediksi laju inflasi di Indonesia menyatakan bahwa dengan kombinasi metode *Grid Search Optimization*, algoritma SVM menghasilkan model dengan performa yang sangat baik, terlihat dari nilai RMSE yang dihasilkannya, yaitu 0.002.

Tabel II. 1 Penelitian Terkait

No	Judul	Penulis	Metode	Hasil
1.	Online Learning Implementation in the Covid-19 Pandemic	(Madya & Abdurahman , 2021)	Analisis media pembelajaran <i>online</i> seperti Rumah Belajar, Meja Kita, Icando, IndonesiaX, Google for education such as Google form, Google Meeting, and Google Classroom, Microsoff Office 365, Quipper School, Ruang Guru, Sekolahmu, Zenius, Cisco Webex, Whatsapp, Edmodo, Zoom, and Quizizz	Pada sistem <i>blended learning</i> , nilai ujian secara <i>online</i> dapat digunakan sebagai salah satu acuan dalam menentukan tingkat minat belajar siswa.
2.	Analisis Model Kepuasan Mahasiswa Terhadap Cara Pengajaran Dosen Menggunakan Algoritma C4.5	(R.H. Zer et al., 2022)	Menganalisis tingkat kepuasan 100 mahasiswa STIKOM Tunas Bangsa terhadap cara pengajaran dosen dengan menggunakan algoritma C4.5	Dengan tingkat akurasi hingga 80%, penelitian ini menyimpulkan bahwa cara pengajaran dosen merupakan faktor yang penting dalam menilai tingkat kepuasan mahasiswa terhadap suatu sistem pembelajaran
3.	Students' Interest in Online Learning in Higher Education During the Covid-19 Pandemic	(Setiana et al., 2021)	Dengan sampel 72 mahasiswa jurusan Matematika di salah satu perguruan tinggi di Yogyakarta, dilakukan kuesioner melalui Google Form untuk menganalisa tingkat ketertarikan mahasiswa terhadap sistem pembelajaran <i>online</i> .	Penelitian ini menyatakan bahwa tingkat kehadiran mahasiswa merupakan faktor penting dalam menilai tingkat ketertarikan terhadap sistem ini, dimana kehadiran pada sistem ini lebih rendah dibandingkan sistem <i>offline</i> .
4.	An Efficient Data Mining Technique for Assessing	(Abdelkader et al., 2022)	Dengan menggunakan <i>Educational Data Mining</i> , dapat dilakukan analisa	Penelitian ini menyatakan bahwa dalam sistem pembelajaran <i>online</i> ,

No	Judul	Penulis	Metode	Hasil
	Satisfaction Level With Online Learning for Higher Education Students during the COVID-19		terhadap proses pembelajaran <i>online</i> , untuk mengukur tingkat kepuasan siswa pada sistem tersebut.	jumlah kesempatan bertanya yang diberikan pengajar merupakan salah satu faktor yang mempengaruhi tingkat kepuasan tersebut.
5.	Analyzing Student's Learning Interests in the Implementation of Blended Learning Using Data Mining	(Ariyanto et al., 2020)	Algoritma: <i>Linear Regression</i> dan <i>Neural Network</i> . Metrik Kinerja: RMSE, SE	RMSE <i>Linear Regression</i> : 0.00 SE <i>Linear Regression</i> : 0.00 RMSE <i>Neural Network</i> : 0.525 SE <i>Neural Network</i> : 0.275
6.	Student Performance Prediction Using Regression Analysis & Feature- Based Opinion Mining on Student Feedback	(Bramhe et al., 2023)	Algoritma: <i>Linear Regression</i> . Metrik Kinerja: RMSE	RMSE <i>Linear Regression</i> : 0.5
7.	Pemodelan Prediktif Konsumsi Energi Bangunan Gedung Komersial Dengan Algoritma Support Vector Machine	(Dabhade et al., 2021)	Algoritma: <i>Linear Regression</i> . Metrik Kinerja: MAE, MSE, RMSE, R^2	MAE <i>Linear Regression</i> : 0.44079 MSE <i>Linear Regression</i> : 0.23040 RMSE <i>Linear Regression</i> : 0.48 R^2 <i>Linear Regression</i> : 0.83278
8.	Enhancing Student Performance Prediction via Educational Data Mining on Academic data	(Alamgir et al., 2023)	Algoritma: <i>Linear Regression</i> . Metrik Kinerja: MAE, RMSE	Rata-rata MAE <i>Linear Regression</i> : 0.1354 Rata-rata RMSE <i>Linear Regression</i> : 0.1758

No	Judul	Penulis	Metode	Hasil
9.	Predicting student satisfaction of emergency remote learning in higher education during COVID-19 using machine learning techniques	(Ho et al., 2021)	Algoritma: <i>Linear Regression</i> . Metrik Kinerja: MAE, RMSE, R^2	MAE <i>Linear Regression</i> : 0.683 RMSE <i>Linear Regression</i> : 0.883 R^2 <i>Linear Regression</i> : 0.721
10.	Prediction of Student's Academic Performance during Online Learning Based on Regression in Support Vector Machine	(Samsudin et al., 2022)	Algoritma: <i>SVM Linear, SVM RBF, SVM Polynomial</i> Metrik Kinerja: RMSE	RMSE <i>SVM Linear</i> : 0.2661 RMSE <i>SVM RBF</i> : 0.2557 RMSE <i>SVM Polynomial</i> : 0.2649
11.	Comparison of random forest and support vector machine regression models for forecasting road accidents	(Gatera et al., 2023)	Algoritma: <i>Random Forest, SVM</i> Metrik Kinerja: MAE, MSE, RMSE, R^2	MAE <i>Random Forest</i> : 6.119 MSE <i>Random Forest</i> : 67.682 RMSE <i>Random Forest</i> : 8.226 R^2 <i>Random Forest</i> : 0.983 MAE <i>SVM</i> : 0.001 MSE <i>SVM</i> : 1.240 RMSE <i>SVM</i> : 0.001 R^2 <i>SVM</i> : 0.999
12.	Machine Learning Regression Models to Predict Particulate Matter (PM2.5)	(Letchumanan & B, 2022)	Algoritma: <i>Linear Regression, SVM</i> Metrik Kinerja: MAE, RMSE, R^2	MAE <i>Linear Regression</i> : 13.28 RMSE <i>Linear Regression</i> : 23.42 R^2 <i>Linear Regression</i> : 0.83 MAE <i>SVM</i> : 13.1 RMSE <i>SVM</i> : 23.94 R^2 <i>SVM</i> : 0.82

No	Judul	Penulis	Metode	Hasil
13.	Comparison between Regression Models, Support Vector Machine (SVM), and Artificial Neural Network (ANN) in River Water Quality Prediction	(Najwa Mohd Rizal et al., 2022)	Algoritma: <i>SVM</i> , <i>Neural Network</i> Metrik Kinerja: MSE, RMSE, R^2	MSE <i>SVM</i> : 0.0167 RMSE <i>SVM</i> : 0.1292 R^2 <i>SVM</i> : -0.03 MSE <i>Neural Network</i> : 0.0001 RMSE <i>Neural Network</i> : 0.0104 R^2 <i>Neural Network</i> : 0.9953
14.	Analisis Support Vector Regression (Svr) Dengan Kernel Radial Basis Function (Rbf) Untuk Memprediksi Laju Inflasi Di Indonesia	(Rais, 2022)	Algoritma: <i>SVM RBF</i> Metrik Kinerja: RMSE	RMSE <i>SVM RBF</i> : 0.002

II.6. Perbedaan Penelitian dengan Penelitian Terdahulu

Penelitian ini mengkombinasikan empat faktor yang mempengaruhi minat belajar siswa, yaitu nilai ujian, cara pengajaran dosen, tingkat kehadiran, serta kesempatan bertanya siswa, yang diperoleh dari hasil penelitian sebelumnya. Penelitian ini menambahkan parameter regulasi *lasso* dan *ridge* pada algoritma *Linear Regression*, serta membandingkan fungsi kernel *Linear* dan *Polynomial* pada algoritma SVM.