

Judul : ANALISIS HUBUNGAN MATA KULIAH KOMPUTASI DASAR DENGAN IPK MAHASISWA TEKNIK INFORMATIKA MENGGUNAKAN *SUPPORT VECTOR REGRESSION*

Penulis :
1. Novario Jaya Perdana
2. Kelvin Ferdinand
3. Carisha Puspa Gozali
4. Dyah Erny Herwindiati

Penerbit : Infotech: Journal of Technology Information
STMIK Widuri, Jakarta

Akreditasi : SINTA 4 (177/E/KPT/2024 Tanggal 15 Oktober 2024)

Tanggal terbit : 24 Juni 2025

Tautan : <https://jurnal.kampuswiduri.ac.id/index.php/infotech/article/view/383>

SERTIFIKAT

Direktorat Jendral Pendidikan Tinggi, Riset dan Teknologi
Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi Republik Indonesia



Kutipan dari Keputusan Direktorat Jendral Pendidikan Tinggi, Riset, dan Teknologi
Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi Republik Indonesia

Nomor: 177/E/KPT/2024 Tanggal: 15 October 2024

Peringkat Akreditasi Jurnal Ilmiah Periode II Tahun 2024

Nama Jurnal Ilmiah

Infotech: Journal of Technology Information

E-ISSN

24602108

STMIK Widuri

Ditetapkan Sebagai Jurnal Ilmiah:

TERAKREDITASI PERINGKAT 4

Akreditasi Berlaku selama 5 (lima) Tahun, yaitu:
Volume 9 Nomor 1 Tahun 2023 sampai Volume 13 Nomor 2 Tahun 2027

Jakarta,

Direktur Riset, Teknologi dan Pengabdian Kepada Masyarakat



M. Faiz Syaib

NIP. 196708311994021001



ANALISIS HUBUNGAN MATA KULIAH KOMPUTASI DASAR DENGAN IPK MAHASISWA TEKNIK INFORMATIKA MENGGUNAKAN *SUPPORT VECTOR REGRESSION*

Novario Jaya Perdana^{1*}, Kelvin Ferdinand², Carisha Puspa Gozali³, Dyah Erny Herwindiati⁴

¹Program Studi Sistem Informasi, Universitas Tarumanagara, Jakarta, Indonesia

^{2,3,4}Program Studi Teknik Informatika, Universitas Tarumanagara, Jakarta, Indonesia

Correspondence email: novariojp@fti.untar.ac.id

Article history: Submission date: May-01-2025 Revised date: June-06-2025 Published date: June-24-2025

ABSTRACT

Student academic performance serves as a key indicator in higher education assessment. For students in Informatics Engineering, foundational computing skills are critical to their academic progression, and these are primarily acquired through first-semester courses. This study proposes a predictive model for Cumulative Grade Point Average (GPA) using the Support Vector Regression (SVR) method with a Radial Basis Function (RBF) kernel. The courses "Introduction to Algorithms," "Computation I," "Computation II," and "Data Structures" were selected as independent variables, as they provide essential computing foundations for subsequent coursework. The dataset comprised 270 records, each containing grades from the aforementioned courses and the corresponding GPA achieved by students in their fourth semester. To ensure data quality, outlier detection was performed using the Z-score method, resulting in a refined dataset of 200 entries. This dataset was then split into 75% for training and 25% for testing. A grid search optimization identified the best hyperparameter combination: $C = 100$, $\gamma = 0.05$, and $\epsilon = 0.05$. Model evaluation yielded promising results, with a Mean Absolute Error (MAE) of 0.0742, a Mean Absolute Percentage Error (MAPE) of 2.19%, a Mean Squared Error (MSE) of 0.012, and an R^2 score of 0.8695—indicating strong predictive accuracy. Furthermore, the F-test produced a value of 74.9440, which exceeds the critical F-value of 2.5787, confirming the statistical significance of the independent variables in predicting GPA. This model has the potential to support academic monitoring and enhancement efforts by delivering actionable predictions and insights for the Informatics Engineering program.

Keywords: Grade Point Average, Prediction, Support Vector Regression, Informatics Engineering.

ABSTRAK

Prestasi akademik mahasiswa merupakan indikator penting dalam evaluasi pada tingkat pendidikan tinggi. Bagi mahasiswa program studi Teknik Informatika, kemampuan dalam bidang komputasi dasar adalah landasan bagi kelanjutan studi mereka. Kemampuan bidang komputasi dasar ini didapatkan pada mata kuliah yang ditawarkan di semester satu. Pada penelitian ini dibangun model prediksi Indeks Prestasi Kumulatif (IPK) dengan metode *Support Vector Regression* (SVR) berbasis kernel *Radial Basis Function* (RBF). Mata kuliah *Introduction to Algorithms*, *Computation I*, *Computation II*, dan *Data Structures* dipilih sebagai variabel bebas karena berperan sebagai mata kuliah di tingkat persiapan yang menjadi dasar kemampuan komputasi dasar bagi mahasiswa di semester lanjutan. Dataset yang digunakan pada penelitian ini sejumlah 270 baris, berisikan nilai keempat mata kuliah tersebut dan IPK yang didapatkan oleh mahasiswa di semester 4. Deteksi *outlier* menggunakan *Z-score* dilakukan untuk memastikan hanya data representatif yang digunakan, sehingga total data yang digunakan adalah sebesar 200 baris. Dataset dibagi menjadi 75% sebagai data latih dan 25% sebagai data uji. Proses *grid search* menemukan kombinasi parameter optimal pada $C = 100$, $\gamma = 0.05$, dan $\epsilon = 0.05$. Hasil evaluasi menunjukkan kinerja model yang baik dengan nilai MAE sebesar 0.0742, MAPE sebesar 2.19%, MSE sebesar 0.012, dan R^2 sebesar 0.8695. Nilai ini mengindikasikan bahwa model mampu memberikan prediksi IPK yang akurat berdasarkan nilai keempat mata kuliah ini. Uji F menghasilkan nilai 74.9440, melebihi nilai kritis (F-tabel) 2.5787, menunjukkan bahwa variabel bebas berpengaruh signifikan terhadap IPK. Model ini dapat membantu dalam memonitor dan meningkatkan kinerja akademik mahasiswa dengan memberikan prediksi dan wawasan yang relevan bagi program studi Teknik Informatika.

Kata Kunci: Indeks Prestasi Kumulatif, Prediksi, *Support Vector Regression*, Teknik Informatika



PENDAHULUAN

Kemampuan dalam bidang komputasi dasar memegang peranan penting bagi mahasiswa Teknik Informatika. Mata kuliah komputasi dasar berfungsi sebagai landasan penting untuk studi lanjutan dalam ilmu komputer dengan memperkenalkan siswa pada langkah pemikiran komputasi. Mata kuliah komputasi dasar tidak hanya memberikan fondasi teoritis, tetapi juga keterampilan praktis yang diperlukan untuk memahami dan mengembangkan teknologi informasi. Mata kuliah ini dirancang untuk membekali siswa dengan keterampilan terkait dengan penanganan tugas berbasis data yang kompleks dan pengembangannya (Silapachote & Srisuphab, 2017). Mata kuliah komputasi dasar biasanya mencakup berbagai topik, termasuk algoritma, struktur data, bahasa pemrograman, dan pemikiran komputasi, yang merupakan dasar untuk memahami topik ilmu komputer yang lebih kompleks di kemudian hari (Harimurti dkk., 2023).

Pada penelitian yang dilakukan oleh Vourletsis & Politis (2020) menunjukkan bahwa kemampuan siswa dalam memecahkan masalah akan meningkat jika diiringi dengan adanya integrasi daya pikir komputasi yang dimiliki oleh para siswa. Daya pikir komputasi memberikan siswa kemampuan untuk berpikir secara sistematis dalam penyelesaian masalah, dimana terdapat pengurutan logis, analisis data, dan persiapan solusi menggunakan alat visualisasi seperti diagram dan grafik (Rana dkk., 2022). Daya pikir komputasi ini perlu diasah melalui mata kuliah komputasi dasar yang perlu diambil oleh mahasiswa di semester awal.

Egon Börger (2004) menekankan bahwa seorang ahli komputer harus memiliki kemampuan prinsip komputasi dan desain sistem, menghubungkan teori komputasi dengan pengembangan dan analisis sistem tingkat tinggi, sehingga mata kuliah tentang algoritma dan desain sistem diperlukan untuk mempersiapkan siswa saat akan mengambil kelas lanjutan (Börger, 2004). Demikian pula Shackelford dan LeBlanc yang menerapkan kelas abstraksi dan keterampilan pengembangan perangkat lunak yang tidak bergantung pada bahasa. Kelas-kelas ini memberikan landasan konseptual yang melibatkan siswa dan mempersiapkan para siswa untuk pemecahan masalah tingkat lanjut menggunakan komputasi (Shackelford & LeBlanc, 1997).

Pentingnya pemikiran komputasi lebih lanjut disorot oleh Lu dan Fletcher, yang berpendapat bahwa cara berpikir komputasi harus diajarkan sebagai keterampilan formatif, mirip dengan membaca dan aritmatika, untuk mempersiapkan siswa untuk pemrograman dan kurikulum ilmu komputer yang lebih luas (Lu & Fletcher, 2009). Selain itu, mata kuliah dasar sering mencakup komponen praktis yang membantu

kemampuan siswa dalam pemecahan masalah sederhana hingga kompleks (Berenguer, 2021).

Mata kuliah komputasi dasar juga mengajarkan mengenai tantangan pedagogis untuk memperkenalkan dasar-dasar ilmu komputer sebelum pemrograman, memungkinkan siswa untuk membangun kerangka konseptual yang kuat sebelum mempelajari bahasa pemrograman (Lu & Fletcher, 2009; Shackelford & LeBlanc, 1997). Selanjutnya, seperti yang dijelaskan oleh Behforooz (1984) dan Tucker dkk (1994) bahwa mata kuliah komputasi dasar seperti kalkulus dan logika informatika mengintegrasikan pemecahan masalah, struktur data, dan abstraksi, yang sangat penting untuk mengasah kemampuan pemecahan masalah besar dan kompleks dalam ilmu komputer (Behforooz & Sharma, 1984; Tucker dkk., 1994). Secara keseluruhan, mata kuliah komputasi dasar dirancang untuk membekali siswa dengan keterampilan dan pengetahuan yang diperlukan untuk unggul dalam studi ilmu komputer tingkat lanjut, menumbuhkan keingintahuan intelektual dan kompetensi praktis di lapangan.

Namun, terdapat tantangan dalam mengevaluasi sejauh mana hubungan antara pemahaman mahasiswa terhadap mata kuliah ini dengan indeks prestasi akademik mahasiswa. Dalam konteks pendidikan tinggi, indeks prestasi kumulatif (IPK) berperan sebagai indikator utama untuk mengukur keberhasilan akademik mahasiswa. IPK tidak hanya mencerminkan kinerja akademik mahasiswa, namun juga memberikan gambaran mengenai efektivitas kurikulum dan metode pengajaran yang diterapkan oleh perguruan tinggi (Marschalkó & Szamosközi, 2017). Penelitian yang dilakukan oleh Nurudeen dkk., (2023) menunjukkan bahwa kinerja akademik mahasiswa pada tahun-tahun awal dapat mempengaruhi keberhasilan akademik seorang mahasiswa di tahun berikutnya secara signifikan.

Berdasarkan pemaparan tersebut, maka pada penelitian ini akan dipelajari hubungan antara mata kuliah komputasi dasar dengan IPK mahasiswa. Proses analisis data memanfaatkan teknik *machine learning*. Dengan menggunakan pendekatan ini, diharapkan dapat mengidentifikasi pola yang signifikan dan memberikan wawasan yang lebih mendalam mengenai pengaruh mata kuliah komputasi dasar terhadap kesuksesan studi mahasiswa. Metode *machine learning* dipilih karena kemampuannya dalam melakukan analisis prediksi berbasis data.

Terdapat beberapa penelitian sebelumnya yang telah membuktikan kemampuan metode *machine learning* ini. Contohnya penelitian yang dilakukan oleh Nayani dkk., (2023) dimana menggunakan metode *machine learning* seperti *Convolutional Neural Network* (CNN) dan *Recurrent Neural Network* (RNN) untuk memprediksi kinerja siswa. Penelitian tersebut

memperkenalkan model *machine learning* secara hibrida yang menggabungkan CNN dan RNN menjadi *Convolutional Recurrent Network* (CRN). Model ini ditingkatkan oleh teori himpunan kasar entropi yang dioptimalkan untuk penambangan fitur, yang membantu dalam memprediksi kinerja siswa secara akurat. Model yang dibangun pada penelitian ini berhasil mendapatkan akurasi mencapai 93%.

Pada penelitian yang dilakukan oleh Abuzinadah dkk. (2023) juga dibuktikan bahwa metode *machine learning* terutama CNN dapat memberikan akurasi yang tinggi untuk langkah prediksi kinerja studi siswa. Penelitian tersebut menggunakan teknik *synthetic minority oversampling technique* (SMOTE) untuk menyeimbangkan kumpulan data, yang sangat penting untuk meningkatkan kinerja model pembelajaran mesin, terutama ketika berhadapan dengan data yang tidak seimbang. Pendekatan ini terbukti membantu dalam mencapai prediksi yang lebih andal, yaitu hingga 99%.

Penelitian-penelitian tersebut memberikan gambaran bahwa metode *machine learning* dapat memprediksi kinerja siswa secara akurat. Oleh karena itu, pada penelitian ini dipilih metode *Support Vector Regression* (SVR) yang merupakan salah satu dari metode *machine learning* yang memiliki kemampuan untuk menganalisis data besar secara akurat (Dash dkk., 2023; R dkk., 2022; Samsudin dkk., 2022). Dengan mempertimbangkan pentingnya akurasi prediksi IPK dan relevansi mata kuliah dasar terhadap kinerja akademik mahasiswa, penelitian ini bertujuan menganalisis pengaruh dari nilai mata kuliah dasar terhadap IPK mahasiswa di semester mendatang serta mengembangkan model prediksi yang dapat diandalkan.

METODE PENELITIAN

Metode *machine learning* yang digunakan pada penelitian ini adalah *Support Vector Regression* (SVR). Model prediksi dilatih menggunakan variabel bebas berupa perolehan nilai mata kuliah komputasi dasar. Pada akhirnya, model diharapkan untuk dapat mengevaluasi sejauh mana mata kuliah tersebut memengaruhi IPK, sehingga memberikan gambaran yang lebih jelas mengenai faktor-faktor yang berpengaruh terhadap kinerja akademik mahasiswa.

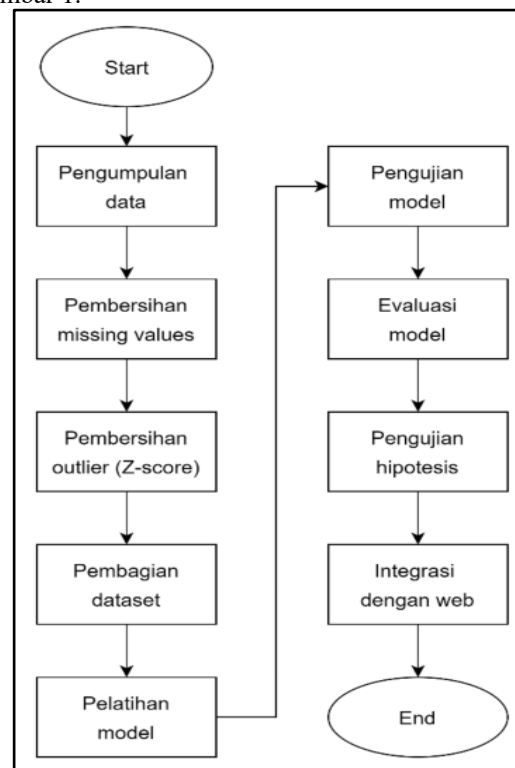
Pada bagian ini akan dibahas mengenai desain penelitian, sumber data dan teknik analisis serta metode evaluasi yang digunakan pada penelitian ini.

Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan fokus pada pengembangan model prediksi IPK berdasarkan nilai mata kuliah. Desain penelitian ini mengimplementasikan analisis data historis dengan metode SVR. Secara garis besar terdapat empat tahap

pada pelaksanaan penelitian ini. Tahap pertama adalah persiapan dataset, dimana data didapatkan dari Program Studi Teknik Informatika Universitas Tarumanagara. Data dipilih terlebih dahulu untuk mata kuliah komputasi dasar. Kemudian dataset ini dipersiapkan melalui langkah preprosesing. Setelah itu, dataset akan dibagi untuk proses pelatihan sebanyak 75% dan pengujian sebanyak 25%. Tahapan kedua adalah proses pembuatan model berdasarkan dataset pelatihan menggunakan SVR. Tahapan ketiga adalah proses pengujian model menggunakan data uji. Tahapan terakhir adalah integrasi model prediksi ke antarmuka aplikasi berbasis web. Alur penelitian dapat dilihat pada

Gambar 1.



Gambar 1. Diagram Alir Penelitian

Sumber Data

Data yang digunakan pada penelitian ini adalah data nilai dari mahasiswa Program Studi Teknik Informatika Universitas Tarumanagara. Data diambil dari mahasiswa yang sudah menyelesaikan empat semester, sehingga sudah mendapatkan mata kuliah komputasi dasar. Subjek yang diambil adalah angkatan 2022 sejumlah 270 mahasiswa. Subjek ini dipilih karena merupakan angkatan pertama yang mengikuti perkuliahan secara tatap muka dan sudah menggunakan kurikulum Merdeka Belajar Kampus Merdeka, yang diterapkan sejak tahun 2020. Keputusan ini diambil karena angkatan ini lebih mencerminkan kinerja akademik mahasiswa yang



sesungguhnya dibandingkan angkatan sebelumnya yang menjalankan perkuliahan secara daring.

Data yang dikumpulkan mencakup IPK mahasiswa pada akhir semester empat serta nilai dari empat mata kuliah komputasi dasar yang terdapat pada Prodi Teknik Informatika Universitas X, yaitu Pengenalan Algoritma (x_1), Komputasi I (x_2), Komputasi II (x_3), dan Struktur data (x_4). Pemilihan empat mata kuliah ini didasarkan pada sifat fundamental dari mata kuliah tersebut, yang perlu ditempuh oleh mahasiswa selama dua semester pertama.

Penggunaan Metode SVR

SVR merupakan metode pembelajaran mendalam yang ditemukan oleh Vapnik, Steven Golowich dan Alex Smola (Vapnik dkk., 1996). Metode ini merupakan pengembangan dari *Support Vector Machine* (SVM) yang lebih dikenal untuk klasifikasi. Pada prinsipnya, SVR bertujuan mencari sebuah fungsi sebagai garis pemisah terbaik yang dapat memprediksi nilai output berdasarkan input yang diberikan. Dari data latih yang diberikan, model SVR akan mencoba menemukan fungsi regresi yang menghasilkan estimasi nilai y dengan kesalahan prediksi sekecil mungkin. Pembuatan model SVR dengan menggunakan kernel RBF sesuai dengan persamaan (1) dan (2).

$$f(x) = \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) k(x_i, x) + b \quad (1)$$

$$k(x_i, x) = \exp(-\gamma \|x_i - x\|^2) \quad (2)$$

Dalam SVR dengan kernel RBF, terdapat parameter gamma (γ) yang mengontrol seberapa jauh pengaruh dari data latih terhadap prediksi (Nariswari & Pudjihastuti, 2023). Nilai gamma yang rendah menunjukkan bahwa pengaruh setiap data latih akan tersebar lebih jauh, memengaruhi area yang lebih luas dalam ruang fitur.

Selain itu, dalam SVR terdapat parameter C yang memengaruhi akurasi prediksi. Parameter ini mengatur trade-off antara kesalahan prediksi dan kompleksitas model (Pande dkk., 2023). Nilai C yang rendah memberikan toleransi lebih besar terhadap kesalahan, sehingga model dapat menghasilkan fungsi regresi yang lebih sederhana. Pendekatan ini dapat meningkatkan kemampuan generalisasi model pada data baru, tetapi berisiko menyebabkan model tidak mampu menangkap pola yang ada dalam data, yang dikenal sebagai *underfitting*. Sebaliknya, nilai C yang tinggi memaksa model untuk meminimalkan kesalahan pada data latih, sehingga menghasilkan fungsi regresi yang lebih kompleks. Meskipun pendekatan ini dapat meningkatkan akurasi pada data latih, terdapat kemungkinan besar bahwa model akan mengalami *overfitting*. Model yang dihasilkan sangat cocok dengan

data latih, namun gagal menunjukkan performanya dengan baik pada data yang belum pernah dilihat sebelumnya.

Oleh karena itu, pemilihan nilai parameter yang tepat sangat penting untuk mencapai keseimbangan antara akurasi prediksi dan kesederhanaan model. Kombinasi yang optimal dari parameter C , γ , dan ε dapat membantu meningkatkan kemampuan generalisasi model pada data baru dan mengurangi risiko *overfitting*.

Evaluasi Model

Evaluasi yang digunakan pada penelitian ini menggunakan matriks *Mean Absolute Error* (MAE), *Mean Absolute Percentage Error* (MAPE), *Mean Square Error* (MSE) dan *R-Squared* (R^2). Ketiganya dapat memberikan gambaran menyeluruh mengenai akurasi dari model prediksi *Support Vector Regression* (SVR). MAE digunakan untuk mengukur seberapa besar kesalahan rata-rata absolut antara nilai prediksi (\hat{y}) dan nilai sebenarnya (y). Rumus untuk mencari MAE, MAPE, MSE dan R^2 dapat dilihat pada persamaan (3), (4), (5), dan (6).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (4)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

$$R^2 = 1 - \frac{SSE}{SST} \quad (6)$$

Rumusan Hipotesis

Berdasarkan rumusan masalah yang telah diidentifikasi, hipotesis dalam penelitian ini dapat dirumuskan sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

Tidak terdapat pengaruh signifikan dari variabel independen yang terdiri dari nilai mata kuliah *Introduction to Algorithms*, *Computation I*, *Computation II*, dan *Data Structures* terhadap variabel dependen yaitu IPK

$$H_1: \text{Minimal terdapat sebuah } \beta_i \neq 0, i = 1, 2, 3, \text{ dan } 4$$

Terdapat minimal 1 variabel independen yang terdiri dari nilai mata kuliah *Introduction to Algorithms*, *Computation I*, *Computation II*, dan *Data Structures*

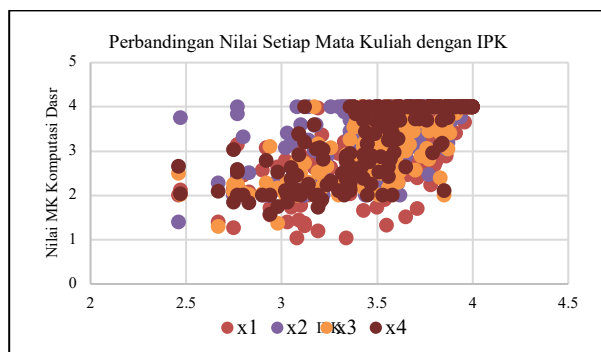
yang berpengaruh signifikan terhadap variabel dependen yaitu IPK.

HASIL DAN PEMBAHASAN

Hasil Prapemrosesan Data

Sebelum proses pelatihan dan pengujian model SVR, data yang didapatkan terlebih dahulu melalui tahapan praproses untuk membersihkan baris data yang mengandung data kosong. Selain itu, dilakukan juga standarisasi data menggunakan metode Z-score untuk mengidentifikasi outlier. Hasil dari tahap ini, didapatkan data sebanyak 200 baris data yang siap untuk diproses lebih lanjut.

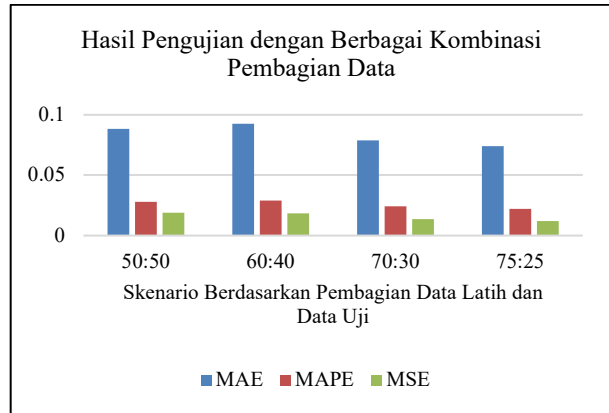
Gambar 2 menampilkan data yang digunakan pada penelitian ini.



Gambar 2. Data Setelah Tahapan Pembersihan *Outlier*

Pelatihan Model SVR

Pada pelaksanaan pelatihan model, dataset dibagi menjadi dua yaitu data latih dan data uji. Untuk menentukan proporsi pembagian data, dilakukan pengujian terhadap beberapa kombinasi untuk melihat pengaruhnya terhadap akurasi model prediksi. Kombinasi yang diuji meliputi proporsi 50:50, 60:40, 70:30, dan 75:25 untuk data latih dan uji. Tujuan dari pengujian ini adalah mengevaluasi stabilitas model dan mengidentifikasi pembagian data yang menghasilkan kinerja terbaik berdasarkan nilai MAE, MAPE, dan MSE. Hasil pengujian dari berbagai kombinasi pembagian data tersebut dapat dilihat pada Gambar 3.



Gambar 3. Evaluasi Pembagian Data Latih dan Data Uji

Nilai MAE, MAPE, dan MSE menunjukkan model akurasi terbaik dihasilkan oleh proporsi 75:25, dimana nilai ketiganya pada proporsi ini adalah yang paling kecil dibandingkan proporsi pembagian data lainnya. Hal ini dikarenakan proporsi data latih yang tinggi memberikan lebih banyak informasi kepada model untuk mempelajari pola dari dataset, sehingga model dapat lebih baik dalam memprediksi variabel terikat. Sehingga untuk tahap selanjutnya, proporsi yang digunakan adalah 75% untuk data latih dan 25% untuk data uji.

Untuk mengoptimalkan kinerja model, metode *grid search* digunakan dalam model SVR dengan kernel RBF untuk mencari kombinasi terbaik dari parameter yang meliputi C, gamma (γ), dan epsilon (ϵ). Rentang nilai parameter yang akan diuji dapat dilihat pada Tabel 1.

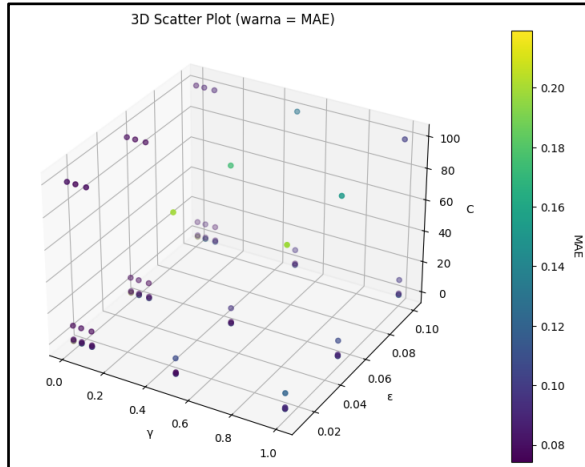
Tabel 1. Rentang Nilai SVR pada Grid Search

Parameter	Rentang Nilai
C	0.01, 0.1, 1, 10, 100
γ	0.01, 0.05, 0.1, 0.5, 1
ϵ	0.01, 0.05, 0.1

Dengan kombinasi parameter yang telah ditentukan, sebanyak 75 kombinasi diuji dalam proses *grid search*. Setiap kombinasi parameter dilatih menggunakan data latih dan kemudian dievaluasi menggunakan data uji yang terdiri dari 50 baris. Setelah model dilatih, nilai MAE dihitung untuk setiap model.

Model dengan nilai MAE terbaik akan dipilih sebagai model akhir yang digunakan untuk prediksi. Proses ini memastikan bahwa model yang dihasilkan memiliki akurasi yang tinggi dan mampu memberikan hasil yang lebih andal dalam memprediksi IPK mahasiswa berdasarkan variabel independen yang telah ditetapkan. Tabel 1 menunjukkan hasil proses *grid search*. Kombinasi model SVR terbaik ditunjukkan pada Gambar 4.





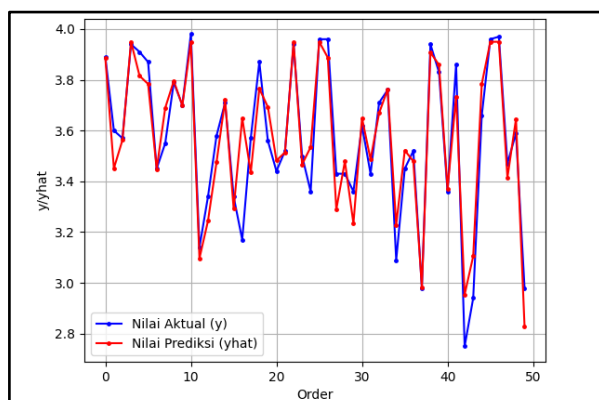
Gambar 4. Kombinasi Model SVR Terbaik

Hasil ini menunjukkan bahwa model mampu memprediksi IPK mahasiswa dengan akurasi yang baik sehingga dapat diandalkan untuk analisis selanjutnya.

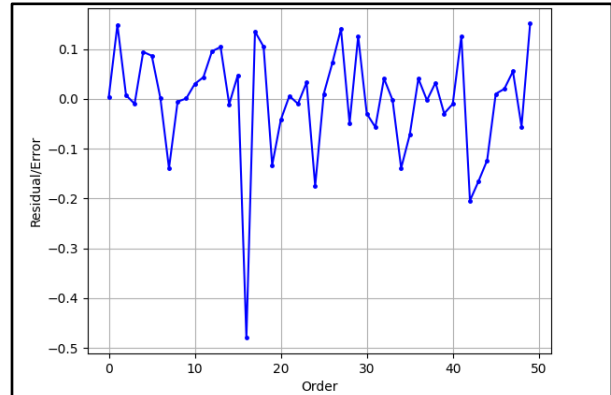
Kombinasi parameter ini memberikan prediksi yang paling optimal. Nilai C yang tinggi memastikan model menangkap hubungan antarvariabel, γ rendah menghindari overfitting dan menyeimbangkan support vector, dan ϵ yang optimal menjaga fleksibilitas dan menangkap pola data.

Pengujian Model SVR

Setelah pelatihan, model SVR diuji menggunakan data uji. Pengujian ini bertujuan untuk mengevaluasi performa model dalam memprediksi IPK berdasarkan data baru. Pengujian dilakukan dengan membandingkan nilai aktual (y) dan hasil prediksi model (\hat{y}). Hasil pengujian ini dapat dilihat pada 5, sedangkan plot residual dapat dilihat Gambar 6.



Gambar 5. Perbandingan Nilai Aktual dan Prediksi



Gambar 6. Plot Residual Hasil Pengujian Model

Nilai aktual adalah nilai asli dari variabel terikat yang ingin diprediksi, sementara nilai prediksi adalah nilai yang dihasilkan oleh model. Residual dihasilkan dari perhitungan selisih nilai aktual dan nilai prediksi, menunjukkan deviasi pada setiap pengamatan. Terlihat sebagian besar nilai prediksi mendekati nilai aktual dengan residual kecil. Hal ini menunjukkan bahwa model mampu memprediksi dengan baik. Sebagian besar nilai residual tersebar di sekitar garis nol, menunjukkan kesalahan prediksi model tanpa pola yang jelas dan acak. Sebaran ini menunjukkan model bekerja konsisten tanpa masalah bias.

Tabel 2 menunjukkan hasil metrik evaluasi MAE, MAPE, dan MSE. MAE sebesar 0,0742 menunjukkan rata-rata kesalahan prediksi yang kecil, mencerminkan model SVR mampu memprediksi yang dekat dengan nilai aktual. Nilai MAPE sebesar 2,19% mengindikasikan akurasi tinggi, dengan rata-rata kesalahan di bawah 10%. Nilai MSE sebesar 0,012 menunjukkan bahwa rata-rata kesalahan relatif kecil dan model konsisten dalam menjaga kesalahan kecil di data uji, meminimalkan perbedaan prediksi dan nilai aktual.

Tabel 2. Hasil Evaluasi Model SVR

MAE	MAPE	MSE	R^2
0.07425265	0.02192259	0.01200884	0.86948052

Berdasarkan nilai kesalahan prediksi yang diukur melalui MAE, MAPE, dan MSE selanjutnya dilakukan perhitungan koefisien determinasi R^2 , seperti yang terlihat pada Tabel 2. Nilai R^2 yang didapatkan menunjukkan bahwa terdapat sekitar 86,95% variabilitas dalam nilai IPK dapat dijelaskan oleh model SVR. Semakin mendekati 1, maka semakin baik model dalam menjelaskan hubungan variabel bebas dan dependen, dan mampu menangkap pola dengan baik.

Pengujian Hipotesis

Setelah model SVR dievaluasi, langkah selanjutnya adalah menguji hipotesis untuk memastikan adanya pengaruh signifikan terhadap variabel terikat, yaitu nilai

IPK, yang diberikan oleh variabel bebas, yaitu mata kuliah komputasi dasar. Pengujian hipotesis menggunakan uji F untuk mengetahui kontribusi semua variabel secara simultan dalam memprediksi variabel terikat. Hasil pengujian hipotesis terlampir pada Tabel 3.

Tabel 3. Hasil Analisis F-Hitung

Sumber Variasi	SS	df	MS	F-hitung
Regresi	3.9999582	4	0.99998955	74.94403215
Error	0.6004418	45	0.01334315	
Total	4.6004			

Nilai F-hitung yang diperoleh sebesar 74,9440 menunjukkan bahwa model signifikan dalam menjelaskan variasi nilai IPK berdasarkan variabel bebas. Selain itu, didapatkan juga bahwa nilai F-hitung jauh lebih besar daripada nilai kritis F-tabel sebesar 2,5787 dengan signifikansi 0,05. Hal ini mengindikasikan adanya pengaruh signifikan dari variabel-variabel bebas terhadap nilai IPK. Dengan demikian, hipotesis nol (H_0) ditolak dan hipotesis alternatif (H_1) diterima. Ini berarti setidaknya terdapat satu variabel independen memberikan kontribusi signifikan dalam model prediksi. Hal ini juga memperkuat validitas model SVR yang telah dibangun, menunjukkan bahwa data yang digunakan relevan dalam memprediksi prestasi akademik mahasiswa.

Kombinasi Variabel

Pengujian tambahan dilakukan dengan membuat kombinasi dua variabel untuk mengevaluasi pengaruh jumlah variabel terhadap performa model. Pendekatan ini bertujuan untuk memberikan wawasan tambahan sekaligus memperkuat kesimpulan bahwa empat variabel tersebut adalah model yang akurat dan optimal dalam memprediksi nilai IPK. Pengujian dilakukan tahapan sama seperti model dengan empat variabel, namun berbeda pada kombinasi variabel yang digunakan, yaitu sebagai berikut:

1. *Introduction to Algorithms* dan *Data Structures* (x1-x4)
2. *Introduction to Algorithms* dan *Computation I* (x1-x2)
3. *Computation I* dan *Computation II* (x2-x3)
4. *Computation II* dan *Data Structures* (x3-x4)
5. Kombinasi seluruh mata kuliah (x1, x2, x3, x4)

Kombinasi nomor 1 berisikan mata kuliah yang menekankan dasar dan logika pemrograman, di mana *Introduction to Algorithms* di semester 1 menjadi prasyarat untuk mata kuliah *Data Structure* di Semester 2. Kombinasi nomor 2 meliputi mata kuliah yang diambil pada semester 1. Kombinasi ketiga merepresentasikan mata kuliah berbasis logika

matematika dan pemrograman, di mana *Computation I* akan menjadi prasyarat untuk *Computation II*. Kombinasi nomor 4 mencakup dua mata kuliah yang diambil pada semester 2.

Hasil pengujian dan evaluasi tiap model kombinasi disajikan pada Tabel 4. Berdasarkan hasil ini didapatkan bahwa penggunaan empat variabel dapat meningkatkan akurasi model, dengan mempertimbangkan keempat variabel membuat model menangkap lebih banyak pola. Model dengan empat variabel ini disarankan sebagai pendekatan utama dalam sistem ini karena terbukti memberikan hasil optimal, memperkuat relevansi penelitian dalam merekomendasikan penggunaan variabel yang tepat.

Tabel 4. Hasil Pengujian Model dengan Kombinasi Variabel

Kombinasi Variabel	Parameter Terbaik			MAE	MAPE	MSE
	C	γ	ϵ			
x1-x4	100	0,01	0,01	0,174389	0,078058	0,0973
x1-x2	100	0,1	0,01	0,285118	0,139771	0,237917
x2-x3	100	0,1	0,01	0,121732	0,035491	0,026786
x3-x4	100	1	0,01	0,211829	0,080621	0,129707
4 variabel	100	0,05	0,05	0,074253	0,021923	0,012009

Pengujian kombinasi empat variabel berbeda juga dilakukan untuk memastikan kombinasi ini menghasilkan model paling akurat. Mata kuliah *Database Systems* menunjukkan nilai korelasi lebih tinggi dari *Computation I*, namun tidak digunakan dalam penelitian karena tidak termasuk mata kuliah prasyarat.

Dilakukan pengujian empat kombinasi variabel, yaitu kombinasi mata kuliah dari semester 1, semester 2, *Database Systems* menggantikan *Computation I*, dan model pada penelitian ini untuk membuktikan keefektifan model. Hasil terlihat pada Tabel 5.

Tabel 5. Hasil Pengujian dengan Kombinasi Empat Variabel

Kombinasi Variabel	MAE	MAPE	MSE
Semester 1	0.30237354	0.18244330	0.29203078
Semester 2	0.12070126	0.03714851	0.03208741
<i>Database</i>	0.12766475	0.04275081	0.03512499
Penelitian	0.07425265	0.02192259	0.01200884

Hasil pengujian pada Tabel 5 menunjukkan bahwa model dengan empat variabel dalam penelitian ini menjadi model yang paling akurat dalam memprediksi IPK. Hal ini ditunjukkan dengan nilai MAE, MAPE, dan MSE terendah. Model dengan seluruh variabel di



semester 1 menjadi model dengan akurasi terendah. Ini dapat disebabkan mata kuliah IT Trends yang nilai korelasinya rendah dengan IPK. Oleh karena itu, model prediksi pada penelitian ini dengan variabel *Introduction to Algorithms*, *Computation I*, *Computation II*, dan *Data Structures* merupakan model yang optimal untuk melakukan prediksi IPK secara akurat.

Pembahasan

Dalam penelitian ini, hasil analisis kontribusi variabel independen menunjukkan bahwa seluruh mata kuliah dasar yang digunakan memiliki pengaruh terhadap IPK mahasiswa. Variabel x_3 , atau mata kuliah *Computation II*, memberikan kontribusi terbesar secara simultan sebesar 0.1565. Temuan ini menegaskan penguasaan materi pada mata kuliah dasar ini sangat penting untuk mendukung kinerja akademik mahasiswa secara keseluruhan. Pemahaman yang baik terhadap mata kuliah tersebut dapat memengaruhi pemahaman mahasiswa dalam mata kuliah lanjutan dan berdampak pada IPK. Hal ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa mata kuliah dasar memiliki peran penting dalam pembentukan dasar pengetahuan yang mendukung pencapaian akademik.

Selain itu, hasil analisis korelasi menunjukkan bahwa variabel x_3 juga memiliki korelasi tertinggi, dengan koefisien korelasi sebesar 0.789, yang mengindikasikan hubungan yang kuat antara variabel ini dan IPK. Hal ini menunjukkan bahwa kemampuan mahasiswa dalam mata kuliah tersebut sangat berpengaruh terhadap performa akademiknya secara keseluruhan.

Selain kontribusi dan korelasi variabel, penelitian ini juga menganalisis pengaruh kombinasi variabel terhadap akurasi model prediksi. Hasil pengujian menunjukkan bahwa model dengan kombinasi seluruh variabel independen memberikan performa terbaik. Hal ini menegaskan bahwa penggunaan keempat variabel secara bersamaan mampu memberikan informasi tambahan yang meningkatkan akurasi prediksi secara signifikan.

Model SVR yang dilatih menggunakan metode *grid search* menghasilkan parameter optimal, yaitu nilai C sebesar 100, γ (gamma) sebesar 0.05, dan ϵ (epsilon) sebesar 0.05. Model dengan konfigurasi ini mencapai nilai MAE sebesar 0.0742, mencerminkan akurasi prediksi yang cukup baik dan menunjukkan bahwa rata-rata kesalahan prediksi berada pada kisaran yang rendah. Selain itu, nilai MAPE sebesar 2.19% menunjukkan bahwa model memiliki tingkat akurasi tinggi dengan rata-rata kesalahan yang berada di bawah 10%. Nilai MSE sebesar 0.012 memperlihatkan bahwa rata-rata kuadrat kesalahan antara nilai aktual dan prediksi juga relatif kecil, sementara nilai R^2 sebesar 0.8695 menunjukkan bahwa model SVR ini mampu

menjelaskan sekitar 86.95% variabilitas dalam IPK mahasiswa.

Pengujian hipotesis menggunakan uji F dilakukan untuk memastikan bahwa variabel independen dalam model benar-benar memberikan pengaruh signifikan terhadap IPK mahasiswa. Hasil pengujian menunjukkan bahwa hipotesis alternatif (H_1) diterima dengan nilai F sebesar 74.9440 yang lebih besar dari nilai kritis. Ini menunjukkan bahwa setidaknya terdapat satu variabel independen yang berkontribusi signifikan dalam model prediksi IPK.

Selain itu, dilakukan perbandingan hasil pengujian menggunakan data dari mahasiswa angkatan 2023. Hasil pengujian yang terlampir di Lampiran 18 menunjukkan bahwa kombinasi variabel yang digunakan dalam penelitian ini tetap menghasilkan korelasi tertinggi dibanding mata kuliah lain pada tingkat persiapan.

Hasil ini konsisten dengan temuan sebelumnya, yang mengindikasikan bahwa mata kuliah *Introduction to Algorithms*, *Computation I*, *Computation II*, dan *Data Structures* memiliki kontribusi signifikan terhadap prediksi IPK mahasiswa. Hal ini menunjukkan bahwa hasil penelitian ini dapat diterapkan secara luas untuk prediksi kinerja akademik mahasiswa di angkatan selanjutnya.

Secara keseluruhan, hasil yang diperoleh menunjukkan bahwa model prediksi IPK menggunakan SVR dengan parameter optimal yang ditemukan melalui *grid search* mampu memberikan prediksi yang cukup akurat dan dapat menjelaskan sebagian besar variabilitas dalam IPK mahasiswa. Meski demikian, model ini masih dapat diperbaiki dengan menyertakan faktor-faktor lain yang mungkin berpengaruh, seperti mata kuliah tambahan atau faktor eksternal yang belum disertakan dalam penelitian ini.

KESIMPULAN

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa pendekatan yang digunakan mampu memberikan solusi yang efektif untuk memprediksi IPK mahasiswa. Penelitian ini berhasil membuktikan bahwa model prediksi berbasis SVR dapat digunakan untuk memprediksi nilai IPK mahasiswa dengan akurasi tinggi. 1. Seluruh variabel independen yang digunakan dalam penelitian ini, yaitu nilai mata kuliah *Introduction to Algorithms* (x_1), *Computation I* (x_2), *Computation II* (x_3), dan *Data Structures* (x_4), terbukti memiliki pengaruh dan berkorelasi terhadap IPK mahasiswa. Nilai mata kuliah *Computation II* memiliki kontribusi dan korelasi tertinggi terhadap IPK, baik secara individu maupun simultan kemudian diikuti oleh mata kuliah *Data Structures*, *Introduction to Algorithms*, dan

Computation I. Secara garis besar, hal ini menunjukkan pentingnya penguasaan mata kuliah dasar untuk mendukung pencapaian akademik secara keseluruhan.

Selain itu, analisis kombinasi variabel menunjukkan bahwa penggunaan seluruh variabel independen secara bersamaan memberikan hasil prediksi yang paling akurat. Sebaliknya, pengurangan jumlah variabel menyebabkan penurunan akurasi model, yang menegaskan pentingnya informasi yang diberikan oleh setiap variabel dalam menjelaskan IPK mahasiswa

DAFTAR PUSTAKA

- Abuzinadah, N., Umer, M., Ishaq, A., Hejaili, A. Al, Alsubai, S., Eshmawi, A. A., Mohamed, A., & Ashraf, I. (2023). Role of convolutional features and machine learning for predicting student academic performance from MOODLE data. *PLOS ONE*, *18*(11 November). <https://doi.org/10.1371/JOURNAL.PONE.0293061>
- Behforooz, A., & Sharma, O. P. (1984). A foundation course in computer science. *ACM SIGCSE Bulletin*, *16*(1), 159–163. <https://doi.org/10.1145/952980.808643>
- Berenguer, I. A. (2021). Sistema Básico de Habilidades para la Algoritmización Computacional. *Revista de investigación, formación y desarrollo: Generando productividad institucional*, *9*(1), 14–14. <https://doi.org/10.34070/RIF.V9I1.255>
- Börger, E. (2004). A Practice-Oriented Course on the Principles of Computation, Programming, and System Design and Analysis. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *3294*, 65–84. https://doi.org/10.1007/978-3-540-30472-2_5
- Dash, R. K., Nguyen, T. N., Cengiz, K., & Sharma, A. (2023). Fine-tuned support vector regression model for stock predictions. *Neural Computing and Applications*, *35*(32), 23295–23309. <https://doi.org/10.1007/S00521-021-05842-W/METRICS>
- Harimurti, R., Anistiyasari, Y., Rusimamto, P. W., & Haryudo, S. I. (2023). *Automatic Programming Assessment to Measure Programming Problem-Solving Skills*. <https://doi.org/10.4108/EAI.20-10-2022.2328840>
- Lu, J. J., & Fletcher, G. H. L. (2009). Thinking about Computational Thinking. *SIGCSE'09 - Proceedings of the 40th ACM Technical Symposium on Computer Science Education*, 260–264. <https://doi.org/10.1145/1508865.1508959>
- Marschalkó, E. E., & Szamosközi, I. (2017). Predictors of Academic Success in College: Actual Learning Skills. *Transylvanian Journal of Psychology*, *18*(1), 35–55. <https://doi.org/10.24193/epsz.2017.1.3>
- Nariswari, R., & Pudjihastuti, H. (2023). Support Vector Machine Method for Predicting Non-Linear Data. *Procedia Computer Science*, *227*, 884–891. <https://doi.org/10.1016/J.PROCS.2023.10.595>
- Nayani, S., Rao, P. S., & Lakshmi, D. R. (2023). Combination of Deep Learning Models for Student's Performance Prediction with a Development of Entropy Weighted Rough Set Feature Mining. *Cybernetics and Systems*, 1–43. <https://doi.org/10.1080/01969722.2023.2166259>
- Nurudeen, A. H., Fakhrou, A., Lawal, N., & Ghareeb, S. (2023). Academic performance of engineering students: A predictive validity study of first-year GPA and final-year CGPA. *Engineering reports*, *6*(5). <https://doi.org/10.1002/ENG2.12766>
- Pande, C. B., Kushwaha, N. L., Orimoloye, I. R., Kumar, R., Abdo, H. G., Tolche, A. D., & Elbeltagi, A. (2023). Comparative Assessment of Improved SVM Method under Different Kernel Functions for Predicting Multi-scale Drought Index. *Water Resources Management*, *37*(3), 1367–1399. <https://doi.org/10.1007/S11269-023-03440-0/METRICS>
- R, I., Sudarmin, S., & Rais, Z. (2022). Analisis Support Vector Regression (SVR) dengan Kernel Radial Basis Function (RBF) untuk Memprediksi Laju Inflasi di Indonesia. *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, *4*(1), 30–38. <https://doi.org/10.35580/VARIANSIUNM13>
- Rana, D. S., Dimri, S. C., Malik, P., & Dhondiyal, S. A. (2022). Impact of Computational Thinking in Engineering and K12 Education. *4th International Conference on Inventive Research in Computing Applications, ICIRCA 2022 - Proceedings*, 697–701. <https://doi.org/10.1109/ICIRCA54612.2022.9985593>
- Samsudin, N. A. M., Shaharudin, S. M., Sulaiman, N. A. F., Ismail, S., Mohamed, N. S., & Husin, N. H. M. (2022). Prediction of Student's Academic Performance during Online Learning Based on Regression in Support Vector Machine. *International Journal of Information and Education Technology*, *12*(12), 1431–1435. <https://doi.org/10.18178/IJJET.2022.12.12.1768>



- Shackelford, R. L., & LeBlanc, R. J. (1997). Introducing computer science fundamentals before programming. *Proceedings - Frontiers in Education Conference, 1*, 285–289. <https://doi.org/10.1109/FIE.1997.644858>
- Silapachote, P., & Srisuphab, A. (2017). Engineering Courses on Computational Thinking Through Solving Problems in Artificial Intelligence. *International Journal of Engineering Pedagogy (iJEP)*, 7(3), 34–49. <https://doi.org/10.3991/IJEP.V7I3.6951>
- Tucker, A. B., James, B. W., Cupper, R. D., & Eipstein, R. G. (1994). *Fundamentals Of Computing II: Abstraction, Data Structures, and Large Software Systems, C++ Edition* (International ed.). McGraw-Hill.
- Vapnik, V., Golowich, S., & Smola, A. (1996). Support Vector Method for Function Approximation, Regression Estimation and Signal Processing. *Advances in Neural Information Processing Systems*, 9.
- Vourletsis, I., & Politis, P. (2020). Effects of a Computational Thinking Experimental Course on Students' Perceptions of Their Problem-Solving Skills. *ACM International Conference Proceeding Series*, 14–20. <https://doi.org/10.1145/3383923.3383935>