



## Peramalan Jumlah Kedatangan Wisatawan Menggunakan Support Vector Regression Berbasis Sliding Window

Ma'idatul Fitriah\*, Inggih Permana, Febi Nur Salisah, Medyantiwi Rahmawita Munzir, Megawati

Fakultas Sains dan Teknologi, Program Studi Sistem Informasi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: <sup>1,\*</sup>12050320499@students.uin-suska.ac.id, <sup>2</sup>inggihpermana@uin-suska.ac.id, <sup>3</sup>febinursalisah@uin-suska.ac.id, <sup>4</sup>medyantiwi.rahmawita@uin-suska.ac.id, <sup>5</sup>megawati@uin-suska.ac.id

Email Penulis Korespondensi: 12050320499@students.uin-suska.ac.id

**Abstrak**—Sebagai kota yang sedang berkembang, Pekanbaru memiliki potensi objek wisata yang menarik bagi wisatawan. Kedatangan wisatawan ini memberikan dampak positif yang besar terhadap perekonomian Kota Pekanbaru. Jumlah kedatangan wisatawan dapat mengalami naik dan turun setiap bulannya, untuk itu diperlukan suatu peramalan kedatangan jumlah wisatawan di masa mendatang. Penelitian ini bertujuan untuk menerapkan aplikasi Orange Data Mining dalam meramalkan jumlah kedatangan wisatawan dengan membandingkan kernel yang ada pada metode Support Vector Regression (SVR) serta menerapkan Sliding Window size 3 hingga window size 13 untuk mentransformasi menjadi data time series. Serta melakukan pembagian data menggunakan metode K-Fold Validation dengan nilai K-10. Kemudian kinerja dari kernel-kernel yang digunakan dapat dilihat menggunakan widget Test and Score yang menyajikan dari hasil Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), dan R-squared(R<sup>2</sup>). Hasil untuk peramalan jumlah kedatangan wisatawan ke Kota Pekanbaru menggunakan metode SVR menunjukkan bahwa Kernel RBF merupakan pilihan optimal dibandingkan dengan Kernel Polinomial dan Linear. Hasil dari widget Test and Score menunjukkan bahwa Kernel RBF dengan window size 10 memiliki nilai MAE, MSE, dan RMSE yang lebih rendah, yaitu 0,118, 0,022, dan 0,147. Selain itu, perbandingan R<sup>2</sup> pada window size 10 untuk Kernel RBF menunjukkan performa yang lebih baik dengan nilai sebesar 0,519.

**Kata Kunci:** Orange Data Mining; Peramalan; Sliding Window; Support Vector Regression; Wisatawan

**Abstract**—As a developing city, Pekanbaru has the potential for attractive tourist attractions for tourists. The arrival of tourists has had a big positive impact on the economy of Pekanbaru City. The number of tourist arrivals can experience ups and downs every month, for this reason it is necessary to forecast the number of tourists in the future. This research aims to apply the Orange Data Mining application in predicting the number of tourist arrivals by comparing the kernels in the Support Vector Regression (SVR) method and applying Sliding Window size 3 to window size 13 to transform into time series data. As well as sharing data using the K-Fold Validation method with a value of K-10. Then the performance of the kernels used can be seen using the Test and Score widget which presents the results of Root Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), and R-squared(R<sup>2</sup>). The results for forecasting the number of tourist arrivals to Pekanbaru City using the SVR method show that the RBF Kernel is the optimal choice compared to the Polinomial and Linear Kernels. The results of the Test and Score widget show that the RBF Kernel with window size 10 has lower MAE, MSE and RMSE values, namely 0.118, 0.022 and 0.147. Apart from that, the comparison of R<sup>2</sup> in window size 10 for Kernel RBF shows better performance with a value of 0.519.

**Keywords:** Forecasting; Orange Data Mining; Sliding Window; Support Vector Regression; Tourist

### 1. PENDAHULUAN

Pertumbuhan industri pariwisata memberikan dampak positif yang signifikan terhadap penerimaan devisa dan kehidupan masyarakat di sekitar destinasi pariwisata[1], [2]. Dampak ekonomi meliputi peningkatan pendapatan masyarakat lokal, peluang bisnis di sektor formal dan informal, serta peningkatan penerimaan pajak oleh pemerintah[3]–[5]. Selain itu, industri pariwisata juga mendorong pembangunan infrastruktur yang memfasilitasi akses ke tempat wisata dan menciptakan lapangan kerja bagi Masyarakat. Oleh karena itu, diperlukan perhatian yang sangat penting dari pemerintah terhadap sektor pariwisata. [1]. Dengan fokus yang tepat pada sektor pariwisata, pemerintah dapat meningkatkan potensi ekonomi dan daya tarik destinasi pariwisata, yang pada gilirannya akan memberikan manfaat besar bagi masyarakat lokal serta kontribusi yang signifikan terhadap pertumbuhan ekonomi regional dan nasional.

Kota Pekanbaru, sebagai pusat kota di Provinsi Riau, memiliki beragam tempat wisata yang menarik, seperti Masjid Raya Pekanbaru, museum Sang Nila Utama, danau Bandar Kahyangan, taman rekreasi Alam Mayang dan lainnya, yang menarik minat wisatawan[6]. Berdasarkan data yang diperoleh dari Dinas Kebudayaan dan Pariwisata Kota Pekanbaru dalam rentang Tahun 2016 hingga Tahun 2023, terdapat tren peningkatan kunjungan wisatawan ke Kota Pekanbaru dari tahun 2016 hingga 2019. Namun, pada tahun 2020 terjadi penurunan jumlah pengunjung, diikuti oleh kenaikan signifikan pada tahun 2021 hingga tahun 2023. Kenaikan jumlah pengunjung wisata disaat tertentu khususnya saat liburan, dapat menjadi tantangan jika tidak didukung oleh infrastruktur yang memadai. Tentunya hal ini dapat mengurangi minat wisatawan untuk kembali berkunjung sehingga berdampak pada perekonomian lokal dan berdampak pada pengambilan keputusan pembangunan oleh pemerintah daerah yang semestinya harus sejalan dengan pertumbuhan kunjungan wisatawan[7], [8].

Untuk mengatasi hal ini, diperlukan peramalan yang dapat memberikan gambaran jumlah wisatawan yang akan datang agar pembangunan dapat dilakukan secara akurat dan efisien[3], [9]–[11]. Agar mendapatkan hasil yang akurat perlu dilakukan sebuah teknik untuk membentuk pola data timeseries yaitu dengan Sliding Window



[12]. Penggunaan teknik Sliding Window menghasilkan varian data baru dengan ukuran window yang ditentukan. Teknik ini secara bertahap menggeser window sesuai dengan ukuran yang telah ditetapkan untuk memperoleh akurasi optimal dari ukuran window tersebut, yang kemudian digunakan sebagai acuan dalam peramalan jumlah wisatawan di masa mendatang [13].

Hasil peramalan jumlah wisatawan yang akurat, dapat dimanfaatkan oleh Dinas Kebudayaan dan Pariwisata Kota Pekanbaru untuk meningkatkan perencanaan pariwisata dengan strategi yang lebih tepat, termasuk alokasi sumber daya dan pengembangan infrastruktur sesuai kebutuhan. Hal ini memungkinkan pengelolaan sumber daya yang lebih efisien, meliputi alokasi anggaran yang bijaksana, optimalisasi pemasaran, dan peningkatan kapasitas pelayanan sesuai permintaan. Strategi pemasaran pun dapat lebih tepat sasaran dengan meramalkan jumlah wisatawan, memungkinkan identifikasi segmen pasar potensial, penyesuaian kampanye promosi, dan peningkatan daya tarik destinasi Kota Pekanbaru. Salah satu metode peramalan yang efektif adalah menggunakan algoritma Support Vector Regression (SVR).

Algoritma SVR memiliki keunggulan signifikan dibandingkan dengan ARIMA, yang umumnya digunakan dalam model time series. ARIMA seringkali tidak dapat menangani data outlier dan memodelkan time series yang nonlinier. Sebaliknya, SVR dapat menangkap nonlinieritas dengan menambahkan fungsi kernel, sehingga dapat mengatasi kelemahan yang dimiliki ARIMA [14]. Kelebihan ini membuat SVR menjadi pilihan yang lebih baik dalam konteks peramalan data yang kompleks dan nonlinear, memberikan kontribusi yang signifikan terhadap analisis deret waktu secara akurat dan efektif [13].

Penelitian yang menerapkan metode SVR telah sering digunakan dalam penelitian sebelumnya, sebagaimana yang terlihat dalam penelitian yang dilakukan oleh [16] menganalisis prediksi harga saham menggunakan metode ARIMA dan SVR menghasilkan bahwa metode SVR lebih unggul karena memiliki nilai RMSE yang rendah. Selanjutnya penelitian oleh [17] mengenai perbandingan metode Artificial Neural Network (ANN) dan SVR untuk peramalan jumlah penumpang bus menghasilkan bahwa model terbaik dalam peramalan ini yaitu metode SVR dengan nilai MAPE yang rendah yaitu 2.46% sementara metode ANN yang memiliki nilai MAPE 14.27%.

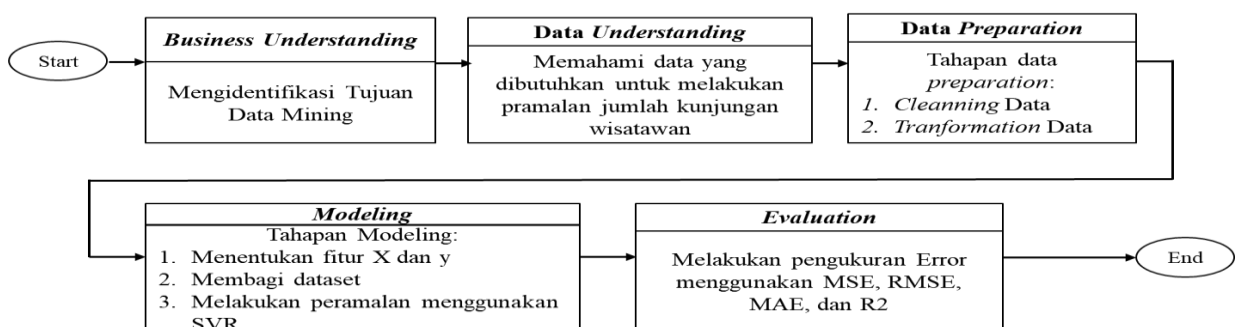
Penelitian lain juga dilakukan oleh [18] untuk memprediksi jumlah penumpang dan penambahan gerbong kereta api yang terbilang metode SVR sangat baik karena hasil evaluasi sangat dekat dengan 0 yaitu MEA sebesar 0.1276, MSE 0.1796 dan MAPE 0,00376. Kemudian penelitian yang dilakukan oleh [19] menggunakan metode SVR untuk memprediksi data kunjungan wisatawan domestic ke Bali dari bulan Januari 2004 sampai Desember 2018 di peroleh nilai MAPE pada data training 11,34% dan untuk data testing diperoleh nilai 7,30%. Selanjutnya penelitian implementasi algoritma SVR Pada Prediksi Jumlah Pengunjung Pariwisata oleh [20] menggunakan 72 data jumlah pengunjung bulanan data pariwisata dari tahun 2010 hingga 2015 yang menghasilkan MAPE terbaik 6.98%.

Berdasarkan penelitian sebelumnya, penelitian ini akan menerapkan metode SVR untuk melakukan peramalan jumlah wisatawan di Kota Pekanbaru dengan data time series dengan teknik Sliding Window. Tujuan utamanya adalah untuk menghasilkan perkiraan jumlah wisatawan pada masa mendatang dengan tingkat kesalahan yang minimal, diukur dengan metrik MAE (Mean Absolute Error), MSE (Mean Squared Error), RMSE (Root Mean Squared Error), dan R2 (R-squared). Dengan demikian, penelitian ini bertujuan untuk menyediakan hasil peramalan yang akurat dan dapat dipercaya mengenai pola kedatangan wisatawan di Kota Pekanbaru untuk membantu dalam perencanaan dan pengelolaan sektor pariwisata.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahap Penelitian

Penelitian ini menerapkan kerangka metodologi CRISP-DM, yang menguraikan enam tahap penting dalam pelaksanaan model data mining [21]. Ke-enam tahapan-tahapan tersebut yaitu Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment [21], [22]. Tetapi dalam penelitian ini hanya menerapkan lima tahapan yang dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian



## 2.1.1 Business Understanding

Pada tahap pertama dari CRISP-DM yang berfokus pada Business Understanding, tujuannya adalah memahami kebutuhan secara menyeluruh berdasarkan evaluasi bisnis. Informasi ini kemudian digunakan untuk merancang rencana awal data mining yang bertujuan mencapai tujuan yang telah ditetapkan[23]. Tujuan dari penelitian ini adalah untuk mengidentifikasi pola kunjungan wisatawan guna mengetahui jumlah wisatawan di masa yang akan datang ke kota Pekanbaru dengan tingkat error yang rendah, menggunakan metode SVR.

## 2.1.2 Data Understanding

Langkah Data Understanding dalam kerangka CRISP-DM melibatkan eksplorasi mendalam dan pemahaman menyeluruh terhadap dataset penelitian. Fleksibilitas dan keterukuran dalam sumber data serta proses persiapannya memudahkan data analis dalam memilih sumber yang valid dan kredibel sebelum memasuki tahap pemodelan[24]. Data yang digunakan pada penelitian ini termasuk dalam kategori Sekunder, yang artinya data tersebut diperoleh dari pihak ketiga. Data yang didapat dari Dinas Kebudayaan dan Pariwisata kota Pekanbaru, yang mencakup data time series mengenai kunjungan wisatawan Nusantara dari tahun 2016 hingga tahun 2023. Dengan memanfaatkan dataset ini, penelitian dapat mengakses sumber informasi yang signifikan dan relevan untuk keperluan analisis serta peramalan jumlah kedatangan wisatawan.

## 2.1.3 Data Preparation

Langkah Data Preparation dalam kerangka CRISP-DM melibatkan serangkaian langkah untuk memilih dan menyiapkan data yang akan digunakan pada tahap berikutnya. Proses dimulai dengan melakukan pembersihan data, yang merupakan tahap penting dalam analisis data, bertujuan untuk memastikan bahwa data yang digunakan dalam analisis atau pemodelan memiliki tingkat akurasi, konsistensi, dan kebebasan dari anomali yang optimal. Pembersihan data yang dilakukan seperti penanganan terhadap missing value, menangani outliers dan menangani data duplikat. Transformasi data adalah proses mengubah atau memanipulasi data asli untuk membuatnya lebih cocok atau lebih berguna untuk analisis atau pemodelan tertentu. Pada penelitian ini akan mentransformasi data menggunakan Normalize to interval [0,1] dengan persamaan (1).

$$X_{\text{norm}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \quad (1)$$

Dalam konteks normalisasi data,  $X_{\text{norm}}$  merujuk pada nilai yang telah disesuaikan ke dalam rentang tertentu, umumnya antara 0 dan 1, untuk memudahkan analisis dan perbandingan data. Nilai  $X$  adalah nilai asli dari data sebelum normalisasi dilakukan.  $X_{\text{min}}$  dan  $X_{\text{max}}$  mewakili nilai terkecil dan terbesar dari data tersebut, yang digunakan sebagai acuan dalam proses normalisasi.

Berikut langkah selanjutnya, yaitu menerapkan teknik Sliding Window dengan membuat window size 3 hingga window size 13. Tujuan dari penggunaan Sliding Window adalah untuk mengurangi kesalahan aproksimasi, seperti jarak Euclidean atau jarak vertikal antara aproksimasi yang diperoleh dengan data time series yang sebenarnya. Jumlah parameter yang digunakan dalam time series merepresentasikan kesalahan batas ini[23].

## 2.1.4 Modeling

Tahap Modeling dalam kerangka CRISP-DM adalah saat Machine Learning menentukan teknik data mining dan tools yang akan digunakan. Dalam penelitian ini, metode yang dipilih adalah SVR untuk melakukan peramalan dan menggunakan tools Orange Data Mining. Langkah pentingnya adalah memisahkan dataset menjadi fitur ( $X$ ) dan target ( $y$ ). Fitur ( $X$ ) adalah variabel input model yang digunakan untuk memprediksi target. Selanjutnya Proses pembagian dataset menjadi data training dan testing menggunakan teknik K-Fold Cross Validation. Pemilihan tipe kernel seperti Kernel Linear, Kernel polinomial dengan parameter gamma, Cost, Degree (auto, 2.00, 1.0), dan RBF dengan parameter gamma (auto) merupakan keputusan krusial dalam SVR. Model terbaik dipilih melalui Grid Search untuk menemukan kombinasi hyperparameter terbaik.

## 2.1.5 Evaluation

Tahap Evaluation untuk memastikan kualitasnya yang baik dari sudut pandang analisis data menggunakan Test and Score untuk menganalisis hasil prediksi model[25]. Test and Score digunakan untuk perhitungan error yaitu MAE, MSE, RMSE dan R2.

## 2.2 Support Vector Regression

Vapnik bersama Boser & Guyon memperkenalkan SVM pada tahun 1992, yang merupakan pengembangan dari teori komputasi sebelumnya[26]. Ketika diterapkan dalam regresi, SVM menjadi SVR[20]. SVR berakar pada prinsip risk minimization dan menghasilkan prediksi berupa angka riil[19]. Rumus umum untuk prediksi SVR dapat dilihat pada persamaan 2.

$$f(x) = w^T x + b \quad (2)$$



Dalam konteks prediksi, fungsi  $f(x)$  mewakili nilai prediksi yang dihasilkan oleh model. Bobot vector  $w$  digunakan untuk memberikan nilai penting pada setiap fitur yang ada dalam fitur vector  $x$ . Setiap elemen dalam vector  $x$  mewakili satu fitur yang digunakan dalam proses prediksi. Sedangkan  $b$  merupakan indeks atau nomor baris tertentu dalam data yang sedang diproses. Dengan menggunakan bobot vektor  $w$  dan fitur vektor  $x$ , nilai prediksi  $x$  untuk baris atau data  $b$  dalam dataset yang sedang diolah dapat dihitung. sementara untuk kasus regresi, Batasan margin didefinisikan pada persamaan 3 dan 4.

$$y - (w^T x + b) \leq \epsilon \tag{3}$$

$$(w^T x + b) - y \leq \epsilon \tag{4}$$

Dalam konteks penelitian, sering kali digunakan notasi untuk menyatakan konsep-konsep penting. Misalnya, dalam sebuah model prediktif, simbol "y" digunakan untuk merepresentasikan nilai target aktual, yaitu nilai yang ingin diprediksi. Selain itu, simbol "ε" digunakan untuk menunjukkan margin kesalahan (epsilon), yang merupakan ukuran seberapa jauh prediksi dari nilai aktual yang sebenarnya. Margin kesalahan ini penting untuk dievaluasi dalam menilai seberapa baik model dapat memprediksi data. Dengan memahami dan menggunakan notasi-notasi ini dengan tepat, analisis dan penelitian di berbagai bidang dapat dilakukan dengan lebih efektif. Tujuan dari SVR adalah untuk meminimalkan norma dari atas bobot  $w$  sekaligus memastikan bahwa seluruh data pelatihan berada di dalam atasan marginal. Fungsi tujuan untuk SVR dapat dilihat pada persamaan 5.

$$\text{Min } \frac{1}{2} \|w\|^2 \tag{5}$$

Fungsi kernel yang umum digunakan dalam SVR antara lain Kernel Linear untuk data yang terpisah secara linear, Kernel Polinomial untuk proyeksi data ke ruang fitur yang lebih tinggi, dan Kernel RBF untuk menangani hubungan nonlinear antar data. Penelitian ini menggunakan fungsi kernel yaitu Linear, RBF dan Polinomial pada persamaan 6, 7 dan 8.

**Kernel Linear**

$$K(x, x') = x^T x' \tag{6}$$

Fungsi kernel, disimbolkan dengan  $K$ , memiliki peran penting dalam analisis data karena memungkinkan pengukuran dan manipulasi hubungan antara vector fitur  $X$  dan  $x'$ . Dalam konteks ini,  $X$  dan  $x'$  mewakili vector-vector fitur yang digunakan dalam proses analisis data.

**Kernel Polinomial**

$$K(x, x') = (x^T x' + c)^d \tag{7}$$

Dalam konteks analisis kernel, simbol  $K$  merujuk pada fungsi kernel yang digunakan untuk mengukur kesamaan atau jarak antara dua vektor fitur, yang disimbolkan sebagai  $x$  dan  $x'$ . Parameter  $c$  merupakan konstanta yang mempengaruhi trade-off antara margin dan kesalahan klasifikasi dalam model. Sementara itu,  $d$  merupakan derajat dari polinomial yang digunakan dalam fungsi Kernel Polinomial, yang menentukan kompleksitas model dalam memodelkan hubungan antara data.

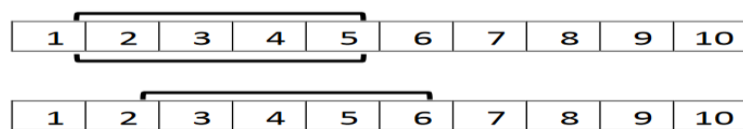
**Kernel RBF**

$$K(x, x') = \exp(-\gamma \|x - x'\|^2) \tag{8}$$

Fungsi kernel  $K$  dalam konteks ini merujuk pada sebuah fungsi matematis yang digunakan dalam berbagai metode machine learning, terutama dalam pembelajaran SVR. Notasi  $x$  dan  $x'$  mewakili vector fitur, yang merupakan representasi dari data dalam ruang fitur. Parameter  $\gamma$ , yang juga dikenal sebagai parameter kernel, mempengaruhi bentuk dan kompleksitas fungsi kernel yang digunakan dalam pemetaan data ke dalam ruang fitur yang lebih tinggi. Dengan mengubah nilai  $\gamma$ , nilai kompleksitas pemetaan tersebut dapat diatur, yang pada akhirnya dapat mempengaruhi kinerja dan generalisasi dari model machine learning yang digunakan.

**2.3 Sliding Window**

Sliding window atau windowing adalah pembentukan struktur data time series dengan ukuran window dan segmen yang disesuaikan untuk meminimalkan kesalahan[27]. Contoh dari sliding window dengan ukuran 5 dapat diilustrasikan dari Gambar 2.



**Gambar 2.** Proses Sliding Window

Proses Sliding Window ukuran 5 pada Gambar 2 adalah metode analisis data berkelanjutan yang menggunakan Sliding Window untuk mengambil kelompok lima titik data pada setiap langkahnya. Misalnya,



dalam deret waktu {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}, jendela pertama akan berisi {1, 2, 3, 4, 5}, jendela kedua berisi {2, 3, 4, 5, 6}, dan seterusnya. Setiap window dimulai dari satu unit observasi dan digunakan untuk memprediksi unit berikutnya. Tujuannya adalah mengurangi kesalahan aproksimasi dengan mengatur parameter pada time series.

## 2.4 K-Fold Validation

Metode K-Fold Cross Validation adalah teknik validasi model yang umum digunakan dalam machine learning untuk mengukur kinerja model secara akurat, terutama pada dataset terbatas. Prosedurnya melibatkan pembagian dataset ke dalam k subset yang sama ukurannya, di mana model dilatih pada k-1 subset dan diuji pada subset yang tersisa dalam setiap iterasi[28]. Proses ini diulang hingga semua subset berfungsi sebagai subset pengujian sekali, sehingga validasi silang K-Fold dapat dilakukan[29]. Kinerja model diukur sebagai nilai rata-rata dari seluruh tahap pelatihan yang dilakukan. [30].

## 2.5 Perhitungan Error

Perhitungan error adalah perbandingan antara hasil prediksi dan data aktual, dengan beberapa matrik evaluasi yang umum digunakan MSE mengukur rata-rata dari selisih kuadrat antara nilai sebenarnya dan prediksi, Metode ini dapat menghasilkan kesalahan yang cenderung lebih kecil, namun ada juga kemungkinan terjadi perbedaan yang signifikan[31]. RMSE adalah akar kuadrat dari MSE akuratnya suatu metode estimasi kesalahan pengukuran ditandai dengan nilai RMSE yang kecil[32], sementara MAE mengukur rata-rata dari selisih absolut antara nilai sebenarnya dan prediksi[33]. R<sup>2</sup> mengukur tingkat keberhasilan regresi dalam memprediksi nilai, dengan nilai mendekati 1 menunjukkan keakuratan yang tinggi[34]. Rumus untuk masing-masing metrik telah diberikan. Penelitian ini menggunakan MSE, RMSE, MAE, dan R<sup>2</sup> pada persamaan 9, 10, 11, 12, 13 dan 14.

### MAE

$$MAE = \frac{1}{n} \sum_{i=1}^n |A_i - F_i| \quad (9)$$

### MSE

$$MSE = \frac{1}{n} \sum_{i=1}^n (A_i - F_i)^2 \quad (10)$$

### RMSE

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (A_i - F_i)^2} \quad (11)$$

### R<sup>2</sup>

$$SSE = \sum (A_i - F_i)^2 \quad (12)$$

$$SST = \sum (A_i - F)^2 \quad (13)$$

$$R^2 = \frac{SSE}{SST} \quad (14)$$

Dalam konteks ini, terdapat variabel A<sub>i</sub> yang mewakili data aktual ke-i, serta variabel F<sub>i</sub> yang merupakan nilai hasil prediksi untuk data ke-i. Selain itu, N digunakan untuk menunjukkan jumlah total data yang tersedia. Dengan menggunakan variabel A<sub>i</sub>, F<sub>i</sub>, dan N, metode analisis yang sesuai dapat diterapkan untuk mengevaluasi keakuratan prediksi terhadap data aktual. Dengan kata lain, A<sub>i</sub>, F<sub>i</sub> adalah dua konsep penting dalam analisis prediksi yang membantu dalam memahami tingkat akurasi suatu model atau metode dalam memprediksi nilai-nilai data.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Data Understanding

Penelitian ini menggunakan data Skunder yang langsung didapatkan dari Dinas Kebudayaan dan Pariwisata Kota Pekanbaru, yang mencakup informasi penting mengenai jumlah kunjungan wisatawan selama periode waktu tertentu. Dataset ini mencakup informasi tentang jumlah kunjungan wisatawan nusantara dari tahun 2016 hingga 2023. Total terdapat 96 entri dalam dataset yang menyimpan informasi penting terkait kunjungan wisatawan. Setiap entri mencakup atribut seperti tahun pencatatan kunjungan, bulan kunjungan, dan data jumlah pengunjung perorangan yang tercatat untuk bulan dan tahun tertentu. Detail lengkap mengenai atribut-atribut ini dapat diidentifikasi dan dipelajari lebih lanjut dalam Tabel 1.

**Tabel 1.** Dataset Jumlah Kunjungan Wisatawan ke Kota Pekanbaru

No	Tahun	Bulan	Pengunjung
1	2016	JANUARI	40.434



No	Tahun	Bulan	Pengunjung
2	2016	FEBRUARI	35.788
3	2016	MARET	39.298
4	2016	APRIL	39.047
5	2016	MEI	38.404
6	2016	JUNI	89.325
...	...	...	...
91	2023	JULI	113.388
92	2023	AGUSTUS	116.489
93	2023	SEPTEMBER	119.925
94	2023	OKTOBER	120.086
95	2023	NOVEMBER	128.228
96	2023	DESEMBER	159.801

**3.2 Data Preparation**

Pada tahap ini, dataset disiapkan untuk analisis selanjutnya dengan memastikan tidak adanya nilai duplikat, outlier, dan missing values yang dapat mengganggu validitas analisis. Hasil analisis menunjukkan bahwa tidak ada nilai duplikat, outlier dan missing values yang terdeteksi dalam dataset. Langkah selanjutnya adalah normalisasi pada kolom Pengunjung untuk memastikan bahwa semua atribut memiliki skala yang seragam dan tidak mendominasi atribut lainnya, dengan menggunakan metode normalize to interval [0, 1]. Hasil dari proses normalisasi ini terdokumentasi pada Tabel 2, yang memungkinkan evaluasi lebih lanjut terhadap konsistensi data dan persiapan yang telah dilakukan sebelum analisis lanjutan dilakukan.

**Tabel 2.** Dataset Hasil Normalisasi

No	Tahun	Bulan	Pengunjung
1	2016	JANUARI	0,213
2	2016	FEBRUARI	0,188
3	2016	MARET	0,207
4	2016	APRIL	0,206
5	2016	MEI	0,202
6	2016	JUNI	0,479
...	...	...	...
91	2023	JULI	0,221
92	2023	AGUSTUS	0,255
93	2023	SEPTEMBER	0,292
94	2023	OKTOBER	0,294
95	2023	NOVEMBER	0,382
96	2023	DESEMBER	0,726

Setelah normalisasi data, langkah berikutnya adalah transformasi data pada kolom pengunjung dengan menerapkan metode Sliding Window, yang merupakan proses pembentukan struktur dari data time series yang tersedia. Ukuran dan segmentasi dari jendela diubah untuk mengurangi kesalahan sekecil mungkin. Penelitian ini akan melakukan uji coba terhadap ukuran jendela yang optimal, mencakup 11 ukuran window size yang berbeda mulai dari window size 3 hingga window size 13. Misalnya, ukuran jendela 3 terdiri dari t2', t1', dan t, sedangkan ukuran jendela 13 terdiri dari t12', t11', t10', t9', t8', t7', t6', t5', t4', t3', t2', t1' dan t. Setelah data dijadikan window-windowing, langkah berikutnya adalah memisahkan data dengan menjadikan kolom t di setiap window menjadi variabel y dan kolom lain menjadi variabel X. Contoh dari penggunaan window size 3 dan pemisahan data dapat ditemukan dalam Tabel 3

**Tabel 3.** Window Size 3 dan Pemisahan Data

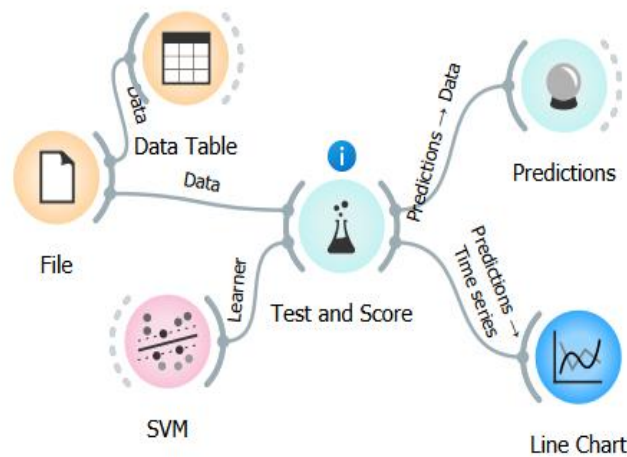
t2'	t1'	t'
X1	X2	y
0,213	0,188	0,207
0,188	0,207	0,206
0,207	0,206	0,202
0,206	0,202	0,479
0,202	0,479	0,197
...	...	...
0,461	1	0,601
1	0,601	0,617
0,601	0,617	0,610
0,617	0,610	0,627



t2'	t1'	t'
X1	X2	y
0,610	0,627	0,646
0,627	0,646	0,647
0,646	0,647	0,691
0,647	0,691	0,863

### 3.3 Modeling

Tahap pemodelan untuk meramalkan jumlah kunjungan wisatawan ke Kota Pekanbaru menerapkan metode SVR untuk membangun model peramalan. Proses ini dilakukan di dalam tools Orange Data Mining, yang menyediakan alat analisis dan teknologi, proses ini menghasilkan peramalan yang berguna untuk menginformasikan keputusan terkait manajemen pariwisata di Kota Pekanbaru. Perancangan widget yang ada dalam tools Orange Data Mining menggunakan model peramalan yaitu SVR seperti terlihat pada Gambar 3.



**Gambar 3.** Design Widget untuk Model Peramalan Jumlah Kedatangan Wisatawan

#### 3.3.1 Pembagian Data Training dan Data Testing

Dataset yang digunakan telah dibagi menjadi dua kelompok utama: data training dan data testing. Pendekatan yang diterapkan adalah metode K-Fold Cross Validation, di mana data kunjungan wisatawan dipisahkan secara proporsional. Untuk penggunaan metode SVR, dibagi menjadi 10 folds yang diatur pada widget Test and score seperti yang terlihat pada Gambar 3. Langkah ini merupakan praktik umum yang memungkinkan model untuk belajar dari sebagian besar data, sementara juga menyediakan data independen untuk menguji kinerjanya.

#### 3.3.2 Implementasi SVR

Setelah melakukan pembagian data training dan data testing selanjutnya akan dilakukan proses peramalan menggunakan metode SVR melalui tools Orange Data Mining, serangkaian percobaan dilakukan pada tiga jenis kernel, yaitu Linear, Polinomial, dan RBF yang diatur pada widget SVM yang dapat dilihat pada Gambar 3. Tujuan utama dari percobaan ini adalah untuk menemukan kernel yang optimal dalam peramalan jumlah kedatangan wisatawan ke Kota Pekanbaru. Melalui proses percobaan yang dilakukan, diharapkan dapat diidentifikasi kernel yang mampu memberikan hasil peramalan yang paling akurat dengan nilai error yang paling rendah, sehingga memungkinkan untuk peramalan yang lebih efektif terkait jumlah wisatawan yang akan datang.

### 3.4 Evaluation

Langkah berikutnya adalah membandingkan hasil evaluasi antara Kernel Linear, Polinomial, dan RBF menggunakan Widget Test and Score dalam tools Orange Data Mining. Proses perbandingan ini diperlukan untuk mengukur tingkat keberhasilan dari masing-masing kernel pada metode SVR dalam konteks analisis data. Evaluasi tersebut mencakup parameter-parameter seperti MAE, MSE, RMSE, dan R2 yang penting dalam mengevaluasi performa model peramalan. Widget Test and Score dapat menampilkan hasil perhitungan dan membandingkan hasil peramalan antara kernel-kernel yang digunakan, sehingga dapat dipilih kernel yang paling optimal untuk tujuan peramalan tertentu. Hasil dari percobaan yang dilakukan terdapat pada Tabel 4.

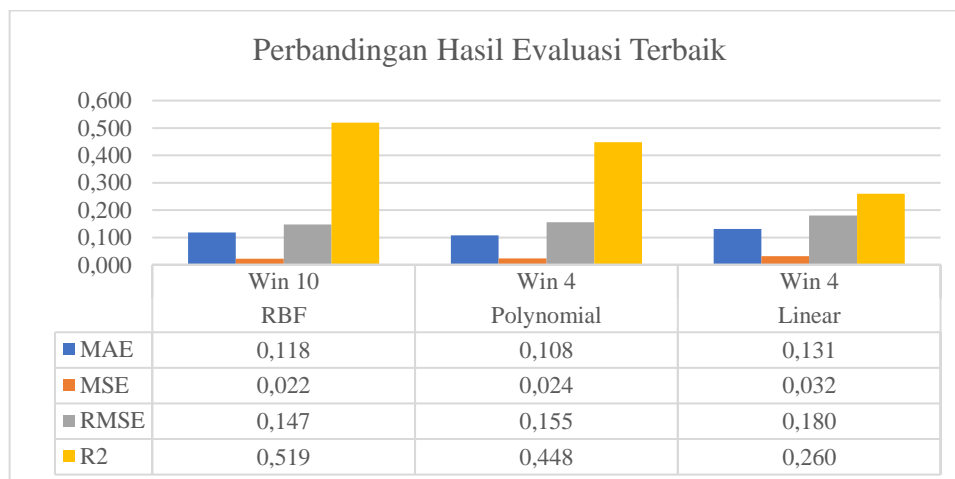
**Tabel 4.** Hasil Percobaan Menggunakan Tiga Jenis Kernel RBF, Polinomial, dan Linear

Kernel	Window	MAE	MSE	RMSE	R2
RBF	Win 3	0,119	0,030	0,174	0,304
	Win 4	0,126	0,031	0,177	0,280



Kernel	Window	MAE	MSE	RMSE	R2	
RBF	Win 5	0,122	0,027	0,122	0,388	
	Win 6	0,130	0,030	0,174	0,309	
	Win 7	0,114	0,022	0,147	0,512	
	Win 8	0,111	0,023	0,152	0,485	
	Win 9	0,133	0,030	0,172	0,338	
	<b>Win 10</b>	<b>0,118</b>	<b>0,022</b>	<b>0,147</b>	<b>0,519</b>	
	Win 11	0,130	0,029	0,170	0,357	
	Win 12	0,123	0,026	0,161	0,430	
	Win 13	0,119	0,026	0,119	0,445	
	Polinomial	Win 3	0,120	0,029	0,170	0,335
		<b>Win 4</b>	<b>0,108</b>	<b>0,024</b>	<b>0,155</b>	<b>0,448</b>
		Win 5	0,115	0,027	0,165	0,377
		Win 6	0,126	0,032	0,179	0,271
Win 7		0,109	0,025	0,159	0,433	
Win 8		0,118	0,026	0,164	0,407	
Win 9		0,117	0,027	0,163	0,404	
Win 10		0,115	0,026	0,115	0,422	
Win 11		0,120	0,027	0,164	0,408	
Win 12		0,122	0,029	0,172	0,354	
Win 13		0,122	0,028	0,122	0,384	
Linear		Win 3	0,139	0,036	0,191	0,164
		<b>Win 4</b>	<b>0,131</b>	<b>0,032</b>	<b>0,180</b>	<b>0,260</b>
	Win 5	0,136	0,033	0,181	0,254	
	Win 6	0,133	0,034	0,184	0,228	
	Win 7	0,146	0,037	0,193	0,160	
	Win 8	0,144	0,038	0,196	0,137	
	Win 9	0,137	0,035	0,187	0,214	
	Win 10	0,132	0,034	0,183	0,253	
	Win 11	0,156	0,040	0,201	0,106	
	Win 12	0,178	0,058	0,240	-0,263	
	Win 13	0,177	0,048	0,219	-0,040	

Setelah melakukan serangkaian uji coba sebanyak 33 kali dengan tiga jenis kernel pada model peramalan jumlah kunjungan wisatawan menggunakan metode SVR. Maka di peroleh hasil analisis menunjukkan keunggulan kernel RBF pada window size 10. Kernel RBF menunjukkan performa yang sangat baik dengan nilai MAE sebesar 0,118, MSE 0,022, dan R2 mencapai 5,19. Di samping itu, Kernel Polinomial menunjukkan kinerja terbaik pada window size 4, dengan nilai MAE 0,108, MSE 0,024, dan R2 0,448. Kernel Linear pada window size 4 juga memberikan hasil yang baik, dengan nilai MAE 0,131, MSE 0,032, dan R2 0,260. Penilaian menyeluruh ini sangat penting untuk memilih kernel terbaik yang cocok dengan kebutuhan model peramalan. Evaluasi ini dilakukan dengan cermat dan didukung oleh visualisasi yang rinci (lihat Gambar 4), yang menjadi dasar untuk menentukan kernel terbaik dalam memperkirakan jumlah wisatawan dengan tingkat akurasi yang diinginkan.



**Gambar 4.** Perbandingan Hasil Evaluasi Terbaik di Setiap Kernel

Gambar 4 menunjukkan bahwa hasil evaluasi MAE pada Kernel Polinomial window size 4 lebih rendah yaitu 0,108 di bandingkan Kernel RBF dan Kernel Linear dengan nilai 0,118 dan 0,131. Namun, saat menggunakan





window size 10 pada Kernel RBF lebih optimal dalam peramalan jumlah kedatangan wisatawan. Hal ini terlihat dari nilai yang lebih rendah untuk nilai MSE, dan RMSE, yaitu masing-masing sebesar 0,022, dan 0,147, dibandingkan dengan penggunaan window lainnya pada kernel RBF dan Linear. Nilai-nilai yang lebih rendah ini menunjukkan tingkat kesalahan yang lebih kecil dalam peramalan, sehingga hasilnya lebih akurat. Selain itu, perbandingan nilai R<sup>2</sup> pada window size 10 untuk RBF menunjukkan kinerja yang lebih unggul, dengan nilai mencapai 0,519. Hal ini mengindikasikan bahwa model menggunakan Kernel RBF mampu menjelaskan variasi data dengan lebih baik, memberikan prediksi yang lebih akurat terhadap jumlah kedatangan wisatawan ke Kota Pekanbaru.

### 3.5 Hasil Peramalan Metode SVR

Setelah menetapkan Kernel terbaik, hasil peramalan menggunakan Metode SVR disusun dengan data pengunjung aktual dan perubahan data pengunjung, yang dapat dilihat pada Tabel 5.

**Tabel 5.** Hasil Peramalan menggunakan Metode SVR

Data Pengunjung	Perubahan Data Pengunjung Hasil Peramalan SVR
0,237	0,256
0,188	0,296
0,185	0,316
0,237	0,255
0,228	0,286
0,185	0,353
...	...
0,491	0,356
0,199	0,257
0,516	0,222
0,445	0,522
0,529	0,570
0,779	0,593
0,610	0,520

Tabel 5 merupakan hasil peramalan perubahan data pengunjung yang diperoleh melalui penerapan metode SVR, yang dapat diakses melalui widget prediction pada Gambar 3. Data yang ditampilkan dalam tabel ini merupakan contoh sampel dari perubahan data pengunjung aktual dan hasil peramalan perubahan data pengunjung menggunakan SVR. Meskipun hanya beberapa data yang ditampilkan di sini, tabel ini memberikan gambaran bahwa perbedaan antara perubahan data pengunjung aktual dan hasil peramalan dari metode SVR tidak signifikan secara statistik. Namun, penting untuk dicatat bahwa hasil peramalan data pengunjung masih dalam bentuk normalisasi. Dengan demikian, langkah berikutnya yang diambil adalah melakukan denormalisasi data guna mendapatkan jumlah pengunjung yang sebenarnya. Hasil dari denormalisasi data peramalan dapat ditemukan pada Gambar 4.



**Gambar 4.** Visualisasi Data Hasil Peramalan

Grafik pada Gambar 4 hasil peramalan menggunakan metode SVR bulan Januari tahun 2024 hingga Maret 2031 yang dijelaskan pada Sumbu Horizontal, sedangkan Sumbu Vertikal akan mewakili jumlah pengunjung



perorangan. Hasilnya menunjukkan jumlah kunjungan pariwisata tertinggi terjadi pada bulan Mei 2028 sementara kunjungan terendah terjadi pada bulan Mei 2027.

#### 4. KESIMPULAN

Hasil Implementasi Orange Data Mining untuk meramalkan jumlah kedatangan wisatawan ke Kota Pekanbaru dengan metode SVR berbasis Sliding Window dapat dilihat pada widget Test and Score yang menggabungkan hasil MAE, MSE, RMSE dan R2 dari masing-masing kernel. Setelah melakukan 33 percobaan menunjukkan bahwa Kernel RBF dengan window size 10 lebih optimal dibandingkan dengan Kernel Polinomial dan Linear, karena memiliki nilai MAE, MSE, dan RMSE yang lebih rendah, yaitu 0,118, 0,022, dan 0,147. Selain itu, perbandingan nilai R2 juga menunjukkan performa yang lebih baik, dengan nilai 0,519. Hasil Peramalan didapat dari model terbaik yaitu data junjungan wisatawan dari bulan Januari tahun 2024 hingga Maret 2031. Dalam konteks penelitian lanjutan, bisa dilakukan dengan menggunakan metode prediksi lainnya untuk membandingkan metode SVR atau melakukan optimasi dengan metode lainnya untuk meramalkan jumlah kedatangan wisatawan menggunakan tools Orange Data Mining atau menerapkan tools analisis lainnya yang relevan. Hal ini bertujuan untuk mengukur dan membandingkan kinerja berbagai metode prediksi dengan lebih komprehensif, serta memperluas pemahaman terhadap penggunaan algoritma-algoritma yang dapat memberikan hasil peramalan yang optimal dalam konteks perencanaan dan pengelolaan pariwisata di Kota Pekanbaru.

#### UCAPAN TERIMAKASIH

Terima kasih disampaikan kepada Dinas Kebudayaan dan Pariwisata Kota Pekanbaru atas kesempatan dan dukungan yang diberikan dalam pelaksanaan penelitian ini. Peneliti juga mengucapkan terima kasih kepada Keluarga besar Sistem Informasi dan teman-teman yang turut mendukung penelitian ini.

#### REFERENCES

- [1] M. K. Aidid and M. Nusrang, "Implementasi Metode Fuzzy Time Series untuk Peramalan Jumlah Pengunjung di Benteng Fort Rotterdam," vol. 2, no. 1, pp. 1–12, 2020, doi: 10.35580/variansium12895.
- [2] A. Ersan, A. Rahmawati, and D. H. Amrina, "Analisis sosial ekonomi masyarakat terhadap pemanfaatan taman lindung hutan mangrove di Desa Sidodadi Kec. Teluk Pandan Kab. Pesawaran Lampung," *Entrep. Bisnis Manaj. Akunt.*, pp. 102–112, 2022.
- [3] C. Rahmad, M. F. Ramadhani, and D. Puspitasari, "Peramalan Jumlah Kedatangan Wisatawan Mancanegara Dengan Menggunakan Metode Time Invariant Fuzzy Time Series (Studi Kasus : Wisata Kabupaten Pasuruan)," *J. Inform. Polinema*, vol. 4, no. 3, p. 195, 2018, doi: 10.33795/jip.v4i3.206.
- [4] N. Uli and W. Sulandari, "Jurnal Ilmiah Matematika Peramalan jumlah kedatangan wisatawan mancanegara ke bali menggunakan metode hibrida SSA-WFTS," vol. 8, no. 1, pp. 19–32, 2021.
- [5] F. Eddyono, *Pengelolaan destinasi pariwisata. uwa is inspirasi indonesia*, 2021.
- [6] Rahmawati, Yuniza, A. N. Rahma, and Zukrianto, "Prediksi Jumlah Wisatawan di Kota Pekanbaru pada Tahun 2019-2023 Dengan Menggunakan Metode Fuzzy Time Series Chen," *J. Pendidik. Mat.*, vol. 2, no. 1, pp. 36–44, 2020.
- [7] M. Wati and N. Puspitasari, "Penerapan Metode Fuzzy Time Series Chen dan Hsu dalam Memprediksi Kunjungan Wisatawan di Museum Mulawarman," vol. 4, no. 2, 2020.
- [8] I. J. Thira, N. A. Mayangky, D. N. Kholifah, I. Balla, and W. Gata, "Peramalan Data Kunjungan Wisatawan Mancanegara ke Indonesia menggunakan Fuzzy Time Series," vol. 5, no. 1, pp. 18–23, 2019.
- [9] D. Febrian, D. Kartika, D. Agnes, and J. Nainggolan, "Peramalan Jumlah Wisatawan Mancanegara Yang Datang Ke Sumatera Utara dengan Fuzzy Time Series," no. January 2019.
- [10] Y. D. W. I. PUTRA, "Implementasi Fuzzy Time Series Cheng Untuk Peramalan Titik Panas Di Provinsi Riau." Universitas Islam Negeri Sultan Syarif Kasim Riau, 2021.
- [11] N. R. N. Rizky, "Analisis Peramalan Tingkat Penghunian Kamar Pada Hotel & Resort Griya Inkoppabri." Fakultas Ekonomi Dan Bisnis Universitas Pakuan, 2023.
- [12] A. Wahyuni, P. Studi, S. Informasi, F. Teknologi, and U. N. Mandiri, "RANCANG BANGUN SISTEM INFORMASI ABSENSI KARYAWAN Perkembangan teknologi informasi berkembang secara cepat , efektif dan efisien . Dalam hal ini perkembangan dalam bidang komputer saat ini telah membuka peluang kepada pengambil keputusan baik dalam institusi," *J. Inform.*, no. February, pp. 27–33, 2022.
- [13] A. E. Radho, P. Sugiartawan, and G. A. Santiago, "Prediksi Jumlah Kasus COVID-19 Menggunakan Teknik Sliding Wondows dengan Metode BPNN," *J. Sist. Inf. dan Komput. Terap. Indones.*, vol. 4, no. 1, pp. 12–23, 2022, doi: 10.33173/jsikti.123.
- [14] X.-Q. Su, J.-L. An, Y.-X. Zhang, J.-S. Liang, J. Liu, and X. Wang, "Application of support vector machine regression in ozone forecasting," *Huan Jing ke Xue= Huanjing Kexue*, vol. 40, no. 4, pp. 1697–1704, 2019.
- [15] M. P. Raharyani, R. R. M. Putri, and B. D. Setiawan, "Implementasi Algoritme Support Vector Regression Pada Prediksi Jumlah Pengunjung Pariwisata," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 4, pp. 1501–1509, 2018.
- [16] A. Kurniawati and A. Arima, "Analisis Prediksi Harga Saham PT. Astra International Tbk Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA) dan Support Vector Regression (SVR)," *J. Ilm. Komputasi*, vol. 20, no. 3, pp. 417–423, 2021, doi: 10.32409/jikstik.20.3.2732.
- [17] M. HALILAH, "Perbandingan Peramalan Jumlah Penumpang Bus Damri Yogyakarta Menggunakan Metode Artificial



- Neural Network Dan Support Vector Regression.” Muhammadiyah University, Semarang, 2019.
- [18] U. N. Aeni, A. L. Prasati, and M. Kallista, “Grafik Jumlah Penumpang 2018,” vol. 7, no. 2, pp. 4919–4926, 2020.
- [19] N. P. N. Hendayanti, I. K. P. Suniantara, and M. Nurhidayati, “Penerapan Support Vector Regression (Svr) Dalam Memprediksi Jumlah Kunjungan Wisatawan Domestik Ke Bali,” *J. Varian*, vol. 3, no. 1, pp. 43–50, 2019, doi: 10.30812/varian.v3i1.506.
- [20] M. P. Rahayani, R. R. M. Putri, and B. D. Setiawan, “Implementasi Algoritme Support Vector Regression Pada Prediksi Jumlah Pengunjung Pariwisata,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 4, pp. 1501–1509, 2018.
- [21] C. Neto, M. Brito, V. Lopes, H. Peixoto, A. Abelha, and J. Machado, “Application of data mining for the prediction of mortality and occurrence of complications for gastric cancer patients,” *Entropy*, vol. 21, no. 12, p. 1163, 2019.
- [22] D. Ferreira, S. Silva, A. Abelha, and J. Machado, “Recommendation system using autoencoders,” *Appl. Sci.*, vol. 10, no. 16, p. 5510, 2020.
- [23] M. A. Hasanah, S. Soim, and A. S. Handayani, “Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir,” *J. Appl. Informatics Comput.*, vol. 5, no. 2, pp. 103–108, 2021, doi: 10.30871/jaic.v5i2.3200.
- [24] Y. A. Singgalen, “Penerapan Metode CRISP-DM dalam Klasifikasi Data Ulasan Pengunjung Destinasi Danau Toba Menggunakan Algoritma Naïve Bayes Classifier (NBC) dan Decision Tree (DT),” *J. Media Inform. Budidarma*, vol. 7, no. 3, pp. 1551–1562, 2023, doi: 10.30865/mib.v7i3.6461.
- [25] M. Fajar, D. Mahdiana, A. Diana, and G. Triyono, “PERAMALAN HARGA SAHAM UBER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE UBER STOCK PRICE FORECASTING USING SUPPORT VECTOR,” vol. 2, no. April, pp. 304–308, 2023.
- [26] A. Emrul, M. Yamin, and L. Surimi, “Analisis Quality of Service (QoS) Kinerja Sistem Hotspot Pada Routerboard Mikrotik 951Ui-2HnD Pada Jaringan Teknik Informatika,” vol. 3, pp. 87–94, 2018.
- [27] R. E. Wahyuni, “Optimasi Prediksi Inflasi Dengan Neural Network Pada Tahap Windowing Adakah Pengaruh Perbedaan Window Size,” *Technol. J. Ilm.*, vol. 12, no. 3, p. 176, 2021, doi: 10.31602/tji.v12i3.5181.
- [28] R. Oktafiani, A. Hermawan, and D. Avianto, “Pengaruh Komposisi Split data Terhadap Performa Klasifikasi Penyakit Kanker Payudara Menggunakan Algoritma Machine Learning,” *J. Sains dan Inform.*, pp. 19–28, 2023.
- [29] P. I. Nainggolan, D. S. Prasvita, and D. S. Bukit, “Klasifikasi Informasi Kesehatan Pada Data Media Sosial Menggunakan Support Vector Machine dan K-Fold Cross Validation,” *Malikussaleh J. Mech. Sci. Technol.*, vol. 5, no. 2, pp. 34–38, 2021.
- [30] M. Syukron, R. Santoso, and T. Widiarhah, “Perbandingan metode smote random forest dan smote xgboost untuk klasifikasi tingkat penyakit hepatitis C pada imbalance class data,” *J. Gaussian*, vol. 9, no. 3, pp. 227–236, 2020.
- [31] A. N. Maulidaniar and E. Widodo, “Perbandingan Metode Peramalan Double Exponential Smoothing dan Triple Exponential Smoothing Pada Penjualan Indihome di Wilayah Telekomunikasi Cirebon,” *Emerg. Stat. Data Sci. J.*, vol. 1, no. 2, pp. 320–330, 2023, doi: 10.20885/esds.vol1.iss.2.art32.
- [32] D. Ramadhani, “Prediksi Harga Cryptocurrency Menggunakan Algoritma Prophet.” Universitas Medan Area, 2023.
- [33] M. Muharrom, “Analisis Komparasi Algoritma Data Mining Naive Bayes, K-Nearest Neighbors dan Regresi Linier Dalam Prediksi Harga Emas,” *Bull. Inf. Technol.*, vol. 4, no. 4, pp. 430–438, 2023, doi: 10.47065/bit.v4i4.986.
- [34] M. Devanda, H. Kusuma, and S. Hidayat, “Penerapan Model Regresi Linier dalam Prediksi Harga Mobil Bekas di India dan Visualisasi dengan Menggunakan Power Abstrak,” vol. 5, no. 2, pp. 1097–1110, 2024.