

---

## Peramalan Arus Kas dengan Pendekatan *Time Series* Menggunakan *Support Vector Machine*

Bella Audina<sup>1</sup>, Mohamat Fatekurohman<sup>2</sup>, dan Abduh Riski<sup>3</sup>  
<sup>1,2,3</sup>Jurusan Matematika, Universitas Jember

bellaaudina27@gmail.com

**Abstract.** Cash flow is a form of financial report that is used as a measure of the company success in the investment world. So that companies need to forecast the cash flow to manage their finances. Statistics can be applied for the forecasting of cash flow using the *Support Vector Machine* (SVM) method on the time series data. The aim of this research is to determine the optimal parameter pair model of the *Radial Basic Function* kernel and to obtain the forecasting results of cash flow using the SVM method on the time series data. The independent variable is needed the data on cash flow from operating income, expenditure and investment expenditure, sum of all cash flow. While the dependent variable is the financial condition based on the *Free Cash Flow*. The result of this research is a model with the best parameter pairs of the SVM tuning results with the greatest accuracy that is 75%, 82%, 88%, 64% and the forecasting financial condition of PT Cakrawala for the next 16 months.

**Keywords:** cash flow, forecasting, time series, support vector machine.

### 1. Pendahuluan

Arus kas adalah bentuk laporan keuangan yang dijadikan tolok ukur keberhasilan perusahaan di dunia investasi, sehingga perusahaan perlu melakukan peramalan arus kas untuk mengatur keuangannya. Statistika dapat diaplikasikan pada peramalan arus kas dengan menggunakan metode peramalan (*forecasting*). Peramalan adalah pengolahan data di masa lalu dengan metode matematika untuk memperkirakan kejadian di masa yang akan datang [1]. Sebagian besar masalah peramalan melibatkan penggunaan data deret waktu (*time series*). Deret waktu adalah pengamatan rangkaian data yang diukur selama kurun waktu tertentu [2]. Pendekatan *time series* sangat baik digunakan untuk menyelesaikan data linier, sedangkan model data non-linier seperti data keuangan arus kas dapat diselesaikan dengan pendekatan *Support Vector Machine* (SVM). SVM merupakan salah satu metode yang dapat menyelesaikan masalah peramalan dan mampu mengatasi *overfitting* sehingga dapat meningkatkan akurasi hasil prediksi. Dalam pemodelan SVM sangat penting memilih tipe kernel terbaik untuk menentukan kinerja prediksi *time series*. Kernel *Radial Basic Function* (RBF) dinilai lebih baik diterapkan jika ketergantungan antara data terdahulu dengan data sekarang bersifat non-linier [3]. Pendekatan *time series* baik dalam pemodelan linier dan SVM dengan kernel RBF bisa digunakan dalam pemodelan non-

---

linier, maka penelitian ini bertujuan untuk menentukan hasil peramalan arus kas dengan menggunakan metode SVM pada data *time series*.

## 2. Tinjauan Pustaka

**2.1. Arus Kas.** Menurut Kamus Besar Bahasa Indonesia (KBBI), arus kas adalah pemasukan dan pengeluaran dana perusahaan berdasarkan harian, mingguan, atau dalam rentang waktu tertentu. Laporan arus kas disusun untuk membantu tugas manajer keuangan perusahaan dan memberikan informasi mengenai perubahan kas perusahaan dengan *relevan* dalam periode waktu tertentu. Berdasarkan keperluan manajerial, laporan arus kas dibagi menjadi tiga kategori yaitu arus kas dari aktivitas operasi, investasi, dan pendanaan [4].

Arus kas aktivitas operasi diperoleh dari transaksi dan peristiwa lain yang memengaruhi penetapan laba atau rugi bersih perusahaan. Arus kas aktivitas investasi diperoleh dari kegiatan investasi yaitu aktiva tetap. Arus kas aktivitas pendanaan adalah arus kas yang berasal dari modal dan pinjaman perusahaan.

**2.2. Peramalan (*Forecasting*).** Peramalan atau *forecasting* adalah perkiraan dengan menggunakan teknik statistika dalam bentuk gambaran masa depan berdasarkan pengolahan angka secara historis. Terdapat dua pendekatan umum yang dilakukan pada saat peramalan yaitu metode kualitatif dan kuantitatif. Peramalan metode kuantitatif terdiri dari model deret waktu dan model asosiatif. Peramalan yang akan digunakan pada penelitian ini yaitu metode SVM dengan sistem klasifikasi. Salah satu metode yang dapat digunakan untuk mengukur kinerja klasifikasi yaitu *confusion matrix*.

Pada *confusion matrix* terdapat empat istilah sebagai representasi hasil klasifikasi yaitu *True Positive* (TP) yang merepresentasikan kelas positif terdeteksi benar, *True Negative* (TN) merepresentasikan data kelas negatif yang terdeteksi benar, *False Positive* (FP) merepresentasikan jumlah kelas negatif yang terdeteksi sebagai kelas positif dan *False Negative* (FN) merepresentasikan data kelas positif yang terdeteksi sebagai kelas negatif. Performa klasifikasi dapat dihitung nilai akurasi. Adapun perhitungan menentukan besar nilai akurasi sebagai berikut [5].

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} 100\% \quad (1)$$

**2.3. Deret Waktu (*Time Series*).** Deret waktu adalah salah satu pemodelan dalam penelitian dinamis yang bertujuan untuk mengumpulkan dan mempelajari pengamatan

deret waktu sebelumnya. Hal ini bertujuan untuk mengembangkan model yang menggambarkan struktur yang melekat dari deret tersebut. Deret waktu dibedakan menjadi analisis deret waktu dan peramalan deret waktu. Terdapat dua hal yang harus diperhatikan untuk kesempurnaan model deret waktu yang dihasilkan yaitu adanya autokorelasi dan data harus stasioner.

Uji stasioneritas bertujuan untuk menguji konsistensi pergerakan data deret waktu. Salah satu jenis uji stasioneritas yaitu uji *Augmented Dickey Fuller* (ADF). Uji ADF dilakukan dengan menggunakan paket “*ur.df*” yang sudah disediakan oleh aplikasi R-Studio. Adapun prosedur uji stasioneritas sebagai berikut.

Hipotesis:

$H_0$ : data tidak stasioner

$H_1$ : data stasioner

Statistik uji antara nilai statistik- $t$  dan nilai kritis ADF sebagai berikut [6].

1. Jika nilai statistik- $t <$  nilai kritis, maka tolak  $H_0$
2. Jika nilai statistik- $t >$  nilai kritis, maka terima  $H_0$

Adapun persamaan umum uji stasioneritas sebagai berikut.

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \alpha_i \sum_{i=1}^m \Delta Y_{t-1} + u_t \quad (2)$$

dengan

$Y_t$  : data *time series*

$t$  : trend deterministic

$\beta_1, \beta_2$  : intersep / konstanta

$m$  : panjang lag

$\delta$  : 0 (jika memiliki akar unit) dan  $\delta$ : 1 (jika tidak memiliki akar unit)

**2.4. Support Vector Machine (SVM) pada Time Series.** SVM adalah salah satu metode *machine learning* yang digunakan untuk memprediksi klasifikasi menggunakan konsep pencarian *hyperplane* dengan margin maksimum. SVM mudah diaplikasikan pada kasus data *non linier* dengan menggunakan konsep kernel pada ruang kerja berdimensi tinggi. Salah satu contoh fungsi kernel yang sering digunakan yaitu *Radial Basic Function* (RBF). Adapun persamaan umum fungsi kernel sebagai berikut.

$$K(\vec{x}_i, \vec{x}_j) = \phi(\vec{x}_i) \cdot \phi(\vec{x}_j) \quad (3)$$

dimana  $\vec{\phi}(x_i)$  adalah fungsi data  $x_i$ ,  $\vec{\phi}(x_j)$  adalah fungsi data  $x_j$ , dan  $j$  merupakan ruang vektor yang lebih tinggi dari  $i$ .

Fungsi kernel di atas dapat digunakan pada fungsi kernel RBF yaitu [7].

$$K(\vec{x}_i, \vec{x}_j) = \exp(-\gamma \| \vec{x}_i - \vec{x}_j \|^2) \quad (4)$$

dimana  $x_i$  dan  $x_j$  adalah sebuah data *time series*, sedangkan  $\gamma$  adalah nilai kernel RBF yang optimal.

**2.5. Metode Grid Search.** Metode *grid search* digunakan untuk menentukan parameter optimal pada model SVM dengan melatih pasangan parameter yang menghasilkan akurasi terbaik. Penentuan parameter SVM terbaik dengan metode *grid search* menggunakan kernel trick. Adapun jenis kernel trick antara lain linier, polynomial, RBF, dan sigmoid. Jenis kernel trik yang digunakan dalam penelitian ini adalah kernel RBF. Jenis parameter yang digunakan pada SVM dengan kernel RBF yaitu *cost* ( $c$ ) dan *gamma* ( $\gamma$ ) dengan persamaan

$$c, \gamma = \{i, j \mid i \in c \text{ dan } j \in \gamma\} \quad (5)$$

dengan

$i$  : anggota himpunan  $c$

$j$  : anggota himpunan  $\gamma$

**2.6. K-Fold Cross Validation.** Uji *Cross-Validation* (CV) merupakan salah satu metode statistika untuk mengevaluasi kinerja algoritma dimana data akan dipisahkan menjadi dua subset yaitu data proses pembelajaran dan data validasi. Berdasarkan ukuran dataset, CV dibagi menjadi beberapa jenis dan salah satunya *k-fold cross-validation*. Salah satu model *k-fold cross-validation* yang sering digunakan yaitu *10 fold cross-validation*. Penelitian [8] membuktikan bahwa *10 fold cross-validation* mampu memberikan estimasi akurasi yang kurang bias dibandingkan dengan model lainnya. Penelitian [9] menghasilkan nilai akurasi terbesar yaitu  $k = 5$  atau *5 fold* untuk hasil pengujian terbaiknya.

### 3. Metode Penelitian

Jenis data yang digunakan pada penelitian ini adalah data primer yang diambil langsung dari PT Cakrawala dalam beberapa periode terakhir. Data bersumber dari data arus kas selama 56 bulan terakhir menggunakan metode kuantitatif dengan penelitian

---

survei. Variabel yang digunakan pada penelitian ini adalah sebagai berikut :

1. Variabel terikat ( $Y$ ) pada penelitian ini adalah kondisi keuangan arus kas perusahaan berdasarkan data *Free Cash Flow*.
2. Variabel bebas ( $X$ ) pada penelitian ini yaitu data keuangan pemasukan arus kas operasi, pengeluaran arus kas operasi, pengeluaran arus kas investasi, dan jumlah dari ketiga jenis arus kas tersebut.

Adapun langkah-langkah yang dilakukan dalam penelitian ini sebagai berikut :

1. Studi literatur dengan cara mencari sumber referensi dari jurnal, skripsi, dan buku.
2. Pengambilan dan pengolahan data ke dalam bentuk matriks melalui program *excel*.
3. *Pre-processing* data dengan dilakukan uji stasioneritas menggunakan uji ADF.
4. Pembagian data yang meliputi data pelatihan (*training*) dan data pengujian (*testing*) dengan persentase 70:30 dan 80:20 menggunakan paket “*caret*” dengan *function* *createDataPartition*.
5. Pemodelan dengan SVM dimulai dari *training* SVM untuk mendapatkan model SVM dengan  $k = 5$  dan  $k = 10$ .
6. Melakukan *tuning* parameter untuk mencari model SVM dengan pasangan parameter terbaik dari hasil uji data *training*.
7. Melakukan pengujian model SVM terbaik pada data *testing*.
8. Melakukan analisis hasil dan menentukan hasil prediksi atau peramalan data keuangan arus kas beberapa periode selanjutnya.

#### 4. Hasil dan Pembahasan

**4.1. Analisis Variabel.** Variabel yang digunakan dalam penelitian ini yaitu variabel bebas ( $X$ ) dan variabel terikat ( $Y$ ). Variabel bebas ( $Y$ ) berupa pemasukan arus kas operasi, pengeluaran arus kas operasi, pengeluaran arus kas investasi, dan jumlah dari ketiga arus kas tersebut. Data pengeluaran arus kas operasi meliputi biaya listrik, Wi-Fi, PDAM, parkir, tol, materai, ATK, ongkos kirim, pajak, zakat, makan siang, dan gaji karyawan. Data pengeluaran arus kas investasi meliputi biaya pembelian kendaraan (mobil dan sepeda motor), mebel, sekat kantor, komputer, mesin jahit, AC dan pembayaran BPJS ketenagakerjaan. Variabel terikat berisi data *Free Cash Flow* (FCF) yang dimiliki PT Cakrawala dengan transformasi *output* 0 untuk kelas kondisi keuangan buruk dan 1 untuk kelas kondisi keuangan baik. Data diinputkan ke dalam program R-Studio dengan paket

“*read.csv*”.

**4.2. Uji Stasioneritas Data.** Data yang telah diinputkan ke dalam program *R-Studio* akan dikelompokkan menjadi empat variabel yaitu *pemasukan\_operasi* mendefinisikan data pemasukan arus kas operasi, *pengeluaran\_operasi* mendefinisikan data pengeluaran arus kas operasi, *pengeluaran\_investasi* mendefinisikan data pengeluaran arus kas investasi, dan *data\_FCF* mendefinisikan data FCF. Kemudian masing-masing data akan diuji kestasioneritasannya dan hasilnya dapat dilihat pada Tabel 1.

Tabel 1. Hasil Uji Stasioneritas Data

| Variabel               | <i>p-value</i>         | Statistik- <i>t</i> | Nilai Kritis ADF |       |       | Keputusan Hipotesis |
|------------------------|------------------------|---------------------|------------------|-------|-------|---------------------|
|                        |                        |                     | 1%               | 5%    | 10%   |                     |
| Pemasukan_ operasi     | $1,724 \times 10^{-8}$ | -4,4692             |                  |       |       | Stasioner           |
| Pengeluaran_ operasi   | $3,145 \times 10^{-5}$ | -3,4387             |                  |       |       | Stasioner           |
| Pengeluaran_ investasi | $4,776 \times 10^{-6}$ | -3,6209             | -3,51            | -2,89 | -2,58 | Stasioner           |
| Arus_kas               | $2,317 \times 10^{-8}$ | -4,613              |                  |       |       | Stasioner           |
| Data_FCF               | $7,301 \times 10^{-7}$ | -4,3844             |                  |       |       | Stasioner           |

Tabel 1 menunjukkan bahwa semua variabel data stasioner karena nilai statistik-*t* < nilai kritis ADF sehingga keputusan hipotesis yaitu tolak  $H_0$  dan terima  $H_1$ . Data dengan variabel *pemasukan\_operasi*, *pengeluaran\_investasi*,  *arus\_kas* dan *data\_out* stasioner pada taraf signifikan 1%, 5%, dan 10%. Sedangkan variabel *pengeluaran\_operasi* stasioner pada taraf signifikan 5% dan 10%.

**4.3. Peramalan SVM.** Pada penelitian ini dibentuk tiga *data frame* dari penggabungan masing-masing variabel bebas dengan variabel terikat. Selanjutnya masing-masing *data frame* akan dipartisi menjadi 70% data *training* dan 30% data *testing* serta 80% data *training* dan 20% data *testing*. Pembagian data menggunakan paket “*caret*” yang sudah disediakan oleh program *R-Studio* dan menghasilkan 40 data *training* dan 16 data *testing* serta 45 data *training* dan 11 data *testing*.

**4.3.1. Peramalan SVM Data Training.** Peramalan SVM pada data *training* menggunakan paket “*e1071*” dan metode *k-fold cross validation* dengan nilai *k* sebesar 5 dan 10. Hasil perhitungan peramalan data pemasukan operasi menghasilkan tingkat akurasi sebesar 85% untuk 40 data *training* dan 84,44% untuk 45 data *training*. Sedangkan hasil perhitungan peramalan data pengeluaran operasi menghasilkan tingkat akurasi sebesar 85% untuk 40 data *training* dan 82,22% untuk 45 data *training*. Hasil perhitungan peramalan data pengeluaran investasi menghasilkan tingkat akurasi sebesar 82,5% untuk 40 data *training* dan 84,44% untuk 45 data *training*. Dan hasil perhitungan jumlah arus kas menghasilkan tingkat akurasi sebesar 80% untuk 40 data *training* dan 86,67% untuk 45 data *training*.

**4.3.2. Pemodelan dan Penentuan Nilai Parameter.** *Ranges* nilai yang digunakan untuk parameter *cost* sebesar  $2^{-5}, 2^{-4}, 2^{-3}, \dots, 2^{13}, 2^{14}, 2^{15}$  dan parameter *gamma* sebesar  $2^{-15}, 2^{-14}, 2^{-13}, \dots, 2^1, 2^2, 2^3$ . *Ranges* parameter tersebut ditentukan berdasarkan *ranges* yang sering digunakan dalam penelitian secara umum [10]. Penentuan nilai parameter pada penelitian ini menggunakan perpaduan metode *grid search* dan *k-fold cross validation* yang disebut juga *tuning SVM*. *Tuning* dilakukan sebanyak lima kali, kemudian dari hasil *tuning* dipilih pasangan parameter dengan nilai error terkecil. Pasangan parameter terbaik pada setiap data dapat dilihat pada Tabel 2.

Tabel 2. Pasangan Parameter Terbaik Setiap Data

| Variabel                     | Fold | Nilai Parameter |         | Nilai Error |
|------------------------------|------|-----------------|---------|-------------|
|                              |      | Cost            | Gamma   |             |
| <i>Pemasukan_operasi</i>     | 5    | 64              | 1       | 0,1494876   |
|                              | 10   | 4               | 2       | 0,1492745   |
| <i>Pengeluaran_operasi</i>   | 5    | 1024            | 0,125   | 0,233167    |
|                              | 10   | 0,5             | 2       | 0,2483137   |
| <i>Pengeluaran_investasi</i> | 5    | 64              | 4       | 0,1326296   |
|                              | 10   | 128             | 4       | 0,1236937   |
| <i>Arus_kas</i>              | 5    | 512             | 0,0625  | 0,1256163   |
|                              | 10   | 4096            | 0,03125 | 0,1221944   |

Dari hasil Tabel 2 hanya akan dipilih satu model SVM saja untuk meramalkan data arus kas. Sehingga model SVM untuk *pemasukan\_operasi* menggunakan *10 fold* dengan nilai parameter *cost* sebesar 4 dan *gamma* sebesar 2, *pengeluaran\_operasi*

menggunakan *5 fold* dengan nilai parameter *cost* sebesar 1024 dan *gamma* sebesar 0,125, dan *pengeluaran\_investasi* menggunakan *10 fold* dengan nilai parameter *cost* sebesar 128 dan *gamma* sebesar 4, serta *arus\_kas* menggunakan *10 fold* dengan nilai parameter *cost* sebesar 4096 dan *gamma* sebesar 0,03125.

**4.3.3. Peramalan Data Testing.** Pengujian ini dilakukan pada setiap data arus kas dengan jumlah data *testing* yang berbeda yaitu sebanyak 16 data untuk *pemasukan\_operasi* dan *pengeluaran\_operasi* serta 11 data untuk *pengeluaran\_investasi*. Pengujian data *testing* bertujuan untuk mengetahui tingkat akurasi model SVM sebelum model tersebut digunakan untuk meramalkan data arus kas. Model SVM dengan metode *10 fold* untuk *pemasukan\_operasi* memberikan kemampuan peramalan sebesar 75%, *pengeluaran\_investasi* mampu memberikan kemampuan peramalan sebesar 82%. Sedangkan model SVM dengan metode *5 fold* untuk *pengeluaran\_operasi* memberikan kemampuan peramalan sebesar 88% dan model SVM dengan metode *10 fold* untuk *arus\_kas* memberikan kemampuan peramalan sebesar 64%. Setelah mengetahui tingkat akurasi, selanjutnya dilakukan peramalan untuk mengetahui kondisi keuangan arus kas beberapa bulan selanjutnya menggunakan model SVM terbaik. Hasil peramalan dapat dilihat pada Tabel 3 dan 4.

Tabel 3. Hasil Peramalan Setiap Data Arus Kas

| Bulan     | Tahun | Hasil Peramalan |
|-----------|-------|-----------------|
| September | 2020  | Baik            |
| Oktober   | 2020  | Baik            |
| November  | 2020  | Baik            |
| Desember  | 2020  | Baik            |
| Januari   | 2021  | Baik            |
| Februari  | 2021  | Baik            |
| Maret     | 2021  | Baik            |
| April     | 2021  | Baik            |
| Mei       | 2021  | Baik            |
| Juni      | 2021  | Baik            |
| Juli      | 2021  | Baik            |
| Agustus   | 2021  | Baik            |
| September | 2021  | Baik            |

Lanjutan Tabel 3. Hasil Peramalan Setiap Data Arus Kas

| Bulan    | Tahun | Hasil Peramalan |
|----------|-------|-----------------|
| Oktober  | 2021  | Baik            |
| November | 2021  | Baik            |
| Desember | 2021  | Baik            |

Tabel 4. Hasil Peramalan Jumlah Arus Kas

| Bulan     | Tahun | Hasil Peramalan |
|-----------|-------|-----------------|
| September | 2020  | Baik            |
| Oktober   | 2020  | Baik            |
| November  | 2020  | Baik            |
| Desember  | 2020  | Baik            |
| Januari   | 2021  | Baik            |
| Februari  | 2021  | Buruk           |
| Maret     | 2021  | Baik            |
| April     | 2021  | Baik            |
| Mei       | 2021  | Baik            |
| Juni      | 2021  | Baik            |
| Juli      | 2021  | Baik            |

## 5. Kesimpulan

Model SVM dengan parameter hasil *tuning* untuk peramalan data pemasukan operasi menggunakan nilai parameter *cost* sebesar 4, *gamma* sebesar 2, dan metode *10 fold cross validation*. Sedangkan pada data pengeluaran operasi menggunakan nilai parameter *cost* sebesar 1024, *gamma* sebesar 0,125, dan metode *5 fold cross validation*. Data pengeluaran investasi menggunakan nilai parameter *cost* sebesar 128, *gamma* sebesar 4, dan metode *10 fold cross validation*. Data arus kas menggunakan nilai parameter *cost* sebesar 4096, *gamma* sebesar 0,03125, dan metode *10 fold cross validation*. Terdapat perbedaan hasil peramalan arus kas kondisi keuangan PT Cakrawala yaitu menghasilkan kondisi keuangan baik untuk peramalan setiap jenis data arus kas dan menghasilkan kondisi keuangan buruk dalam satu bulan untuk peramalan jumlah data arus kas.

---

**Daftar Pustaka**

- [1] Chapman and Hall. *Time-Series Forecasting*. United States of America: CRC Press. 2000.
- [2] Montgomery, D. C., C. L. Jennings, dan M. Kulahci. *Introduction to Time Series Analysis and Forecasting*. New Jersey: John Wiley & Sons, Inc. 2007.
- [3] Adhikari, R. *An Introductory on Time Series Modeling and Forecasting*. Germany: Lambert Academic Publishing. 2013.
- [4] Kakani, R. M. and N. Ramachandran. *How To Read A Cash Flow Statement*. India: Mcgraw Hill Edu Private Limited. 2010.
- [5] Han, J., M. Kamber and J. Pei. *Data Mining: Concept and Techniques, Third Edition*. Waltham: Morgan Kaufmann Publisher. 2012.
- [6] Rusdi. Deteksi Stasioneritas Data Runtun Waktu Melalui Uji Akar-Akar Unit. *Jurnal Saintek*. 3(1): 78-89. 2011.
- [7] Karaa dan Krichene. Credit Risk Assessment using Support Vector Machine and Multilayer Neural Network Models: A Comparative Study Case of a Tunisian Bank. *Journal of Accounting and Management Information System*. 11(4): 587-620. 2012.
- [8] Karyadiputra, E. dan N. Hijriana. Analisis Penerapan Algoritma Naive Bayes untuk Klasifikasi Prioritas Pengembangan Jalan di Provinsi Kalimantan Selatan. *Jurnal Teknologi*. 9(2): 105-108. 2018.
- [9] Eka, P. Implementasi *Support Vector Machine* (SVM) pada Prediksi Harga Saham Gabungan (IHSG). *Jurnal Teknologi dan Rekayasa*. 25(1): 24-38. 2020.
- [10] Andita, A. dan W. Sulistijanti. Perbaikan Peramalan Produksi Padi di Kabupaten Kendal dengan Menggunakan Metode *Support Vector Machine* (SVM). *The 7<sup>th</sup> URECOL2018*. 2018.