

PERAMALAN JUMLAH PERMINTAAN KORAN MENGGUNAKAN METODE SUPPORT VECTOR REGRESSION (SVR)

Alfin Karunia S.¹⁾, Rekyan Regasari M.P, S.T.,M.T²⁾, Rizal Setya P., S.Kom, M.Kom³⁾

Program Studi Informatika
Fakultas Ilmu Komputer

Universitas Brawijaya, Malang 65145, Indonesia

email : alfinkarunia[at]gmail.com¹⁾, rekyan.rmp [at]ub.ac.id²⁾, rizalespe[at]ub.ac.id³⁾

Abstrak

Kebutuhan masyarakat akan informasi semakin tahun semakin meningkat dan membuat beberapa media cetak, elektronik dan digital makin kompetitif dalam menyampaikan informasi. Dalam beberapa tahun ini, media digital berpengaruh pula pada jumlah pembaca media cetak. Hal ini menyebabkan semakin sulit bagi industri koran untuk meramalkan jumlah eksemplar yang akan diproduksi. Dengan adanya perencanaan produksi yang tepat akan dihasilkan efisiensi dalam produksi koran dan meningkatkan keuntungan perusahaan. Peramalan dapat menjadi acuan industri koran untuk menentukan jumlah produksi yang optimum. Banyak metode peramalan yang telah digunakan untuk meramalkan data yang bersifat non linier seperti Support Vector Regression (SVR) yang terbukti pada beberapa kasus peramalan dapat memberikan hasil peramalan yang baik, seperti pada kasus peramalan konsumsi air, harga real estate, nilai tukar Dollar dan peramalan beban listrik. Sesuai dengan pengujian yang dilakukan dengan menggunakan data jumlah penjualan koran Jawa Pos Radar Madura tahun 2015 menghasilkan MAPE dengan nilai 1.7916

Kata Kunci : Koran, SVR, MAPE.

Abstract

A community needs for information is increasing by years and make some print media, electronic media and digital media more competitive to deliver the information. During recent years, digital media also affects the number of readers of print media. This causes more difficult for the newspaper industry to predict the number of copies to be produced. With the right production planning it will be generated in newspaper production efficiency and improve the profitability. Forecasting can be a reference for newspaper industry to determine the optimum amount of production. There are many forecasting methods that have been used to predict the data that is non-linear as Support Vector Regression (SVR) which is proven in some cases of forecasting and it can provide that the forecasting results are good, such as in the case of the forecasting water consumption, real estate prices, the exchange rate of the dollar and forecasting electrical load. According to the tests performed by using data on the number of sales in Jawa Pos Radar Madura newspaper in 2015 resulted in MAPE with a value of 1.7916.

Keywords: Newspapers, SVR, MAPE.

1. PENDAHULUAN

Pers melakukan komunikasi kepada masyarakat dengan menggunakan media massa. Pers menyediakan informasi bagi masyarakat yang berada dalam intitusi sosial^[7]. Kebutuhan masyarakat akan informasi semakin tahun semakin meningkat dan membuat beberapa media cetak, elektronik dan digital pada era digital makin kompetitif dalam menyampaikan informasi. Dalam bentuk media cetak, koran merupakan media massa yang ada sejak ratusan tahun lalu^[8]. Koran dibutuhkan masyarakat karena mengandung banyak informasi dari berbagai bidang dan harganya pun terjangkau.

Semakin pesatnya era digital terutama adanya internet menjadi salah satu tantangan bagi industri koran dalam beberapa tahun ini. Keberadaan internet juga mempengaruhi strategi bisnis industri koran^[8]. Menurut hasil penelitian Roy Morgan 2015, tren global di Indonesia memperlihatkan peningkatan

pengguna media digital sebanyak 43 persen sedangkan koran sebanyak 34 persen. Namun minat pembaca koran di Kota Surabaya dan sekitarnya mencapai angka 52 persen yang mengalahkan jumlah pengguna media digital yang hanya 36 persen. Beberapa industri koran mengalami gulung tikar dalam sepuluh tahun terakhir^[4]. Kini industri koran banyak yang membuat konten berita melalui media digital yang isinya kurang lebih sama dengan media cetak. Menurut Taufiq Rizqon selaku direktur koran Suara Madura, terkadang sumber dari media digital yang disampaikan belum dikroscek secara mendalam.

Dengan adanya media digital berpengaruh pula pada jumlah pembaca media cetak, semakin sulit juga bagi industri koran untuk meramalkan jumlah eksemplar yang akan diproduksi. Dalam industri koran terdapat 2 macam permintaan, yaitu

permintaan tetap dan permintaan eceran. Industri koran sulit menentukan jumlah produksi dari permintaan eceran yang merupakan permintaan tidak tetap atau selalu berubah setiap harinya. Jumlah produksi suatu perusahaan menentukan laba atau rugi dari produk yang dihasilkan. Jika jumlah produksi lebih besar dari jumlah permintaan akan mengakibatkan kerugian pada industri koran, Dengan adanya perencanaan produksi yang tepat akan dihasilkan efisiensi dalam produksi koran dan meningkatkan keuntungan perusahaan. Peramalan dapat menjadi acuan industri koran untuk menentukan jumlah produksi yang optimum.

Banyak metode telah dikemukakan untuk mendapatkan hasil peramalan yang akurat. Dalam bidang informatika terdapat beberapa algoritma yang digunakan untuk peramalan atau yang biasa disebut regresi. Salah satu algoritma regresi yaitu menggunakan *Support Vector Regression*. *Support Vector Regression* termasuk metode *time series* yang menggunakan data di masa lampau untuk meramalkan data di masa mendatang. Dengan menggunakan komputasi, algoritma *Support Vector Regression* dapat mengeluarkan hasil peramalan. Penelitian sebelumnya tentang *Support Vector Regression* digunakan untuk meramalkan konsumsi listrik. Dalam penelitian tersebut, model SVR memungkinkan untuk peramalan karena *error RMS* relatif kecil. (Kavaglioglu, 2011). SVR juga digunakan untuk meramalkan harga real estate. Nilai error yang digunakan adalah MAPE. Hasil dari penelitian tersebut menunjukkan bahwa SVR lebih baik dan lebih efisien dari model BPNN^[13].

Berdasarkan permasalahan tersebut, solusi yang ditawarkan adalah peramalan menggunakan metode SVR dengan objek jumlah permintaan koran. Harapan dari implementasi ini dapat membantu industri koran dalam meramalkan jumlah permintaan yang mempengaruhi jumlah produksi koran sehingga mendapat keuntungan yang maksimal.

2. DASAR TEORI

2.1 Koran

Sesuai dengan Kamus Besar Bahasa Indonesia koran adalah kumpulan lembaran kertas yang berisi berita. Koran berfungsi sebagai media informasi dan sarana edukasi bagi masyarakat. Koran dicetak dan diterbitkan setiap hari. Ruang lingkup koran sangat luas mulai dari daerah, nasional dan internasional. Koran dijual dengan harga yang mudah dijangkau masyarakat. Isi dari koran terbagi dalam beberapa topik, diantaranya politik, ekonomi, budaya, kriminalitas, olahraga, cuaca dan hiburan. Koran tidak hanya mendapat pemasukan dari penjualan koran

tetapi juga berasal dari iklan-iklan yang dipasang di koran tersebut.

2.2 Support Vector Regression (SVR)

Support Vector Regression (SVR) merupakan pengembangan metode dari *Support Vector Machine* dalam model non linear untuk kasus regresi dan prediksi^[5]. SVR merupakan metode yang dapat mengatasi overfitting, sehingga akan menghasilkan performansi yang bagus (Scholkopf et al., 2004). *Support Vector Regression* (SVR) terbukti dalam menyelesaikan masalah peramalan karena *error RMS* relatif kecil^[6]. Dalam penelitian lain parameter SVR berpengaruh besar terhadap hasil peramalan^[2].

Parameter utama dalam SVR adalah ϵ dan C . ϵ mendefinisikan kesalahan sensitif di sekitar fungsi regresi, dan mengendalikan seberapa baik fungsi sesuai dengan data pelatihan^[1]. Parameter C mengontrol *trade off* antara kesalahan pelatihan dan kerumitan model. Nilai ϵ yang kecil terkait dengan nilai yang tinggi pada variabel *slack* dan akurasi yang tinggi. Sebaliknya, nilai ϵ yang tinggi berkaitan dengan nilai akurasi yang kecil.

Metode sekuensial dapat mempercepat proses iterasi^[12]. Berikut merupakan langkah – langkah *sequential latih* untuk kasus non linier:

1. Inisialisasi parameter cLR , C , iterasi maksimum ϵ dan sigma.
2. Hitung Matriks *Hessian*

$$D_{ij} = (K(x_i, x_j) + \lambda^2) \quad (1)$$

Keterangan:

D_{ij} = Elemen Matriks *Hessian* pada baris ke- i dan kolom ke- j

K = *Kernel* yang digunakan

λ^2 = Variabel Skalar

3. Untuk setiap data latih, hitung :

$$E_i = y_i - \sum_{j=1}^l (\alpha_j^* - \alpha_j) D_{ij} \quad (2)$$

Keterangan:

y_i = Nilai aktual data latih yang digunakan

E_i = Nilai error ke- i

$$\delta \alpha_i^* = \min \{ \max [\gamma (E_i - \epsilon), -\alpha_i^*], C - \alpha_i^* \} \quad (3)$$

$$\delta \alpha_i = \min \{ \max [\gamma (-E_i - \epsilon), -\alpha_i], C - \alpha_i \} \quad (4)$$

Keterangan:

$\delta \alpha_i^*$, $\delta \alpha_i$ = variabel tunggal

ϵ = Nilai epsilon

C = Nilai kompleksitas

γ = Nilai *Learning Rate* (LR) yang didapatkan dari *constant learning rate* (cLR).

$$\alpha_i = \alpha_i + \delta\alpha_i \quad (5)$$

$$\alpha_i^* = \alpha_i^* + \delta\alpha_i^* \quad (6)$$

4. Ulangi step 3 sampai iterasi maksimum.

5. Hitung Fungsi regresi:

$$f(x) = \sum_{i=1}^l (\alpha_i^* - \alpha_i)(K(x_i, x) + \lambda^2) \quad (8)$$

6. Selesai

2.3 Metode Kernel

Secara umum, kasus kasus di dunia nyata adalah kasus yang tidak linear (Santosa,2006). Bila fitur berada di ruang *input* / ruang asli, hubungan non linear pada data tidak dapat ditangani dengan mudah [10]. Metode kernel dapat diterapkan dalam SVR karena metode kernel fleksibel untuk mengatasi masalah non-linear [9][3]. Dengan metode kernel suatu data x di input dipetakan ke dimensi yang lebih tinggi sehingga kasus nonlinier dapat ditangani dengan baik.

Kernel yang paling digunakan adalah Kernel *Gaussian Radial Basic Function* (RBF). RBF memiliki kinerja terbaik saat dibandingkan dengan fungsi kernel yang lain. Keluaran fungsi ini tergantung pada jarak Euclidean antara Xi dan Xj [11]. RBF akan dijabarkan pada Persamaan 9

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (9)$$

Keterangan:

x_i = Nilai dari fitur data pelatihan ke-*i*

x = Nilai dari fitur data untuk menentukan peramalan

2.4 Nilai Evaluasi

Nilai evaluasi yang digunakan adalah *Mean Absolute Percentage Error* (MAPE), dimana MAPE memberikan indikasi seberapa besar prosentase kesalahan dalam meramal dibandingkan dengan nilai aktual. Penjabaran rumus MAPE akan dijelaskan pada Persamaan 16

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \dots \dots \dots (10)$$

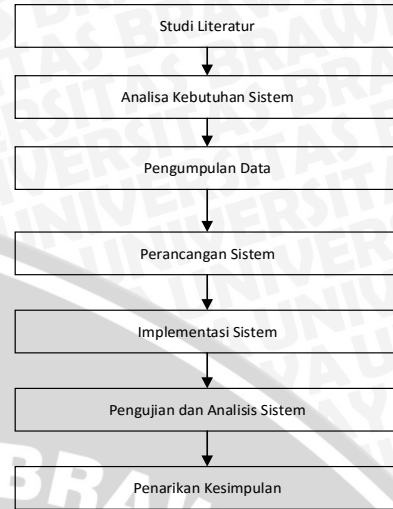
Keterangan:

\hat{y}_i = Nilai hasil ramalan

y_i = Nilai aktual

3. METODOLOGI PENELITIAN

Diagram Alir Metode Penelitian seperti ditunjukkan pada **Error! Reference source not found.**:



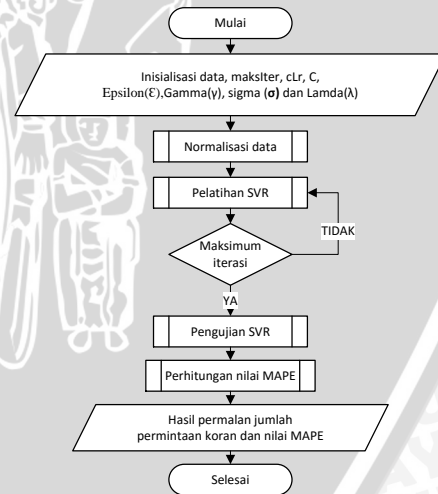
Gambar 1 Diagram Alir Tahapan Penelitian

3.1 Data Penelitian

Data yang digunakan pada penelitian ini adalah data penjualan koran Radar Madura tahun 2015 (dengan rentang waktu harian) yang didapatkan dari bagian Akuntansi Jawa Pos Radar Madura.

4. PERANCANGAN

Perancangan peramalan jumlah permintaan koran menggunakan metode SVR dapat dilihat pada Gambar 2



Gambar 2 Diagram Alir Peramalan jumlah permintaan koran

Langkah 1 :
Inisialisasi data dan parameter SVR

Langkah 2 :
Normalisasi data

Langkah 3 :
Melakukan pelatihan SVR hingga iterasi maksimum

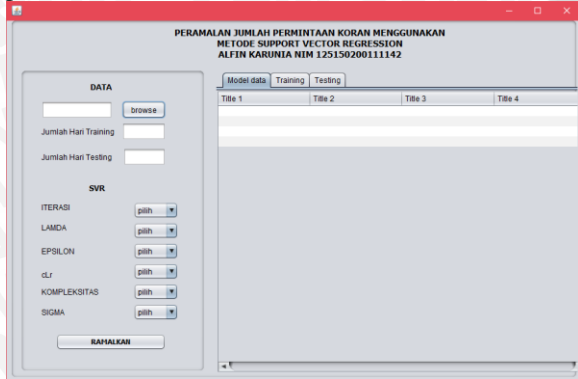
Langkah 4 :
Pengujian SVR

Langkah 5 :

Perhitungan nilai MAPE

5. IMPLEMENTASI

Implementasi antarmuka peramalan jumlah permintaan koran menggunakan SVR berisi masukan data dan parameter SVR yang diposes dan ditampilkan pada tabel model data, pelatihan dan pengujian seperti pada Gambar 3

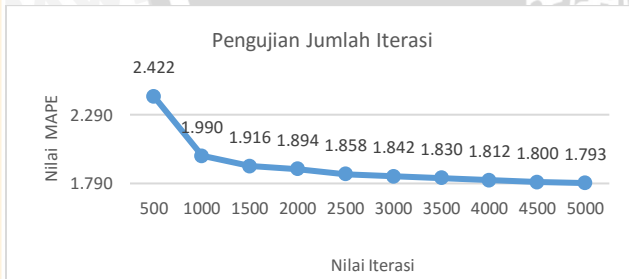


Gambar 3 Tampilan Sistem

6. PENGUJIAN DAN ANALISIS

6.1 Pengujian Jumlah Iterasi

Jumlah iterasi optimum dievaluasi berdasarkan rata-rata nilai MAPE terbaik dan waktu komputasi. Menurut Ari Wijono selaku kepala Akuntansi Radar Madura, batas nilai kesalahan jumlah produksi yang dapat ditolerir Radar Madura adalah 15%. Pada pengujian ini jumlah iterasi yang digunakan adalah kelipatan 500. Grafik hasil pengujian jumlah iterasi terhadap nilai MAPE ditunjukkan pada Gambar 6.1 sebagai berikut :



Gambar 4 Grafik Uji Coba Jumlah Iterasi SVR terhadap Nilai MAPE

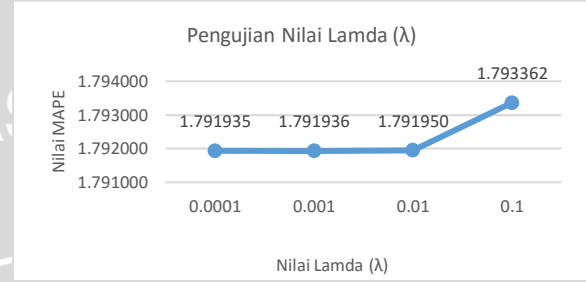
Dari grafik tersebut dapat dilihat semakin besar jumlah iterasi, semakin optimal nilai MAPE yang dihasilkan. Namun berbanding terbalik terhadap waktu komputasi yang dibutuhkan. Pada jumlah iterasi yang banyak menyebabkan waktu komputasi semakin meningkat.

6.2 Pengujian Variasi Parameter

Pengujian variasi parameter digunakan untuk menentukan batas pencarian parameter yang terbaik untuk menghasilkan solusi yang terbaik dalam kasus peramalan ini. Pengujian batas pencarian parameter dilakukan untuk tiap parameter yang dioptimasi yaitu *Lambda* (λ), *Epsilon* (ϵ), konstanta *Learning Rate* (cLR) dan Kompleksitas (*C*).

a. Pengujian Parameter Lamda

Grafik hasil pengujian parameter lamda terhadap nilai MAPE ditunjukkan pada Gambar 5 sebagai berikut :

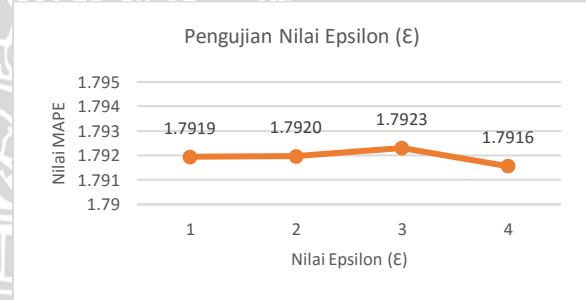


Gambar 5 Grafik Uji Coba Nilai Lamda (λ)

Dari grafik tersebut, semakin kecil nilai dari λ akan memberikan akurasi yang baik pada regresi.

b. Pengujian Parameter Epsilon (ε)

Grafik pengujian parameter *Epsilon* (ϵ) ditunjukkan pada Gambar 6

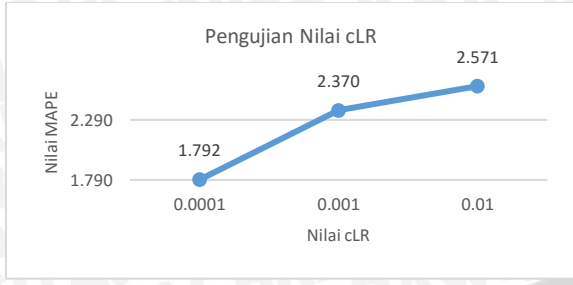


Gambar 6 Grafik Uji Coba Batas Parameter Epsilon (ε)

Dari grafik tersebut nilai *Epsilon* (ϵ) yang kecil belum tentu menghasilkan nilai error yang minimum. Terbukti nilai MAPE minimum didapatkan *Epsilon* (ϵ) 0.001.

c. Pengujian Parameter cLR

Berikut Gambar 7 yang berisi grafik pengujian nilai *cLR*.

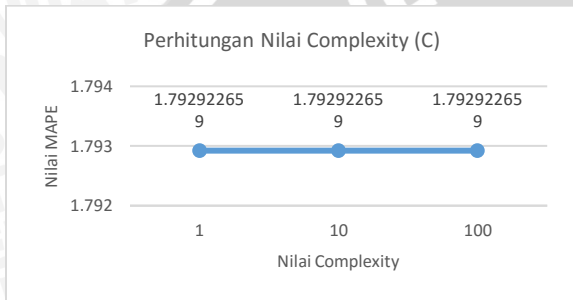


Gambar 7 Grafik Uji Coba Batas Parameter cLR

Berdasarkan grafik diatas, semakin kecil nilai cLR, menghasilkan nilai MAPE minimum. Nilai MAPE terbaik didapatkan pada nilai cLR 0.0001.

d. Pengujian Parameter Complexity (C)

Berikut hasil pengujian Complexity (C) berupa grafik pada Gambar 8

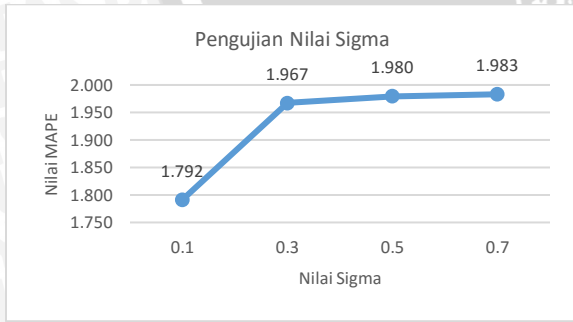


Gambar 8 Grafik Uji Coba Batas Parameter Complexity (C)

Berdasarkan Gambar 8 nilai MAPE yang dihasilkan cenderung tetap pada ketiga nilai Complexity (C).

e. Pengujian Parameter Sigma

Berikut Gambar 9 yang berisi grafik pengujian nilai sigma.



Gambar 9 Grafik Uji Coba Batas Parameter Sigma

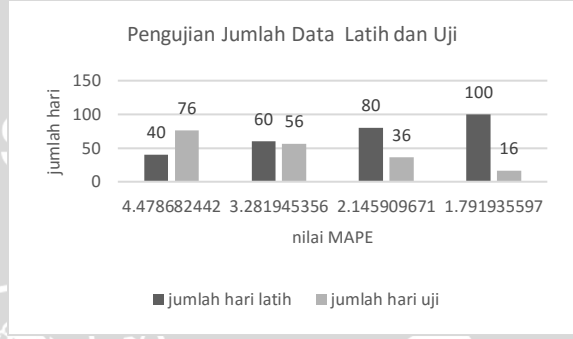
Dari grafik di atas dapat dilihat semakin kecil nilai sigma akan menghasilkan nilai MAPE minimum.

6.3 Pengujian Variasi Data Latih Dan Data Uji

Uji coba variasi data latih dan data uji untuk melihat data latih dan data uji yang optimal untuk peramalan jumlah permintaan koran. Terdapat 2 pengujian yaitu pengujian variasi bulan data latih dan uji dan pengujian variasi jumlah data latih dan uji.

a. Pengujian Jumlah Data Latih dan Data Uji

Berikut grafik pengujian variasi jumlah data latih dan uji pada Gambar 10

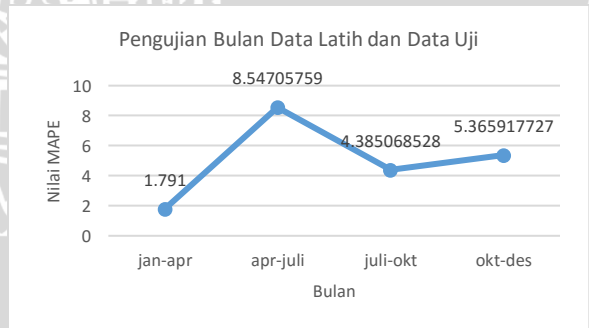


Gambar 10 Grafik Uji Coba Variasi Jumlah Data Latih dan Data Uji

Nilai MAPE terkecil dihasilkan pada jumlah data latih 100 dan jumlah data uji 16. Terlihat pada grafik semakin banyak jumlah data latih akan menghasilkan nilai MAPE minimum.

b. Pengujian Bulan Data Latih dan Data Uji

Berikut grafik pengujian variasi bulan data latih dan uji pada Gambar 11



Gambar 11 Grafik Uji Coba Variasi Bulan Data Latih dan Data Uji

Berdasarkan grafik di atas, nilai MAPE terkecil dihasilkan pada bulan januari sampai april. Pada bulan April sampai Juli menghasilkan nilai MAPE terbesar. Hal ini terjadi karena beberapa faktor yang mempengaruhi fluktuasi jumlah penjualan. Mulai dari konten berita yang disajikan, atau terdapat event pada Radar Madura yang membagikan koran gratis.

7. PENUTUP

7.1 Kesimpulan

Berikut adalah kesimpulan yang dapat diambil dari hasil yang telah didapatkan dari perancangan, implementasi dan pengujian yang telah dilakukan

1. Proses awal yang harus dilakukan dalam mengimplementasikan metode SVR untuk peramalan jumlah permintaan koran adalah menentukan data latih dan data uji yang akan digunakan, peramalan yang digunakan adalah peramalan secara sekuensial.
2. Untuk mengukur solusi dari permasalahan peramalan jumlah permintaan koran ini digunakan perhitungan nilai *error* dari peramalan yang berupa MAPE. Berdasarkan pengujian yang telah dilakukan, nilai *error* terkecil sebesar 1.7916
3. Berdasarkan pengujian yang telah dilakukan, pengaruh parameter terhadap peramalan bervariasi. Semakin kecil nilai *Lambda* (λ), *cLR* dan *sigma* semakin kecil pula MAPE yang dihasilkan. Untuk nilai *Epsilon* (ϵ) yang kecil belum tentu menghasilkan nilai *error* yang minimum. Sedangkan hasil pengujian untuk parameter *Complexity* (C) cenderung konstan sehingga tidak terlalu mempengaruhi hasil peramalan.

7.2 Saran

Untuk penelitian selanjutnya, peneliti dapat menambahkan parameter SVR yang dioptimasi untuk hasil yang lebih optimal lagi, selain itu penggunaan faktor – faktor yang mempengaruhi jumlah permintaan koran.

8. DAFTAR PUSTAKA

- [1] A. J. Smola and B. Scholkopf, 2003. A tutorial on support vector Regression.
- [2] Amanda,R.,2014. Analisis *Support Vector Regression* (SVR) Dalam Meramalkan Kurs Rupiah Terhadap Dollar Amerika Serikat. UNDIP.
- [3] Braga,I.,Maria, Monard ,C., 2015. Improving the kernel regularized least squares method for small sample regression. Brazil : University of São Paulo
- [4] Dian, R.,2014. Pertarungan Antara Media Konvensional dan Digital, Siapakah Pemenangnya? Tersedia di:<
<http://www.kompasiana.com/rosalitudian.com/pertarungan-antara-media-konvensional-dan-digital-siapakah->

[pemenangnya_54f7b346a33311747a8b4b17](#)

>

[Diakses 1 April 2016]

- [5] Drucker, H., Burges, C., 1997. Support vector regression machines. *Adv. Neural Inf. Process. Syst.*, vol. 9, pp. 155–161,
- [6] Kavaklioglu, K.,2011. Modeling and prediction of Turkey's electricity consumption using Support Vector Regression. Turkey : Pamukkale University.
- [7] Siregar,A., 2000. Media Pers Dan Negara: Keluar Dari Hegemoni.
- [8] Puspaningrum, B.I., 2010. Analisis perubahan media habit dalam mengkonsumsi media di era teknologi informasi. Universitas Indonesia.
- [9] Schölkopf, B., A.J. Smola, 2002. Learning With Kernels: Support Vector Machines, Regularization, Optimization and Beyond, MIT Press, Cambridge, Mass., USA
- [10] Rajkumar, N., Jaganathan., P., 2013. *A New RBF Kernel Based Learning Method Applied to Multiclass Dermatology Diseases Classification*. Proceedings of 2013 IEEE Conference on Information and Communication Technologies (ICT 2013).
- [11] Ravale, U., Marathe, N., Padiya, P., 2015. Feature Selection Based Hybrid Anomaly Intrusion Detection System Using K Means and RBF Kernel Function. Navi Mumbai, India
- [12] Vijayakumar, S., Wu S., 1999. Sequential Support Vector Classifiers and Regression. RIKEN Brain Science Institute 1999, pp.610 – 619.
- [13] Ying Li, D., Xu,W., Zhao, H., Qiu Chen, R., 2009. A SVR BASED FORECASTING APPROACH FOR REAL ESTATE PRICE PREDICTION. Huazhong University of Science and Technology, Wuhan, Hubei, 430074, China