

PENGARUH JUMLAH DATA LATIH SVM PADA PERAMALAN BEBAN LISTRIK BULANAN DI SEKTOR INDUSTRI

The Total of SVM Train Data Effect On Monthly Electric Load Forecasting
Monthly Charges In Industrial Sector

Luqman Assaffat

Program Studi Teknik Elektro, Fakultas Teknik
Universitas Muhammadiyah Semarang
Jl. Kasipah No. 12 Semarang
email : assaffat@unimus.ac.id

ABSTRAK

Peramalan terhadap beban listrik bulanan di industri adalah sangat penting dalam perencanaan dan pengawasan serta monitoring penggunaan listrik di sektor industri, sehingga tingkat keakuratan sistem peramalan sangat diperlukan. Support Vector Machine (SVM) merupakan salah satu metode peramalan yang berbasis kecerdasan buatan dapat memberikan nilai akurasi yang baik, dengan hasil nilai error yang kecil. Sebagai sistem cerdas, SVM memerlukan pelatihan pada sistem nya. Penelitian ini mengkaji pengaruh jumlah data latih yang diberikan pada SVM sebagai metode peramalan beban listrik. Pengujian dilakukan dengan memberikan variabel jumlah data dari 12 data sampai 72 data latih, dan SVM menggunakan fungsi Kernel Gaussian RBF. Hubungan antara jumlah data latih dengan nilai kesalahan peramalan (yang diukur dengan MAPE) dianalisis dengan korelasi Person's. Penelitian ini menghasilkan nilai Korelasi Person's $r = -0,96$ memberikan arti bahwa antara jumlah data latih SVM dengan nilai MAPE yang didapatkan dalam prakiraan SVM terdapat hubungan negatif dengan kekuatan hubungan yang kuat.

Kata kunci : Peramalan Beban Listrik, SVM, Data Latih, Korelasi Person's

ABSTRACT

Forecasting of the monthly electricity load in the industry is very important in the planning, supervision and monitoring of electric energy use of industrial sector, so the level of forecasting system accuracy is required. Support Vector Machine (SVM) is one forecasting method based on artificial intelligence can give good accuracy, with the result that a small error value. As an intelligent system, SVM requires training on his system. This research is studied the influence of training data amount given on SVM as a method of the electrical load forecasting. Testing of system by given variable amount of data, from 12 datas to 72 datas, and SVM using the Gaussian kernel function of RBF. The relationship between the amount of training data with forecasting error value (as measured by MAPE) was analyzed by Person's correlation. This research resulted in the value of Person's correlation $r = -0.96$ gives conclusion that the amount of training data SVM with MAPE values obtained in the forecast SVM negative associations with the strength of relationship

Key-word: forecasting, Electric Load, SVM, Train Data, Person's Correlation

PENDAHULUAN

Peramalan adalah suatu perkiraan tingkat permintaan yang diharapkan untuk suatu produk atau beberapa produk dalam periode waktu tertentu di masa yang akan datang. (John E. Biegel, 1999). Oleh karena itu, peramalan pada dasarnya merupakan suatu taksiran, tetapi dengan menggunakan cara-cara tertentu peramalan dapat lebih dari pada hanya satu taksiran. Dapat dikatakan bahwa peramalan adalah suatu taksiran yang ilmiah meskipun akan terdapat sedikit kesalahan yang disebabkan oleh adanya keterbatasan kemampuan manusia.

Peramalan merupakan bagian integral dari kegiatan pengambilan keputusan manajemen. Organisasi selalu menentukan sasaran dan tujuan, berusaha menduga faktor-faktor lingkungan, lalu memilih tindakan yang diharapkan akan menghasilkan pencapaian sasaran dan tujuan tersebut. Kebutuhan akan peramalan meningkat sejalan dengan usaha manajemen untuk mengurangi ketergantungannya pada hal-hal yang belum pasti. Peramalan menjadi lebih ilmiah sifatnya dalam menghadapi lingkungan manajemen. Karena setiap organisasi berkaitan satu sama lain, baik buruknya ramalan dapat mempengaruhi seluruh bagian organisasi. (Makridakis, 1988)

Penggunaan beban listrik bulanan merupakan salah satu aspek yang perlu diprediksikan dalam suatu industri. Hal ini berkaitan dengan pengambilan keputusan manajemen dalam penyediaan besarnya energy listrik, serta penjadwalan perawatan dan perbaikan peralatan dan sistem tenaga listrik di industri tersebut. Sehingga tingkat keakuratan peramalan beban listrik di industry sangat diperlukan.

METODE PENELITIAN

Metode Peramalan Beban Listrik

Peramalan beban listrik dapat diklasifikasikan berdasarkan waktu, yaitu peramalan beban jangka pendek (*Short Term Load Forecasting*, STLF) dengan dimensi kurang dari 24 jam, dan biasanya memberikan informasi dalam harian, jam, atau beban puncak. Peramalan beban jangka menengah

(*Medium Term Load Forecasting*, MTLF) dengan dimensi waktu mingguan atau bulanan dalam satu tahun, dan biasanya selain menyediakan informasi beban listrik juga terdapat faktor tambahan seperti demografi dan pengaruh ekonomi. Peramalan beban jangka panjang (*Long Term Load Forecasting*, LTLF) merupakan peramalan di atas satu tahun dan biasanya untuk jangka 25 – 30 tahun. Peramalan beban listrik juga dipengaruhi oleh kondisi daerah dan iklim yang terjadi seperti cuaca dan temperatur (Hahn, 2009).

Terdapat beberapa pendekatan yang digunakan di dalam peramalan beban listrik, antara lain yaitu metode regresi dan *time series* klasik dan metode kecerdasan buatan. Salah satu metode kecerdasan buatan yang sering digunakan dan terkenal yaitu metode jaringan syaraf tiruan (*Artificial Neural Network*, ANN), sedangkan metode kecerdasan buatan yang relatif baru yaitu metode *Support Vector Machine* (SVM) (Hahn, 2009).

Ukuran kesalahan yang paling sering digunakan untuk menilai kinerja dari model peramalan beban listrik adalah MAPE (*Mean Absolute Percentage Error*) yang merupakan nilai kesalahan persentase absolut rata-rata (Hong, 2009) yaitu :

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - f_i}{y_i} \right|$$

(1)

Keterangan :

N : jumlah data dalam periode prakiraan

y_i : data aktual pada periode ke i

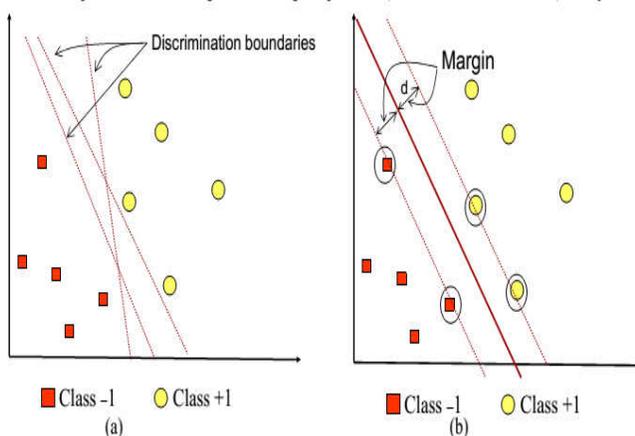
f_i : data prakiraan pada periode ke i

Support Vector Machine

Support Vector Machine (SVM) merupakan salah satu algoritma *machine learning* yang paling populer untuk klasifikasi dan regresi (Maali dkk, 2013). Konsep dasar SVM sebenarnya merupakan kombinasi harmonis dari teori-teori komputasi yang telah ada puluhan tahun sebelumnya. Konsep Kernel diperkenalkan tahun 1950, dan demikian juga dengan konsep-konsep pendukung yang lain. Akan tetapi hingga tahun 1992, dan pernah ada

upaya merangkaikan komponen-komponen tersebut (Vapnik, 1999). Dalam proses pembelajarannya, SVM memperkenalkan strategi baru dengan menemukan *hyperplane* yang terbaik pada *input space*, lewat strategi yang disebut *Structural Risk Minimization*. Prinsip dasar SVM yaitu sebagai *linear classifier*, dan selanjutnya dikembangkan agar dapat bekerja pada problem non-linear dengan memasukkan konsep *Kernel trick* pada ruang kerja berdimensi tinggi (Cristianini, 2000).

Konsep SVM dapat dijelaskan secara sederhana sebagai usaha mencari *hyperplane* terbaik yang berfungsi sebagai pemisah dua buah kelas pada *input space*. Gambar 2.1a memperlihatkan beberapa pola yang merupakan anggota dari dua buah kelas : positif (dinotasikan dengan +1) dan negatif (dinotasikan dengan -1). Pola yang tergabung pada kelas negatif disimbolkan dengan kotak, sedangkan pola pada kelas positif, disimbolkan dengan lingkaran. Proses pembelajaran dalam problem klasifikasi diterjemahkan sebagai upaya menemukan garis (*hyperplane*) yang memisahkan antara kedua kelompok tersebut. Berbagai alternatif garis pemisah (*discrimination boundaries*) ditunjukkan pada Gambar 1.



Gambar 1– SVM Berusaha Menemukan Hyperplane Terbaik Yang Memisahkan Kedua Class Negatif Dan Positif

Hyperplane pemisah terbaik antara kedua class dapat ditemukan dengan mengukur *margin hyperplane* tersebut dan mencari titik maksimalnya. Margin merupakan jarak antara *hyperplane* tersebut dengan data terdekat dari

masing-masing class. Subset data training set yang paling dekat ini disebut sebagai *support vector*. Garis solid pada Gambar 2.1b menunjukkan *hyperplane* yang terbaik, yaitu yang terletak tepat pada tengah-tengah kedua kelas, sedangkan titik kotak dan lingkaran yang berada dalam lingkaran hitam merupakan *support vector*. Upaya mencari lokasi *hyperplane* optimal ini merupakan inti dari proses pembelajaran pada SVM .

Koefisien Korelasi Person's

Koefisien korelasi (KK) merupakan indeks atau bilangan yang digunakan untuk mengukur derajat hubungan, meliputi kekuatan hubungan dan arah/bentuk hubungan. Kekuatan hubungan, nilai koefisien korelasi berada di antara -1 dan +1. Bentuk atau arah hubungan ditentukan oleh tanda positif (+) atau negatif (-). Tabel 1. memperlihatkan interval nilai Koefisien Korelasi dan Kekuatan hubungan (Misbahuddin dkk, 2013).

Tabel 1. Interval Nilai Koefisien Korelasi dan Kekuatan Hubungan

No	Interval Nilai	Kekuatan Hubungan
1	KK = 0,00	Tidak ada korelasi
2	0,00 < KK ≤ 0,20	Sangat rendah atau lemah sekali
3	0,20 < KK ≤ 0,40	Rendah atau lemah, tapi pasti
4	0,40 < KK ≤ 0,70	Cukup berarti atau sedang
5	0,70 < KK ≤ 0,90	Tinggi atau kuat
6	0,90 < KK < 1,00	Sangat tinggi atau kuat sekali
7	KK = 1,00	Sempurna

Rumus koefisien korelasi Pearson's (r) digunakan pada analisis korelasi sederhana untuk variabel interval/rasio dengan variabel/rasio. Koefisien Korelasi Pearson's (r) dirumuskan sebagai berikut (Misbahuddin dkk, 2013).

$$r = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}} \dots\dots\dots(2)$$

Keterangan :

- r = koefisien korelasi Pearson's
n = jumlah variabel
X = variabel bebas
Y = variabel terikat

Sistem peramalan beban listrik bulanan dibangun menggunakan bahasa Matlab, berdasarkan metode Support Vector Machine dengan Kernel Gaussian RBF sebagai fungsi Kernel yang digunakan. Data yang digunakan pada penelitian adalah data kapasitas produksi dan data beban listrik bulanan pada PT. Phapros Indonesia, Semarang untuk periode tahun 2008 sampai tahun 2014 (84 bulan/data). Variabel nama bulan dalam setahun sebagai input vektor x pada SVM, dikonversi menjadi angka 1 (satu) sampai 12 (dua belas).

Dalam penelitian ini jumlah data pelatihan SVM dibuat variabel, dari 1 (satu)tahun sampai 6 (enam) tahun. Sedangkan waktu prakiraan dibuat dengan data kelompok selama 1 (satu)tahun waktu prakiraan. Rentang waktu antara data pelatihan dan data validasi mempunyai jarak 1 (satu)tahun untuk bulan yang sama. Tabel 2. memperlihatkan skema pelatihan dan prakiraan beban listrik dengan SVM.

Output dari hasil peramalan berupa data perkeriaan beban listrik bulanan, dibandingkan dengan data validasi sebagai data real atau data nyata beban listrik yang terjadi. Nilai kesalahan dihitung menggunakan nilai MAPE. Koefisin Person's digunakan untuk menganalisis hubungan antara jumlah data latih yang diberikan pada SVM dengan nilai MAPE hasil peramalan.

HASIL DAN PEMBAHASAN

Dari hasil pengujian sistem peramalan beban listrik bulanan pada sector industry menggunakan SVM dengan Fungsi Kernel Gaussian RBF di perlihatkan pada tabel 3 berikut ini.

Tabel 2. Skema Waktu Pelatihan dan Prakiraan SVM

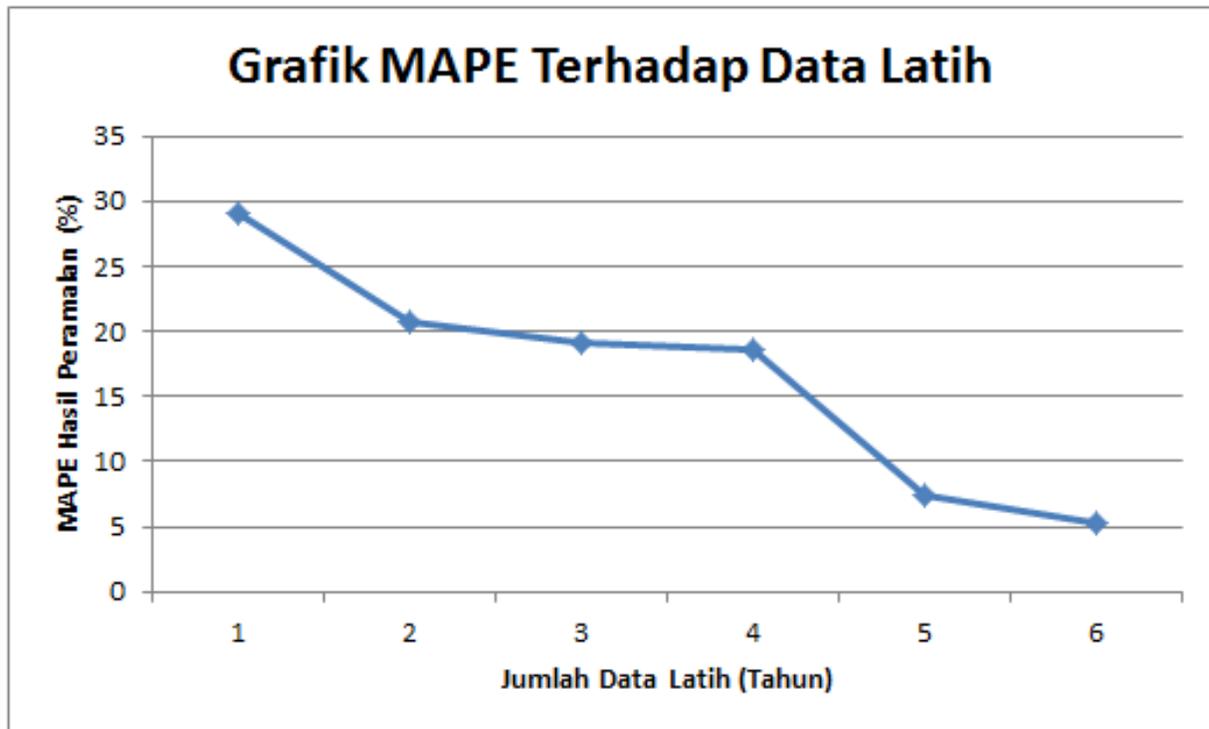
	DATA LATIH SVM	DATA VALIDASI PERAMALAN
1 th :	Data 2008 (12 data)	Data 2009 (12 data)
2 th :	Data 2008-2009 (24 data)	Data 2010 (12 data)
3 th :	Data 2008-2010 (36 data)	Data 2011 (12 data)
4 th :	Data 2008-2011 (48 data)	Data 2012 (12 data)
5 th :	Data 2008-2012 (60 data)	Data 2013 (12 data)
6 th :	Data 2008-2013 (72 data)	Data 2014 (12 data)

Tabel 3. MAPE Hasil pengujian Peramalan

Jumlah Data	MAPE Hasil Peramalan
Latih (X)	SVM (Y)
1 tahun : 12 data	29,13%
2 tahun : 24 data	20,66%
3 tahun : 36 data	19,15%
4 tahun : 48 data	18,61%
5 tahun : 60 data	7,45%
6 tahun : 72 data	5,33%

Dari Tabel 3 tersebut terlihat, semakin banyak jumlah data latih yang diberikan pada SVM, maka nilai MAPE semakin kecil. Sehingga nilai error peramalannya semakin kecil, nilai akurasi peramalannya semakin baik. Hal ini dapat diperlihatkan pada Gambar 2, yang memperlihatkan Grafik MAPE terhadap jumlah data latih menunjukkan trend menurun..

Dari tabel 3, hasil pengujian sistem peramalan dengan SVM, dilakukan analisis korelasi Person's dalam table 4. :



Gambar 2. Grafik MAPE terhadap jumlah data latih

Tabel 4. Analisis Korelasi Person's

X	Y	XY	X ²	Y ²
12	29,13	349,56	144	848,56
24	20,66	495,84	576	426,84
36	19,15	689,40	1.296	366,72
48	18,61	893,28	2.304	346,33
60	7,45	447,00	3.600	55,50
72	5,33	383,76	5.184	28,41

$$\begin{aligned}
 n &= 6 \\
 \sum X &= 252 \\
 \sum Y &= 100,33 \\
 \sum XY &= 3.258,84 \\
 \sum X^2 &= 13.104 \\
 \sum Y^2 &= 2.072,36
 \end{aligned}$$

$$r = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

$$r = \frac{(6 \times 3.258,84) - (252)(100,33)}{\sqrt{[6(252) - 13.104][6(100,33) - 2.072,36]}}$$

$$r = \frac{5.730,12}{5.983,71} = 0,96$$

Dari analisis tersebut di atas diperoleh nilai koefisien korelasi Pearson's $r = -0,96$ yang memberikan arti bahwa antara jumlah data latih yang diberikan pada SVM berbanding terbalik dengan MAPE hasil peramalannya, sehingga kedua variabel tersebut terdapat hubungan negatif dengan kekuatan hubungan yang kuat. Artinya, jika jumlah data latih pada SVM semakin banyak, maka nilai MAPE peramalan SVM akan semakin kecil. Dengan semakin kecilnya nilai MAPE, maka sistem peramalan akan semakin baik dan hasilnya akan semakin akurat.

KESIMPULAN

Dari penelitian ini dapat disimpulkan bahwa jumlah data latih yang diberikan pada SVM mempengaruhi hasil peramalan beban listrik bulanan dengan metode SVM. Semakin banyak jumlah data latih, maka hasil peramalan beban listrik yang menggunakan SVM semakin baik dan akurat, dengan nilai MAPE yang semakin kecil.

DAFTAR PUSTAKA

- Biegel, John E., 1999, Pengendalian Produksi Suatu Pendekatan Kuantitatif, Akademika Presindo, Jakarta.
- Cristianini N., Taylor J.S., 2000, An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods, Cambridge Press University, 2000
- Hahn, H., Nieberg, S.M., Pickl, S., 2009, Electric Load Forecasting Methods: Tools for Decision Making, European Journal of Operational Research 199, 902–907

Hong, W.C., 2009, Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model, Energy Conversion and Management, 50, 105–117

Maali, Y., Al-Jumaily, A., 2013, Self-Advising Support Vector Machine, Knowledge-Based Systems 52, 214–222.

Makridakis, Spyross, Steven C., Whellwright, Victore Mcgee, Untung Sus Adriyanto, Abdul Basith, 1988, Metoda dan Aplikasi Peramalan, Jilid Satu Edisi Kedua, Erlangga.

Misbahuddin, Hasan, I., 2013, Analisis Data Penelitian dengan Statistik, Bumi Aksara, Jakarta

Vapnik, V.N., 1999, The Nature of Statistical Learning Theory 2nd edition, Springer-Verlag, New York Berlin Heidelberg