

PREDIKSI HARGA KOMODITAS EMAS DAN BATUBARA DI PASAR DUNIA DENGAN ALGORITMA *SUPPORT VECTOR MACHINE*

Eko Pudjianto¹, Purwanto², Catur Supriyanto³

^{1,2,3}Pascasarjana Teknik Informatika Universitas Dian Nuswantoro

Abstract

Changes in commodity prices of gold and coal in the world market is very influential on the Indonesian government's policy, especially in the country's revenue in the foreign exchange sector. By predicting the price of gold and coal in the world market expected the government to determine important strategy especially in the fields of mining, trade (exports), Energy and Mineral Resources in Indonesia. By applying the method of SVM (Support Vector Machine) can be found a configuration that is able to predict the prediction of gold and coal prices in the coming period. Data processing using SVM algorithm based on k - fold validation , C (cost) and its kernel , then searched the level RMSE (root mean square error) is the smallest. RMSE is the smallest design that is used in predicting the price of gold and coal. Gold commodity price prediction method with RMSE (root mean square error) is at best 43 509 + / - 37 487 with data input 7 (seven) months earlier , k - fold 10 , C (cos) of 0.3 and using a kernel -type dot . So the commodity price forecast gold in the world market for the period December 2013 amounted to U.S. \$ 1,298.33 and for coal commodities with RMSE (root mean square error) is best at 3,185 + / - 3,591 with data input 2 (two) months earlier , k - 10 fold , C (cost) of 0.3 and using a kernel-type dot. So the prediction of coal commodity prices on the world market for the period from December 2013 is U.S. \$ 81.58

Keywords : SVM (Support Vector Machine), RMSE (Root Mean Square Error), Predicted Price

1. PENDAHULUAN

Perubahan harga komoditas emas dan batubara di pasar dunia sangat berpengaruh pada kebijakan pemerintah dalam bidang pertambangan khususnya bagi kementerian ESDM (Energi dan Sumber Daya Mineral), dengan memprediksi harga komoditas emas dan batubara di pasar dunia diharapkan pemerintah bisa mengambil kebijakan tentang pola produksi dan peningkatan mutu serta kuantitas produksi emas dan batubara di Indonesia, yang nantinya dapat menghasilkan devisa bagi negara.

Tujuan dari penelitian ini adalah menemukan konfigurasi prediksi dan memprediksi harga komoditas emas dan batu bara berdasarkan data harga sebelumnya dengan menggunakan algoritma SVM (*Support Vector Machine*).

1.1. Penelitian Terkait

Guohao Zhao dkk. [1] memprediksi harga pasar batubara China dengan menggunakan *Time Sequence Models*, yaitu model urutan waktu yang telah diterapkan secara luas dalam berbagai urutan waktu situasi, prediksi data dan mencapai hasil yang lebih tinggi. Penelitian ini menunjukkan bahwa model urutan waktu secara akurat dapat memprediksi proses perubahan harga batubara di Cina. Zhao dkk membangun model target prediksi berdasarkan ARMA (*autoregressive moving average*) yaitu mensyaratkan bahwa urutan acak menjadi nol nilai rata-rata dan stabil. Namun seiring waktu, target prediksi urutan waktu acak akan terus memiliki tren goyah (naik atau turun), maka waktu yang goyah berurutan. Sebagai hasil dari ini,

sebelum menerapkan model, peneliti harus mengatasi dengan nilai rata-rata waktu urutan nol dan perbedaan kemandapan. Peneliti menggabungkan proses pembentukan harga batubara reformasi mekanisme domestik Cina, menggunakan model urutan waktu data historis untuk mengungkapkan fitur data time dari waktu ke waktu dan menetapkan model prediksi yang menganalisis tren harga batubara domestik Cina. Hasil penelitian menunjukkan bahwa model urutan waktu bisa digunakan relatif baik ke dalam proses perubahan harga batu bara nasional di Cina

Mahsa Kargarani Farahani dan Soheil Mehralian [2] memprediksi harga emas dengan menyajikan perbandingan *Artificial Neural Network* (ANN) dan *Adaptive Neural Fuzzy Inference System* (ANFIS) untuk memprediksi harga emas. Selain itu Farahani dan Mehralian juga membandingkan model *hybrid* baru yang merupakan rata-rata tertimbang ANN dan Model ANFIS. Tujuan utamanya adalah untuk memprediksi harga emas di pasar Forex. Farahani dan Mehralian menggunakan dua mesin prediksi model dalam JST, model yang umpan balik output jaringan sebagai input dan model lain yang tidak melakukannya. Hasil penelitian menunjukkan bahwa lebih dari dua kesalahan dari kinerja model terdahulu, dan juga kinerja ANFIS lebih baik dari model ANN. Untuk mengevaluasi metode tiga kinerja pengukuran yang digunakan : *Root Mean Squared Error* (RMSE), *error persentase* dan *Mean Tendensi Error* (MTE) yang diajukan dalam penelitian ini. Titik kekuatan metode ini adalah prediksi model mesin yang merupakan salah satu yang paling tinggi. Model mesin prediksi JST. Hasilnya sebuah *Denoising Wavelet* algoritma diterapkan pada data, tetapi karena struktur menjadi kacau dari harga emas yang mengakibatkan rusaknya data dan menyebabkan untuk kinerja hasil prediksi berkurang. Dalam penelitian ini, perbandingan antara ANN dan ANFIS untuk prediksi harga emas disajikan. Data harga emas berasal dari Pasar Forex, dengan menggunakan 20 data untuk menguji algoritma. Tiga pengukuran kinerja, termasuk RMSE, MTE, dan *error persentase* digunakan untuk analisis dan evaluasi pertunjukan model. Hasil menunjukkan bahwa ANFIS dan metode ANN keduanya alat yang kuat untuk memodelkan harga emas dan ANFIS adalah sedikit lebih baik dan kuat dari ANN. Metode *hybrid* di samping akurasi yang sebanding dengan ANN dan ANFIS memiliki hasil yang lebih dapat diandalkan.

Sementara Z. Ismail dkk. [3] memprediksi harga emas dengan menggunakan *Multiple Linear Regression*. Dengan menggunakan data harga emas dari London PM Fix, dan delapan variable yang diidentifikasi memiliki pengaruh terhadap harga emas sebagai variabel independen dalam model regresi. Hasil penelitian menyebutkan bahwa model ini tampaknya berguna untuk memprediksi harga emas dengan 85,2% dari variasi sampel harga emas bulanan. Dalam hal prediksi, model ini mencapai tingkat akurasi yang tinggi prediksi. Jumlah perbedaan dijelaskan adalah sekitar 70% dan koefisien regresi juga menyediakan sarana untuk menilai kepentingan relatif dari variabel individu dalam prediksi keseluruhan harga emas.

Heru Supriyanto [4] memprediksi arah pergerakan harga harian valuta asing menggunakan *Support Vector Machines* dengan metode *Kernel Trick* menggunakan fungsi *Kernel Radial Basis Function*. Dalam penelitian tersebut dikemukakan bahwa SVM dengan metode *kernel trick* menggunakan fungsi kernel RBF dapat digunakan untuk memprediksi arah pergerakan harga harian perdagangan valuta asing dengan akurat. Hal ini ditunjukkan dengan akurasi prediksi terhadap USD/JPY untuk periode bulan Juni 2012 yang mencapai 100% (21 data dapat diprediksi dengan benar seluruhnya) dan pada EUR/USD maupun GBP/USD yang mencapai 95.24% (20 dari 21 data dapat diprediksi dengan benar).

Melihat dari penelitian-penelitian tersebut di atas dapat diketahui bahwa algoritma *Support Vector Machine* (SVM) memiliki tingkat akurasi yang lebih baik dibandingkan dengan metode algoritma lainnya dalam memprediksi harga.

1.2. Komoditas Emas dan Batubara Indonesia

Indonesia memiliki berbagai macam bahan tambang yang terdapat di berbagai daerah. Minyak bumi, gas alam, emas, batubara, bijih besi, dan aspal merupakan jenis-jenis bahan tambang yang dimiliki oleh Indonesia. Salah satu jenis bahan tambang yang cukup banyak dan tersebar ketersediaannya di Indonesia adalah emas. Emas merupakan salah satu jenis bahan tambang yang memiliki nilai ekonomis sangat tinggi.

Emas hampir dipasarkan dan diperdagangkan hampir di semua pasar perdagangan bahan tambang di seluruh dunia. Nilai investasi emas meningkat setiap terjadi perdagangan emas dalam jumlah yang cukup besar. Bahkan emas memberikan kontribusi berupa devisa yang sangat besar bagi negara-negara pengekspor emas.

Batu bara adalah salah satu bahan bakar fosil. Pengertian umumnya adalah batuan sedimen yang dapat terbakar, terbentuk dari endapan organik, utamanya adalah sisa-sisa tumbuhan dan terbentuk melalui proses pembatubaraan. Unsur-unsur utamanya terdiri dari karbon, hidrogen dan oksigen. Batubara juga adalah batuan organik yang memiliki sifat-sifat fisika dan kimia yang kompleks yang dapat ditemui dalam berbagai bentuk.

Himpunan Pemerhati Lingkungan Hidup Indonesia [5] menyebutkan bahwa cadangan batubara Indonesia hanya 0,5 % dari cadangan dunia namun dari segi produksi Indonesia menempati posisi ke 6 dengan jumlah produksi mencapai 246 juta ton, peringkat pertama ditempati China dengan jumlah produksi 2.761 juta ton, disusul USA 1007 juta ton, dan India 490 juta ton, Australia 325 juta ton, Rusia 247 juta ton.

1.3. Harga Komoditas Emas dan Batubara di Pasar Dunia

Secara potensi dan prediksi, harga emas untuk tahun 2014 diperkirakan akan mengalami kondisi yang cukup sulit. Potensi rebound diperkirakan tidak akan signifikan. Prediksi mengenai harga emas diperkirakan akan menyentuh level 1500 dollar per troy ons di akhir tahun depan. Membaiknya kondisi perekonomian di Eropa memicu adanya kenaikan permintaan emas pada bank sentral pada tahun ini.

Harga batu bara dunia saat ini cenderung turun, selain karena kelebihan pasokan, China juga mengurangi kebutuhan batu bara. Selain itu Amerika yang biasanya tidak pernah mengekspor juga melakukan ekspor batu bara. Harga batu bara dunia saat ini sebesar US\$88-92 per ton, diprediksikan akhir tahun akan mencapai US\$95-100 per ton.

Kenaikan harga emas belakangan ini disebabkan oleh ekspektasi kebijakan pelonggaran moneter yang masif oleh bank sentral AS dan optimisme bank sentral Eropa (ECB) akan mengintervensi pasar obligasi pemerintah. Kebijakan stimulus oleh bank sentral AS akan melemahkan dollar AS karena meningkatkan likuiditas dollar di pasar sehingga akan terjadi penguatan harga emas. Pasar keuangan sangat rentan dengan ekspektasi atau rumor. Jadi meskipun baru sebatas harapan, hal ini sudah mengubah persepsi para pelaku pasar dan membuka posisi beli pada perdagangan emas. Ekspektasi terhadap dikeluarkannya kebijakan stimulus yang dikenal dengan pelonggaran kuantitatif ke-3 ini sudah muncul sejak awal tahun karena para pelaku pasar melihat bahwa AS masih belum bisa bangkit dari krisis keuangan 2008. Namun para pelaku pasar kembali dikecewakan dengan tidak diambilnya keputusan tersebut dalam beberapa rapat kebijakan moneter bank sentral AS hingga saat ini.

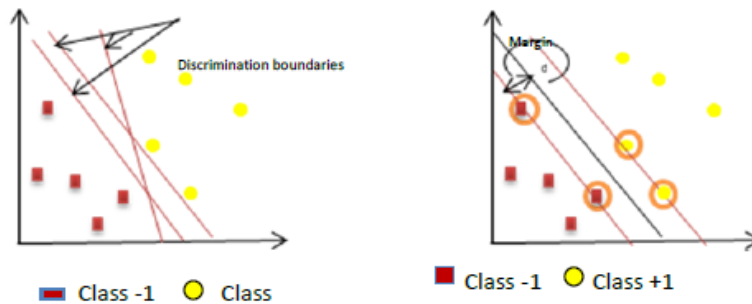
Dampak dari fluktuasi harga komoditas emas dan batubara di pasar dunia adalah bisa mempengaruhi ekonomi global seperti tersendatnya produksi dan konsumsi barang dan jasa, inflasi, dan krisis ekonomi di beberapa wilayah, khususnya Indonesia.

1.4. SVM (Support Vector Machine)

Support Vector Machine (SVM) dikembangkan oleh Boser, Guyon, Vapnik, dan pertama kali dipresentasikan pada tahun 1992 di Annual Workshop on Computational Learning Theory. Konsep dasar SVM sebenarnya merupakan kombinasi harmonis dari teori-teori komputasi yang telah ada puluhan tahun sebelumnya, seperti *margin hyperplane* (Duda & Hart tahun 1973, Cover tahun 1965, Vapnik 1964), kernel diperkenalkan oleh Aronszajn tahun 1950, dan demikian juga dengan konsep-konsep pendukung yang lain. Akan tetapi hingga tahun 1992, belum pernah ada upaya merangkaikan komponen-komponen tersebut. Berbeda dengan strategi *Artificial Neural Network* yang berusaha mencari *hyperplane* pemisah antar-class, SVM berusaha menemukan *hyperplane* yang terbaik pada *input space*. Prinsip dasar SVM

adalah *linear classifier*, dan selanjutnya dikembangkan agar dapat bekerja pada problem non-linear. dengan memasukkan konsep *kernel trick* pada ruang kerja berdimensi tinggi. Perkembangan ini memberikan rangsangan minat penelitian di bidang *pattern recognition* untuk investigasi potensi kemampuan SVM secara teoritis maupun dari segi aplikasi. Dewasa ini SVM telah berhasil diaplikasikan dalam problema dunia nyata (*real-world problems*), dan secara umum memberikan solusi yang lebih baik dibandingkan metode konvensional seperti misalnya *artificial neural network* [6].

Konsep SVM dapat dijelaskan secara sederhana sebagai usaha mencari *hyperplane* terbaik yang berfungsi sebagai pemisah dua buah *class* pada input *space*. Untuk n-dimensional *space*, input data x_i ($i=1, \dots, k$), dimana milik kelas 1 atau kelas 2 dan label yang terkait menjadi -1 untuk kelas 1 dan +1 untuk kelas 2. Gambar 1. memperlihatkan beberapa *pattern* yang merupakan anggota dari dua buah *class*: positif (dinotasikan dengan +1) dan negatif (dinotasikan dengan -1). Pattern yang tergabung pada *class* negatif disimbolkan dengan kotak, sedangkan pattern pada *class* positif, disimbolkan dengan lingkaran. Jika data input dapat dipisahkan secara linear, pemisahan *hyperplane* dapat diberikan dalam: 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 2 [6]

Hyperplane pemisah terbaik antara kedua *class* dapat ditemukan dengan mengukur *margin hyperplane* tersebut dan mencari titik maksimalnya. *Margin* adalah 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 1. menunjukkan *hyperplane* yang terbaik, yaitu yang terletak tepat pada tengah-tengah kedua *class*, sedangkan titik kotak dan lingkaran yang berada dalam lingkaran hitam adalah *support vector*. Upaya mencari lokasi *hyperplane* optimal ini merupakan inti dari proses pembelajaran pada SVM [6].

Data yang tersedia dinotasikan sebagai $x \in R^d$ sedangkan label masing-masing dinotasikan $y_i \in \{-1, +1\}$ untuk $i = 1, 2, \dots, l$ yang mana l adalah banyaknya data. Diasumsikan kedua *class* -1 dan +1 dapat terpisah secara sempurna oleh *hyperplane* berdimensi d , yang didefinisikan:

Diasumsikan kedua *class* -1 dan +1 dapat terpisah secara sempurna oleh *hyperplane* berdimensi d , yang didefinisikan:

$$\vec{w} \cdot x + b = 0 \dots\dots\dots (1)$$

Sebuah *pattern* x_i yang termasuk *class* -1 (sampel negatif) dapat dirumuskan sebagai *pattern* yang memenuhi pertidaksamaan:

$$\vec{w} \cdot x + b = -1 \dots\dots\dots (2)$$

sedangkan *pattern* yang termasuk *class* +1 (sampel positif):

$$\vec{w} \cdot \vec{x} + b = +1 \dots\dots\dots (3)$$

Margin terbesar dapat ditemukan dengan memaksimalkan nilai jarak antara hyperplane dan titik terdekatnya, yaitu $1/||w||$. Hal ini dapat dirumuskan sebagai *Quadratic Programming* (QP) problem, yaitu mencari titik minimal dengan memperhatikan constraint :

$$\min_w \tau(w) = \frac{1}{2} - ||w||^2 \dots\dots\dots (4)$$

$$y_1 - (x_1 \cdot w + b) - 1 \geq 0 \cdot \forall_1 \dots\dots\dots (5)$$

Problem ini dapat dipecahkan dengan berbagai teknik komputasi, diantaranya *Lagrange Multiplier* sebagaimana ditunjukkan pada persamaan di bawah ini :

$$L(w, b, \alpha) = \frac{1}{2} - ||w||^2 - \sum_{i=1}^l \alpha_i (y_i((x_i \cdot w + b) - 1)) \quad (i=1,2,\dots,l) \dots\dots\dots (6)$$

α_i adalah *Lagrange multipliers*, yang bernilai nol atau positif ($\alpha_i \geq 0$). Nilai optimal dari persamaan di atas dapat dihitung dengan meminimalkan L terhadap w dan b, dan memaksimalkan L terhadap α_i . Dengan memperhatikan sifat bahwa pada titik optimal gradient L=0, persamaan dapat dimodifikasi sebagai maksimalisasi problem yang hanya mengandung α_i saja, sebagaimana persamaan :

Maximize :

$$\sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i \cdot y_j \cdot \vec{x}_i \cdot \vec{x}_j \dots\dots\dots (7)$$

$$\alpha_i \geq 0 \quad (i = 1, 2, \dots, l) \quad \sum_{i=1}^l \alpha_i y_i = 0 \dots\dots\dots (8)$$

Dari hasil dari perhitungan ini diperoleh α_i yang kebanyakan bernilai positif. Data yang berkorelasi dengan α_i yang positif inilah yang disebut sebagai *support vector*.

1.5. Komponen Penghitungan Prediksi dalam SVM

a. *K-fold cross-validation*

Dalam pendekatan *cross-validation*, setiap *record* digunakan beberapa kali dalam jumlah yang sama untuk *training* dan tepat sekali untuk *testing*. Untuk mengilustrasikan metode ini, anggaplah kita mempartisi data ke dalam dua *subset* yang berukuran sama. Pertama, kita pilih satu dari kedua *subset* tersebut untuk *training* dan satu lagi untuk *testing*. Kemudian dilakukan pertukaran fungsi dari *subset* sedemikian sehingga *subset* yang sebelumnya sebagai *training set* menjadi *test set* demikian sebaliknya. Pendekatan ini dinamakan *two-fold crossvalidation*. Total *error* diperoleh dengan menjumlahkan *error-error* untuk kedua proses tersebut. Dalam contoh ini, setiap *record* digunakan tepat satu kali untuk *training* dan satu kali untuk *testing*. Metode *k-fold cross-validation* mengeneralisasi pendekatan ini dengan mensegmentasi data ke dalam k partisi berukuran sama. Selama proses, salah satu dari partisi dipilih untuk *testing*, sedangkan sisanya digunakan untuk *training*. Prosedur ini diulangi k kali sedemikian sehingga setiap partisi digunakan untuk *testing* tepat satu kali. Total *error* ditentukan dengan menjumlahkan *error* untuk semua k proses tersebut [7].

Metode *k-fold cross-validation* menetapkan $k = N$, ukuran dari *data set*. Metode ini dinamakan pendekatan *leave-one-out*, setiap *test set* hanya mengandung satu *record*. Pendekatan ini memiliki keuntungan dalam penggunaan sebanyak mungkin data untuk *training*. *Test set* bersifat *mutually exclusive* dan secara efektif mencakup keseluruhan *data set*. Kekurangan dari pendekatan ini adalah banyaknya komputasi untuk mengulangi prosedur sebanyak N kali.

K-fold cross-validation adalah salah satu teknik untuk mengevaluasi keakuratan model, dengan ciri-

ciri [6]:

1. Mempartisi data secara random ke dalam k buah himpunan/fold yaitu $D_1, D_2, ..D_k$. Setiap kelompok mempunyai jumlah yang hampir sama.
2. Pada perulangan i, gunakan D_i sebagai data uji dan himpunan lainnya sebagai data pelatihan
 Contoh :
 - Pada perulangan ke-1 : D_1 sebagai data uji dan D_2 s.d. D_k sebagai data pelatihan
 - Pada perulangan ke-2 : D_2 sebagai data uji dan D_1, D_3 s.d. D_k sebagai data pelatihan dan seterusnya
3. Melakukan training dan pengujian sebanyak k kali
4. Menghitung keakuratan dengan rumus.

b. C (cost)

Pada penghitungan SVM diberikan penalti dengan menambahkan nilai *cost* C. Nilai *cost* C dipilih untuk mengontrol keseimbangan antara nilai *margin* dan *error* prediksi. Semakin besar nilai C, maka penalti yang diberikan terhadap data *error* juga semakin besar.

c. Kernel

Banyak teknik data mining atau *machine learning* yang dikembangkan dengan asumsi kelinieran. Sehingga algoritma yang dihasilkan terbatas untuk kasus-kasus yang linier. Karena itu, bila suatu kasus klasifikasi memperlihatkan ketidaklinieran, algoritma seperti *perceptron* tidak bisa mengatasinya. Secara umum, kasus-kasus di dunia nyata adalah kasus yang tidak linier. Semisal data yang sulit dipisahkan secara linier, metode *kernel* adalah salah satu untuk mengatasinya. Dengan metode *kernel* suatu data x di *input space* dimapping ke *feature space* F dengan dimensi yang lebih tinggi melalui *map* ϕ sebagai berikut $\phi : x \rightarrow \phi(x)$. Karena itu data x di *input space* menjadi $\phi(x)$ di *feature space*.

Fungsi *kernel* yang biasanya dipakai pada penelitian ini adalah:

Linier : $K(x,x_i) = (x^T x_i)..... (9)$

Polynomial : $K(x,x_i) = (x^T x_i + 1)^d,..... (10)$

Fungsi kernel yang harus digunakan untuk substitusi *dot product* di *feature space* sangat bergantung pada data. Biasanya metode *cross-validation* digunakan untuk pemilihan fungsi *kernel* ini. Pemilihan fungsi *kernel* yang tepat adalah hal yang sangat penting. Karena fungsi *kernel* ini akan menentukan *feature space* untuk fungsi prediksi yang akan dicari [8].

d. RMSE (Root Mean Square Error)

Keakuratan keseluruhan dari setiap model peramalan baik itu rata-rata bergerak, *eksponensial smoothing* atau lainnya dapat dijelaskan dengan membandingkan nilai yang diproyeksikan dengan nilai aktual atau nilai yang diamati. Untuk tingkat akurasi peramalan dapat diukur dari nilai berikut, adapun rumus dari RMSE adalah sebagai berikut :

$$RMSE = \sqrt{MSE}..... (11)$$

MSE (*Mean Squared Error*) Merupakan rata-rata jumlah kuadrat kesalahan peramalan, yang rumusnya adalah sebagai berikut :

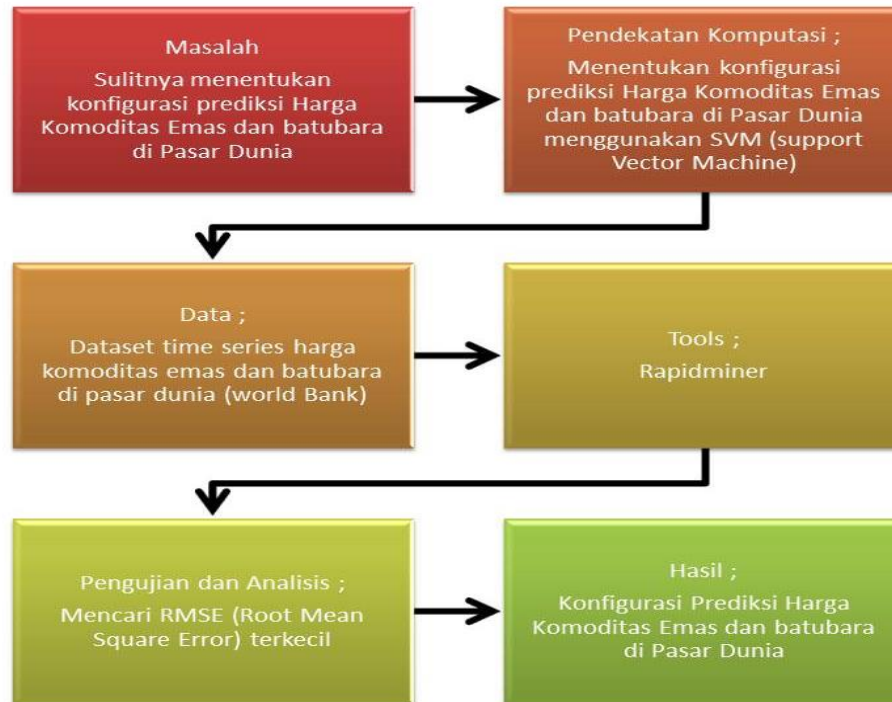
$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - Y'_t)^2..... (12)$$

Semakin kecil nilai RMSE, semakin baik tingkat akurasi prediksinya. Keakuratan sebuah model peramalan dalam melakukan prediksi ditentukan oleh nilai terkecil dari masing-masing metode akurasi data, semakin kecil nilai tersebut semakin akurat sebuah model melakukan prediksi. Untuk mengetahui teknik mana yang terbaik untuk data tertentu maka biasanya dilakukan dengan mencoba beberapa teknik

berbeda dan memilih salah satu yang terbaik.

1.6. Prediksi Harga Komoditas Emas dan Batubara di Pasar Dunia dengan Algoritma *Support Vector Machine*

Pengolahan data dengan menggunakan Algoritma SVM didasarkan pada *k-fold validation*, *c (cost)* dan *kernel*-nya, kemudian dicari tingkat RMSE (*root mean square error*) yang terkecil. RMSE yang terkecil itulah yang dijadikan desain dalam memprediksi harga komoditas emas dan batubara. Untuk lebih memperjelas kerangka pikir, penulis membuat langkah-langkah berupa gambar di bawah ini :



Gambar 2. Tahapan Penelitian

Prediksi harga komoditas emas dan batubara di pasar dunia sangat membantu pemerintah dalam menentukan kebijakan tentang pertambangan khususnya pola produksi dan pengelolaannya.

Dalam memprediksi harga komoditas tersebut, dapat menggunakan teknologi komputer, dalam hal ini peneliti membuat desain prediksi harga dengan menggunakan algoritma SVM (*Support Vector Machines*) berdasarkan data time series yang diambil dari World Bank.

Dalam mendesain prediksi tersebut langkah yang dilakukan peneliti dengan eksperimen dataset yang ada, yaitu dengan mengubah bentuk data set yang semula *univariate* diubah menjadi *multivariate*, kemudian dataset *multivariate* tersebut di pisahkan menjadi inputan data yang bertujuan sebagai data training dalam algoritma SVM, data training yang digunakan adalah inputan data berdasarkan harga periode yang lalu, yaitu harga 2 bulan yang lalu sampai harga 11 bulan yang lalu. Yang kemudian dataset tersebut diolah dengan Algoritma *Support Vector Machine* menggunakan tool rapidminer, dalam pengolahannya ditentukan terlebih dahulu pada *k-fold validation*, *c (cost)* dan *kernel*-nya, kemudian dicari tingkat RMSE (*root mean square error*) yang terkecil. RMSE yang terkecil itulah yang dijadikan desain dalam memprediksi harga komoditas emas dan batubara.

2. METODE PENELITIAN

Data yang diolah pada penelitian ini adalah harga komoditi emas dan batubara di pasar dunia (global) dari bulan Januari 1970 sampai bulan Nopember 2013. Data set bersifat *public* yang diambil dari World Bank. Data tersebut ditampilkan dalam bentuk bulanan dengan satuan dollar sebanyak 527 baris.

Sebelum mengolah data dengan menggunakan *Support Vector Machine* (SVM), data *time series* tersebut terlebih dahulu diubah bentuknya dari *univariate* menjadi *multivariate*. Hal ini dilakukan untuk menghitung prediksi harga yang akan datang / target (Y_t), penentuan harga yang akan datang dipengaruhi oleh harga-harga pada periode-periode sebelumnya,

Pengolahan data awal menggunakan aplikasi Microsoft Excel dimana data univariat yang terdiri satu kolom harga dibuat menjadi beberapa kolom yang termasuk variabel yang mempengaruhi kolom harga periode berikutnya. Untuk lebih jelasnya dapat dilihat pada gambar di bawah ini.

Dataset Univariate		Dataset Multivariate				
Periode	Harga	Yt-4	Yt-3	Yt-2	Yt-1	Yt
1970M01	35.95	34.94	34.99	35.09	35.62	35.95
1970M02	35.42	34.99	35.09	35.62	35.95	35.42
1970M03	35.32	35.09	35.62	35.95	35.42	35.32
1970M04	35.38	35.62	35.95	35.42	35.32	35.38
1970M05	36.19	35.95	35.42	35.32	35.38	36.19
1970M06	37.52	35.42	35.32	35.38	36.19	37.52
1970M07	37.44	35.32	35.38	36.19	37.52	37.44
1970M08	37.44	35.38	36.19	37.52	37.44	37.44
1970M09	37.88	36.19	37.52	37.44	37.44	37.88
1970M10	38.74	37.52	37.44	37.44	37.88	38.74
1970M11	38.87	37.44	37.44	37.88	38.74	38.87
1970M12	39.01	37.44	37.88	38.74	38.87	39.01
1971M01	40.52	37.88	38.74	38.87	39.01	40.52
1971M02	40.10	38.74	38.87	39.01	40.52	40.10
1971M03	40.95	38.87	39.01	40.52	40.10	40.95
1971M04	42.71	39.01	40.52	40.10	40.95	42.71

Gambar 2. Perubahan Data Univariat ke Data Multivariate

Dari gambar di atas dapat dilihat perubahan bentuk *univariate* ke bentuk *multivariate*. Data *multivariate* tersebut menjelaskan bahwa harga yang akan datang / target (Y_t) dipengaruhi oleh harga dari 4 periode sebelumnya.

Adapun input data dalam bentuk *multivariate* yang digunakan dalam penghitungan *Support Vector Machine* (SVM) ini adalah : data input harga 2 bulan sebelumnya sampai dengan data input harga 11 bulan sebelumnya

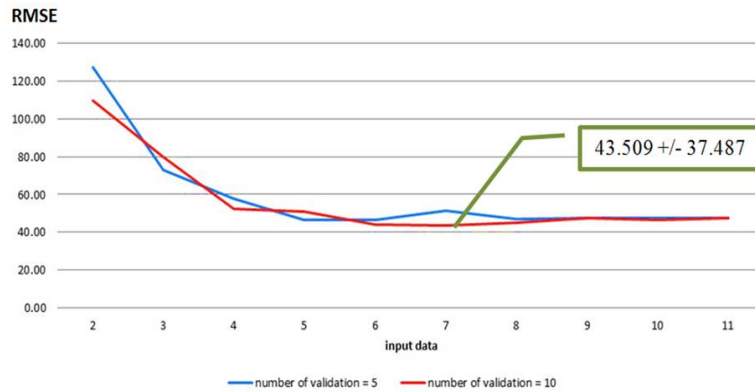
Data input tersebut di atas digunakan sebagai data training dalam penghitungan SVM, jadi tidak semua dataset digunakan untuk penghitungan. Data yang tidak digunakan tersebut secara otomatis menjadi data testing dalam penghitungan SVM.

Pengolahan data pada penelitian ini menggunakan Algoritma SVM dengan konstruksi penghitungan data sebagai berikut : Penghitungan data menggunakan tools rapidminer, yang terlebih dahulu menentukan data inputan (sebanyak 10 inputan) sebagai data training dalam memproses data tersebut, setelah itu eksperimen berlanjut ke penentuan type kernel, nilai C (cost) dan k-fold-nya.

Type kernel yang digunakan adalah : type kernel dot, radial dan polynomial, dan nilai C (cost) berdasarkan 0,1, 0,2, dan 0,3 serta besarnya k-fold (number of validation) range k =5 dan k = 10. Berikut ini hasil dari percobaan yang telah dilakukan dengan beberapa fungsi *kernel* dan memasukkan nilai C (*cost*) serta nilai *range* (*k-fold*) yang telah ditentukan untuk menguji masing-masing kelompok tersebut.

3. HASIL PENELITIAN DAN PEMBAHASAN

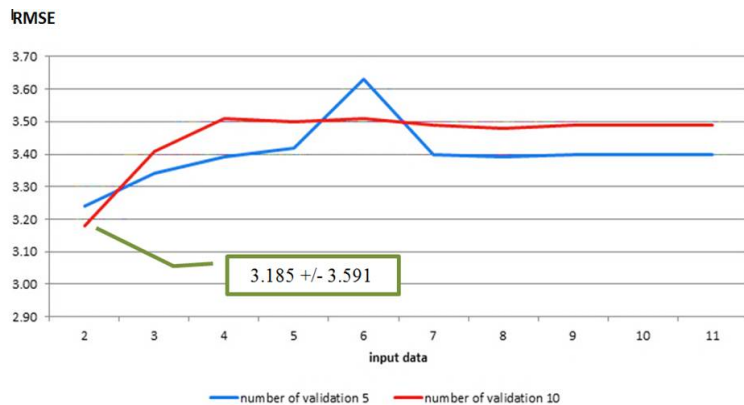
Pengolahan data pada penelitian ini menggunakan rapidminer dan diolah dengan Algoritma *Support Vector Machine* dengan konstruksi penghitungan data sebagai berikut : jumlah data inputan sebanyak 11 dataset *multivariate*, yang telah ditentukan tipe kernel, nilai C (*cost*) dan k-fold-nya, sehingga menghasilkan RMSE terkecil dan dapat dilihat pada gambar di bawah ini :



Gambar 3. RMSE untuk Dataset Harga Komoditas Emas

Dari gambar di atas diketahui bahwa RMSE (*root mean square error*) terkecil ada di *k-fold validation* 10 pada inputan data bulan ke 7 dengan *c (cost) = 0,3* dan *kernel Dot*.

Jadi dapat disimpulkan bahwa metode prediksi harga komoditas Emas dengan RMSE terbaik adalah sebesar **43.509 +/- 37.487** dengan input data **7(tujuh)** bulan sebelumnya, ***k-fold* 10**, **C (*cost*) sebesar 0,3** dan menggunakan ***type kernel dot***. merupakan konfigurasi terbaik untuk memprediksi harga komoditas emas periode berikutnya.



Gambar 4. RMSE untuk Dataset Harga Komoditas Batubara

Dari gambar di atas diketahui bahwa RMSE (*root mean square error*) terkecil ada di *k-fold validation* 10 pada inputan data bulan ke 2 dengan *c (cost) = 0,3* dan *kernel Dot*.

Jadi dapat disimpulkan bahwa metode prediksi harga komoditas batubara dengan RMSE terbaik adalah sebesar **3.185 +/- 3.591** dengan input data **2(dua)** bulan sebelumnya, ***k-fold* 10**, **C (*cost*) sebesar 0,3**

dan menggunakan *type kernel dot* merupakan konfigurasi terbaik untuk memprediksi harga komoditas batubara periode berikutnya.

Konfigurasi terbaik tersebut digunakan untuk prediksi harga komoditi emas dan batubara di pasar dunia untuk periode yang akan datang dengan hasil sebagai berikut ;

Tabel 1 ; prediksi harga komoditas emas

No	Periode	Harga Prediksi (\$)	Harga Riil (\$)	Selisih (\$)	Tingkat Akurasi
1	2013M12	1293.55	1221.51	72.04	94,43%
2	2014M01	1264.97	1244.27	20.70	98.36%
3	2014M02	1251.05	1319.40	68.35	94.81%

Tabel 2 ; prediksi harga komoditas batubara

No	Periode	Harga Prediksi (\$)	Harga Riil (\$)	Selisih (\$)	Tingkat Akurasi
1	2013M12	81.58	84.34	2.76	96,72%
2	2014M01	83.94	81.61	2.33	97.22%
3	2014M02	82.85	81.88	0.97	98.82%

4. KESIMPULAN DAN SARAN

4.1. Kesimpulan

Kesimpulan yang dapat diambil dari penelitian tentang prediksi harga komoditas emas dan batubara ini adalah sebagai berikut :

- RMSE terbaik untuk prediksi harga emas adalah sebesar **43.509 +/- 37.487** sehingga konfigurasi atau model prediksinya adalah dengan input data **7 (tujuh)** bulan sebelumnya, *k-fold 10*, *C (cost)* sebesar **0,3** dan menggunakan *type kernel dot*. Dan untuk batubara RMSE terbaik adalah sebesar **3.185 +/- 3.591** sehingga konfigurasi atau model prediksinya adalah dengan input data **2 (dua)** bulan sebelumnya, *k-fold 10*, *C (cost)* sebesar **0,3** dan menggunakan *type kernel dot*
- Tingkat akurasi prediksi harga komoditas emas untuk periode yang akan datang sebesar 94,43 % dan tingkat akurasi prediksi harga komoditas batubara adalah sebesar 96,72%.

4.2. Saran

Dalam penelitian ini telah menghasilkan prediksi yang optimal dan akurat, namun demikian untuk penelitian selanjutnya agar mendapatkan hasil yang lebih baik, perlu memperhatikan hal-hal sebagai berikut :

- Data observasi untuk simulasi prediksi jumlahnya perlu diperbanyak, agar proses training mampu mendapatkan model yang lebih, sehingga menghasilkan tingkat akurasi prediksi lebih akurat.
- Untuk mendapatkan hasil prediksi yang lebih baik dapat dilakukan penelitian lebih lanjut dengan melakukan optimasi dengan menggunakan algoritma yang lain seperti Naïve Bayes dan lainnya.

UCAPAN TERIMA KASIH

Terselesaikannya penelitian ini berkat adanya bantuan dan dukungan dari berbagai pihak. Pada kesempatan ini perkenankan penulis menyampaikan ucapan terima kasih yang setulus-tulusnya kepada Bapak Dr. Ir. Edi Noersasongko, M.Kom selaku Rektor Universitas Dian Nuswantoro, Bapak Dr. Abdul Syukur, MM selaku Direktur MTI Universitas Dian Nuswantoro, Bapak Purwanto, SSi., M.Kom dan Catur Supriyanto, M.C.S selaku dosen dan pembimbing. Semua pihak yang telah membantu penulis dan tidak dapat disebutkan satu persatu.

PERNYATAAN ORIGINALITAS

“Saya menyatakan dan bertanggung jawab dengan sebenarnya bahwa ArtikeLL ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya”

[Eko Pudjianto – P31.2011.01136]

DAFTAR PUSTAKA

- [1] Guohao Zhao, Kangmo Che, Yushan Zhao, *A Study on Predicting Coal Market Price in China Based on Time Sequence Models*, International Journal of Business and Social Science Vol. 3 No. 4 [Special Issue - February 2012], 2012.
- [2] Mahsa KargaraniFarahani, Soheil Mehralian, *Comparison Between Artificial Neural Network and Neuro-Fuzzy for Gold Price Prediction*, 13th Iranian Conference on Fuzzy Systems (IFSC), 2013.
- [3] Z. Ismail, A. Yahya and A. Shabri, *Forecasting Gold Prices Using Multiple Linear Regression Method*, American Journal of Applied Sciences 6 (8): 1509-1514, 2009 ISSN 1546-9239 2009 Science Publications, 2009.
- [4] Heru Supriyanto, *Implementasi Support Vector Machines untuk Memprediksi Arah Pergerakan Harga Harian Valuta Asing (EUR/USD, GBP/USD, Dan USD/JPY) dengan Metode Kernel Trick Menggunakan Fungsi Kernel Radial Basis Function*. Tesis, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Brawijaya, Malang, Indonesia, 2012.
- [5] Himpunan Pemerhati Lingkungan Hidup Indonesia (HPLI) [online] <http://www.hpli.org/tambang.php> [accessed; 27/01/2014]
- [6] Nugroho, *SVM : Paradigma Baru dalam Softcomputing dan Aplikasinya*, Konferensi Nasional Sistem & Informatika 2008, Bali, 2008.
- [7] J. Han, *Data Mining : Concepts and techniques*, Morgan Kaufman, 2006.
- [8] N. Cristianini, J. Shawe-Taylor, *An Introduction to Support Vector Machines and other kernel-based learning method*, Cambridge University Press, New York, 2000.