

SUPPORT VECTOR MACHINE BERBASIS CHI SQUARE UNTUK PREDIKSI HARGA BERAS ECER KABUPATEN POHUWATO

Sunarto Taliki¹, Ivo Colanus Rally Drajana², Andi Bode³

^{1,3}Teknik Informatika, Universitas Ichsan Gorontalo

²Informatika, Universitas Pohuwato

e-mail: ¹sunartotaliki@gmail.com, ²ivocolanusrally@gmail.com,

³andibode22@gmail.com

Abstract: *One of the staple foods for most Indonesians is rice. Rice is one of the staple foods of the Indonesian people, and the need for rice is also increasing, given the large population of Indonesia. The rise and fall of rice prices also affects farmers because of higher production. The solution to overcome concerns about changes in rice prices is to make price predictions. One way to find out the status of rice sales is to make predictions using the Support Vector Machine algorithm that uses Chi Square. The results of the experiment succeeded in predicting the price of rice. The minimum error for the Support Vector Machine algorithm is RMSE 733,061. Now the planned model is almost complete, because the comparison of the results of the rice price estimation results in an average value of 95.82%. Thus, the proposed method was published successfully.*

Keywords: *support vector machine; chi square; prediction; rice price*

Abstrak: Salah satu bahan pokok wajib sebagian besar penduduk Indonesia adalah beras. Beras merupakan salah satu makanan pokok masyarakat Indonesia, dan kebutuhan akan beras juga semakin meningkat, mengingat jumlah penduduk Indonesia yang besar. Naik turunnya harga beras juga mempengaruhi petani karena produksi yang lebih tinggi. Solusi untuk menghadapi perubahan kekhawatiran harga beras adalah dengan memprediksi harga. Salah satu cara untuk mengetahui status penjualan beras adalah dengan melakukan prediksi menggunakan algoritma Support Vector Machine yang menggunakan Chi Square. Hasil eksperimen memprediksi harga beras berhasil. Kesalahan minimum untuk algoritma Support Vector Machine adalah RMSE 733.061. Sekarang model yang direncanakan hampir selesai, karena perbandingan hasil estimasi harga beras menghasilkan nilai rata-rata 95,82%. Dengan demikian, metode yang diusulkan berhasil diterbitkan.

Kata kunci: support vector machine; chi square; predikdi; harga beras

PENDAHULUAN

Beras merupakan salah satu makanan pokok Indonesia, dan beras merupakan salah satu sumber gizi utama. Untuk itu, suplai dan distribusi beras menjadi sangat penting. Penduduk Indonesia, permintaan beras semakin meningkat, mengingat jumlah penduduk Indonesia yang besar dan terus bertambah, namun harga beras yang beredar dipasaran semakin meningkat karena permintaan beras semakin meningkat, pedagang banyak yang

menjual beras. kualitas buruk dan tidak berwujud. Dari segi ekonomi, harga beras merupakan faktor penting yang perlu mendapat perhatian karena kenaikan harga beras dapat mempengaruhi perekonomian lokal. Permasalahan yang dihadapi masyarakat akibat naik turunnya harga beras berdampak pada petani karena produksi yang tinggi. Oleh karena itu, melihat permasalahan tersebut, perlu adanya respon terhadap perubahan harga jual beras agar dapat memberikan solusi pengambilan keputusan berdasarkan prediksi kondisi harga..

Cara lain untuk mendapatkan ide jualan beras adalah dengan mencari keberuntungan atau kejelian. Memprediksi atau memprediksi adalah metode memperkirakan dengan mengukur kuantitas sesuatu yang akan mengikuti di masa lalu yang telah diperiksa secara ilmiah dengan cara-cara tertentu. (Iku, 2019). Data mining didefinisikan sebagai data mining atau mencoba untuk mengekstrak data yang berharga dan berguna dari organisasi yang sangat besar. Dalam data mining terdapat beberapa metode prediksi yaitu Linear Regression, C-45, Naïve Bayes, K-Nearest Neighbor, Support Vector Machine dan masih banyak lagi. Dalam tutorial ini menggunakan Algoritma Support Vector Machine dengan menambahkan Seleksi Fitur Chi Square.

Algoritma Support Vector Machine merupakan metode yang didasarkan pada studi statistik dan memberikan hasil yang lebih menjanjikan dibandingkan metode lainnya. Vector Support Engine juga bekerja dengan baik untuk data tingkat tinggi yang menggunakan sistem kernel dan memiliki keunggulan dalam memberikan kinerja yang sangat baik untuk visi kelompok waktu. (Bode, 2019). Support Vector Machinemenawarkan performa terbaik dibandingkan sistem lainnya (Laia, 2020). Metode SVM tidak memberikan hasil yang akurat ketika beberapa komponen tidak berfungsi, tidak semua komponen diperlukan untuk operasi. Seleksi segmen bekerja dengan secara langsung mengurangi jumlah bagian dan memilih bagian yang benar-benar memberikan informasi, jumlah bagian sangat berkurang dan masalah ekstrim terpecahkan. Metode SVM memberikan operasi yang efisien, ketika komponen yang tidak perlu dihilangkan (Drajana, 2017). Seleksi Fitur adalah proses memilih satu set komponen kunci dengan menghapus komponen yang tidak aktif. Chi Square merupakan pemilihan komponen yang digunakan untuk menghitung tingkat ketergantungan kelompok (Pratama, 2018). Dalam

penelitian ini, chi square yang digunakan akan dicoba untuk memprediksi harga beras.

METODE

Support Vector Machine memungkinkan sistem operasi komputer untuk menggunakan nomor yang berbeda untuk melakukan tugas kernel. Svm memiliki konsep sentral, yaitu menentukan *hyperplane* mana yang terbaik untuk menyediakan ruang atau pemisahan antara dua grup yang telah diprogram sebelumnya. Cara kerja mesin *Support Vector* adalah dengan menemukan *hyperplane* yang tepat yang memberikan ruang atau pemisahan antara dua kelompok (Tineges, 2020). Metode ini menggunakan pendekatan teori sistem linier dengan penggunaan komponen lanjutan, menggunakan studi matematis. Tingkat akurasi desain yang dihasilkan oleh proses transisi dan svm sangat bergantung pada kinerja kernel dan parlemen yang digunakan. (Lumbanraja, 2020).

Seleksi bagian adalah proses pemilihan komponen yang terlibat, dengan menemukan komponen yang tepat, menemukan komponen yang tepat adalah tugas yang sulit. Manfaatkan pemilihan komponen, kurangi persyaratan penyimpanan, kurangi pelatihan, dan tingkatkan kinerja. *Chi square* menjadi pilihan tim untuk melihat saling ketergantungan kata dan kelompok (Nisa, 2019). Beberapa tingkatan uji *chi square* yang dapat digunakan, yaitu:

1. Tidak ada sel yang memiliki bilangan real atau panjang gelombang real (F_0) yaitu 0 (nol);
2. Jika tabel darurat dalam mode 2×2 , maka tidak akan ada sel dengan jumlah yang diharapkan atau frekuensi yang diharapkan (F_h) kurang dari 5;
3. Jika bentuk tabel melebihi 2×2 , misalnya 2×3 , maka jumlah sel dengan frekuensi harapan (F_h)

kurang dari 5 tidak boleh melebihi 20%.

Uji persamaan uji *chi square* dapat dilakukan sebagai berikut::

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$$

Dengan

O_i = nilai observasi ke- i

E_i = nilai ekspektasi ke- i

HASIL DAN PEMBAHASAN

Objek penelitian adalah Harga Beras Ecer Di Kabupaten Pohuwato. Penelitian berlokasi di Dinas Perindustrian Perdagangan dan UMK Kabupaten Pohuwato.

Pengumpulan Data

Data harga beras yang dikumpulkan dengan teknik wawancara dan observasi.

Pengolahan Data

Dari Data yang diperoleh dilakukan pengolahan data khusus buat harga beras. Data yang ada merupakan data harian yang kemudian di rekap menjadi data mingguan dengan jumlah 55 record data.

Tabel. Sampel Data Harga Beras

Harga Besar Ecer/Minggu		
Bulan	Tgl	Harga
Maret	7	8,000
Maret	21	8,000
Maret	28	8,000
April	4	8,000
April	11	8,500
April	18	9,000
April	25	9,000
Mei	2	9,000
Mei	9	9,000
Mei	16	9,000

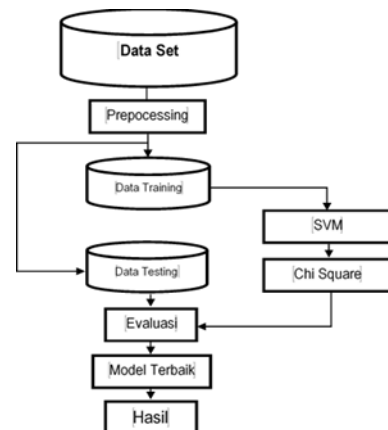
Dari data di atas kemudian diubah menjadi data askending ke deskending, dan diubah menjadi data periode. Sampel data periode:

Tabel. Multi Variabel 4 Periode

xt-4	xt-3	xt-2	xt-1	xt
9,50 0	9,50 0	10,00 0	10,00 0	10,00 0
9,00 0	9,50 0	9,500	10,00 0	10,00 0
9,00 0	9,00 0	9,500	9,500	10,00 0
9,00 0	9,00 0	9,000	9,500	9,500
9,50 0	9,00 0	9,000	9,000	9,500
9,50 0	9,50 0	9,000	9,000	9,000
9,50 0	9,50 0	9,500	9,000	9,000
9,50 0	9,50 0	9,500	9,500	9,000
9,50 0	9,50 0	9,500	9,500	9,500
8,50 0	9,50 0	9,500	9,500	9,500

Pemodelan

Berikut ini adalah model pada penelitian ini:



Gambar. Diagram Metode Penelitian

Parameter Support Vector Machine

Eksperimen dilakukan dengan penentuan parameter *Support Vector Machine* seperti variabel input

independen atau variabel periode $xt-i$ dan pemilihan *type* kernel, dilakukan uji coba yakni untuk memperoleh model terbaik dengan nilai *error* terkecil.

Tabel. Hasil Eksperimen SVM

Periode	Validation	Kernel	RMSE
1	10	dot	737,626
2	10	dot	737,670
3	10	dot	733,061
4	10	dot	753,534

Sesuai dengan tabel 3 merupakan tabel keseluruhan eksperimen metode *Support Vector Machine* dengan variabel periode 1 sampai 4 untuk data harga beras menggunakan *number of validation* 10 dengan *type* kernel *dot* yakni di ambil berdasarkan nilai *root means square error*, maka ditemukan model terbaik 733,061 pada algoritma *support vector machine* pada periode 3.

Parameter SVM dan Chi Square

Tahap selanjutnya adalah melakukan eksperimen uji coba dengan penambahan seleksi fitur *chi square* pada algoritma *support vector machine* dengan tujuan untuk mendapatkan model terbaik dengan variabel periode 1 sampai 4.

Tabel. SVM dan Chi Square

Periode	Validation	Kernel	RMSE
1	10	dot	737,626
2	10	dot	737,670
3	10	dot	733,061
4	10	dot	753,534

Sesuai dengan tabel 4 yaitu tabel rangkuman eksperimen algoritma *support vector machine* dan *chi square*. Dari hasil yang dilihat, model terbaik yakni variabel periode 3 dengan *number of validation* 10 dan *type* kernel *dot*.

Evaluasi

Pada tahapan ini, yaitu evaluasi dari kedua uji coba, antara algoritma *svmdan svm* menggunakan *chi square*. Perbandingan tersebut dilihat pada tabel berikut :

Tabel 5. Perbandingan Hasil Eksperimen

Algoritma	Periode	RMSE
SVM	3	733,061
SVM dan Chi Square	3	733,061

Sesuai dengan tabel 5 menjelaskan bahwa model terbaik ada pada variabel periode 3 dengan nilai *error* 733,061. Dari tabel diatas menjelaskan bahwa penambahan seleksi fitur *chi square* tidak berpengaruh terhadap hasil nilai *error* atau performa dari algoritma *support vector machine*.

Implementasi

Tahapan ini yakni tahapan penerapan model yang diperoleh dari beberapa tahapan sebelumnya untuk prediksi harga beras. Hasil prediksi harga beras tersebut ada pada tabel berikut:

Tabel. Hasil Prediksi

April	Harga Beras
Minggu-1	Rp. 8.500
Minggu-2	Rp. 8.500
Minggu-3	Rp. 8.499
Minggu-4	Rp. 8.499

Sesuai dengan tabel 6 merupakan hasil prediksi 4 minggu kedepan pada bulan April 2022 untuk harga beras pada Kabupaten Pohuwato. Perbandingan jumlah harga yang sebenarnya dan harga hasil prediksi terdapat pada tabel berikut :

Tabel. Persentasi Keakuratan

April	Hasil Prediksi	Data Aktual	Persentase
Minggu -1	Rp. 8.500	Rp. 9.000	94,44%
Minggu -2	Rp. 8.500	Rp. 9.000	94,44%
Minggu -3	Rp. 8.499	Rp. 9.000	94,43%
Minggu -4	Rp. 8.499	Rp. 8.500	99,98%

Pada tabel 7 yakni merupakan perbandingan antara hasil prediksi harga

beras Kabupaten Pohuwato April 2022. Hasil prediksi menghasilkan nilai keakuratan rata-rata 95,82%, Dengan demikian, hasil prediksi berhasil karena akurasi hasil mendekati nilai penyelesaian.

Sampel perhitungan persentasi:

$$\frac{\text{hasil prediksi}}{\text{nilai aktual}} \times 100$$

$$\frac{8.500}{9.000} \times 100 = 94,44\%$$

SIMPULAN

Berdasarkan hasil pengujian, prediksi harga beras menggunakan SVM dan chi square telah selesai di lakukan. Hasil estimasi kemudian dapat digunakan untuk memperkirakan atau membuat keputusan kebijakan. Tingkat kesalahan kecil dalam desain algoritma Support Vector Machine RMSE 733,061 pada algoritma Support Vector Machine dan chi square tingkat RMSE 733,061. Dengan demikian dari hasil eksperimen menunjukkan bahwa chi square tidak memberikan dampak bagi algoritma support vector machine, karena hasil yang didapatkan tidak mengalami perubahan hasil nilai error. Kemudian dapat disimpulkan Model yang direncanakan hampir sempurna, karena analogi hasil tes memprediksi harga beras Kabupaten Pohuwato April 2022 menghasilkan nilai keakuratan dengan rata-rata 95,82%. Dengan demikian metode yang diusulkan dinyatakan berhasil, meskipun tanpa harus adanya penambahan fitur seleksi. Saran bagi peneliti selanjutnya, agar dapat mencoba menggunakan seleksi fitur yang lainnya pada algoritma Support Vector Machine dengan tujuan agar dapat memberikan hasil nilai error lebih kecil untuk data harga beras.

DAFTAR PUSTAKA

Bode, A. (2019). “Perbandingan Metode Prediksi Support Vector Machine

Dan Linear Regression Menggunakan Backward Elimination Pada Produksi Minyak Kelapa”. *Simtek: jurnal sistem informasi dan teknik komputer*, 4(2), 104-107.

Drajana, I. C. R. (2017). : “Metode support vector machine dan forward selection prediksi pembayaran pembelian bahan baku kopra”. *ILKOM Jurnal Ilmiah*, 9(2), 116-123.

Iku, M. H., Mustofa, Y. A., & Kumala, I. S. (2019). “Metode K-Nearest Neighbor untuk Memprediksi Harga Eceran Beras di Pasar Tradisional Gorontalo”. *Jurnal Cosphi*, 3(2).

Laia, M. L., & Setyawan, Y. (2020). “Perbandingan Hasil Klasifikasi Curah Hujan Menggunakan Metode SVM dan NBC”. *Jurnal Statistika Industri dan Komputasi*, 5(02), 51-61.

Lumbanraja, F. R., Ira Hariati, B. S., Kurniawan, D., & Aristoteles, A. (2020). “Prediksi Jumlah Penderita Penyakit Tuberkulosis Di Kota Bandar Lampung Menggunakan Metode Svm (Support Vector Machine)”. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, 7(3), 320-330.

Nisa, A., Darwiyanto, E., & Asror, I. (2019). “Analisis Sentimen Menggunakan Naive Bayes Classifier dengan Chi-Square Feature Selection Terhadap Penyedia Layanan Telekomunikasi”. *eProceedings of Engineering*, 6(2).

Pratama, N. D., Sari, Y. A., & Adikara, P. P. (2018). “Analisis Sentimen Pada Review Konsumen Menggunakan Metode Naive Bayes Dengan Seleksi Fitur Chi Square Untuk Rekomendasi Lokasi Makanan Tradisional”. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN*, 2548, 964X.

Tineges, R., Triayudi, A., & Sholihati, I. D. (2020). “Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)”. *Jurnal Media*

Informatika Budidarma, 4(3), 650-658.