

OPTIMASI K-NEAREST NEIGHBOR DENGAN PARTICLE SWARM OPTIMIZATION UNTUK MEMPREDIKSI HARGA TEMBAKAU

Luluk Suhartini¹, M. Burhanis Sulthan², Imam Wahyudi³

¹²³Teknologi Informasi, Fakultas Teknik, Institut Sains dan Teknologi Annuqayah

¹lulukkhafi@gmail.com, ²burhan.sulthan33@gmail.com, ³wahyudigo94@gmail.com

ABSTRAK

Komoditas tembakau merupakan salah satu penyumbang pendapatan keuangan terbesar bagi negara, menjadi alasan penting dalam menempatkan komoditas tembakau dan produk olahannya sebagai komoditas yang strategis. namun dalam bidang pertanian, hampir semua aktivitas produksi sering dihadapkan dengan permasalahan fluktuasi harga dan hasil produksi. Untuk itu perlu dilakukan peramalan harga pada masa yang akan datang untuk dijadikan acuan dalam mengambil keputusan agar dapat mengurangi risiko serta meningkatkan keuntungan. Dalam penelitian ini digunakan metode K-NN dengan optimasi PSO. Metode ini merupakan salah satu metode yang baik untuk diterapkan dalam peramalan harga dan hasil penelitian menunjukkan bahwa performa K-NN meningkat setelah dioptimasi dengan algoritma PSO. Hal ini dibuktikan dengan penurunan nilai eror (RMSE) pada metode K-NN yang semula 0.093 menjadi 0.072 setelah dioptimasi dengan algoritma PSO. Hasil ini membuktikan bahwa metode K-NN dengan optimasi PSO lebih baik dibandingkan dengan metode K-NN saja.

Kata Kunci: Harga tembakau, Prediksi harga, K-NN, Optimasi, *Particle Swarm Optimization*.

ABSTRACT

Tobacco commodity is one of the biggest contributors to financial income for the country, an important reason in placing tobacco commodities and their processed products as strategic commodities. but in agriculture, almost all production activities are often faced with problems of price fluctuations and production results. For this reason, price forecasting needs to be done in the future to be used as a reference in making decisions in order to reduce risk and increase profits. The method used in this research is K-NN with PSO optimization. The K-NN method is one of the good methods to be applied in price forecasting and the results of the study show that the K-NN performance increases after being optimized with the PSO algorithm. This is evidenced by the decrease in the error value (RMSE) in the K-NN method, which was originally 0.093 to 0.072 after being optimized with the PSO algorithm. These results prove that the K-NN method with PSO optimization is better than the K-NN method.

Keyword: Price of tobacco, K-NN, Optimation, Particle Swarm Optimization.

I. PENDAHULUAN

Pertanian tembakau beserta hasil olahannya menjadi komoditas paling signifikan dalam perdagangan dan industri dalam negeri. Salah satu faktor penentunya yaitu terdapat keunggulan perbandingan (*comparative advantage*) pada komoditas tembakau yang merupakan ciri khas Indonesia dan tidak diproduksi oleh negara lain, khususnya pada produk utamanya yakni *kretek* dengan pangsa pasar tradisional (*captive market*) yang luar biasa di dalam negeri. Di Jawa Timur, wilayah Madura menjadi pusat terbesar pertanian tembakau. Lebih dari 50% pasokan tembakau nasional berasal dari Jawa Timur. Itu artinya produksi tembakau Madura memiliki peran penting sebagai penentu produksi tembakau dalam skala nasional. Sentra produksi tembakau di Madura masih berada di Kabupaten Pamekasan. Produksi tembakau rajangan Kabupaten Pamekasan pada tahun 2008 mencapai 16.384 ton [1].

Namun dalam bidang pertanian, hampir semua aktivitas produksi sering dihadapkan dengan berbagai permasalahan yakni hasil produksi dan harga yang tidak tetap dan terus berubah setiap tahunnya sehingga mengakibatkan naik turunnya pendapatan bagi para petani dan meningkatkan resiko bagi pelaku usaha yang berinvestasi di bidang pertanian [2]. Untuk itu perlu dilakukan peramalan harga pada masa yang akan datang agar mempunyai gambaran mengenai kondisi tentang keadaan dimasa depan untuk dijadikan acuan dalam mengambil keputusan agar dapat mengurangi risiko serta meningkatkan keuntungan.

Terdapat beberapa penelitian yang sudah dilakukan sebelumnya dalam memprediksi harga tembakau yaitu penelitian [3] yang menggunakan metode peramalan *Moving Average*, *Exponential Smoothing* dan *Regresi Linier (Trend)* untuk memprediksi harga tembakau. Dalam penelitian [4] dilakukan prediksi harga tembakau dengan menggunakan metode *ARIMA Box-Jenkins*, sedangkan penelitian [5] menggunakan beberapa metode

untuk memprediksi harga tembakau yaitu metode *Box-Jenkins*, *Winters*, *Decomposition*, *Trend* dan *Double Exponential Smoothing*. Namun beberapa penelitian diatas masih menghasilkan nilai eror yang cukup tinggi.

Metode *K-Nearest Neighbour* adalah salah satu metode *Machine Learning* yang pada beberapa tahun terakhir ini banyak digunakan untuk peramalan [6]. Beberapa penelitian yang menerapkan metode *k*-NN untuk melakukan prediksi adalah penelitian yang dilakukan oleh Shatnawi, Mohammed K. Ali dkk [7] yang memprediksi harga saham. dalam penelitian tersebut dijelaskan bahwa algoritma *k*-NN memiliki kinerja yang bagus dengan rasio kesalahan yang kecil. Selanjutnya penelitian yang dilakukan oleh Imandoust Sadegh Bafandeh dan Bolandraftar Mohammad [8] yang menerapkan *k*-NN untuk memprediksi Ekonomi di Iran.

Metode *K-Nearest Neighbour* dianggap memiliki kelebihan salah satunya adalah dapat digunakan pada data dalam jumlah banyak dan tangguh terhadap suatu data latih yang noise. Namun disini lain juga terdapat kelemahan pada metode *K*-NN yaitu optimalisasi dalam penentuan nilai *k* tetangga terdekat masih kurang dan harus menyeleksi atribut yang akan dipilih untuk mendapatkan hasil paling baik [9].

Untuk mengatasi kelemahan *K*-NN maka dapat digunakan *Particle Swarm Optimization* (PSO) yang memiliki keunggulan kinerja dalam banyak topik optimasi dengan *speed* yang lebih baik dan stabil dalam tingkat konvergensi [10]. Terdapat beberapa penelitian yang menerapkan metode *K*-NN dan PSO salah satunya adalah penelitian yang dilakukan oleh Mohammad Efendi Lasulika [2]. Dalam penelitian tersebut disimpulkan bahwa kinerja algoritma *K*-NN meningkat melalui fitur seleksi yang terdapat di dalamnya. Jika dibandingkan antara hasil metode *K*-NN dengan optimasi PSO dengan metode *K*-NN tanpa dioptimasi maka algoritma pertama menunjukkan hasil yang lebih baik dengan nilai RMSE 0,06 serta nilai akurasi metode prediksi mencapai 98,7%.

Penelitian lain yang menerapkan metode *K*-NN dengan optimasi *Particle Swarm Optimization* dilakukan oleh Sri Rahmiyajuida Haras [11] untuk memprediksi harga *Crude Palm Oil*. Prediksi yang menambahkan *Particle Swarm Optimization* tersebut bertujuan untuk mengoptimasi nilai RMSE setelah dilakukan pengujian menggunakan *K-Nearest Neighbor*. Dari hasil pengujian *K-Nearest Neighbor* dengan penambahan PSO tersebut dihasilkan nilai Error sebesar 0.072. Nilai tersebut lebih rendah dibanding percobaan yang hanya menggunakan *K-Nearest Neighbor* saja yakni sebesar 0.078. Sedangkan pada penelitian yang dilakukan oleh Andi Nirfah [12] yang mengklasifikasi masalah kredit sepeda motor dengan menerapkan algoritma *K*-NN berbasis *Particle Swarm Optimization* menunjukkan peningkatan akurasi dari 91,18% ke 94,92%. Hal ini menunjukkan bahwa penambahan *Particle Swarm Optimization* pada *K-Nearest Neighbor* mampu meningkatkan akurasi atau mengurangi nilai RMSE dengan baik. Pada penelitian ini diusulkan optimasi *K*-NN dengan *Particle Swarm Optimization* (PSO) untuk memprediksi harga tembakau di Kabupaten Pamekasan. Penambahan *Particle Swarm Optimization* (PSO) disini berperan mengoptimalkan kerja *K*-NN sehingga akurasi dalam memprediksi harga tembakau bertambah.

2. TINJAUAN PUSTAKA

2.1 Penelitian Terkait

Berikut ini merupakan beberapa penelitian terkait yang telah dilakukan:

Penerapan PSO dan *K*-NN dalam penelitian [13] untuk memonitoring pengendalian hama pada tanaman jeruk menyebutkan bahwa akurasi KNN yang kurang maksimal dapat diatasi dengan menambahkan PSO sebagai fitur seleksi dan optimasi nilai *k*. Terjadi peningkatan akurasi pada sebelum dan sesudah optimasi yaitu akurasi tertinggi pada KNN mencapai 90%, dan akurasi tertinggi pada PSO-KNN mencapai 96,25%. Adanya peningkatan akurasi menunjukkan bahwa PSO mampu memperbaiki kekurangan yang ada pada KNN.

Sedangkan menurut peneliti [2] dalam memprediksi harga jagung, penerapan *Particle Swarm Optimazation* dalam algoritma *K*-NN menunjukkan hasil yang lebih baik dari pada algoritma *K*-NN tanpa seleksi fitur dengan nilai RMSE yang paling rendah yakni 0.06 dengan nilai akurasi mencapai 98,7 % yang ditunjukkan pada data jagung dengan 4 variabel dan *k* 7 sebagai parameter, serta 5 *max Of Generation* 40 untuk nilai populasinya.

2.2 Landasan Teori

2.2.1 Data Mining

Data Mining atau *Knowledge Discovery in Databases* (KDD) merupakan istilah untuk menemukan pengetahuan dalam sebuah *database* baik berupa pengetahuan atau data dari tahun-tahun sebelumnya atau peramalan untuk tahun-tahun yang akan datang dengan teknik-teknik tertentu seperti matematika, statistika, kecerdasan buatan serta *machine learning* guna mendapatkan informasi serta pengetahuan yang bermanfaat untuk mengambil keputusan pada tahun-tahun yang akan datang [14]

2.2.2 Normalisasi Data

Pada umumnya Dataset mempunyai atribut dengan skala yang berbeda-beda sehingga akan menimbulkan ketimpangan data apabila atribut dengan skala yang berbeda tersebut digunakan untuk menghitung

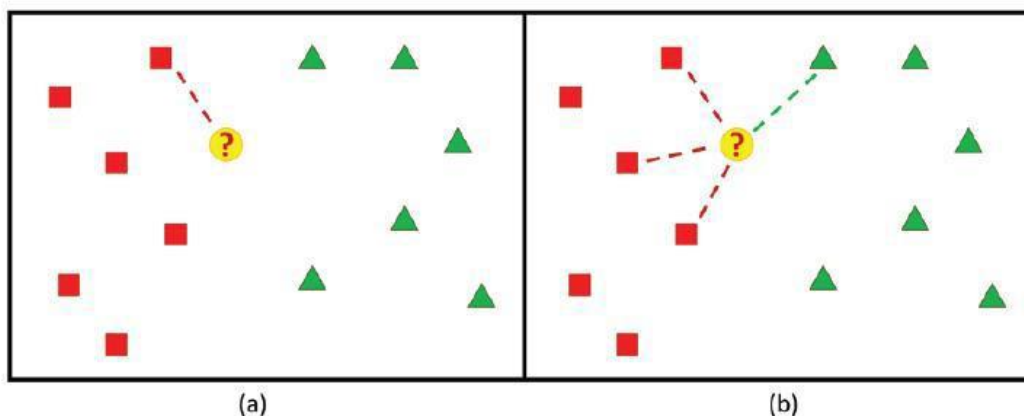
dataset maupun menghitung jarak antar atribut pada K-NN [13]. Maka dari itu agar tidak terjadi ketimpangan data perlu dilakukan normalisasi data antara atribut 0 sampai 1 melalui persamaan berikut:

$$\text{Normalisasi} = (X - \text{minimal}) / (\text{maximal} - \text{minimal}) \quad (2.1)$$

Ket: X = Data asli / Data awal
 maximal = Data maksimum
 minimal = Data minimum

2.2.3 K-Nearest Neighbor (K-NN)

Algoritma K-NN adalah metode Algoritma dengan pengawasan pengetahuan yang bertujuan untuk menemukan pola yang baru pada sebuah data [2]. K-NN adalah teknik klasifikasi yang mendasar dan paling sederhana ketika ada sedikit atau tidak ada pengetahuan sebelumnya tentang distribusi data. *Nearest Neighbor* (NN) mempunyai aturan paling sederhana dalam K-NN ketika nilai K-nya sama dengan 1. Setiap sampel yang digunakan dalam metode KNN harus diklasifikasikan seperti sampel sekitarnya. Oleh sebab itu, apabila tidak diketahui klasifikasi sampelnya, maka klasifikasi sampel tetangga terdekat dapat dipertimbangkan untuk melakukan prediksi. Selain dapat digunakan untuk klasifikasi, metode KNN juga dapat digunakan untuk regresi, hanya dengan menentukan nilai properti dari objek untuk mencari nilai rata-rata k tetangga terdekatnya. Hal ini berguna untuk mempertimbangkan kontribusi para tetangga, sehingga tetangga yang lebih dekat lebih berkontribusi dari pada yang lebih jauh [8].



Gambar 2.1. (a) Aturan keputusan 1-NN: inti ? ditugaskan ke kelas di sebelah kiri; (b) Aturan keputusan KNN, dengan $K = 4$: inti ? ditugaskan ke kelas di sebelah kiri

1) Jarak Matrik

Seperi disebutkan sebelumnya KNN membuat prediksi berdasarkan hasil dari K tetangga yang terdekat dengan titik itu. Oleh karena itu, untuk membuat prediksi dengan KNN, kita perlu mendefinisikan metrik untuk mengukur jarak antara titik kueri dan dari data sampel. Salah satu pilihan paling populer untuk mengukur jarak ini dikenal sebagai Euclidean. Langkah-langkah lain termasuk Euclidean kuadrat, City-blok, dan Chebychev, dengan model sebagai berikut :

$$D(x, p) = \sqrt{(x_1 - p_1)^2 + (x_2 - p_2)^2} \quad [8] \quad (2.2)$$

dimana:

x = data sampel

p = data uji

D = Jarak antara titik pada data uji y dengan data latih x yang akan diprediksi, yang mana $y = y_1, y_2, \dots, y_i$; $x = x_1, x_2, \dots, x_i$; i merupakan nilai atribut sedangkan n adalah dimensi atribut.

2) Regresi K-Nearest Neighbor

K-Nearest Neighbor merupakan dasar dari pendekatan *Unsupervised Nearest Neighbor* (UNN). Masalah dalam regresi adalah untuk memprediksi nilai output $y \in \mathbb{R}^d$ untuk memberikan nilai input $x \in \mathbb{R}^q$ berdasarkan set input-output N contohnya $((x_1, y_1), \dots, (x_N, y_N))$. Tujuannya adalah untuk mempelajari fungsi $f: x \rightarrow y$ yang dikenal sebagai fungsi regresi. Dataset terdiri dari pasangan (x_i, y_i) dengan $(x_i, y_i) \in X \times Y$. Untuk pola baru x' , regresi KNN menghitung rata-rata nilai fungsi k terdekatnya seperti rumus berikut ini:

$$f_{knn}(x') = \frac{1}{K} \sum_{i \in Nk(x')} y_i \quad (2.3)$$

dimana:

x' : Perkiraan
 K : Jumlah dari semua *nearest neighbor*
 $Nk(x')$: *nearest neighbor*
 y_i : Output tetangga terdekat

Dengan set $NK(x')$ yang berisi indeks K - tetangga terdekat dari x' . Gagasan KNN diasumsikan pada lokalitas dalam ruang data: Nilai output y diharapkan mempunyai nilai yang sama dengan nilai $f(x)$ dalam lingkungan lokal pola x . Sehingga untuk x' yang tidak diketahui, labelnya harus sama dengan label pada pola paling dekat, yang dimodelkan dengan rata-rata nilai output dari sampel K terdekat. KNN telah terbukti dengan baik di berbagai aplikasi, misalnya, dalam mendeteksi quasar di dataset antar bintang [15].

Penghitungan dalam Algoritma K-NN dijelaskan dalam langkah-langkah berikut:

- Menentukan parameter k
- Menghitung jarak data yang baru dengan seluruh data latih
- Mengurutkan jarak minimum ke- k dan menetapkan tetangga yang paling dekat
- Menghitung rata-rata jarak semua tetangga
- Menjumlahkan jarak dari tetangga terdekat untuk dijadikan hasil prediksi data baru

2.2.4 K-Fold Cross Validation

K-Fold Cross Validation adalah suatu algoritma yang digunakan agar rata-rata keberhasilan dapat diketahui melalui pengulangan atribut masukan secara acak dengan menggunakan metode validasi melalui cara membagi data menjadi k bagian, selanjutnya bagian-bagian tersebut akan diklasifikasi [2].

K-Fold Cross Validation dimulai dengan memisah data sebanyak n -fold sesuai keinginan. Data akan dibagi menjadi n bagian dengan jumlah yang sama $a1, a2, a3...an$ untuk dijadikan data uji dan data latih yang akan diproses sesuai dengan banyaknya n . Pada *iterasi* ke- i , $a1$ akan dijadikan data *testing* sedangkan sisanya dijadikan data *training*. Penggunaan jumlah *fold* paling baik untuk menguji validitas, maka dianjurkan untuk menggunakan 10-fold *cross validation* pada model

2.2.5 Particle Swarm Optimization (PSO)

Algoritma *Particle Swarm Optimization* (PSO) merupakan sebuah metode untuk mencari solusi optimal secara *global* berdasarkan simulasi perilaku sosial kawanan burung. Kawanan tersebut terdiri dari satu set partikel, di mana setiap partikel mewakili suatu solusi potensial. Semua partikel ini terbang dalam n -dimensi mencari ruang untuk mendapatkan solusi optimal (terbaik) *Particle Swarm Optimization* (PSO) serupa dengan Algoritma Genetika (GA) yang memanfaatkan fungsi *fitness* untuk mengevaluasi kualitas solusi dari suatu permasalahan. PSO memiliki beberapa kelebihan yaitu sedikit parameter, mudah diterapkan, konvergensi yang cepat, dan sederhana sehingga PSO banyak diterapkan pada optimasi fungsi, optimasi metode konvensional dan pola klasifikasi [16].

Untuk mendapatkan solusi terbaik, setiap partikel bergerak menuju posisi paling baik secara global dan menuju posisi sebelumnya. Contohnya, partikel $ke - i$ adalah $xi = (xi1, xi2, ..., xid)$ didalam ruangan dimensi. Sedangkan posisi paling baik pada partikel $ke - i$ disimpan dan dijadikan $pbest_i = (pbest_{i,1}, pbest_{i,2}, ..., pbest_{i,d})$, indeks pada partikel paling baik dari semua populasi partikel dijadikan $gbest_d$ sedangkan $vi = (vi,1, vi,2, ..., vi,d)$ adalah kecepatan partikel. Kita bisa menggunakan kecepatan sekarang, jarak $pbest_i$ dan $gbest_d$ untuk menghitung besar kecepatan dan letak partikel seperti persamaan dibawah ini:

$$v_{i,d} = w * v_{i,d} + c1 * R * (pbest_{i,d} - x_{i,d}) + c2 * R * (gbest_d - x_{i,d}) \quad (2.4)$$

$$x_{i,d} = x_{i,d} + v_{i,d} \quad (2.5)$$

Ket:

Vi, d : Iterasi ke- i dari kecepatan partikel ke- i
 w : Bobot inersia
 $c1, c2$: (*learning rate*) konstanta akselerasi
 R : Angka acak (0-1)
 Xi, d : Iterasi ke- i dari partikel ke- i pada posisi saat ini
 $pbest_i$: Partikel ke- i pada posisi terbaik sebelumnya
 $gbest_i$: Partikel terbaik dari semua populasi
 n : Banyaknya partikel dalam populasi
 d : Dimensi

Persamaan (2.4) untuk mengetahui kecepatan yang baru pada setiap partikel mengacu pada (V_i, m) ; $pbest$; $gbest$. Sedangkan Persamaan (2.5) untuk memperbaiki posisi setiap partikel pada ruang solusi [9].

Penggunaan algoritma PSO dijelaskan dalam langkah-langkah berikut ini:

- Mengatur parameter X
- Setiap $x_i \in X$
- Mengevaluasi $fitness$ menggunakan fungsi dari $fitness f$ dan posisi $x_{i(t)}$
- Membandingkan $fitness x_i$ dengan $fitness$ yang paling baik; jika $f x_{i(t)} > f Pbest_i$, Kemudian $Pbest_i = x_{i(t)}$ dan $f Pbest_i = f x_{i(t)}$
- membandingkan $fitness x_i$ dengan $fitness$ global yang terbaik sehingga:
Jika
 $f x_{i(t)} > f gbest$, maka $gbest = x_{i(t)}$ dan $f gbest = f x_{i(t)}$
- menghitung kecepatan pada saat berikutnya x_i menurut
 $v_{i(t)+1} = W.v_{i(t)} + c1\psi1 Pbest_i - x_{i(t)} + c2\psi2 gbest_i - x_{i(t)}$
- Memindahkan x_i ke posisi yang baru dengan menggunakan persamaan
 $x_{i(t)+1} = x_{i(t)} + v_{i(t)+1}$
- Melanjutkan ke langkah berikutnya hingga mencapai kriteria.

2.2.6 Particle Swarm Optimization - K-Nearest Neighbour

PSO-KNN digunakan untuk memperbaiki akurasi K-NN pada hasil prediksi. Pencarian pada PSO akan mendapatkan nilai k pada K-NN yang dikodekan melalui angka biner sejumlah $n1$ bit, sedangkan panjang fitur dikodekan dengan $n2$ bit hingga berjumlah $(n1 + n2)$ bit gen dalam setiap partikel. Menurut teori, nilai k mempunyai nilai 1 hingga banyaknya nilai sampel. Nilai k dijadikan substring $s1$ sedangkan $n1$ bit dan panjang fitur dijadikan substring $s2$ dengan $n2$ bit dimana panjang totalnya akan merepresentasikan partikel pada PSO. Untuk mendapatkan nilai k dapat menggunakan persamaan berikut ini:

$$k = 1 + decimal(s1) \times a - 12n1 - 1 \quad (2.6)$$

Berikut penjelasan prosedur penggunaan PSO-KNN:

- Membangkitkan populasi PSO
Terdapat 3 proses untuk membangkitkan populasi PSO yaitu menginisialisasi kecepatan awal, menginisialisasi posisi awal, serta menginisialisasi $Pbest$ dan $Gbest$. Hal ini dilakukan pada biner sejumlah N string dalam panjang partikel. N merepresentasikan jumlah dari $(n1 + n2)$. Selanjutnya menghitung $fitness$ untuk memperoleh $Gbest$ pada langkah ke- 4.
- Mengupdate Kecepatan PSO
- Mengupdate Posisi PSO
- Menghitung $fitness$ (evaluasi akurasi KNN)
- Menghitung nilai $fitness$ pada setiap individu dalam populasi
- Mengupdate $Pbest$ PSO
- Test Terminasi
- Ulangi langkah ke-2 sampai ke-5 hingga kondisi terpenuhi.

2.2.7 Root Mean Square Error (RMSE)

Untuk mengetahui hasil akurasi prediksi dalam mengevaluasi data sebenarnya digunakan metode salah satunya dengan menghitung kesalahan peramalan yang disebut dengan Root Mean Squared Error (RMSE). Metode ini diketahui mampu dalam mengevaluasi kinerja prediksi. Adapun cara kerjanya yakni dengan mengetahui error yang berada untuk menunjukkan besar hasil antara estimasi dengan data yang akan diprediksi sehingga diketahui perbedaan datanya. RMSE juga mengakarkan nilai dari MSE yang sudah diketahui sebelumnya. Dengan menunjukkan nilai hasil akurasi peramalan data dalam histroy dengan rumus yang terdapat pada nilai kesalahan sehingga RMSE menghasilkan akurasi data [3]. Akurasi data dikatakan bagus jika nilai RMSE nya kecil nilai yang dihasilkan. Rumus RMSE sebagai berikut :

$$RMSE = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n}} \quad (2.7)$$

Dimana ;

y_t = Nilai actual indeks

\hat{y}_t = Nilai prediksi indeks

n = Banyaknya sampel MSR

2.1 Kerangka pemikiran

Adapun kerangka pemikiran pada penelitian ini tersusun sebagai berikut:

a) Permasalahan

- Terjadinya fluktuasi pada harga tembakau
- Penelitian sebelumnya masih menghasilkan nilai RMSE yang tinggi

b) Sumber Data

Data yang digunakan adalah data harga tembakau dari tahun 2008 sampai Agustus 2018

c) Metode/ Pendekatan

Metode yang diusulkan dalam penelitian ini K-NN dengan Optimasi PSO

d) eksperimen/ Pengujian

Pengujian data menggunakan metode yang telah dipilih yakni K-NN dengan optimasi PSO untuk mendapatkan hasil RMSE terkecil.

e) Tujuan

Untuk mengetahui kinerja PSO sebagai optimasi Pada K-NN dalam menghasilkan RMSE

f) Hasil

Hasil dari pengujian ini prediksi harga tembakau pada bulan yang akan datang

3. METODE PENELITIAN

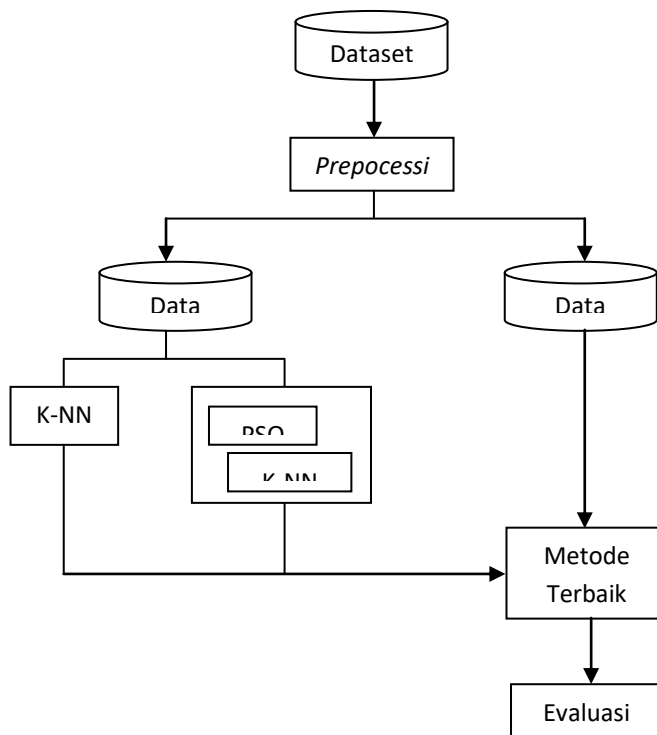
Pada penelitian ini digunakan data sekunder berupa dataset harga bulanan komoditas tembakau mutu I dari tahun 2008 sampai bulan Agustus 2018 yang didapat dari UD Sumber Rejeki yang merupakan ranting perusahaan rokok PT Gudang Garam. Dataset berupa data *univariate* harga tembakau dengan satu variabel yaitu variabel harga sebanyak 128 record seperti contoh pada tabel berikut:

Tabel 3.1 Data harga Tembakau

No	Tahun	Bulan ke	Harga
1	2008	1	26000
....
.....
121	2018	1	50000
122		2	53000
123		3	56000
124		4	50000
125		5	46000
126		6	43000
127		7	40000
128		8	36000

Data ini merupakan data *univariate*, yaitu data yang hanya memiliki 1 atribut (harga tembakau) tanpa ada atribut lain yang saling mempengaruhi.

Dalam penelitian ini untuk memprediksi harga tembakau metode penyelesaiannya menggunakan model K-NN dan PSO sebagai optimasi. Tahap penyelesaian pada penelitian ini ditunjukkan pada gambar berikut:



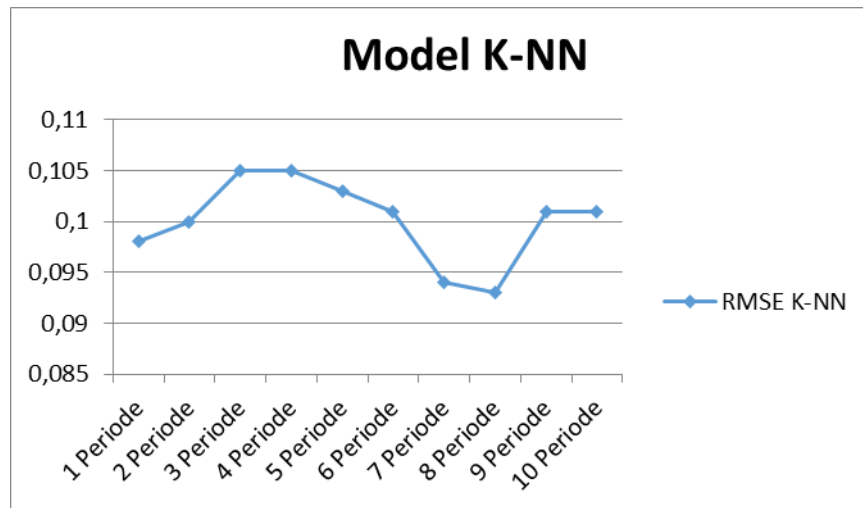
Gambar 3.1 Metode yang diusulkan

4. HASIL DAN PEMBAHASAN

4.1 Hasil Pengujian Metode K-NN

Tabel 4.2 RMSE hasil pengujian K-NN semua periode (1 s/d 10)

Periode (input)	Cross V	Parameter k	RMSE
xi s/d xi-1	10	14	0.098
xi s/d xi-2	10	13	0.100
xi s/d xi-3	10	12	0.105
xi s/d xi-4	10	9	0.105
xi s/d xi-5	10	6	0.103
xi s/d xi-6	10	7	0.101
xi s/d xi-7	10	4	0.094
xi s/d xi-8	10	3	0.093
xi s/d xi-9	10	4	0.101
xi s/d xi-10	10	3	0.101

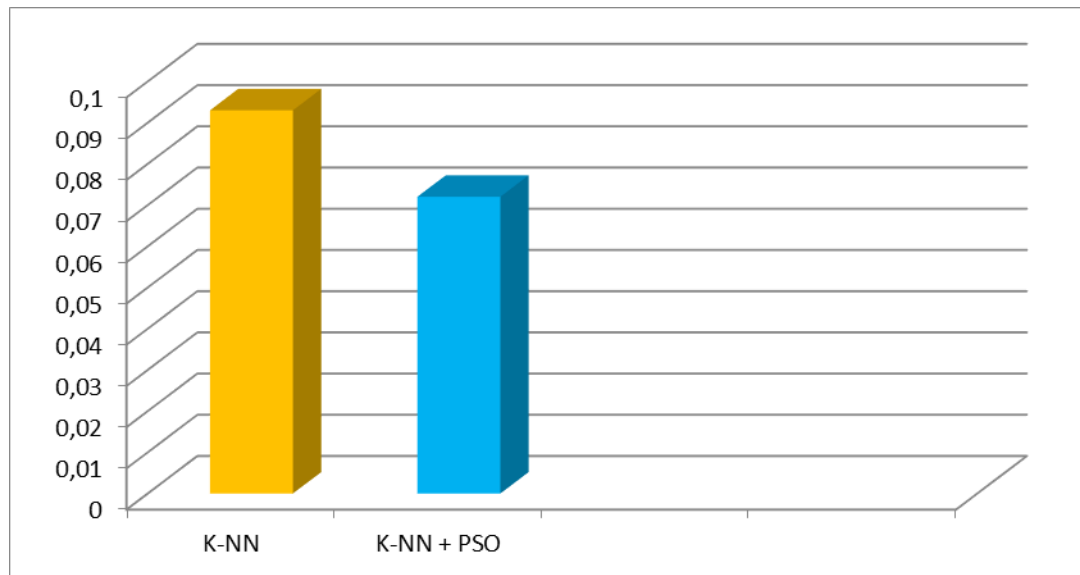


Gambar 4.1 Grafik model K-NN

4.2 Hasil Pengujian Metode K-NN Berbasis PSO

Tabel 4.3 Parameter model terbaik optimasi K-NN dengan PSO

	Parameter		Nilai
K-NN	Variabel input		Periode 8 (x_{i-1} s/d x_{i-8})
	K		3
K-NN dan PSO	Populasi		9
	<i>max.number of generation</i>		90
	<i>Inersia weight</i>		0.9
	<i>local best</i>		0.9
	<i>global best</i>		0.9
	Weight	x_{i-1}	1
		x_{i-2}	0.37279
		x_{i-3}	1
		x_{i-4}	0
		x_{i-5}	0
		x_{i-6}	0.75516
		x_{i-7}	0
		x_{i-8}	1



Gambar 4.2. Perbandingan nilai RMSE K-NN dan K-NN dengan PSO

4.3 Hasil Perbandingan RMSE

Tabel 4.9 Hasil perbandingan RMSE K-NN dan K-NN berbasis PSO

Hasil	K-NN	K-NN berbasis PSO
RMSE	0.093	0.072

5. SARAN DAN KESIMPULAN

Berdasarkan uji coba metode yang dilakukan, dapat diambil kesimpulan bahwa parameter terbaik pada model K-NN dalam penelitian ini yaitu pada periode ke 8 dengan parameter $K = 3$ dengan nilai RMSE 0.093. Sedangkan parameter terbaik dari K-NN berbasis PSO ditunjukkan oleh nilai populasi = 9, maksimal nomor generasi, bobot inersia, bobot lbest, bobot gbest yang masing-masing bernilai 0.9 dengan nilai eror menurun menjadi 0.072. Dengan demikian performa optimasi PSO untuk pembobotan atribut pada K-NN terbukti lebih baik dari pada K-NN tanpa optimasi

Harapan untuk penelitian selanjutnya agar dapat menggunakan dataset dalam jumlah yang lebih besar supaya memperoleh hasil yang lebih baik dengan menggunakan metode optimasi lain seperti GA, C Squere, dll untuk memperoleh hasil yang lebih baik.

DAFTAR PUSTAKA

- [1] K. Kholilurrahman, *Tembakau Madura: Tantangan dan Prospek*, 1st ed. Kencana Jaya Promosindo, 2010.
- [2] M. E. Lasulika, "PREDIKSI HARGA KOMODITI JAGUNG MENGGUNAKAN K-NN DAN PARTICLE SWARM OPTIMAZATION," vol. 9, pp. 233–238, 2017.
- [3] O. J. Anggraeni, "Peramalan Harga dan Permintaan Komoditas Tembakau di Kabupaten Jember," *J. Ilm. Inov.*, vol. 1, p. 13, 2013.
- [4] D. Eliawan and S. Y. J. P. S. Si, "Peramalan Harga dan Produksi Tembakau Indonesia Menggunakan ARIMA Box-Jenkins," 2012.
- [5] D. W. I. M. Sari, "PERAMALAN HARGA DAN PRODUKSI TEMBAKAU," pp. 1–10, 2008.
- [6] D. Sinta, "Metode ensemble k-nearest neighbor untuk prediksi harga beras di indonesia," p. 39, 2015.
- [7] M. K. A. Shatnawi, "Stock Price Prediction Using K -Nearest Neighbor (k NN) Algorithm," vol. 3, no. 3, pp. 32–44, 2013.
- [8] S. B. Imandoust and M. Bolandraftar, "Application of K-Nearest Neighbor (KNN) Approach for

- Predicting Economic Events : Theoretical Background,” vol. 3, no. 5, pp. 605–610, 2013.
- [9] N. Bhatia and C. Author, “Survey of Nearest Neighbor Techniques,” vol. 8, no. 2, pp. 302–305, 2010.
- [10] J. Pseudocode, I. Muzakkir, A. Syukur, and I. N. Dewi, “BACKPROPAGATION DENGAN SELEKSI FITUR PARTICLE SWARM OPTIMIZATION DALAM PREDIKSI PELANGGAN TELEKOMUNIKASI,” vol. 1, pp. 1–10, 2014.
- [11] S. R. Haras, “Metode K-Nearest Neighbor Berbasis Particle Swarm Optimization Untuk Prediksi Harga Komoditi Crude Palm Oil,” p. 51.
- [12] A. Nirfah, “Klasifikasi Resiko Kredit Sepeda Motor Menggunakan Algoritma K-Nearest Neighbor Berbasis Particle Awarm Optimization,” p. 66, 2016.
- [13] A. A. Yuita Arum Sari, “Optimasi K-Nearest Neighbour Menggunakan Particle Swarm Optimization Optimasi K-Nearest Neighbour Menggunakan Particle Swarm Optimization pada Sistem Pakar untuk Monitoring Pengendalian Hama pada Tanaman Jeruk,” *J. Teknol.*, vol. 2, no. July, p. 13, 2018.
- [14] K. Dengan, A. K. Neighbour, and E. S. Y. Pandie, “PROGRAM PASCASRAJANA UNIVERSITAS DIPONEGORO,” 2012.
- [15] O. Kramer, “Unsupervised K-Nearest Neighbor Regression,” pp. 8–11.
- [16] T. Park, “Optimization for Artificial Neural Network with Adaptive Inertial Weight of Particle Swarm Optimization,” pp. 481–485.
- [17] N. A. Gugun El Guyanie, Hifdzil Alim, Badruddin, Ibi Syatibi, *Ironi Cukai Tembakau: Carut-marut Hukum & Pelaksanaan Dana Bagi Hasil Cukai Hasil Tembakau di Indonesia*, 1st ed. Jakarta: Indonesia Berdikari, 2013.
- [21] S. Jain, S. Shukla, and R. Wadhvani, “Dynamic selection of normalization techniques using data complexity measures,” *Expert Syst. Appl.*, vol. 106, pp. 252–262, 2018.
- [22] S. C. Nayak, “Index Prediction with Neuro-Genetic Hybrid Network : A Comparative Analysis of Performance.”
- [23] M. Nanja and purwanto, “Metode K-Nearest Neighbor Berbasis Forward Selection Untuk Prediksi Harga Komoditi Lada,” Pseudocode, vol. 2, no. 1-ISSN 2355 – 5920, pp. 53–64, 2015.