



Perbandingan Metode *Forecasting* K-NN, NN dan SVM Untuk Peramalan Jumlah Produksi *Coconut Oil*

Ivo Colanus Rally Drajana¹, Marniyati H. Botutihe²

Teknik Informasi¹, Sistem Informasi², Universitas Pohuwato
ivocolanusrally@email.com¹, marniyatihusainbotutihe@gmail.com²

Abstrak

Tanaman pohon kelapa memiliki banyak bagian yang dimanfaatkan, sehingga tumbuhan ini dianggap tumbuhan serbaguna. Minyak kelapa (*coconut oil*) dihasilkan oleh buah pohon kelapa salah satunya adalah buah kelapa yang diolah menjadi minyak kelapa (*coconut oil*). Peramalan sangat diperlukan untuk meramalkan jumlah produksi minyak kelapa (*coconut oil*) pada sebuah perusahaan untuk mencapai target produksi. Penelitian ini memiliki tujuan untuk membandingkan metode *forecasting* untuk mendapatkan model terbaik. Dari hasil eksperimen menggunakan data *sales order* (SO) di peroleh model terbaik untuk peramalan menunjukkan bahwa model yang terbaik dihasilkan oleh algoritma *Support Vector Machine* (SVM) dilihat dari hasil RMSE terkecil yaitu 0,172 jika di bandingkan dengan model *K-Nearest Neighbor* (K-NN) dan model *Neural Network* (NN).

Kata Kunci: *K-Nearest Neighbor, Neural Network, Support Vector Machine, Produksi Coconut Oil*

ABSTRACT

Tree plants coconut have many parts that are used, so this plant is considered a versatile plant. Coconut oil (coconut oil) produced by the fruit of the coconut tree, one of which is coconut fruit which is processed into coconut oil (coconut oil). Forecasting is very necessary to predict the amount of coconut oil production (coconut oil) in a company to achieve production targets. This study aims to compare forecasts to get the best model. From the experimental results using sales order (SO) data, the best model for forecasting shows that the best model produced by the Support Vector Machine (SVM) algorithm is

seen from the smallest RMSE result, which is 0.172 when compared to the K-Nearest Neighbor (K-NN) model and Neural Network (NN) models..

Keywords: *K-Nearest Neighbor Neural Network Support Vector Machine Coconut Oil Production*

A. PENDAHULUAN

Pohon kelapa memiliki ciri tersendiri yang dapat dikenali dari bentuk fisiknya. Kelapa termasuk dalam genus *Cocos* dan dapat tumbuh di daerah tropis tanpa masalah. Tumbuhan Pohon kelapa banyak ditemukan di daerah pantai karena tumbuhan ini membutuhkan kelembapan yang tinggi. Tanaman kelapa memiliki banyak manfaat, hampir seluruh bagian pohonnya dimanfaatkan, sehingga tanaman ini dianggap sebagai tanaman serbaguna bagi masyarakat. *Coconut oil* terbuat dari buah kelapa. Pengolahan hasil kelapa terutama dilakukan oleh masyarakat yang berprofesi sebagai petani berupa produk kopra. Produk kopra ditanam langsung oleh petani di sekitar perkebunan kelapa. Olahan kelapa adalah kopra, bahan pengolahan untuk perusahaan minyak kelapa berupa minyak kelapa mentah yang kemudian diolah menjadi produk yang lebih berkualitas seperti minyak goreng yang telah mengalami proses pemurnian., *Bleaching*, dan *Deodorize* (RBD). Produk olahan minyak goreng inipun masih dapat di proses menjadi produk yang bernilai tambah. Produksi *coconut oil* yang terjadi mengalami perubahan atau fluktuatif sehingga diperlukan prediksi. Produksi minyak kelapa yang terjadi mengalami perubahan atau fluktuatif sehingga diperlukan peramalan sehingga jumlah produksi mampu mengimbangi jumlah permintaan dan untuk mendukung kegiatan kinerja bagian Produksi dalam meramalkan

Jumlah Produksi Minyak Kelapa dengan memperoleh model terbaik dengan nilai error terkecil. Penelitian ini bertujuan untuk membandingkan metode forecasting untuk mendapatkan model terbaik, maka metode dalam penelitian ini *K-Nearest Neighbor*, *Neural Network* dan *Support Vector Machine* pada data *sales order* (SO) untuk prediksi jumlah produksi *coconut oil*. Penelitian yang dilakukan oleh Marniyati H. Botutihe, 2017 Berdasarkan pengujian yang telah dilakukan untuk prediksi jumlah produksi minyak kelapa, Dengan hasil kesimpulan bahwa untuk prediksi jumlah produksi minyak kelapa menggunakan algoritma *Neural Network* (NN) dan *forward selection* sukses dilakukan. Dari persentase perbandingan hasil prediksi dan nilai aktual mencapai 91,01%. Hasil dari penelitian ini dapat dijadikan sebagai bahan acuan bagi pihak perusahaan untuk pengambilan keputusan. Ivo Colanus Rally Drajana, 2018 melakukan penelitian dengan Algoritma *k-Nearest Neighbor* adalah metode penerapan algoritma *supervised* di mana algoritma ini terbagi menjadi dua bagian yaitu *supervised learning* dan *unsupervised learning*. *Feature selection* yakni *Backward Elimination* meningkatkan performa yang lebih baik. *Backward Elimination* diterapkan untuk mengeliminasi atribut atau variabel yang dianggap tidak relevan, variabel yang tidak berpengaruh serta tidak signifikan dalam model dihapuskan dari dalam model. Algoritma *k-Nearest Neighbor* dihasilkan model terbaik yang dilihat berdasarkan nilai error terkecil yaitu 0.111 . Andi Bode, 2019 melakukan penelitian dengan hasil penelitian pada dataset *sales order* dengan menggunakan metode *Support Vector Machine* didapatkan RMSE 0,127, dengan menggunakan metode *Support Vector Machine* dan *Backward Elimination* didapatkan RMSE 0,115, dengan metode *Linear Regression* didapatkan RMSE 0,118 dan dengan menggunakan metode

Linear Regression dan *Backward Elimination* didapatkan RMSE 0,118. Hijratul Aini, dkk, 2019 Berdasarkan hasil percobaan, metode BPNN dengan parameter arsitektur 5-10-11-12-13-1; fungsi pembelajaran adalah *trainlm*; fungsi aktivasi adalah *logsig* dan *purelin*; laju pembelajaran adalah 0.7 mampu menghasilkan tingkat kesalahan prediksi yang baik dengan nilai MSE sebesar 0.0069. Hal ini menunjukkan bahwa metode BPNN dapat menjadi alternative metode dalam memprediksi produksi CPO dengan data yang berjenis *time series*. Prediksi atau *forecasting* adalah aktivitas memprediksi atau memperkirakan yang terjadi di masa mendatang dengan waktu yang relatif lama. Ada beberapa metode *forecasting* di antaranya yaitu *Algoritma k-Nearest Neighbor* adalah metode penerapan algoritma terawasi, yang mana algoritma ini dibagi menjadi dua bagian, yaitu pembelajaran terawasi dan pembelajaran tidak terawasi. Algoritma ini memiliki tujuan *supervised learning* dalam menciptakan pola baru, sedangkan tujuan *unsupervised learning* adalah membuat pola pada data. Kelebihan dari *algoritma k-Nearest Neighbor* adalah *over training* dataset dengan banyak noise dan sangat efektif dengan jumlah data training yang banyak, namun *algoritma k-Nearest Neighbor* memiliki kekurangan dalam menentukan nilai k dan memilih atribut yang terbaik. Algoritma *neural network* merupakan model yang didasarkan pada tiga lapisan, seperti lapisan input yang terhubung ke lapisan tersembunyi dari lapisan tersembunyi dan lapisan tersembunyi dari lapisan tersembunyi yang juga terhubung ke lapisan output. Masing-masing lapisan ini dihubungkan oleh neuron atau node. Jaringan saraf dilatih untuk mengenali hubungan antara lapisan input dan output. Algoritma *Support Vector Machine* adalah metode yang didasarkan pada teori belajar statistik dan memberikan hasil terbaik daripada metode lainnya. *Support Vector Machine* sangat baik

bekerja dengan data berdimensi tinggi memakai teknik kernel. Penelitian ini bertujuan untuk membandingkan model dari metode prediksi agar mendapatkan model yang lebih baik dalam memprediksi atau memprediksi deret waktu, maka metode dalam penelitian ini *K-Nearest Neighbor*, *Neural Network* dan *Support Vector Machine* pada data *sales order (SO)* untuk prediksi jumlah produksi *coconut oil*.

B. METODE

Jenis penelitian ini menggunakan metode deskriptif, subjek pada penelitian ini yaitu pada PT. Multi Nabati Sulawesi Unit Maleo. Teknik analisis data yang pertama data set dilakukan *preprocessing* menggunakan Ms. Excel, dilakukan pengolahan data menggunakan *tools rapid miner*.

Metode yang digunakan pada penelitian ini yaitu metode *forecasting* k-nn, nn dan svm untuk peramalan jumlah produksi *coconut oil*, dari ketiga metode ini nantinya dibandingkan metode mana yang terbaik berdasarkan hasil eksperimen.

1. K-Nearest Neighbor

Algoritma k-Nearest Neighbor adalah metode klasifikasi dan regresi non-parametrik. *K-Nearest Neighbor* menghitung jarak datanya untuk semua sampel. Jarak Euclidean lebih sering digunakan. *Algoritma k-Nearest Neighbor* merupakan salah satu metode penerapan *algoritma supervised*. Algoritma ini terbagi menjadi dua jenis yaitu *supervised learning* dan *unsupervised learning*. Tujuan dari algoritma ini adalah *supervised learning* untuk mendapatkan pola baru, sedangkan tujuan *unsupervised learning* adalah untuk mendapatkan pola pada data. *Algoritma k-Nearest Neighbor* memiliki kelebihan dibandingkan data latih,

yaitu banyak *noise* dan sangat efektif dengan data latih yang banyak atau banyak.

Persamaan dapat dilihat dibawah ini:

$$D = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad \text{Pers. 1}$$

Keterangan:

x= sampel data

y= data uji

D= jarak

$$(\hat{x}') = 1K \sum(x') \quad \text{Pers. 2}$$

Keterangan:

x'= Perkiraan atau estimasi

K= Jumlah tetangga terdekat

(x^)= Tetangga terdekat

2. *Neural Network*

Neural network sering disebut sebagai jaringan saraf tiruan, yang merupakan metode komputasi yang meniru sistem jaringan saraf biologis. Metode ini menggunakan elemen komputasi nonlinier dasar, juga disebut neuron, diorganisasikan sebagai jaringan yang saling berhubungan mirip dengan jaringan saraf manusia. Model *neural network* memiliki tiga bagian, yaitu bagian input yang terhubung dengan bagian tersembunyi, bagian tersembunyi yang juga terhubung dengan bagian *output*. Kemudian setiap bagian dihubungkan melalui sekitar *neuron* atau *node*. Jaringan syaraf tiruan dilatih untuk menentukan hubungan antara lapisan *input* dan lapisan *output*. Jaringan syaraf tiruan ini dibentuk untuk memecahkan masalah tertentu seperti pengenalan pola atau klasifikasi berdasarkan proses pembelajaran

3. *Support Vector Machine*

Algoritma Support Vector Machine (SVM) Ini pertama kali diperkenalkan oleh Vapnik pada tahun 1992 sebagai rangkaian konsep kunci yang harmonis di bidang pengenalan pola. Konsep dasar SVM sebenarnya merupakan kombinasi harmonis dari beberapa dekade teori komputer, seperti batas *hyperplane*. Secara sederhana, konsep SVM dapat dijelaskan sebagai upaya untuk menemukan *hyperplane* terbaik yang memiliki fungsi sebagai pemisah antara dua kelas dalam ruang input. Aktivitas menemukan tempat *hyperplane* yaitu inti terpenting dari proses pembelajaran di SVM.

Proses pembelajaran SVM yaitu memutuskan *support vector* dengan pengetahuan fungsi kernel yang dipakai dan tidak harus mengetahui bentuk fungsi nonlinier SVM persamaan.

$$f(x) = w^T \phi(x) + b \quad \text{Pers. 3}$$

Dimana:

- b = Bias
- $x = (x_1, x_2, \dots, x_D)^T$ = Variabel Input
- $w = (w_0, w_1, \dots, w_D)^T$ = Parameter Bobot
- $\phi(x)$ = Fungsi Transformasi fitur

C. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mencari model terbaik untuk memprediksi deret waktu terkait jumlah produksi minyak kelapa. Sumber data untuk penelitian ini saat ini adalah PT. Diambil dari Unit Maleo Multi-Nabati Sulawesi. Dataset tersebut berupa pendataan SO harian (pesanan pelanggan), yang kemudian diubah menjadi data bulanan sehingga diperoleh 208 dataset. Jika ada transaksi pada hari Sabtu dan Minggu dan

kalender merah dalam catatan. Data record terdiri dari variabel tanggal dan tahun.

Sumber pengumpulan data dalam penelitian ini adalah dari PT. Multi Nabati Sulawesi Maleo Unit, Kabupaten Pohuwato pada tanggal 17 Mei 2018. Data yang dikumpulkan adalah data time series kuantitatif univariat harian. Data ini merupakan data SO (*Sales Order*) dari tahun 2014 – 2017. Pendataan pada hari libur dan hari minggu tidak dilakukan setiap tahun sesuai dengan peraturan waktu kerja perusahaan. Data record yang diambil dari variabel order berisi 208 record data dari setiap order customer.

Proses awal pada tahap ini adalah proses pengumpulan data, data yang diolah berupa data *time series univariat* berupa data numerik harian. Data yang diolah adalah data *univariat* hingga *multivariat*, yang terdiri dari 1 variabel periode atau 4 variabel bebas dengan masing-masing 1 variabel terikat. Langkah proses ini dilakukan untuk menentukan model terbaik menggunakan *Algoritma K-Nearest Neighbor*, *Neural Network* dan *Support Vector Machine*. Sebelum mengonversi data ke beberapa variabel, urutan data diubah dari menaik ke menurun.

Ketika urutan data diubah dari *ascending* ke *descending*, data dinormalisasi, fungsi normalisasi dalam *database science* adalah untuk menghindari terjadinya berbagai anomali data dan inkonsistensi data. Tujuan dari normalisasi *database* adalah untuk mendapatkan data yang lebih kecil untuk mewakili data asli tanpa kehilangan propertinya. Data yang semula *univariat* dijadikan *multivariat*, sedangkan data yang digunakan sebagai *multivariat* adalah data turunan. Denormalisasi rekaman mengembalikan ukuran rekaman yang dinormalisasi untuk memulihkan

data asli. Denormalisasi diterapkan pada hasil eksperimen uji data berupa prediksi volume produksi *coconut oil*.

Tabel 1. Multivariat data SO

xt-3	xt-2	xt-1	Xt
0.7881	0.5168	0.7175	0.4530
0.6167	0.7881	0.5168	0.7175
0.6417	0.6167	0.7881	0.5168
0.7204	0.6417	0.6167	0.7881
0.8565	0.7204	0.6417	0.6167
0.5725	0.8565	0.7204	0.6417
0.8969	0.5725	0.8565	0.7204
0.4481	0.8969	0.5725	0.8565
0.8530	0.4481	0.8969	0.5725
0.7165	0.8530	0.4481	0.8969

Tabel 1 adalah sampel data multivariat yang terdiri dari empat variabel periode. Disini xt adalah variabel terikat dan xt-1, xt-2 dan xt-3 adalah variabel bebas.

1. *Parameter K-Nearest Neighbor*

Untuk menentukan jarak atau jumlah tetangga terdekat, pada tahap ini data terstandarisasi diubah menjadi bentuk pelatihan, yaitu membagi data latih dan data uji kemudian mengujinya dengan metode K-NN.

Tabel 2. Nilai RMSE SO Coconut Oil K-NN Periode 1

Of Validation	10	10	10	10
K	1	3	5	7
RMSE	0.229	0.194	0.188	0.184

Tabel 2 merupakan tabel hasil percobaan SO (*Sales Order*) minyak kelapa dengan penentuan model menggunakan algoritma *k-Nearest Neighbor* yang terdiri dari variabel 1 periode. Menggunakan *cross-validation* 10 dan k-values 1, 3, 5, 7 berdasarkan *least square mean*

(RMSE). Hasil eksperimen memberikan model terbaik berdasarkan nilai error terkecil 0,148 dengan angka validasi 10 dan nilai k 7.

Tabel 3. Nilai RMSE SO Coconut Oil K-NN Periode 2

Of Validation	10	10	10	10
K	1	3	5	7
RMSE	0.224	0.182	0.178	0.174

Tabel 3 menunjukkan tabel dengan hasil eksperimen SO (*Sales Order*) minyak kelapa dengan menentukan model menggunakan algoritma *k-Nearest Neighbor*, validasi silang 10 dari 2 variabel periode dan nilai k1,3,5 yang digunakan. 7 terlihat terkecil berdasarkan *root mean square error* (RMSE). Hasil eksperimen memberikan model terbaik berdasarkan nilai error terkecil 0,174 dengan jumlah validasi 10 dan k-value 7.

Tabel 4. Nilai RMSE SO Coconut Oil K-NN Periode 3

Of Validation	10	10	10	10
K	1	3	5	7
RMSE	0.228	0.193	0.184	0.179

Tabel 4 merupakan tabel hasil percobaan SO (*Sales Order*) minyak kelapa dengan penentuan model menggunakan *algoritma k-Nearest Neighbor*, yang terdiri dari 3 variabel periode validasi silang 10 dan k-values 1, 3, 5 menggunakan 7 yang terlihat terkecil berdasarkan *root mean square error* (RMSE). Hasil eksperimen memberikan model terbaik berdasarkan nilai error terkecil sebesar 0,179 dengan jumlah validasi sebanyak 10 kali dan nilai k sebesar 7.

Tabel 5. Nilai RMSE SO Coconut Oil K-NN Periode 4

Of Validation	10	10	10	10
K	1	3	5	7
RMSE	0.230	0.193	0.184	0.180

Tabel 5 merupakan tabel hasil percobaan SO (*Sales Order*) minyak kelapa dengan model menggunakan *algoritma k-Nearest Neighbor* yang terdiri dari 4 variabel periode dengan *cross-validation* 10 dan nilai k 1, 3, 5, 7 dianggap terkecil berdasarkan *root mean square error* (RMSE). Hasil eksperimen memberikan model terbaik berdasarkan nilai error terkecil 0,180 dengan validasi 10 dan nilai k 7. Model terbaik untuk data SO minyak kelapa pada variabel periode 2 nilai validasi 10 dengan K 7, dengan root mean square error 0,174.

2. Parameter Neural Network

Level ini menentukan parameter *neural network* dengan variabel periode 1-4, *hidden layer* 8, *learning rate* 0.3, *momentum* 0.2, nilai *training cycle* dengan range 100 - 1000.

Tabel 6. Nilai RMSE SO Coconut Oil NN

Variabel Periode	1	2	3	4
Hidden Layer	8	8	8	8
Learning Rate	0.3	0.3	0.3	0.3
Momentum	0.2	0.2	0.2	0.2
Training Cycle	500	500	500	500
RMSE	0.196	0.181	0.189	0.181

Tabel 6 diatas merupakan tabel ringkasan hasil Percobaan SO Minyak Kelapa (*Sales Order*) dengan menentukan model menggunakan *algoritma neural network*. Parameter terbaik didapatkan dari *neural network* dengan memasuki *layer* 8, *learning rate* 0.3, *impulse* 0.2, *training cycle value* 500. Hasil RMSE terkecil yang diterima *neural network* pada

periode 2 dan 4 adalah 0,181. Model terbaik untuk data SO minyak kelapa yang model terbaiknya berasal dari *neural network* dengan *input layer* 8, dengan *input layer* 8, *learning rate* 0.3, *pulse* 0.2, nilai *training cycle* 500. RMSE terkecil. Hasil diperoleh *neural network* pada periode 2 dan 4 sebesar 0,181.

3. *Parameter Support Vector Machine*

Pada fase ini dilakukan eksperimen untuk mendapatkan model terbaik untuk data dari 1 periode ke 4 variabel dengan menguji 10 angka validasi tipe kernel polinomial.

Tabel 7. Nilai RMSE SO Coconut Oil SVM

Variabel Periode	Number of Validation	Type Kernel	RMSE
1	10	Polynomial	0,172
2	10	polynomial	0,173
3	10	polynomial	0,183
4	10	polynomial	0,173

Sesuai dengan Tabel 7, tabel ringkasan hasil percobaan percobaan pesanan pelanggan (SO) dengan pemilihan model menggunakan *algoritma Support Vector Machine* (SVM), dengan asumsi variabel periode 1-4 untuk data SO memiliki rentang validasi dari . gunakan 10 dan jenis polinomial kernel diambil berdasarkan *least square root mean error* (RMSE). Kemudian mencari model terbaik berdasarkan tingkat kesalahan terkecil sebesar 0,172 dengan jumlah variabel 1 periode. Model terbaik untuk data pesanan pelanggan, model *algoritma Support Vector Machine* (SVM) menggambarkan bahwa model terbaik berada pada variabel periode 1 dengan jumlah validasi 10 dan polinomial tipe kernel. yaitu tingkat RMSE sebesar 0,172.

D. PENUTUP

Simpulan dan Saran

Model yang dipilih berdasarkan nilai *root mean square error* (RMSE) terkecil yang diperoleh. Berdasarkan hasil percobaan yang dilakukan, maka dapat ditentukan model yang akan digunakan untuk melakukan prediksi jumlah produksi *coconut oil* yaitu model dengan tingkat *error* terkecil, dapat dilihat pada tabel berikut:

Tabel 8. Perbandingan K-NN, NN dan SVM

<i>Sales Order</i>		
Algoritma	Variabel Periode	RMSE
K- Nearest Neighbor	2	0.174
Neural Network	2 & 4	0.181
Support Vector Machine	1	0.172

Pada Tabel 8 Yang merupakan tabel hasil perbandingan antara algoritma *K-Nearest Neighbor*, *Neural Network* dan *Support Vector Machine*, menunjukkan bahwa model yang terbaik dihasilkan oleh algoritma *Support Vector Machine* (SVM) dilihat dari hasil RMSE terkecil yaitu 0,172. Model terbaik untuk data *sales order* adalah *model algoritma support vector machine* (SVM) dibandingkan dengan model *k-nearest neighbor* dan *neural network*.

DAFTAR PUSTAKA

Bode, A., 2017. K-nearest neighbor dengan feature selection menggunakan backward elimination untuk prediksi harga komoditi kopi arabika. *ILKOM Jurnal Ilmiah*, 9(2), pp.188-195.

- Botutihe, M.H., 2017. Model Neural Network Berbasis Forward Selection Untuk Prediksi Jumlah Produksi Minyak Kelapa. *ILKOM Jurnal Ilmiah*, 9(3), pp.239-243.
- Drajana, I.C.R., 2017. Metode support vector machine dan forward selection prediksi pembayaran pembelian bahanbaku kopra. *ILKOM Jurnal Ilmiah*, 9(2), pp.116-123.
- Drajana, I.C.R., 2018. Prediksi Jumlah Produksi Coconut Oil Menggunakan k-Nearest Neighbor dan Backward Elimination. *JURNAL TECNOSCIENZA*, 3(1), pp.51-64.
- Jefika, M., Kosasi, H., Prayogi, G. and Dharma, A., 2020. Prediksi Gelombang Corona Dengan Metode Neural Network. *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, 3(2), pp.102-107.
- Rahmawati, N. and Lestari, T.E., 2019. Implementasi Model Fungsi Transfer Dan Neural Network untuk Meramalkan Harga Penutupan Saham (Close Price)(Studi Kasus: PT. Bank Central Asia, Tbk). *Jurnal Matematika Vol*, 9(1), pp.11-25.
- Sukmaya, S.G. and Perwita, A.D., 2019. Daya Saing Komoditas Kelapa Indonesia dan Produk Turunannya. *UNEJ e-Proceeding*.