

Implementasi Algoritme *K-Nearest Neighbor* (KNN) untuk Prediksi Hasil Produksi

Farkhina Dwi Utari
Universitas Buana Perjuangan
Karawang, Indonesia
If15.farkhinautari@mhs.ubpkarawang.ac.id

Amril Mutoi Siregar
Universitas Buana Perjuangan
Karawang, Indonesia
amrilmutoi@ubpkarawang.ac.id

Deden Wahiddin
Universitas Buana Perjuangan
Karawang, Indonesia
deden.wahiddin@ubpkarawang.ac.id

Abstract—

*Penjadwalan sebuah produksi merupakan proses penting dalam produksi. Penjadwalan produksi bertujuan untuk menghasilkan produk yang akan dijual dan memperoleh keuntungan bagi perusahaan. Agar produksi berjalan lancar diperlukan sebuah departemen perencanaan produksi dan kontrol persediaan. Permasalahan yang ada saat ini yaitu metode yang digunakan dalam penjadwalan produksi tidak tentu, karena metode tersebut tidak cocok digunakan, sehingga menyebabkan kendala yang mempengaruhi perencanaan produksi. Salah satu faktor yang mempengaruhi adalah besarnya permintaan pelanggan per harinya (cycle time urgent). Maka solusi untuk mengolah data hasil produksi perusahaan adalah dengan teknik klasifikasi-prediksi menggunakan algoritme *K-Nearest Neighbor* (KNN). Sehingga mampu membantu memberikan prediksi bagi departemen perencanaan dan kontrol persediaan dalam penjadwalan produksi. Berdasarkan hasil perhitungan klasifikasi prediksi data hasil produksi pada PT. SKI dengan algoritme *K-Nearest Neighbor* (KNN) dengan data latih sebanyak 130 dan satu data uji diperoleh nilai akurasi sebanyak 100% dengan menentukan $K=5$.*

Kata kunci — *K-Nearest Neighbor* (KNN), Perencanaan Produksi, Penjadwalan Produksi

I. PENDAHULUAN

Sebuah perusahaan berdiri dengan tujuan menghasilkan barang atau jasa untuk memenuhi kebutuhan pelanggan dan memperoleh keuntungan bagi perusahaan. Perusahaan memiliki proses yang panjang agar menjadi sebuah perusahaan yang berkembang di masa mendatang. Agar perusahaan mampu berkembang di masa mendatang, salah satu yang menentukan yaitu produk yang akan diproduksi dan dijual. Dalam menentukan sebuah keputusan perencanaan serta penjadwalan produksi, diperlukan sebuah departemen perencanaan dalam sebuah perusahaan. Cara untuk menganalisis keputusan dengan tepat, hal yang dilakukan adalah dengan mengatur dan menjalankan perencanaan dan mengawasi produksi. Di samping fungsi perencanaan produksi yang dimiliki, departemen perencanaan juga berperan sebagai kontrol produksi. Perencanaan produksi merupakan hal yang sangat penting dalam menjalankan proses produksi, karena jika perencanaan produksi tidak kuat maka dalam menjalankan produksi tidak teratur. Lemahnya perencanaan bisa menyebabkan kehilangan waktu dalam produksi, pengiriman barang terlambat karena stok di gudang tidak ada dan harus menunggu barang diproduksi selesai. Serta mempengaruhi hasil produksi yang ada. Salah satu solusi untuk menyelesaikan permasalahan yang ada yaitu dengan teknik komputasi [1,2,3,4].

Penelitian yang dilakukan dengan Algoritma *K-Nearest Neighbor* (KNN) telah berhasil untuk memprediksi kelulusan mahasiswa [1]. Penelitian tersebut melakukan prediksi kelulusan tepat waktu mahasiswa program studi ilmu komputer pada salah satu universitas di Indonesia berdasarkan indeks prestasi hingga Semester 4. Dalam penerapan prediksi kelulusan tersebut, bukti bahwa akurasi mencapai 80.00%. Kemudian Penelitian Yustanti [2] tentang topik Prediksi harga jual tanah menggunakan Algoritme *K-Nearest Neighbor* untuk bukti bahwa akurasi mencapai 80.00%. Penelitian Rufaidha, et al [3] tentang topik Penerapan Metode *K-Nearest Neighbor* untuk prediksi penyakit dengan bukti bahwa akurasi diatas 75%. Kemudian penelitian Areif, et al [4] tentang topik Metode *K-Nearest Neighbor* untuk analisis prediksi tingkat pengunduran diri Mahasiswa Metode bukti bahwa akurasi 79%.

Bersumber pada permasalahan yang sudah dijelaskan, bukti bahwa algoritma KNN mampu menyelesaikan beberapa masalah, maka penelitian ini mengimplementasikan algoritma KNN untuk melakukan prediksi pada data hasil produksi PT. SKI sehingga berhasil membantu departemen perencanaan dalam penjadwalan produksi.

II. ALGORITME K-NEAREST NEIGHBOR (KNN)

Algoritme KNN di temukan oleh P. Hart, et al [5] Kelebihan algoritme KNN bersifat nonparametrik yaitu model yang tidak mengkonsumsikan mengenai distribusi instance kedalam dataset. KNN adalah salah satu metode yang menerapkan algoritme *supervised learning* yang bertujuan untuk mendapatkan pola baru [6]. Tujuan dari algoritme ini adalah mengklasifikasi objek baru berdasarkan atribut dan sampel tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan

titik latih. Pengklasifikasian uji, akan ditemukan sejumlah K objek (titik training) yang paling dekat dengan titik uji. Klasifikasi menggunakan voting dari sample uji yang baru. Dekat atau jauhnya tetangga biasanya dihitung berdasarkan jarak *Euclidian*. Menentukan terbanyak di antara klasifikasi dari K objek. Algoritme K-NN menggunakan klasifikasi ketetanggaan sebagai nilai prediksi Parameter K (Jumlah tetangga paling dekat). Langkah - langkah untuk menghitung algoritme KNN yang pertama menghitung kuadrat jarak *Euclid (queri instance)* masing-masing objek terhadap data sampel yang diberikan. Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai nilai jarak *Euclid* terkecil. Selanjutnya mengumpulkan kategori YA (Klasifikasi *Nearest Neighbor*). Yang terakhir dengan menggunakan kategori KNN mayoritas maka dapat diprediksi nilai *queri distance* yang telah dihitung.

Persamaan 1 merupakan persamaan untuk Algoritme KNN.

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2} \tag{1}$$

Di mana :

$d(x_i, x_j)$ = Jarak Eucliden

(x_i) = record ke-i

(x_j) = record ke-j

(a_r) = Data ke-r

i, j = 1,2,3,...n

III. IMPLEMENTASI K-NEAREST NEIGHBOR

A. Data

PT. Showa Katou Indonesia merupakan perusahaan yang menyediakan data untuk penelitian ini. Rincian data yang diperoleh adalah hasil produksi pada bulan Desember 2018 sampai dengan bulan Februari 2019. Data yang digunakan berjumlah 130 data dan sembilan atribut. Kemudian di proses pembersihan data. Dari proses pembersihan didapatkan data sebesar 130 data dan sembilan atribut menjadi empat yaitu nama produk, aktual produksi shift satu, aktual produksi shift dua dan kategori yang akan digunakan dalam proses klasifikasi. Seperti atribut no, tanggal, nama operator, nama proses, shift. Atribut-atribut tersebut tidak digunakan dalam prediksi karena tidak memiliki pengaruh terhadap penentuan kategori hasil produksi. Sehingga nantinya tidak mengganggu pada saat perhitungan klasifikasi. *Shift* (peralihan) yaitu waktu penempatan jam kerja dari jam kerja pada umumnya.

Tabel 1 Data yang akan diproses klasifikasi

No.	Nama Produk	Aktual produksi		Kategori
		Shift 1	Shift 2	
1	Half Elbow (DPM)	3810	2831	TERCAPAI
2	Cap Rear (SDI)	280	0	TIDAK TERCAPAI
3	End Plate (SDI)	0	1160	TERCAPAI
4	Stopper U5 (FTI)	0	612	TIDAK TERCAPAI
5	Half Elbow (DPM)	6218	0	TERCAPAI
6	Cap End (SDI)	2536	0	TERCAPAI
7	Brkt Brake LHD R (FTI)	0	630	TIDAK TERCAPAI
8	Protector (YMI)	2475	0	TERCAPAI
9	Half Elbow (DPM)	4082	5358	TERCAPAI
10	Stay Middle K64F (SDI)	1460	0	TERCAPAI
.				
.				
.				
130	Half Elbow	2617	15689	TERCAPAI

B. Implementasi KNN pada prediksi hasil produksi

Sebelum melakukan proses perhitungan algoritme KNN perlu disiapkan data kelas (label), data atribut, data lama sebagai data latih dan data baru sebagai data uji. Kategori hasil produksi berdasarkan rentang nilai untuk menentukan hasil produksi disajikan pada Tabel 2.

Tabel 2 Kategori hasil produksi

Kategori	Rentang Nilai
Tidak Tercapai	0-999 atau (< 1000)
Tercapai	>= 1000

Hasil perhitungan manual dilakukan menggunakan Microsoft Excel 2013. Pada perhitungan manual proses dilakukan berdasarkan rumus algoritme KNN dengan menambahkan data uji pada data latih. Kemudian menghitung jarak dan memberikan peringkat pada hasil perhitungan jarak, sehingga akan dihasilkan kategori baru dengan menentukan nilai k=5. Contoh perhitungan dengan algoritme KNN yang merujuk pada persamaan 1 adalah:

$$\begin{aligned}
 d1 &= \sqrt{(3810 - 1500)^2 + (2831 - 1000)^2} \\
 &= \sqrt{(2310)^2 + (1831)^2} \\
 &= \sqrt{(5336100) + (335261)} \\
 &= \mathbf{2947.6534735277}
 \end{aligned}$$

$$\begin{aligned}
 d2 &= \sqrt{(280 - 1500)^2 + (0 - 1000)^2} \\
 &= \sqrt{(1220)^2 + (-1000)^2} \\
 &= \sqrt{1488400 + 1000000} \\
 &= \mathbf{1577.4663229369}
 \end{aligned}$$

(dan seterusnya sampe d ke 130).

C. Evaluasi

Evaluasi dilakukan untuk mengetahui seberapa baik penelitian dilakukan. Pada penelitian ini evaluasi dihitung dengan persamaan 2 [7].

$$\text{Akurasi} = \frac{\text{Jumlah pengujian sesuai}}{\text{Jumlah total pengujian}} \times 100\% \tag{2}$$

IV. HASIL DAN PEMBAHASAN

Hasil klasifikasi yang digunakan merupakan data hasil produksi pada PT. Showa Katou Indonesia pada bulan Desember sampai dengan Februari 2019 dengan jumlah data sebanyak 130 dan sembilan atribut. Proses klasifikasi dilakukan dengan menentukan k=5, seperti yang ditunjukkan pada Algoritme 1. Kemudian dihitung jarak dan menentukan *rangking* berdasarkan data uji yang sudah ditentukan. Setelah menghitung jarak dan *rangking* maka berhasil menentukan klasifikasi baru sesuai standart kategori yang sudah ditentukan oleh PT. Showa Katou Indonesia. Hasil dari klasifikasi baru dapat ditentukan dengan kategori mayoritas.

Algoritme 1, Algoritme k-Nearest Neighbor
Input :
 d=(xi,xj) // untuk menghitung jarak
 k=5 // jumlah klasifikasi yang digunakan
Output:
 Prediksi K=5
K-Nearest Neighbor
 Menentukan jarak antara data latih dan data uji
 Menentukan *rangking*
 Menentukan klasifikasi kategori
 Menentukan Y mayoritas
Sampai
 Prediksi terpenuhi

Hasil yang diperoleh pada penelitian ini ditunjukkan pada Tabel 3. Penjelasan dari pengujian pada nomor 1 penelitian ini dilakukan dengan memasukan nama produk Half elbow (DPM) dengan jumlah hasil produksi shift 1 = 3810 dan shift 2 = 2831 dengan kategori “TERCAPAI” dengan distance 2790.

Tabel 3 Hasil perhitungan jarak.

No.	Product Name	Shift 1	Shift 2	Kategori	Distance
1	Half Elbow (DPM)	3810	2831	TERCAPAI	2790,979
2	Cap Rear (SDI)	280	0	TIDAK TERCAPAI	1753,682
3	End Plate (SDI)	0	1160	TERCAPAI	90
4	Stopper U5 (FTI)	0	612	TIDAK TERCAPAI	638
5	Half Elbow (DPM)	6218	0	TERCAPAI	6342,399

No.	Product Name	Shift 1	Shift 2	Kategori	Distance
6	Cap End (SDI)	2536	0	TERCAPAI	2827,33
7	Brkt Brake LHD R (FTI)	0	630	TIDAK TERCAPAI	620
8	Protector (YMI)	2475	0	TERCAPAI	2772,747
9	Half Elbow (DPM)	4082	5358	TERCAPAI	5791,234
10	Stay Middle K64F (SDI)	1460	0	TERCAPAI	1922,004
	-				
	-				
	-				
130	Half Elbow	2617	15689	TERCAPAI	14920,3

Hasil perhitungan jarak pada Tabel 3 dibagi menjadi dua data, data uji dan data latih. Hasil perhitungan jarak diurutkan berdasarkan peringkat nilai terkecil hingga terbesar. Langkah selanjutnya berdasarkan susunan nilai Tabel 4.

Tabel 4 Hasil perhitungan *rangking*

No.	Product Name	Shift 1	Shift 2	Kategori	Distance	Rangking
1	Half Elbow (DPM)	3810	2831	TERCAPAI	2790,979	93
2	Cap Rear (SDI)	280	0	TIDAK TERCAPAI	1753,682	66
3	End Plate (SDI)	0	1160	TERCAPAI	90	2
4	Stopper U5 (FTI)	0	612	TIDAK TERCAPAI	638	19
5	Half Elbow (DPM)	6218	0	TERCAPAI	6342,399	116
6	Cap End (SDI)	2536	0	TERCAPAI	2827,33	95
7	Brkt Brake LHD R (FTI)	0	630	TIDAK TERCAPAI	620	17
8	Protector (YMI)	2475	0	TERCAPAI	2772,747	92
9	Half Elbow (DPM)	4082	5358	TERCAPAI	5791,234	115
10	Stay Middle K64F (SDI)	1460	0	TERCAPAI	1922,004	74
	-					
	-					
	-					
130	Half Elbow	2617	15689	TERCAPAI	14920,3	126

Dari hasil perhitungan klasifikasi Tabel 5 berdasarkan data uji yang diberikan, maka hasil kategori yang dihasilkan adalah “TERCAPAI” karena jumlah mayoritas TERCAPAI = 3. Hasil dari klasifikasi baru dapat ditentukan dengan kategori mayoritas, sehingga dengan menggunakan persamaan (2) nilai akurasi adalah 100 %. Pada perhitungan prediksi data hasil produksi menggunakan Algoritme KNN terdapat perbedaan hasil yaitu pada saat menentukan prediksi dengan nilai K yang berbeda. Prediksi yang dibuat untuk menentukan nilai K sebaiknya menggunakan jumlah nilai K yang ganjil agar tidak terjadi kesamaan hasil kategori.

Tabel 5 Hasil Klasifikasi manual dengan nilai K=5

Distance	Rangking	Kategori Setelah Klasifikasi	Klasifikasi
111	1	TERCAPAI	YA
152	2	TERCAPAI	YA
153	3	TIDAK TERCAPAI	YA
160	4	TERCAPAI	YA
163	5	TIDAK TERCAPAI	YA

V. KESIMPULAN

Dari hasil perhitungan yang telah dilakukan dengan teknik klasifikasi prediksi menggunakan Algoritme KNN. Maka berhasil ditarik kesimpulan bahwa data latih sebanyak 130 data dan satu data uji dapat menghasilkan kategori baru untuk memprediksi hasil produksi pada bulan berikutnya. Selanjutnya hasil kategori k=5 yang ada menunjukkan bahwa hasil produksi pada PT. Showa Katou Indonesia di bulan berikutnya adalah “TERCAPAI” dan hasil penelitian ini menghasilkan tingkat akurasi 100% . Maka, KNN mampu melakukan klasifikasi untuk memprediksi hasil produksi barang.

Penelitian yang akan datang disarankan menggunakan data dari beberapa perusahaan dengan produk yang sama, maka penelitian tersebut dapat memberikan solusi secara global. Sehingga pemanfaatan teknik komputasi dapat dilakukan secara optimal.

VI. PENGAKUAN

Naskah ilmiah ini adalah sebagian penelitian Tugas Akhir milik Farkhina Dwi Utari dengan judul Penerapan Algoritme K-Nearest Neighbor (KNN) untuk prediksi hasil produksi PT. Showa Katou Indonesia yang dibimbing oleh Amril Mutoi Siregar dan Deden Wahiddin.

DAFTAR PUSTAKA

- [1] M. A. Banjarsari, H. I. Budiman, and A. Farmadi, "Penerapan K-Optimal Pada Algoritma Knn untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan IP Sampai Dengan Semester 4," vol. 02, no. 02, pp. 50–64, 2015.
- [2] W. Yustanti, "Algoritma K-Nearest Neighbour untuk Memprediksi Harga Jual Tanah," vol. 9, no. 1, pp. 57–68, 2012.
- [3] H. Rufaidha, I. Komputasi, F. Informatika, and U. Telkom, "PREDIKSI PENYAKIT MENGGUNAKAN ALGORITMA K-NEAREST TINGGI DISEASE PREDICTION USING K-NEAREST NEIGHBOUR AND GENETIC ALGORITHM FOR HIGH DIMENSIONAL DATA," vol. 3, no. 2, pp. 3771–3777, 2016.
- [4] F. I. Komputer and U. D. Nuswantoro, "PERAMALAN PENJUALAN MOBIL PADA PT BENGAWAN."
- [5] P. Hart and T. Cover, "Nearest neighbor pattern classification," *IEEE*, vol. 13, no. 1, pp. 21–27.
- [6] A. Bode, "K-NEAREST NEIGHBOR DENGAN FEATURE SELECTION MENGGUNAKAN BACKWARD ELIMINATION UNTUK PREDIKSI HARGA KOMODITI KOPI ARABIKA," vol. 9, pp. 188–195, 2017.
- [7] T. Djatna, M. Kusuma, D. Hardhienata, A. Fitri, and N. Masruriyah, "An intuitionistic fuzzy diagnosis analytics for stroke disease," *J. Big Data*, 2018.