



KLASIFIKASI HASIL PRODUKSI DAUN TEH BASAH (DTB) MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR (KNN)

Siti Andra Azahra Purba¹, Sriani²

^{1,2}Program Studi Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara

Jl. Lapangan Golf, No 120, Kp. Tengah, Kecamatan Pancurbatu, Kabupaten Deli Serdang, Sumatera Utara, Indonesia, 20353

e-mail : sitiandraazahrapurba@gmail.com

ABSTRAK

Daun teh basah (DTB) merupakan bahan baku utama dalam produksi teh yang menjadi salah satu komoditas unggulan di PTPN IV. Salah satu masalah utama adalah ketidaksesuaian antara target produksi yang ditetapkan dalam RKAP (Rencana Kerja dan Anggaran Perusahaan) dengan hasil yang direalisasikan. Ketidaksesuaian ini disebabkan oleh berbagai faktor, seperti kondisi cuaca yang tidak mendukung, perawatan tanaman yang kurang optimal, dan keterbatasan sumber daya. Untuk mengklasifikasikan produksi agar lebih efisien dengan menggunakan Algoritma K-Nearest Neighbor (KNN). Tujuan penelitian ini adalah mengklasifikasikan hasil produksi DTB. Dataset DTB yang diperoleh sebanyak 116 data, data testing sebanyak 96 data dan data training sebanyak 20 data, yang terdiri dari 4 atribut. Hasil proses pengujian menunjukkan akurasi k-1 sebesar 95.00%, akurasi k-2 sebesar 90.00%, akurasi k-3 sebesar 95.00%, perbandingan k = 1 dan k = 3 adalah hasil precision dan recall. class precision k = 1 sebesar 94.44% dan class recall sebesar 100% sedangkan di k = 3 class precision sebesar 100% dan class recall sebesar 94.12%. Maka penelitian ini dinilai berhasil menerapkan metode KNN untuk melakukan klasifikasi terhadap hasil produksi.

Kata kunci : Daun Teh Basah (DTB), Data Mining, Klasifikasi, K-Nearest Neighbor.

ABSTRACT

Wet tea leaves (DTB) are the main raw material in tea production which is one of the leading commodities at PTPN IV. One of the main problems is the mismatch between the production targets set in the RKAP (Company Work Plan and Budget) and the results realized. This discrepancy is caused by various factors, such as unfavorable weather conditions, less than optimal plant care, and limited resources. To classify production more efficiently using the K-Nearest Neighbor (KNN) algorithm. The aim of this research is to classify DTB production results. The DTB dataset obtained was 116 data, 96 testing data and 20 training data, consisting of 4 attributes. The results of the testing process show that k-1 accuracy is 95.00%, k-2 accuracy is 90.00%, k-3 accuracy is 95.00%, the comparison of k = 1 and k = 3 is the result of precision and recall. class precision k = 1 is 94.44% and class recall is 100%, while at k = 3 class precision is 100% and class recall is 94.12%. So this research is considered successful in applying the KNN method to classify production results.

Keywords: Wet Tea Leaves (DTB), Data Mining, Classification, K-Nearest Neighbor.

1. PENDAHULUAN

Seiring dengan perkembangan jaman, industri teh milik swasta maupun perseorangan terus berkembang. Tingginya minat masyarakat terhadap teh karena kualitasnya yang baik dengan bahan baku dan proses produksinya dilaksanakan dengan baik pula (Putra, 2019). PT Perkebunan Nusantara IV (PTPN IV) merupakan perusahaan yang bergerak dibidang usaha agroindustry merupakan

sebuah perusahaan perkebunan yang terletak di daerah sumatra utara dengan bidang pengolahan Daun Teh Basah (DTB) dan Daun Teh Kering (DTK). Qulsum mengatakan Kualitas bahan baku dan proses produksi berpengaruh signifikan terhadap kualitas produk. Dalam kegiatan Proses produksi, pabrik mengelola pucuk teh yang dipetik dari perkebunan sendiri maupun perkebunan lain (Putra, 2019).

Daun teh basah (DTB) merupakan bahan baku



utama dalam produksi teh yang menjadi salah satu komoditas unggulan di Indonesia. Produksi daun teh basah memegang peran penting dalam memenuhi kebutuhan industri teh, baik di pasar domestik maupun internasional. Namun, proses produksi daun teh basah sering kali menghadapi berbagai tantangan, seperti ketidaksesuaian antara target yang direncanakan dan hasil yang dicapai, serta kendala dalam memastikan efisiensi operasional. Salah satu masalah utama adalah ketidaksesuaian antara target produksi yang ditetapkan dalam RKAP (Rencana Kerja dan Anggaran Perusahaan) dengan hasil yang direalisasikan. Ketidaksesuaian ini disebabkan oleh berbagai faktor, seperti kondisi cuaca yang tidak mendukung, perawatan tanaman yang kurang optimal, dan keterbatasan sumber daya. Di sisi lain, data produksi yang dihasilkan dari proses ini sebenarnya memiliki potensi besar untuk dianalisis. Dengan cara ini, perusahaan dapat lebih mudah memonitor kinerja produksi, mengidentifikasi faktor penyebab ketidaksesuaian target, serta merancang strategi yang lebih efektif untuk mencapai hasil yang diinginkan. Algoritma KNN dapat mengelompokkan data produksi berdasarkan target, efisiensi, atau hasil aktual. Dengan menggunakan data historis, algoritma ini mampu memberikan prediksi atau klasifikasi yang membantu perusahaan memahami apakah produksi tergolong "Mencapai target" "Tidak Mencapai Target" Salah satu metode yang digunakan untuk mengklasifikasi produksi daun teh basah adalah penerapan data mining dengan menggunakan algoritma k-Nearest Neighbor (Zai, 2022).

Dalam proses pengolahan, PTPN IV memiliki 2 Unit Pabrik Pengolahan Teh yaitu : Pabrik Bah Butong dengan kapasitas 100 ton perhari, Tobasari 55 ton perhari. Produksi daun teh basah tentunya tidak selalu mengalami peningkatan, Beberapa faktor yang mempengaruhi produktivitas daun teh basah mulai dari iklim, pekerja serta pengelolaan daun teh basah.

Komoditas teh adalah salah satu unit usaha yang membuat PTPN IV memiliki keunikan dibanding perkebunan lain di Sumatera Utara. Walaupun hanya memiliki 3 unit kebun teh, namun PTPN IV tetap diperhitungkan sebagai salah satu produsen teh hitam yang kompetitif di Indonesia. Hal ini menjadi sebuah kebanggaan sekaligus tantangan.

Produksi daun teh basah dan daun teh kering dari tahun 2012 sampai 2016 terus mengalami peningkatan dengan semakin luas nya lahan pertahun.

REALISASI PRODUKSI DTB & DTK KEBUN TEH PTPN IV					
Tahun	Luas (Ha)	Produksi DTB		Produksi DTK	
		Jumlah	Per Ha	Jumlah	Per ha
2012	1,691.38	18,450.41	10.91	4,064.10	2.40
2013	1,883.28	21,736.67	11.54	4,498.84	2.39
2014	2,534.24	27,022.74	10.66	5,698.98	2.25
2015	2,813.14	33,394.50	11.87	6,947.41	2.47
2016	3,483.54	36,571.88	10.50	7,741.32	2.22

Gambar 1. Produksi DTB dan DTK (Ptpn4.co.id)

Untuk melakukan klasifikasi produksi daun teh basah ini dengan menggunakan data mining. Dalam konteks ini, metode data mining seperti algoritma K-Nearest Neighbors (KNN) dapat digunakan untuk membantu mengolah data produksi daun teh basah secara lebih efektif. Elmayanti dalam (Fahreza, 2022) mengatakan Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual . Data mining adalah suatu teknik menggali informasi berharga yang terpendam atau tersembunyi pada suatu koleksi data (database) yang sangat besar sehingga ditemukan suatu pola yang menarik yang sebelumnya tidak diketahui. Teknologi Data Mining dapat digunakan untuk menghasilkan suatu prediksi dengan menggunakan data historis untuk menyimpulkan suatu kejadian di masa depan (Nozomi, 2023). Ada Beberapa algoritma yang digunakan, salah satunya KNN. Abode mengatakan konsep mendasar dari algoritma KNN yaitu mencari jarak paling terdekat diantara data yang terevaluasi dengan sejumlah K tetangga (neighbor) paling dekat dengan data uji. KNN bekerja dengan cara membandingkan data uji dan data training. KNN mencari jarak data training yang paling mendekati dengan data uji. Perhitungan jarak menggunakan rumus Euclidean Distance . Bhatia dalam (Basriati, M.Sc & Safitri, M.Mat, 2021) mendefenisikan K-Nearest Neighbor merupakan teknik klasifikasi yang sederhana, efektif pada data latih dengan jumlah yang besar dan menghasilkan perfoma yang cukup baik .

Data Mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu database dengan melakukan penggalian pola-pola dari data dengan tujuan untuk memanipulasi data menjadi informasi yang lebih berharga yang diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam database (Shantika Hutapea & Fauzi, 2021).

Teknologi Data Mining dapat digunakan untuk menghasilkan suatu prediksi dengan menggunakan data historis untuk menyimpulkan suatu kejadian di masa depan (Nozomi, 2023). Knowledge Discovery in Databases (KDD) serta Data Mining sering



dipergunakan bergantian untuk menerangkan proses penggalian informasi-informasi yang tersembunyi dalam database dengan kapasitas besar. Meskipun sesungguhnya sebutan tersebut mempunyai konsep yang berbeda, namun saling berhubungan satu sama lain. Salah satu tahapan dalam keseluruhan proses pada Knowledge Discovery in Databases adalah Data Mining (Ginatra, et al., 2021).

Fayyad dalam (Sari Pakpahan, 2021) mengatakan dalam Knowledge Discovery in Databases atau sering disingkat dengan istilah KDD sendiri dapat diartikan sebagai rangkaian proses non-trivial yang berguna untuk menemukan serta mengidentifikasi pola (pattern) di dalam data, di mana pola yang ditemukan bersifat baru, valid, berguna dan mudah dimengerti. Proses KDD berkaitan dengan teknik integrasi serta penemuan ilmiah, interpretasi serta visualisasi pola dari himpunan data.

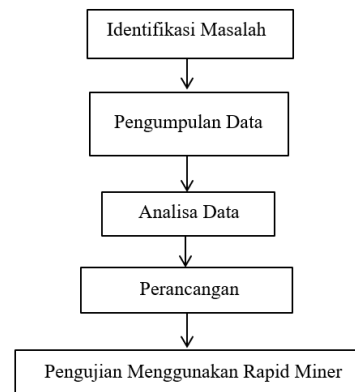
Metode K-Nearest Neighbor (KNN) digunakan untuk melakukan klasifikasi objek berdasarkan data pembelajaran yang memiliki jarak paling dekat dengan objek tersebut. Tujuan dari metode ini adalah untuk mengklasifikasikan objek baru berdasarkan atribut dan sampel pembelajaran. Ketika diberikan sebuah titik query, metode KNN akan mencari sejumlah K objek atau titik pembelajaran yang memiliki jarak paling dekat dengan titik query tersebut. Kemudian, nilai prediksi dari query akan ditentukan berdasarkan klasifikasi tetangga (Muhadi & Octaviano, 2023).

Kursini dalam (Syukri Mustafa & Wayan Simpen, 2019) mengatakan K Nearest Neighbor adalah suatu pendekatan untuk menghitung kedekatan antara kasus baru dengan kasus lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada. Ilustrasi kedekatan kasus pada Gambar 2.2 memberikan gambaran tentang proses mencari solusi terhadap seorang pasien baru dengan menggunakan mengacu pada solusi dari pasien terdahulu. Untuk mencari kasus pasien mana yang akan digunakan, maka dihitung kedekatan antara kasus pasien baru dengan semua kasus pasien lama. Kasus pasien lama dengan kedekatan terbesar - lah yang akan diambil solusinya untuk digunakan pada kasus pasien baru.

J.Han dalam (Santoso & Kusumaningsih, 2018) mendefinisikan Euclidean Distance mengukur kedekatan antara dua buah objek yang digambarkan sebagai pengukuran garis lurus atau secara langsung. Metode pengukuran ini cocok apabila diimplementasikan terhadap 14 data yang memiliki nilai atribut bersifat numerikal, khususnya dengan atribut kontinyu.

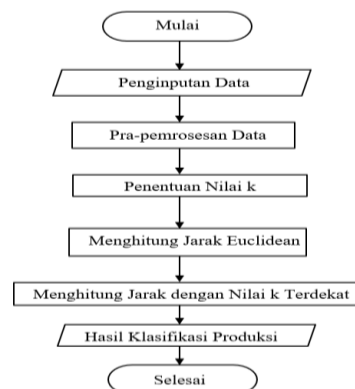
2. METODE PENELITIAN

Pendekatan penelitian yang digunakan dalam penerapan data mining untuk klasifikasi daun teh basah adalah pendekatan kuantitatif. Metode penelitian kuantitatif adalah sebuah metode penelitian yang didalamnya menggunakan banyak angka Mulai dari proses pengumpulan data hingga penafsirannya (Ibrahim N. S., 2023).



Gambar 2. Kerangka Penelitian

Teknik Pengumpulan data yaitu dengan melakukan wawancara cara langsung kepada asisten kepala kebun teh di ptpn IV sidamanik Dan menggunakan studi pustaka yaitu materi jurnal atau buku yang sesuai dengan permasalahan judul penelitian ini. Data yang akan digunakan yaitu data produksi bulanan pada tahun 2020 – 2024. Pada penelitian ini pengolahan data mining yang dilakukan adalah Knowledge discovery in database (KDD). Data yang sudah diperoleh selanjutnya akan dianalisis menggunakan KNN untuk menentukan kelas tertentu yaitu meningkat dan menurun. Perancangan flowchart K-Nearest Neighbors (KNN) dengan pembuatan diagram alur yang menggambarkan langkah-langkah dalam proses klasifikasi data mining menggunakan algoritma KNN.



Gambar 3. Flowchart Klasifikasi Dengan Algoritma (KNN)



Teknik analisis data yang dilakukan yaitu menggunakan proses data mining Knowledge discovery in database (KDD), KNN daer. Penerapan awal dilakukannya klasifikasi dengan kelas Melebihi Target, Sesuai Target, Di Bawah Target. Selanjutnya dilakukan perhitungan knn dengan menggunakan perhitungan jarak euclidean distance. Selanjutnya, dilakukan pengujian k dengan range yang ditentukan 1-3. Dimana, dari range tersebut akan mendapatkan tingkat akurasi yang paling tinggi yang akan diambil. Perhitungan manual dengan excel dan pengujian akurasi menggunakan tools rapid miner.

3. HASIL DAN PEMBAHASAN

Pada penelitian ini diperoleh data hasil produksi Daun Teh Basah (DTB) secara langsung data yang diperoleh yaitu data produksi dari tahun 2020 – 2024. Data terdiri dari atribut, Realisasi, Rkap, Rko, dan Protas. Selanjutnya akan melewati proses data cleaning dan data processing adalah proses mengolah data mentah agar siap digunakan untuk analisis atau pemodelan. Data mentah seringkali mengandung kesalahan, data yang tidak relevan, atau informasi yang tidak terstruktur, sehingga perlu dibersihkan dan diorganisasi terlebih dahulu. Terdapat 3 penghasil produksi daun teh basah yaitu DTB SID, DTB BUT, dan DTB TOB maka yang akan diambil datanya yaitu DTB TEH yang mencakup seluruhnya. Terdapat 116 data dimana, 96 data merupakan data training dan 20 data merupakan data testing.

Dalam transformasi data akan diubah dengan bentuk yang disesuaikan dalam proses data mining. Data yang tadi diproses menggunakan excel akan diproses dan diolah lebih lanjut menggunakan rapid miner.

Tabel 1. Dataset

DTB TEH					
No	REAL	RKAP	RKO	Protas	Kelas
1	3.527.881	4.271.000	4.271.000	861	Tidak Mencapai Target
2	3.527.881	4.271.000	4.271.000	861	Tidak Mencapai Target
3	3.849.504	3.846.000	3.846.000	939	Mencapai Target
4	7.377.385	8.117.000	8.117.000	1.800	Tidak Mencapai Target
...
116	3.306.106	4.139.000	4.139.000	806	Tidak Mencapai Target

Data akan dibagi menjadi 2 bagian yaitu data latih (data training) dan data uji (data testing). data training merupakan Data yang digunakan untuk melatih model atau algoritma sehingga model dapat belajar mengenali pola, hubungan, atau aturan yang ada dalam data. Sedangkan data testing merupakan Data yang digunakan untuk menguji kinerja model setelah dilatih menggunakan data training.

Tabel 2. Data Training

No	REAL	RKAP	RKO	Protas	Kelas
1	3.527.881	4.271.000	4.271.000	861	Tidak Mencapai target
2	3.527.881	4.271.000	4.271.000	861	Tidak Mencapai Target
3	3.849.504	3.846.000	3.846.000	939	Mencapai Target
4	7.377.385	8.117.000	8.117.000	1.800	Tidak Mencapai Target
...
20	3.306.106	4.139.000	4.139.000	806	Tidak Mencapai Target



Tabel 3. Data Testing

4.085.623	3.711.000	3.711.000	1.160	2	D8	20.714.953	14	Tidak Mencapai Target
					D9	414.777	2	Mencapai Target
					D10	27.067.867	15	Tidak Mencapai Target
					D11	583.239	4	Tidak Mencapai Target
					D12	33.699.669	16	Tidak Mencapai Target
					D13	1.057.088	10	Tidak Mencapai Target
					D14	40.734.168	17	Tidak Mencapai Target
					D15	997.000	9	Tidak Mencapai Target
					D16	47.270.950	18	Tidak Mencapai Target
					D17	968.917	7	Tidak Mencapai Target
					D18	54.313.088	19	Tidak Mencapai Target
					D19	1.247.553	11	Tidak Mencapai Target
					D20	61.296.490	20	Tidak Mencapai Target

- Langkah Langkah dalam Perhitungan KNN :
1. Tentukan nilai k yang akan digunakan dengan k=3
 2. Hitung jarak antara data baru dan data training
 3. Perhitungan yang digunakan yaitu Euclidean Distance,

$$d(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}$$

Ket :

d = jarak

n = banyak data

i = variabel data

xi = sampel data / data training

yi = data uji / testing

$$d1 = \sqrt{(3.527.881 - 4.085.623)^2 + (4.271.000 - 3.711.000)^2 + (4.271.000 - 3.711.000)^2 + (861 - 1.160)^2}$$

$$= \sqrt{((-557.742)^2 + (560.000)^2 + (560.000)^2 + (-299)^2)}$$

$$= \sqrt{938.276.277.9965}$$

$$= 968.647$$

Hal yang sama dilakukan hingga jarak d20.

4. Mengurutkan jarak tersebut dan menentukan tetangga mana yang dekat berdasarkan jarak ke k
5. Berdasarkan Perhitungan jarak diatas maka, didapat perhitungan jarak euclidean dengan urutan nilai d1 - d20 . Penelitian ini menggunakan k = 3
6. Menentukan kategori tetangga terdekat
7. Menggunakan kategori mayoritas sederhana dari tetangga yang terdekat sehingga nilai prediksi dari data baru.

Tabel 4. Euclidean Distance

No	Euclidean	Rank	Klasifikasi
D1	968.647	5	Tidak Mencapai Target
D2	968.647	6	Tidak Mencapai Target
D3	303.648	1	Mencapai Target
D4	7.047.082	12	Tidak Mencapai Target
D5	986.922	8	Tidak Mencapai Target
D6	13.768.294	13	Tidak Mencapai Target
D7	430.177	3	Mencapai Target

Hasil perhitungan kelompok data uji berdasarkan mayoritas dari data k tetangga terdekat. Maka pada penelitian ini yaitu nilai k = 1 k = 2 dan k = 3, dengan tiga jarak terkecil yaitu , d3, d7 dan d9 dengan klasifikasi Mencapai Target, Mencapai Target dan Mencapai Target.

Perhitungan manual k = 1 dari tingkat akurasi yang dihasilkan terlihat seperti dibawah ini:

Perhitungan Accuracy pada counfision matrix data uji

$$\frac{TP + TN}{TP + TN + FP + FN}$$

$$\frac{17 + 2 + 0 + 1}{17 + 2}$$

$$\frac{19}{20} = 95.00\%$$

Perhitungan Recall pada counfision matrix data uji

$$\frac{TP}{TP + FP}$$

$$\frac{17 + 0}{17}$$

$$\frac{17}{17} = 100\%$$

Perhitungan Precision pada counfision matrix data uji

$$\frac{TP}{TP + FN}$$

$$\frac{17 + 1}{17 + 1}$$



$$\frac{17}{18} = 94.44\%$$

Perhitungan manual k = 2 dari tingkat akurasi yang dihasilkan terlihat seperti dibawah ini:

Perhitungan Accuracy pada counfision matrix data uji

$$\frac{TP + TN}{TP + TN + FP + FN}$$

$$\frac{15 + 3}{15 + 3 + 2 + 0}$$

$$\frac{18}{20} = 90.00\%$$

Perhitungan Recall pada counfision matrix data uji

$$\frac{TP}{TP + FP}$$

$$\frac{15}{15 + 2}$$

$$\frac{15}{17} = 88.24\%$$

Perhitungan Precision pada counfision matrix data uji

$$\frac{TP}{TP + FN}$$

$$\frac{15}{15 + 0}$$

$$\frac{15}{15} = 100\%$$

Perhitungan manual k = 3 dari tingkat akurasi yang dihasilkan terlihat seperti dibawah ini:

Perhitungan Accuracy pada counfision matrix data uji

$$\frac{TP + TN}{TP + TN + FP + FN}$$

$$\frac{16 + 3 + 1 + 0}{16 + 3}$$

$$\frac{19}{20} = 95.00\%$$

Perhitungan Recall pada counfision matrix data uji

$$\frac{TP}{TP + FP}$$

$$\frac{16 + 0}{16 + 0}$$

$$\frac{16}{16} = 100\%$$

Perhitungan Precision pada counfision matrix data uji

$$\frac{TP}{TP + FN}$$

$$\frac{16}{16 + 1}$$

$$\frac{16}{17} = 94.12\%$$

Berikut accuracy, class recall, dan class precision dengan Nilai k = 1, k = 2, k = 3 sebagai perbandingan dibawah ini

Dari proses parameter k – 1 pada Gambar 4.20 algoritma K-Nearest Neighbor terbentuk hasil matrix Accuracy sebesar 95.00% class recall sebesar 100% dan class precision sebesar 94.44%.

	true Mencapai Target	true Tidak Mencapai Target	class precision
pred. Mencapai Target	17	1	94.44%
pred. Tidak Mencapai Target	0	2	100.00%
class recall	100.00%	66.67%	

Gambar 4. Hasil Accuracy, Class Recall, dan Class Precion k = 1

Dari proses parameter k – 2 pada gambar 4.21 algoritma K-Nearest Neighbor terbentuk hasil matrix Accuracy sebesar 90.00% class recall sebesar 88.24% dan class precision sebesar 100%.

	true Mencapai Target	true Tidak Mencapai Target	class precision
pred. Mencapai Target	15	0	100.00%
pred. Tidak Mencapai Target	2	3	60.00%
class recall	88.24%	100.00%	

Gambar 5. Hasil Accuracy, Class Recall, dan Class Precion k = 2

Dari proses parameter k – 3 pada gambar 4.22 algoritma K-Nearest Neighbor terbentuk hasil matrix Accuracy sebesar 95.00% class recall sebesar 94.12% dan class precision sebesar 100%.

	true Mencapai Target	true Tidak Mencapai Target	class precision
pred. Mencapai Target	16	0	100.00%
pred. Tidak Mencapai Target	1	3	75.00%
class recall	94.12%	100.00%	

Gambar 6. Hasil Accuracy, Class Recall, dan Class Precion k = 3

Sehingga algoritma knn memiliki kinerja yang baik dalam memprediksi hasil produksi daun teh basah (DTB).

Data dengan model knn yang digunakan berjumlah 96 data dan data uji dengan menggunakan nilai k = 1, k = 2, k = 3



4. KESIMPULAN

Pada penelitian ini dilakukan menggunakan algoritma k-Nearest Neighbor dengan menggunakan data yang diolah berdasarkan tahapan Knowledge Discovery in Databases (KDD). Berdasarkan hasil perhitungan data mining menggunakan teknik klasifikasi dan algoritma k-nearest neighbor, didapatkan hasil prediksi produksi daun teh basah dengan dataset sebanyak 116 data dan data testing sebanyak 96 data, yang terdiri dari 4 atribut. Kemudian hasil implementasi algoritma KNN pada aplikasi rapidminer dilakukan pergantian nilai k, dengan accuracy k = 1 sebesar 95% , k = 2 sebesar 90% dan k = 3 sebesar 95%. perbandingan k = 1 dan k = 3 adalah hasil precision dan recall. class precision k =1 sebesar 94.44% dan class recall sebesar 100% sedangkan di k = 3 class precision sebesar 100% dan class recall sebesar 94.12%.

Hasil prediksi yang diperoleh menunjukkan bahwa metode ini layak digunakan untuk prediksi dimasa mendatang. Dengan menerapkan prediksi tersebut, perusahaan dapat memperoleh hasil produksi secara efisien.

5. REFERENSI

- Basriati, M.Sc, S., & Safitri, M.Mat, E. (2021). Penerapan Metode Fuzzy Tsukamoto dalam Menentukan Jumlah Produksi Tahu. *Jurnal Sains, Teknologi Dan Industri*, 18(1), 120. <https://doi.org/10.24014/sitekin.v18i1.11022>
- Fahreza, A. (2022). Penerapan Data Mining dengan Metode Single Moving Average dalam Pengolahan Data Penerimaan Siswa Baru. *Proceeding Seminar Nasional Ilmu Komputer*, 2(1), 25–34.
- Muhadi, A., & Octaviano, A. (2023). Penerapan Data Mining Untuk Prediksi Hasil Keuntungan Lelang Mesin X-Ray Tahun 2020 Dengan Metode K-Nearest Neighbor (Studi Kasus: PT.Ramadika Mandiri). *Jurnal Informatika MULTI*, 1(2), 2985–8860.
- Nozomi, I. (2023). Penerapan Data Mining Untuk Peringatan Dini Banjir Menggunakan Metode Klastering K-Means (Studi Kasus Kota Padang). *Jurnal Sains Informatika Terapan*, 2(2), 39–44. <https://doi.org/10.62357/jsit.v2i2.165>
- Putra, A. D. (2019). Pengawasan mutu proses produksi teh hitam (studi kasus di pabrik teh Sumber Daun di Desa Hegarmanah Kecamatan Takokak Kabupaten Cianjur). *Jurnal Penelitian Dan Pengembangan Sains Dan Teknologi*, 13(1), 1–6.
- Santoso, P. Y., & Kusumaningsih, D. (2018). Algoritma K-Nearest Neighbor Dengan Menggunakan Metode Euclidean Distance Untuk Memprediksi Kelulusan Ujian Nasional Berbasis Desktop Pada Sma Negeri 12 Tangerang. *Skanika*, 1(1), 123–129.
- Sari Pakpahan, N. (2021). Implementasi Data Mining Menggunakan Algoritma J48 Dalam Menentukan Pola Itemset Belanja Pembeli (Study Kasus: Swalayan Brastagi Medan). *Journal of Computing and Informatics Research*, 1(1), 7–13. <https://doi.org/10.47065/comforch.v1i1.111>
- Shantika Hutapea, Y., & Fauzi, R. (2021). Data mining sistem tata letak material di PT. Batam Cyclelect. *Jurnal Comasie*, 04(06), 72–79.
- Syukri Mustafa, M., & Wayan Simpen, I. (2019). Implementasi Algoritma K-Nearest Neighbor (KNN) Untuk Memprediksi Pasien Terkena Penyakit Diabetes Pada Puskesmas Manyampa Kabupaten Bulukumba. *Februari*, 2019(1), 1–10.
- Zai, C. (2022). IMPLEMENTASI DATA MINING SEBAGAI PENGOLAHAN DATA. 2.