



## Perencanaan Jumlah Bahan Baku Produksi Soda Caustic PT. XK Dengan Metode Algoritma K-Nearest Neighbour

Abdul Rohman<sup>1</sup>, Tedjo Sukmono, ST., MT.<sup>2</sup>

<sup>1</sup>Jurusan Teknik Industri, Fakultas Sains dan Teknologi, Universitas Muhammadiyah Sidoarjo, Jl. Raya Gelam 250, Candi Sidoarjo,

<sup>2</sup>Jurusan Teknik Industri, Fakultas Sains dan Teknologi, Universitas Muhammadiyah Sidoarjo, Jl. Raya Gelam 250, Candi Sidoarjo,

### INFORMASI ARTIKEL

**Halaman:**

12 – 22

**Tanggal penyerahan:**

21 Juni 2020

**Tanggal diterima:**

19 September 2020

**Tanggal terbit:**

13 Oktober 2020

### ABSTRACT

*In the process of caustic soda production activities PT. XK, So it is necessary to plan the readiness of the main raw materials in the form of granule salt (NaCl), which is expected to be appropriate planning by making predictions on the availability of the main raw materials using the K-Nearest Neighbor algorithm method, so that it can assist companies in planning the number of main raw materials for the smooth running of industrial activities. That way PT.XK can guarantee the smooth production activities and production capacity to serve consumer demand for caustic soda (NaOH) products. K-Nearest Neighbor algorithm is a data mining process that is classified in the category of machine learning (supervised learning) based on instance base or proximity between training data and test data, with the euclidian distance formula to find the closest distance results. In doing predictive for the desired planning it is necessary to measure the level of accuracy of data, recall data, precision data and receiver operating characteristic (ROC) curves. So based on the research conducted to get predictive results in order to plan, the amount of raw material for one period to come with the accuracy of predictive data of 92.67%, with a comparison of data recall and data precision, resulting in a ROC (receiver operating characteristic) curve of 0.75 where predictions are quite good.*

**Keywords:** k-Nearest Neighbor, accuracy, predictive, raw material, supervised learning, data mining

### EMAIL

<sup>1</sup>rohman.16@outlook.com

<sup>2</sup>thedjoss@gmail.com

### ABSTRAK

Dalam proses kegiatan produksi soda caustic PT. XK, Maka diperlukannya perencanaan kesiapan bahan baku utama berupa garam granul (NaCl) yang diharapkan tepat perencanaan dengan melakukan prediksi pada ketersediaan bahan baku utama menggunakan metode algoritma K- Nearest Neighbour hingga bisa membantu perusahaan dalam merencanakan jumlah bahan baku utama untuk kelancaran kegiatan industri. Dengan begitu PT.XK dapat menjamin kelancaran kegiatan produksi dan kapasitas produksi untuk melayani permintaan konsumen pada produk soda caustic (NaOH). Algoritma K- Nearest Neighbour merupakan proses mining data yang tergolong dalam kategori machine learning (supervised learning) yang berdasarkan pada basis instance atau kedekatan antara data latih dan data uji dengan rumus euclidian distance guna mencari hasil jarak terdekat. Dalam melakukan prediktif untuk perencanaan yang diinginkan maka perlu mengukur tingkat keakurasian data, data recall, data presisi dan kurva ROC (receiver operating characteristic). Maka berdasarkan penelitian yang dilakukan mendapatkan hasil prediktif guna merencanakan jumlah bahan baku selama satu periode yang akan datang dengan keakuratan data prediksi sebesar 92,67%, dengan perbandingan data recall dan presisi data, sehingga menghasilkan kurva ROC (receiver operating characteristic) sebesar 0,75 dimana prediksi tergolong cukup baik.

**Kata kunci:** K-Nearest Neighbor, akurasi, prediktif, bahan baku, supervised learning, data mining

## PENDAHULUAN

Prediksi atau *forecasting* salah satu seni dan ilmu yang diimplementasikan untuk mengetahui atau menerka rencana apa yang akan terjadi pada masa mendatang. Prediksi menjadi penting karena penyusunan suatu rencana perusahaan dalam mengoptimalkan proses kegiatan produksi, dengan didasarkan pada suatu proyeksi atau ramalan. Maka pemilik perusahaan atau badan usaha penting untuk memprediksikan apa yang akan terjadi pada masa berikutnya yang digunakan sebagai acuan dalam menentukan pembuatan keputusan atau kebijakan untuk kelangsungan efektifitas dan efisiensi perusahaan[1].

PT. XK merupakan *industry chemical* yang menggunakan bahan baku garam (NaCl), berdasarkan data permintaan produk oleh konsumen pada periode dua tahun sebelumnya mengalami permintaan yang bersifat cukup fluktuatif, maka perlu perencanaan tepat dalam mempersiapkan bahan baku untuk produksi mengingat proses pengiriman bahan baku yang dikirim langsung import dari Australia maka banyak mengalami transit moda transportasi dan antrian di pelabuhan Surabaya. PT.XK mampu memproduksi berbagai bahan kimia yang diproduksi pada departement soda plant adalah *chlorine gas*, *sodium hydroxide*, *hydrochloric acid*, dan *sodium hypochlorite*, dan pada berbagai produk bahan kimia yang diantaranya menjadi produk utama yaitu *sodium hydroxide* (NaOH). Dalam proses produksi bahan kimia tersebut terjadi proses utama yaitu elektrolisis terjadi ketika aliran arus listrik melalui senyawa ionik dan mengalami reaksi kimia. Larutan elektrolit dapat menghantar listrik karena mengandung ion-ion yang dapat bergerak bebas. Ion-ion tersebut yang menghantarkan arus listrik melalui larutan. Hantaran listrik melalui larutan elektrolit terjadi ketika sumber arus searah memberi muatan yang berbeda pada kedua elektroda. Katoda (elektroda yang dihubungkan dengan kutub negatif) bermuatan negatif, sedangkan anoda (elektroda yang dihubungkan dengan kutub positif) bermuatan positif [2].

Permasalahan utama dalam studi kasus yang diambil pada penelitian ini merupakan bagaimana merencanakan sebuah kegiatan produksi untuk penggunaan bahan baku pada periode yang akan datang dengan melakukan prediksi menggunakan data penggunaan bahan baku mulai periode 2015-2019, berdasarkan metode penelitian *k- Nearest Neighbour* dimana sebuah *database* akan dibagi menjadi dua jenis data yaitu data latih dan data uji berdasarkan pembagian tersebut maka penggunaan bahan baku periode 2015-2018 dijadikan sebagai data latih dan periode tahun 2019 sebagai data uji yang didapat pada PT.XK.

Maka dalam penelitian perencanaan penggunaan bahan baku ini menggunakan konsep data mining dengan metode *K- Nearest Neighbour* merupakan metode yang menggunakan pemrograman dan algoritma yang tergolong *supervised* dimana data uji yang baru diklasifikasi berdasarkan mayoritas kelas pada *k-NN*. Tujuan dari pemrograman dan algoritma ini bermaksud mengklasifikasikan objek baru sebagai parameter dan sebagai data latih. Prinsip dari *k-NN* adalah menghasilkan *k- objek* dari sebuah data latih yang memiliki jarak paling dekat dengan data uji. *k-NN* menggunakan tolak ukuran kemiripan guna mengklasifikasikan data uji yang diberikan dengan data latih. Salah satu tolak ukuran kemiripan yang dibuat adalah jarak *euclidean* antara dua titik salah satunya titik pada data latih ( $x_{train}$ ) dan titik data uji ( $x_{test}$ ) [3].

Data mining merupakan gabungan dari ragam bidang ilmu antara lain basis data, *information retrieval*, statistika, algoritma dan *machine learning*. Bidang keilmuan ini sudah berkembang telah lama namun mempunyai peran penting pada masa sekarang ini dimana keluar kepentingan untuk memperoleh informasi yang lebih dari sebuah data transaksi maupun fakta yang telah dikumpulkan selama bertahun-tahun. Data mining salah satu cara mendapatkan informasi tersembunyi pada sebuah basis data dan tergolong bagian pada proses *Knowledge Discovery in Databases* (KDD) guna mendapatkan informasi atau pola yang berfungsi pada data [1].

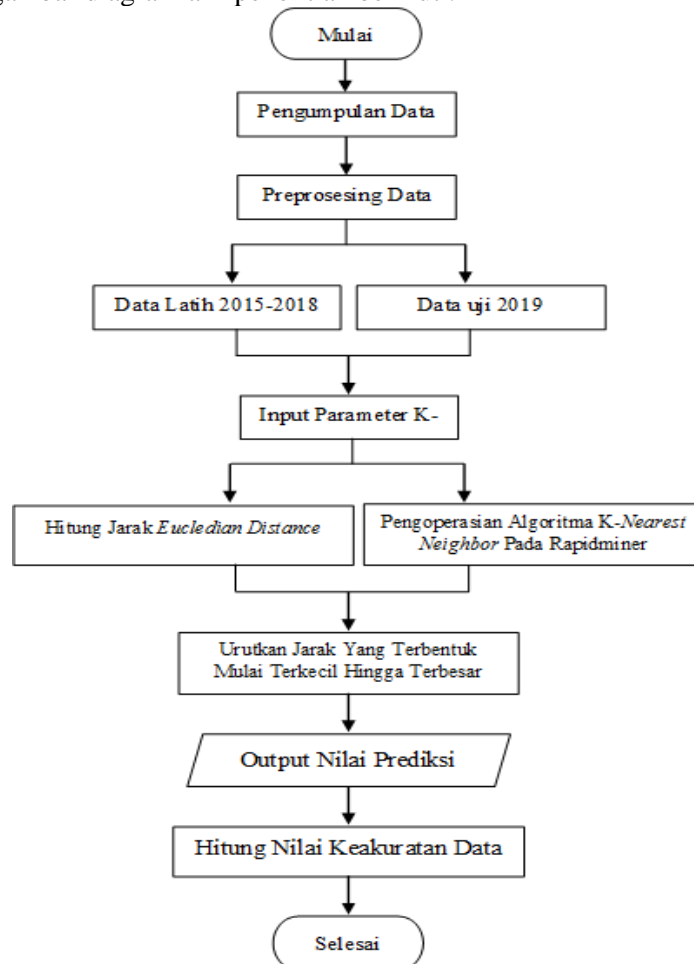
Dengan metode *k-NN* maka output yang dihasilkan berupa *Confusion matrix* yang memberikan sebuah keputusan diperoleh dalam sebuah data *training* dan data *testing*, *confusion matrix* memberikan sebuah penilaian *performance* untuk mengklasifikasi berdasarkan objek dengan benar atau salah. *Confusion matrix* berisi informasi data aktual (*actual*) dan prediktif (*predicted*) pada sistem klasifikasi [4]. Dan dituangkan dalam bentuk Kurva ROC (*Receiver Operating Characteristic*) digunakan dalam mengevaluasi akurasi sebuah *classifier* dan dapat membandingkan klasifikasi yang berbeda model. Kurva ROC juga digunakan untuk mengukur

AUC (*Area Under Curve*). Kurva ROC bisa membagi hasil positif dalam bentuk sumbu y dan hasil negatif dalam sumbu x [5].

Proses pemrograman data mining yang akan dilakukan menggunakan software RapidMiner merupakan salah satu perangkat lunak (*software*) yang memiliki sifat terbuka (*open source*). *Rapid miner* adalah sebuah *software* pemecahan masalah dalam melakukan analisa pada *data mining*, *text mining* dan analisa prediktif. *Rapid miner* dapat digunakan untuk berbagai teknik yang bersifat deskriptif dan prediksi guna memberikan pengetahuan kepada pengguna *software* sehingga mampu mengambil sebuah keputusan yang sangat tepat. *Rapid miner* mempunyai lebih dari 500 operator dalam proses data mining, tergolong dalam operator *output*, *input*, visualisasi dan data *preprocessing*. *Rapid miner* adalah *software* yang mampu berdiri sendiri untuk menganalisa data beserta sebagai mesin *data mining* terintegrasi pada produknya sendiri [6].

## METODE

Metode penelitian yang akan dilakukan pada penelitian ini menggunakan jenis metode penelitian deskriptif analitis yang berfungsi dalam mendiskripsikan dan memberi sebuah gambaran pada objek penelitian yang akan dilakukan melalui pengumpulan data. Berikut alur penelitian yang dilakukan dalam memprediksi jumlah bahan baku proses produksi *chemical industry* pada PT. XK dapat dilihat pada gambar diagram alir penelitian berikut :



Gambar 1. Contoh penggunaan *software* Dia Diagram Editor untuk pembuatan *flowcart*.

## HASIL DAN PEMBAHASAN

Berdasarkan diagram alir penelitian yang telah dijelaskan pada gambar 1, Maka berikut langkah-langkah dalam melakukan penelitian perencanaan jumlah bahan baku produksi *soda caustic* PT.XK dengan metode *K- Nearest Neighbour* :

**Pengumpulan Data**

Data penelitian yang didapatkan dari staf admin gudang bahan baku PT. XK yaitu data persediaan bahan baku berupa garam (NaCl) pada tahun 2016 sampai 2018, informasi data yang didapat berdasarkan syarat-syarat yang telah ditentukan perusahaan, data yang diperoleh untuk memerlukan sebuah pengolahan data lebih lanjut.

**Preprocessing Data**

*Preprocessing* data merupakan dimana tahap ini terbagi menjadi 4 jenis tahapan yaitu data *integration*, data *cleaning*, data *reduction* dan data *transformation*. Data *cleaning* merupakan tahap yang berupaya untuk mengisi nilai-nilai yang telah hilang, mengidentifikasi *outlier*, menghaluskan *noisy data*, dan inkonsistensi yang benar dalam data, dalam penelitian ini data penggunaan bahan baku produksi PT. XK akan dieliminasi dari beberapa data yang telah mengalami *noisy* atau kapasitas penggunaan bahan baku yang sama pada kurun waktu 5 tahun ke belakang.

**Data Latih**

Data *training* yang digunakan pada penelitian ini hanya diambil dari 4 tahun sebelumnya yaitu data penggunaan bahan baku tahun 2015 dan 2018 yang menjadi data *training*. Data *training* ditransformasikan menjadi 2 kelompokan yaitu data input dan target. Dimana data input merupakan data penggunaan dari bulan ke-1 sampai bulan ke-7, sedangkan data target menggunakan data bulan ke-8. dan seterusnya sampai batas data yang ada. Berikut data latih yang dihasilkan dapat dilihat pada tabel.

Tabel 1. Transformasi Data Latih

AT-1	AT-2	AT-3	AT-4	AT-5	AT-6	AT-7	Target	Class
1334	789	1598	1110	1162	1515	1214	1090	Hemat
789	1598	1110	1162	1515	1214	1090	1013	Hemat
1598	1110	1162	1515	1214	1090	1013	1349	Hemat
1110	1162	1515	1214	1090	1013	1349	1429	Hemat
1162	1515	1214	1090	1013	1349	1429	1670	Hemat
1515	1214	1090	1013	1349	1429	1670	1327	Hemat
1214	1090	1013	1349	1429	1670	1327	1300	Hemat
1090	1013	1349	1429	1670	1327	1300	1672	Hemat
1013	1349	1429	1670	1327	1300	1672	1573	Hemat
1349	1429	1670	1327	1300	1672	1573	1648	Hemat
1429	1670	1327	1300	1672	1573	1648	1449	Hemat
1670	1327	1300	1672	1573	1648	1449	1345	Hemat
1327	1300	1672	1573	1648	1449	1345	1855	Hemat
1300	1672	1573	1648	1449	1345	1855	1935	Hemat
1672	1573	1648	1449	1345	1855	1935	1142	Hemat
1573	1648	1449	1345	1855	1935	1142	1619	Hemat
1648	1449	1345	1855	1935	1142	1619	449	Hemat
1449	1345	1855	1935	1142	1619	449	1640	Hemat
1345	1855	1935	1142	1619	449	1640	1582	Hemat
1855	1935	1142	1619	449	1640	1582	1692	Hemat
1935	1142	1619	449	1640	1582	1692	1234	Hemat
1142	1619	449	1640	1582	1692	1234	1683	Hemat
1619	449	1640	1582	1692	1234	1683	1731	Hemat
449	1640	1582	1692	1234	1683	1731	721	Hemat
1640	1582	1692	1234	1683	1731	721	1235	Hemat
1582	1692	1234	1683	1731	721	1235	1339	Hemat
1692	1234	1683	1731	721	1235	1339	1750	Hemat
1234	1683	1731	721	1235	1339	1750	1570	Hemat
1683	1731	721	1235	1339	1750	1570	1511	Hemat
1731	721	1235	1339	1750	1570	1511	1611	Hemat
721	1235	1339	1750	1570	1511	1611	1653	Hemat
1235	1339	1750	1570	1511	1611	1653	1822	Hemat

Tabel 1. Transformasi Data Latih (Lanjutan...)

AT-1	AT-2	AT-3	AT-4	AT-5	AT-6	AT-7	Target	Class
1339	1750	1570	1511	1611	1653	1822	1646	Boros
1750	1570	1511	1611	1653	1822	1646	1795	Boros
1570	1511	1611	1653	1822	1646	1795	1345	Boros
1511	1611	1653	1822	1646	1795	1345	1585	Boros
1611	1653	1822	1646	1795	1345	1585	1737	Boros
1653	1822	1646	1795	1345	1585	1737	1717	Boros
1822	1646	1795	1345	1585	1737	1717	1707	Boros
1646	1795	1345	1585	1737	1717	1707	1450	Boros
1795	1345	1585	1737	1717	1707	1450	1767	Boros

**Data Uji**

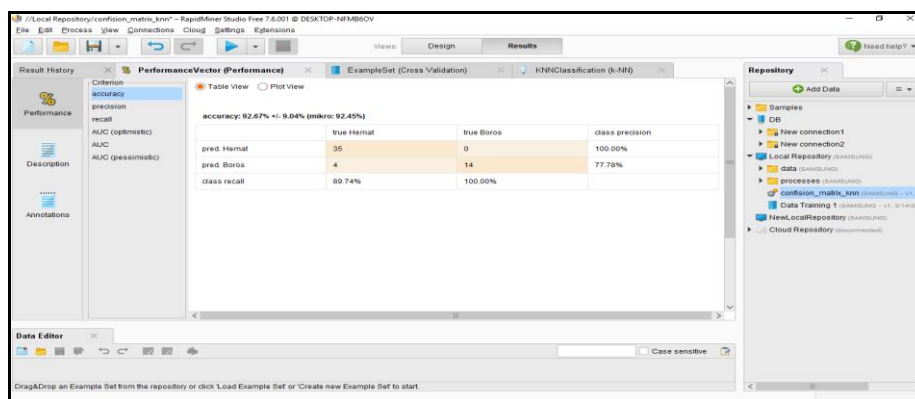
Sedangkan untuk data *testing* (uji) yang digunakan dalam transformasi untuk prediksi adalah penggunaan bahan baku dimulai pada bulan juli tahun 2018 hingga akhir tahun 2019 seperti pada tabel berikut :

Tabel 2. Transformasi Data Uji

Bulan - 1	Bulan - 2	Bulan - 3	Bulan - 4	Bulan - 5	Bulan - 6	Bulan - 7
1585	1737	1717	1707	1450	1767	1665
1737	1717	1707	1450	1767	1665	1682
1717	1707	1450	1767	1665	1682	1768
1707	1450	1767	1665	1682	1768	1618
1450	1767	1665	1682	1768	1618	1728
1767	1665	1682	1768	1618	1728	1030
1665	1682	1768	1618	1728	1030	1060
1682	1768	1618	1728	1030	1060	1291
1768	1618	1728	1030	1060	1291	1608
1618	1728	1030	1060	1291	1608	1724
1728	1030	1060	1291	1608	1724	1729
1030	1060	1291	1608	1724	1729	1889

**Input Parameter K-**

penentuan sebuah nilai K- merupakan hal yang penting karena dapat mempengaruhi kinerja metode algoritma k- *nearest neighbour* pada nilai k yang sangat kecil, nilai k umumnya ditentukan dalam jumlah ganjil untuk menghindari munculnya jumlah jarak yang sama dalam proses pengklasifikasian *Clustering* membagi item menjadi kelompok-kelompok berdasarkan yang ditemukan oleh *tool* data mining. Berdasarkan penggunaan nilai input parameter K- sebesar (3) maka langkah berikutnya dengan cara klik run hingga pemograman dapat mengeluarkan hasil *accuracy* dari proses mining dengan nilai K- sebesar (3) , seperti gambar 2 sebagai berikut :



Gambar 2. Output Accuracy Parameter K-

Berdasarkan hasil gambar diatas menunjukkan hasil akurasi data mining K- nearest neighbour dengan true positive 35 dan false positif 14 hingga menunjukkan nilai akurasi input proses mining sebesar 92% dimana keakuratan tersebut menggunakan nilai input parameter K- sebesar 3.

#### Hitung Jarak *Euclidian Distance*

Setelah mengetahui input parameter K- optimal dapat melakukan perhitungan jarak terdekat pada metode K- *Nearest Neighbour* yaitu *euclidian distance* guna menghasilkan jarak terdekat antara data latih dan data uji yang nantinya menghasilkan prediksi pada periode berikutnya, berikut hasil perhitungan jarak yang dapat dilihat pada tabel 3 selama satu periode 2020 dengan rumus *euclidian distance* :

$$Euclidian Distance = \sqrt{\sum_{i,j=1}^n (x_{train,i} - x_{test,j})^2} \dots(1)$$

Tabel 3. Hasil Perhitungan *Euclidian Distance*

Januari		Februari		Maret		April			
Distance	Rank	Target Distance	Rank	Target Distance	Rank	Target Distance	Rank	Target	
1297	35	1090	1326	32	1090	1425	37	1090	1187
1399	37	1013	1403	37	1013	1433	38	1013	1470
1295	34	1349	1331	33	1349	1281	30	1349	1249
1280	33	1429	1337	34	1429	1389	36	1429	1303
1133	25	1670	1215	29	1670	1229	28	1670	1263
1130	24	1327	1051	20	1327	1075	21	1327	1097
1142	26	1300	1182	27	1300	1117	22	1300	1095
1165	28	1672	1146	25	1672	1159	26	1672	1048
892	18	1573	1055	21	1573	950	19	1573	977
580	12	1648	693	13	1648	791	16	1648	640
659	15	1449	531	7	1449	587	12	1449	701
653	14	1345	681	12	1345	538	8	1345	539
721	16	1855	730	17	1855	799	17	1855	602
569	9	1935	696	16	1935	611	13	1935	737
445	5	1142	554	9	1142	566	9	1142	547
823	17	1619	694	15	1619	833	18	1619	738
936	19	449	825	18	449	691	14	449	821
1356	36	1640	1546	41	1640	1553	41	1640	1356
1485	41	1582	1350	35	1582	1525	40	1582	1530
1215	30	1692	1468	39	1692	1305	32	1692	1479
1463	40	1234	1184	28	1234	1466	39	1234	1299
1425	39	1683	1489	40	1683	1284	31	1683	1499
1423	38	1731	1354	36	1731	1368	35	1731	1150
1173	29	721	1423	38	721	1350	34	721	1369
1095	21	1235	1005	19	1235	1210	27	1235	1010
1262	32	1339	1181	26	1339	1133	23	1339	1266
1089	20	1750	1308	31	1750	1247	29	1750	1159
1155	27	1570	1086	22	1570	1308	33	1570	1254
1116	23	1511	1107	23	1511	983	20	1511	1216
1255	31	1611	1125	24	1611	1135	24	1611	988
1107	22	1653	1248	30	1653	1136	25	1653	1135
574	11	1822	693	14	1822	734	15	1822	547
429	4	1646	475	6	1646	481	5	1646	590
389	2	1795	354	2	1795	287	2	1795	299
485	7	1345	370	3	1345	354	3	1345	336

425	3	1585	591	11	1585	537	7	1585	423	5	1585
571	10	1737	429	5	1737	575	11	1737	497	6	1737
272	1	1717	573	10	1717	411	4	1717	578	11	1717
473	6	1707	266	1	1707	569	10	1707	418	4	1707
497	8	1450	410	4	1450	259	1	1450	566	10	1450
581	13	1767	545	8	1767	511	6	1767	300	2	1767
Result			Result			Result			Result		
Rank	Distance	Target	Rank	Distance	Target	Rank	Distance	Target	Rank	Distance	Target
1	272	1717	1	266	1707	1	259	1450	1	299	1795
2	389	1795	2	354	1795	2	287	1795	2	300	1767
3	425	1585	3	370	1345	3	354	1345	3	336	1345
Jumlah	43466	Class	Jumlah	43497	Class	Jumlah	43525	Class	Jumlah	43556	Class
Rata-rata	1699	Boros	Rata-rata	1616	Boros	Rata-rata	1530	Hemat	Rata-rata	1636	Boros

Tabel 3. Hasil Perhitungan *Euclidian Distance* (Lanjutan...)

Mei			Juni			Juli		
Distance	Rank	Target	Distance	Rank	Target	Distance	Rank	Target
1394	38	1090	1297	29	1090	1331	32	1090
1295	31	1013	1390	34	1013	1223	27	1013
1353	35	1349	1114	21	1349	989	10	1349
1301	32	1429	1382	33	1429	1139	24	1429
1196	29	1670	1318	30	1670	1271	28	1670
1141	23	1327	1326	31	1327	1320	30	1327
1152	26	1300	1177	25	1300	1329	31	1300
1067	21	1672	1160	24	1672	1067	21	1672
848	19	1573	1193	26	1573	1123	23	1573
705	16	1648	908	13	1648	1057	20	1648
537	9	1449	932	16	1449	997	13	1449
702	15	1345	680	6	1345	951	8	1345
660	14	1855	737	7	1855	731	3	1855
484	6	1935	1050	20	1935	991	11	1935
648	13	1142	1015	19	1142	1284	29	1142
801	18	1619	619	3	1619	1014	14	1619
734	17	449	986	18	449	811	4	449
1518	41	1640	915	14	1640	1154	25	1640
1334	34	1582	1636	41	1582	1036	18	1582
1494	40	1692	1440	36	1692	1665	41	1692
1472	39	1234	1583	40	1234	1570	38	1234
1371	36	1683	1405	35	1683	1583	39	1683
1390	37	1731	1487	38	1731	1404	36	1731
1147	25	721	1547	39	721	1625	40	721
1143	24	1235	639	4	1235	879	6	1235
1122	22	1339	1145	22	1339	650	2	1339
1319	33	1750	1156	23	1750	1164	26	1750
1159	27	1570	1483	37	1570	1344	34	1570
1172	28	1511	1261	28	1511	1475	37	1511
1235	30	1611	1245	27	1611	1336	33	1611

995	20	1653	1336	32	1653	1361	35	1653
566	10	1822	920	15	1822	1021	15	1822
293	1	1646	950	17	1646	1070	22	1646
468	5	1795	673	5	1795	1031	17	1795
303	2	1345	845	12	1345	996	12	1345
491	8	1585	420	1	1585	870	5	1585
401	3	1737	738	8	1737	622	1	1737
487	7	1717	797	9	1717	989	9	1717
576	11	1707	817	11	1707	1026	16	1707
403	4	1450	807	10	1450	1041	19	1450
628	12	1767	548	2	1767	888	7	1767
Result			Result			Result		
Rank	Distance	Target	Rank	Distance	Target	Rank	Distance	Target
1	293	1646	1	420	1585	1	622	1737
2	303	1345	2	548	1767	2	650	1339
3	401	1737	3	619	1619	3	731	1855
Jumlah	43586	Class	Jumlah	43617	Class	Jumlah	43647	Class
Rata-rata	1576	Hemat	Rata-rata	1657	Boros	Rata-rata	1644	Boros

Tabel 3. Hasil Perhitungan *Euclidian Distance* (Lanjutan...)

Agustus			September			Oktober		
Distance	Rank	Target	Distance	Rank	Target	Distance	Rank	Target
1300	32	1090	1056	19	1090	1254	33	1090
1305	34	1013	1356	36	1013	1152	27	1013
897	5	1349	1123	24	1349	1177	29	1349
989	11	1429	930	11	1429	1171	28	1429
1004	13	1670	827	4	1670	721	4	1670
1227	28	1327	861	6	1327	564	2	1327
1312	35	1300	1248	31	1300	913	13	1300
1250	29	1672	1264	32	1672	1196	31	1672
976	9	1573	1103	22	1573	1068	19	1573
957	8	1648	712	2	1648	818	7	1648
1070	19	1449	898	7	1449	581	3	1449
982	10	1345	1052	18	1345	876	10	1345
953	7	1855	1016	17	1855	1113	22	1855
860	3	1935	918	10	1935	925	14	1935
1124	23	1142	837	5	1142	817	6	1142
1293	31	1619	1217	30	1619	1015	17	1619
1062	17	449	1288	34	449	1206	32	449
1169	25	1640	1571	40	1640	1808	41	1640
1186	26	1582	1146	26	1582	1541	39	1582
1027	14	1692	1137	25	1692	1074	20	1692
1702	41	1234	1016	16	1234	1134	24	1234
1548	40	1683	1724	41	1683	1115	23	1683
1545	39	1731	1452	37	1731	1607	40	1731
1471	37	721	1549	39	721	1445	38	721
1220	27	1235	1195	29	1235	1290	35	1235
880	4	1339	1275	33	1339	1285	34	1339
646	1	1750	915	9	1750	1318	37	1750
1251	30	1570	662	1	1570	912	12	1570
1303	33	1511	1169	28	1511	420	1	1511
1491	38	1611	1307	35	1611	1185	30	1611
1372	36	1653	1490	38	1653	1311	36	1653
1045	15	1822	981	14	1822	1063	18	1822



1067	18	1646	969	12	1646	829	8	1646
1079	21	1795	1011	15	1795	872	9	1795
1144	24	1345	1093	21	1345	1013	16	1345
993	12	1585	1169	27	1585	1140	25	1585
905	6	1737	979	13	1737	1148	26	1737
763	2	1717	912	8	1717	966	15	1717
1077	20	1707	771	3	1707	904	11	1707
1096	22	1450	1073	20	1450	769	5	1450
1053	16	1767	1107	23	1767	1100	21	1767
Result			Result			Result		
Rank	Distance	Target	Rank	Distance	Target	Rank	Distance	Target
1	646	1750	1	662	1570	1	420	1511
2	763	1717	2	712	1648	2	564	1327
3	860	1935	3	771	1707	3	580	1449
Jumlah	43678	Class	Jumlah	43709	Class	Jumlah	43739	Class
Rata-rata	1801	Boros	Rata-rata	1642	Boros	Rata-rata	1429	Hemat

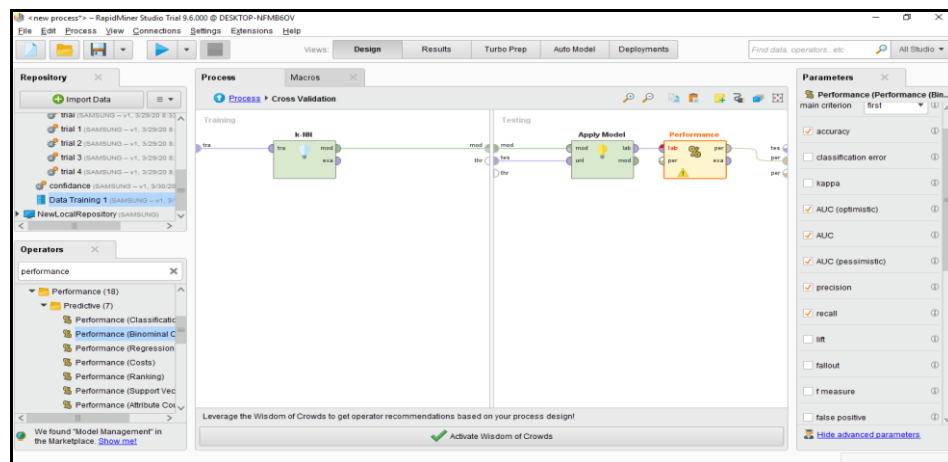
Tabel 3. Hasil Perhitungan *Euclidian Distance* (Lanjutan...)

November			Desember		
Distance	Rank	Target	Distance	Rank	Target
1021	21	1090	1151	25	1090
1378	36	1013	1235	31	1013
1074	23	1349	1336	33	1349
1237	32	1429	1195	29	1429
1097	27	1670	1166	28	1670
561	2	1327	965	15	1327
685	4	1300	765	5	1300
925	13	1672	744	4	1672
1075	24	1573	704	3	1573
894	12	1648	863	11	1648
777	5	1449	845	10	1449
616	3	1345	842	9	1345
955	14	1855	820	7	1855
1082	25	1935	868	12	1935
893	11	1142	995	18	1142
1005	18	1619	1163	27	1619
1020	20	449	1029	21	449
1758	40	1640	1762	40	1640
1803	41	1582	1754	39	1582
1524	38	1692	1788	41	1692
1048	22	1234	1531	38	1234
1196	31	1683	1218	30	1683
1010	19	1731	1063	22	1731
1607	39	721	1016	20	721
1318	34	1235	1520	37	1235
1382	37	1339	1469	35	1339
1343	35	1750	1480	36	1750
1317	33	1570	1349	34	1570
842	8	1511	1261	32	1511
469	1	1611	922	13	1611
1185	30	1653	545	1	1653
959	15	1822	669	2	1822
996	17	1646	826	8	1646
785	6	1795	948	14	1795

864	9	1345	790	6	1345
1083	26	1585	1009	19	1585
1144	28	1737	1103	23	1737
1148	29	1717	1148	24	1717
966	16	1707	1159	26	1707
881	10	1450	978	16	1450
817	7	1767	981	17	1767
Result			Result		
Rank	Distance	Target	Rank	Distance	Target
1	469	1611	1	1653	1653
2	562	1327	2	1822	1822
3	616	1345	3	1573	1573
Jumlah	43770	Class	Jumlah	43800	Class
Rata-rata	1428	Hemat	Rata-rata	1683	Boros

### Pengoperasian Algoritma K-Nearest Neighbor Pada Rapidminer

Berikut pengoperasian algoritma K-NN pada Rapidminer dapat dilihat pada gambar proses pemrograman yang dikerjakan dan dapat dirunning :



Gambar 3. Pemograman K- Nearest Neighbour

Dengan pemograman pada gambar 3 menghasilkan output running untuk menganalisa *accuracy*, *confusion matrix*, *precision*, *recall* dan *area under curve* (AUC). Masing-masing analisa dapat dilihat pada gambar 4 sebagai berikut :

```

PerformanceVector

PerformanceVector:
accuracy: 92.67% +/- 9.53% (micro average: 92.45%)
ConfusionMatrix:
True:  Hemat  Boros
Hemat:  35    0
Boros:  4    14
precision: 81.67% +/- 24.15% (micro average: 77.78%) (positive class: Boros)
ConfusionMatrix:
True:  Hemat  Boros
Hemat:  35    0
Boros:  4    14
recall: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: Boros)
ConfusionMatrix:
True:  Hemat  Boros
Hemat:  35    0
Boros:  4    14
AUC (optimistic): 1.000 +/- 0.000 (micro average: 1.000) (positive class: Boros)
AUC: 0.750 +/- 0.264 (micro average: 0.750) (positive class: Boros)
AUC (pessimistic): 1.000 +/- 0.000 (micro average: 1.000) (positive class: Boros)

```

Gambar 4. Performance Vector

Didapatkan hasil dengan tingkat akurasi 92,67 % yang artinya akurasi datanya sudah bagus dan baik, terdapat nilai pada tabel *Confusion Matrix* true positif sebesar 35 yang mengartikan bahwa prediksi benar penggunaan hemat itu benar, false positif sebesar 4 mengartikan bahwa kesalahan prediksi penggunaan hemat sebanyak 4, false negatif sebesar 0 bahwa kesalahan prediksi penggunaan boros tidak terjadi, dan true negatif sebesar 14. Berikut perhitungan akurasi pada prediksi :

$$\text{Akurasi} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \times 100\%$$

$$\text{Akurasi} = (35 + 14) / (35 + 14 + 4 + 0) \times 100\%$$

$$\text{Akurasi} = 92,67\%$$

Hasil *area under curve* (AUC) 0.750 yang termasuk dalam kategori prediksi yang cukup baik (*Fair Classification*). Dengan *class precission* prediksi kecocokan antara data asli dan Prediksi hemat sebesar 100% dan Kecocokan dengan Prediksi boros sebesar 77,78%. *Recall* sebagai rasio dari item relevan yang dipilih dengan total jumlah item relevan yang tersedia. Jadi dari data diatas tingkat relevan *true* hemat sebesar 89,74% dan *true* boros sebesar 100%.

## KESIMPULAN

Perencanaan jumlah bahan baku produksi soda caustic PT.XK telah dihasilkan rencana kebutuhan bahan baku untuk periode yang akan datang tiap bulan dengan dimulai pada bulan januari hingga desember menggunakan metode K- *Nearest Neighbour* dengan perhitungan kedekatan antara data uji dan latih pada rumus *euclidian distance*. Hingga menghasilkan analisa AUC (*area under curve*) memperlihatkan bahwa menunjukkan tingkat performansi sebesar 0,75 yang tergolong dalam *fair classification* atau cukup baik. Berdasarkan analisa AUC dapat diketahui bahwa prediksi dalam perencanaan penggunaan bahan baku kegiatan produksi soda caustic PT. XK dapat diaplikasikan meskipun prediksi dalam katagori cukup baik. Dengan diketahui tingkat keakurasian prediksi sebesar 92,67 % yang artinya akurasi datanya sudah bagus dan baik.

## DAFTAR PUSTAKA

- [1] Hasmawati, "Aplikasi prediksi penjualan barang menggunakan metode," vol. 3, no. 2, pp. 151–160, 2017.
- [2] Y. Wahyono, H. Sutanto, and E. Hidayanto, "Produksi Gas Hydrogen Menggunakan Metode Elektrolisis Dari Elektrolit Air Dan Air Laut Dengan Penambahan Katalis Naoh," *Youngster Phys. J.*, vol. 6, no. 4, pp. 353–359, 2017.
- [3] S. Satriya, R. H. D., Santoso, E., & Sutrisno, "Implementasi Metode Ensemble K-Nearest Neighbor untuk Prediksi Nilai Tukar Rupiah Terhadap Dollar Amerika," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 4, pp. 1718–1725, 2018.
- [4] Indrayanti, D. Sugianti, and M. A. Al Karomi, "Optimasi Parameter K Pada Algoritma K-Nearest Neighbour Untuk Klasifikasi Penyakit Diabetes Mellitus," *Pros. SNATIF*, pp. 823–829, 2017, doi: 10.1007/s10115-007-0114-2.
- [5] R. Aryanti, S. Atang, F. Eka, and P. Rifky, "Komparasi Algoritma Naive Bayes Dengan Algoritma Genetika Pada Analisis Sentimen Pengguna Busway," vol. V, no. 2, pp. 227–234, 2019, doi: 10.31294/jtk.v4i2.
- [6] D. Ardiansyah and W. Walim, "Algoritma c4.5 untuk klasifikasi calon peserta lomba cerdas cermat siswa smp dengan menggunakan aplikasi rapid miner," *J. Inkofar*, vol. 1, no. 2, pp. 5–12, 2018.