

DATA MINING METODE K-NEAREST NEIGHBOUR UNTUK REGRESI DATA PENJUALAN KAIN TEKSTILE

Fauzi Ilham, Arnes Sembiring, Rosyidah Siregar

Teknik Informatika, Universitas Harapan Medan

Jalan HM. Joni No.70 C, Teladan Baru, Kec. Medan Kota, Kota Medan, Sumatera Utara 20216

Mrfauziilham01@gmail.com

ABSTRAK

Dalam perindustrian kain tekstil di era serba teknologi pada perkembangan zaman ini sangatlah berperan penting dalam menentukan kesuksesan bisnis dalam menganalisis penjualan kain tekstile untuk memahami strategi pemasaran yang efektif. Maka diperlukan penggunaan pengolahan data (*data mining*) dengan aplikasi orange untuk me-regresi data penjualan dominan antara kain lapis 1 dan kain lapis 2 dalam melakukan perhitungan prediksi dari sebuah metode *K-Nearest Neighbour* yang dapat menghasilkan prediksi penjualan dari data yang terkumpul selama 90 hari yang dimana data tersebut terbagi dari data latih dan data uji yang memiliki tingkat akurasi sebesar 0,611 atau 61% sebagai tingkat persentase yang berarti metode ini berada diatas 0 sampai 0,568 sehingga memiliki tingkat akurasi yang baik untuk dipergunakan. Penelitian ini dilakukan untuk prediksi penjualan selama 30 hari kedepan dengan data kain penjualan *offline* untuk kain lapis 1 sebanyak 90 data dan kain lapis 2 sebanyak 90 data, juga secara *online* untuk kain lapis 1 sebanyak 90 data dan kain lapis 2 sebanyak 90 data, sehingga hasil dari sebuah prediksi pada tampilan bar plot bahwa tingkat penjualan kain lapis 1 lebih tinggi dibanding penjualan kain lapis 2 setelah dilakukan perhitungan. Hasil ini membuat perusahaan dapat mengetahui bagaimana strategi penjualan dan penyediaan barang untuk selanjutnya.

Kata kunci: *Data mining, Regresi, K-Nearest Neighbour, Orange*

1. PENDAHULUAN

Dalam perindustrian kain tekstil terdapat hal yang penting dan beragam, yang melibatkan sistem produksi, distribusi dan penjualan berbagai jenis kain. Permintaan pasar ataupun berbagai jenis mode yang tren pada perkembangan zaman sangatlah berperan penting dalam menentukan kesuksesan bisnis perindustrian dalam menganalisis penjualan kain tekstile yang baik untuk memahami pola permintaan dan merancang strategi pemasaran yang efektif. Melalui analisis penjualan, perusahaan akan sangat terbantu untuk dapat memahami tren musiman, perubahan permintaan konsumen, serta faktor lain yang mempengaruhi dalam jual beli kain tekstil. Dengan pemahaman yang baik mengenai pola permintaan konsumen, perusahaan dapat mengambil keputusan yang lebih baik dalam hal produksi, promosi, serta penyesuaian harga.

Sebab persaingan dalam industri kain tekstil sangatlah kompetitif, persaingan yang ketat sehingga memerlukan wawasan yang mendalam untuk mengetahui kekuatan dan kelemahan pesaing, serta peluang yang mungkin ada. Untuk mengetahui berbagai macam tekstile yang terjual banyak dan dominan dibeli dalam penjualan diperlukan adanya penggunaan data mining atau pengolahan data yang dapat memberi informasi yang tepat untuk mengetahui permintaan konsumen. Ada berbagai macam teknik yang dapat kita gunakan untuk keberlangsungan pengolahan data agar dapat dilakukan secara akurat dan mudah. *Data mining* merupakan suatu teknik atau metode tertentu untuk mengambil keputusan berdasarkan model atau informasi terpilih, data mining juga termasuk salah satu pola analisis dalam bidang ilmu yang memprediksi data dalam penerapan data di

masa depan, data mining juga mempunyai berbagai macam teknik dan metode yang sangat bervariasi. [1] Salah satu metode yang dapat dikembangkan adalah regresi. Regresi adalah teknik membangun model yang digunakan untuk memprediksi nilai input data yang diberikan. Regresi adalah suatu ukuran statistik yang digunakan untuk menentukan kekuatan hubungan antara variabel terikat (*dependent*) dengan variabel bebas (*independent*). [2]

Pada penelitian sebelumnya penggunaan metode *K-Nearest Neighbour* (KNN) ini digunakan untuk memprediksi penjualan buah dan sayur yang dipakai dalam penggunaan pengolahan data (*data mining*) agar dapat mengetahui metode mana yang dapat dipergunakan secara mudah dan akurat dalam mencari yang dominan terjual serta mengetahui minat beli pasar konsumen. Algoritma *K-Nearest Neighbour* ini digunakan karena memiliki tingkat akurasi yang baik dan rasio kesalahan yang kecil.[3] Metode *K-Nearest Neighbour* ini ini juga dikenal dengan penjumlahan, yaitu dengan mengambil dari titik data maksimum atau yg terbanyak dari setiap kategori yang kemudian dijumlahkan dengan data yg baru.[4] Dalam penerapan metode ini pengumpulan data yang digunakan, antara lain yaitu berdasarkan riwayat hari penjualan, jenis kain penjualan, jenis penjualan dan yang terjual.

2. TINJAUAN PUSTAKA

2.1. Pengertian Data Mining

Data mining adalah sebuah analisa dari observasi data dalam jumlah besar untuk menemukan hubungan yang tidak diketahui sebelumnya dan dua metode baru untuk meringkas data agar mudah dipahami. Data mining merupakan suatu langkah dalam knowlegde discovery in databases (KDD) yang berkaitan dengan

pengumpulan data dan memiliki teknik menganalisa data untuk digali informasi tersembunyi dalam jumlah besar dan kompleks, sehingga menghasilkan output berupa karakteristik atau pola dari data tersebut. [5]

2.2. Regresi

Regresi atau *regression* adalah suatu cara atau teknik analisis data yang dilakukan dengan perhitungan statistik untuk menguji seberapa erat hubungan antar variabel mengidentifikasi hubungan dan pola dalam data. [6]

Dalam regresi, obyek data dibagi menjadi dua yaitu data mentah dan data uji. Tujuan membagi kedua data tersebut ialah guna mendapatkan pengetahuan karakteristik kelas berdasarkan data mentah dan mengetahui keakuratan data uji. [7]

Regresi Linier adalah suatu metode statistik yang memiliki fungsi sebagai metode pengujian untuk mengetahui sejauh mana hubungan sebab akibat antar beberapa variabelnya yaitu sebagai variabel bebas (*independent variable*) dan sebagai variabel tidak bebas (*dependent variable*) atau juga variabel sebab dan variabel akibat. Regresi linier merupakan metode yang sangat kuat untuk menganalisis data lebih akurat. [8]

2.3. Prediksi

Prediksi hampir sama seperti klasifikasi dan estimasi, namun dalam sebuah prediksi nilai dari hasil yang berupa gambaran dimasa yang mendatang dengan hasil dari prediksi tidak akan benar-benar akurat. [9] Proses dalam prediksi ini sangatlah berhubungan erat dengan perhitungan matematik dan statistik. [10]

Pada penelitian ini prediksi yang berarti untuk memperkirakan bagaimana penjualan dimasa selanjutnya agar tidak terjadinya penumpukan barang yang dimana beresiko akan kerusakan akibat rayap ataupun bahkan kekurangan barang yang dapat mengakibatkan kerugian. [3]

2.4. Metode K-Nearest Neighbour (KNN)

Metode KNN atau *K-Nearest Neighbour* ini merupakan salah satu metode data *mining* yang digunakan dalam klasifikasi dan regresi. Dalam konteks klasifikasi, KNN mengidentifikasi kelas mayoritas dari data terdekat sebagai prediksi kelas untuk titik data baru dan sedangkan dalam konteks regresi ialah menghitung nilai rata-rata dari data terdekat sebagai prediksi nilai target untuk titik data baru. [11]

Berikut langkah-langkah untuk menghitung metode KNN, yaitu. [12]

- Menemukan parameter K.
- Menghitung jarak antara data yang akan dievaluasi dengan semua pelatihan.
- Mengurutkan jarak terbentuk dan jarak terdekat sampai urutan K.
- Memasangkan kelas yang bersesuaian.

Salah satu jarak yang digunakan dalam *K-Nearest Neighbors* adalah jarak *Euclidian*:

$$deuclidian(xy) = \sqrt{\sum (xi - xi')^2 + (yi - yi')^2} \quad (1)$$

Keterangan:

$deuclidian(xy)$: jarak euclidian antara dua titik x dan y

Jarak euclidian : jarak dalam bentuk garis lurus antara dua titik dalam ruang fitur.

$\sum$: simbol penjumlahan

xi : nilai koordinat i dari titik x pada data latih

xi' : nilai koordinat i dari titik x pada data prediksi

yi : nilai koordinat ke-i dari titik y pada data latih

yi' : nilai koordinat ke-i dari titik y pada data prediksi

2.5. Orange

Aplikasi Orange adalah aplikasi yang digunakan untuk mengelola data dalam menganalisa suatu data. Aplikasi orange ini juga bisa disebut dengan aplikasi yang berbasis *open source*. [13]

Orange menyediakan beberapa *widget* atau teknik yang diletakkan pada *board* yang nantinya akan terhubung ke *widget* lainnya guna dalam membuka isi data, membaca data dan kemudian mengklasifikasi data.

3. METODE PENELITIAN

3.1. Analisis Perhitungan Data

Analisis perhitungan data adalah tahap dimana penulis menganalisis data mentah secara manual. Berikut merupakan data hasil penjualan tekstile yang didapat penulis dari mitra yang ditunjukkan pada gambar 1 berikut.

1	HARI	KATEGORI KAIN LAPIS	PENUJUALAN	BANYAK TERJUAL
2	1	KAIN LAPIS 1 (ASIANTEX)	OFFLINE	60
3	1	KAIN LAPIS 1 (ASIANTEX)	ONLINE	15
4	1	KAIN LAPIS 2 (KIRARA)	OFFLINE	20
5	1	KAIN LAPIS 2 (KIRARA)	ONLINE	5
6	2	KAIN LAPIS 1 (ASIANTEX)	OFFLINE	150
7	2	KAIN LAPIS 1 (ASIANTEX)	ONLINE	30
8	2	KAIN LAPIS 2 (KIRARA)	OFFLINE	27
9	2	KAIN LAPIS 2 (KIRARA)	ONLINE	15
10	3	KAIN LAPIS 1 (ASIANTEX)	OFFLINE	25
11	3	KAIN LAPIS 1 (ASIANTEX)	ONLINE	8
12	3	KAIN LAPIS 2 (KIRARA)	OFFLINE	12
345	26	KAIN LAPIS 2 (KIRARA)	ONLINE	4
346	27	KAIN LAPIS 1 (ASIANTEX)	OFFLINE	34
347	27	KAIN LAPIS 1 (ASIANTEX)	ONLINE	22
348	27	KAIN LAPIS 2 (KIRARA)	OFFLINE	15
349	27	KAIN LAPIS 2 (KIRARA)	ONLINE	7
350	28	KAIN LAPIS 1 (ASIANTEX)	OFFLINE	6
351	28	KAIN LAPIS 1 (ASIANTEX)	ONLINE	2
352	28	KAIN LAPIS 2 (KIRARA)	OFFLINE	10
353	28	KAIN LAPIS 2 (KIRARA)	ONLINE	8
354	29	KAIN LAPIS 1 (ASIANTEX)	OFFLINE	7
355	29	KAIN LAPIS 1 (ASIANTEX)	ONLINE	25
356	29	KAIN LAPIS 2 (KIRARA)	OFFLINE	10
357	29	KAIN LAPIS 2 (KIRARA)	ONLINE	15
358	30	KAIN LAPIS 1 (ASIANTEX)	OFFLINE	27
359	30	KAIN LAPIS 1 (ASIANTEX)	ONLINE	15
360	30	KAIN LAPIS 2 (KIRARA)	OFFLINE	27
361	30	KAIN LAPIS 2 (KIRARA)	ONLINE	10

Gambar 1. Data penjualan

Sumber: Data Penjualan Toko Family Usaha Textil

Untuk menguji perhitungan secara manual antar kedua metode *K-Nearest Neighbour* berdasarkan data penjualan yang diberikan, maka sebagai berikut penjelasannya.

3.2. Perhitungan Metode K-Nearest Neighbour

Pada perhitungan ini dibutuhkan data uji dari gambar 1 yang dimana penulis mencoba Sebagian data untuk di uji pada metode *KNN* pada gambar 1. Untuk menghitung sebuah data yang secara *offline* termasuk jenis kain penulis juga membutuhkan sebuah data *blackbox* prediksi yang dimana seperti table 1 dibawah ini.

Tabel 1. Data Prediksi

Hari	Penjualan	Banyak Terjual/m	Kategori Kain
1	OFFLINE	10	?

Pada data tabel 1 dapat kita hitung berdasarkan penjualan secara *offline* pada kolom pertama data

Tabel 2. Perhitungan nilai K *Offline* berdasarkan jarak

Hari	Kategori Kain Lapis	Penjualan	Banyak Terjual/m	Perhitungan Jarak (d)
1	Kain Lapis 2 (KIRARA)	OFFLINE	20	10
3	Kain Lapis 1 (ASIANTEX)	OFFLINE	25	15
4	Kain Lapis 1 (ASIANTEX)	OFFLINE	18	8,5
4	Kain Lapis 2 (KIRARA)	OFFLINE	24	14
5	Kain Lapis 1 (ASIANTEX)	OFFLINE	15	6,4

Dari kelima data tersebut terdapat kain lapis 2 dengan nilai yang terdekat dengan data prediksi maka data kolom yang kosong pada data prediksi yaitu kain lapis 2.

Maka selanjutnya untuk menghitung sebuah data yang secara *online* termasuk jenis kain penulis juga membutuhkan sebuah data *blackbox* prediksi yang dimana seperti tabel 3 berikut.

Tabel 3. Tabel data prediksi

Hari	Penjualan	Banyak Terjual/m	Kategori Kain
1	ONLINE	30	?

Pada data tabel 3 dapat kita hitung berdasarkan penjualan secara *online* pada kolom pertama.

Tabel 4. Perhitungan nilai K *Online* berdasarkan jarak

Hari	Kategori Kain Lapis	Penjualan	Banyak Terjual/m	Perhitungan Jarak (d)
1	Kain Lapis 1 (ASIANTEX)	ONLINE	15	15
1	Kain Lapis 2 (KIRARA)	ONLINE	5	25
2	Kain Lapis 2 (KIRARA)	ONLINE	15	15
3	Kain Lapis 1 (ASIANTEX)	ONLINE	8	22
4	Kain Lapis 1 (ASIANTEX)	ONLINE	6	24
5	Kain Lapis 2 (KIRARA)	ONLINE	12	18

Dari kelima data tersebut terdapat kain lapis 2 dengan nilai yang terdekat dengan data prediksi maka data kolom yang kosong pada data prediksi yaitu kain lapis 2.

4. HASIL DAN PEMBAHASAN

4.1. Pengujian Sistem

Langkah awal untuk melakukan pengujian data secara sistem dengan metode *K-Nearest Neighbour* yaitu dengan menyiapkan data lengkap dan memasukkan ke dalam sistem aplikasi Orange.

tersebut termasuk dalam kategori kain apa. Maka dapat dimulai perhitungan dengan persamaan.

Hari ke-1, Kain Lapis 1 (Asiantex) *Offline* dengan banyak terjual 60

$$d = \sqrt{\sum (xi - xi')^2 + (yi - yi')^2}$$

$$d = \sqrt{\sum (1 - 1)^2 + (60 - 10)^2}$$

$$d = 50 \quad (2)$$

Maka dari perhitungan diatas, nilai K (nilai rentang) yang digunakan yaitu-5. Sehingga perlu mengambil 5 data jarak terdekat (d) dengan data prediksi pada tabel 1.

Hari ke-1, Kain Lapis 1 (Asiantex) *Online* dengan banyak terjual 15

$$d = \sqrt{\sum (xi - xi')^2 + (yi - yi')^2}$$

$$d = \sqrt{\sum (1 - 1)^2 + (15 - 30)^2}$$

$$d = 15 \quad (12)$$

Maka dari perhitungan diatas, nilai K (nilai rentang) yang penulis gunakan yaitu-5. Sehingga perlu mengambil 5 data jarak terdekat (d) dengan data prediksi pada tabel 3. Jarak yang terdekat yaitu dapat dilihat pada tabel 4 berikut.

Data mentah yang digunakan adalah data hasil penjualan tekstile untuk menguji seberapa akurat metode dalam memprediksi dari kedua kategori kain lapis.



Gambar 4. Tampilan awal orange

Pada gambar 4 menunjukkan tampilan pada aplikasi orange saat pertama kali dibuka. Setelah pengguna mengklik perintah *new* maka tampilan selanjutnya akan menampilkan seperti gambar 5 berikut.



Gambar 5. Tampilan lembaran baru

Pada gambar 5 diatas merupakan tampilan awal setelah pengguna mengklik perintah *new*, maka tahap awal yaitu dengan menggunakan *widget* dari menu paling atas yaitu Data.



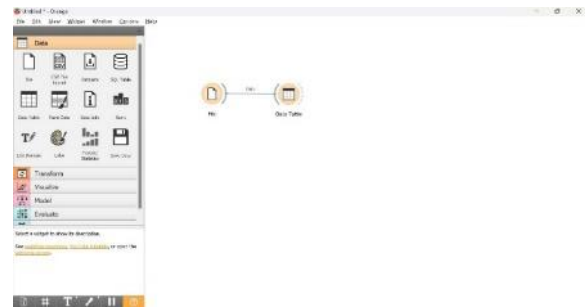
Gambar 6. Tampilan data file

Tampilan gambar 6 merupakan tampilan menu data setelah di klik oleh pengguna lalu menarik *widget* file ke halaman kerja. Setelah itu masukkan data mentah ke dalam *widget* file dengan *double* klik.



Gambar 7. Tampilan masukan data

Pengguna dapat memasukkan data dan mengubah role dengan menyesuaikan target kemudian *file* dapat disambungkan ke dalam *widget* data table yang ada di menu data.



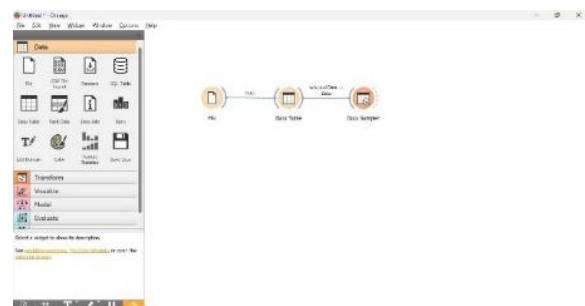
Gambar 8. Tampilan data table

Pada *widget* data table gambar 8 berfungsi untuk menampilkan data yang dimasukkan ke *widget* file agar pengguna mengetahui ada berapa *total* data dalam *file* seperti pada gambar 9 berikut.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

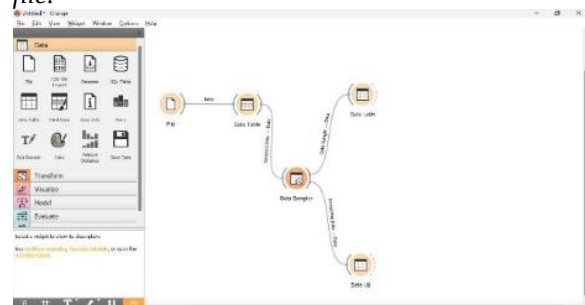
Gambar 9. Tampilan double klik data table

Ketika *double* klik pada data table, maka data table akan menampilkan isi pada file yang telah dimasukkan. Dengan begitu terdapat data yang diketahui sebanyak 360 data yang berisi data penjualan kain lapis 1 dan kain lapis 2.



Gambar 10. Tampilan data sample

Pada *widget* data sample di gambar 10 merupakan pembagian data dari keseluruhan data mentah yang dimasukkan pengguna ke dalam *widget* file.

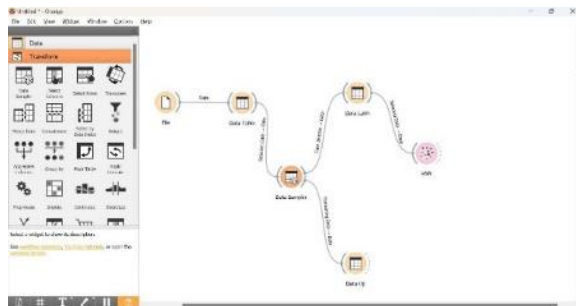


Gambar 11. Data uji dan data latih

Pembagian data latih digunakan untuk prediksi dalam penggunaan metode yang digunakan sedangkan data uji yang dihubungkan memakai data sisa data agar dapat digunakan untuk menguji tingkat akurasi antar metode.

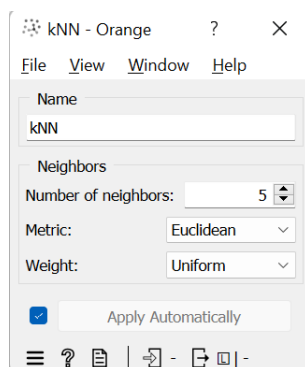
4.2. Pengujian Data K-Nearest Neighbour

Pada pengujian metode *k-nearest neighbour* secara sistem yang dilakukan dengan cara menghubungkan dari data latih ke *widget* metode *KNN* seperti pada gambar 12 berikut.



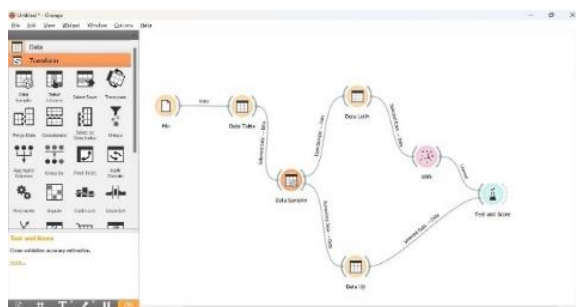
Gambar 12. Penghubung ke metode KKN

Setelah data latih dihubungkan ke metode maka data akan diproses dengan perhitungan metode yang dihubungkan secara sistematis, ditunjukkan pada gambar 13 berikut.



Gambar 13. Tampilan *widjet* KKN

Pada gambar 13 menghubungkan kedua *widget* yaitu *widget* data uji dan *widget* metode *KNN* dengan menambahkan *widget* baru yaitu *test and score* untuk melihat hasil perhitungan keakurasian metode yang dihubungkan, seperti pada gambar 14 berikut.



Gambar 14. Visualisasi *test and score KKN*

Setelah itu perhitungan keakurasian metode knn dapat dilihat dengan cara double klik pada *widget test and score* seperti pada gambar 15 berikut.

Test and Score - Orange

Cross validation

Number of folds: 5

☒ Stratified

☐ Cross validation by feature

☐ Random sampling

Repeat train/test: 10

Training set size: 66 %

☒ Stratified

☐ Leave one out

☐ Test on train data

☐ Test on test data

Evaluation results for target (None, show average over classes)

Model	AUC	CA
kNN	0.	0.
	568	611

Compare models by: Area under ROC curve

☐ Negligible diff.

0.1

kNN	
-----	--

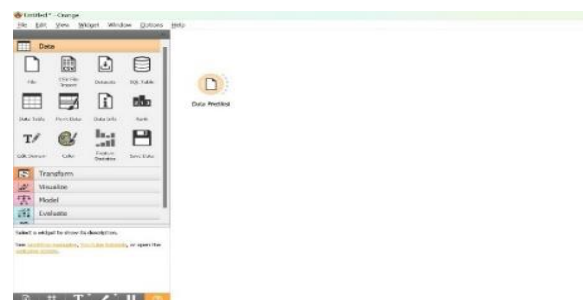
Table shows probabilities that the score for the model in the row is higher than that of the model in the columns. Small numbers show the probability that the difference is negligible.

Gambar 15. Hasil akurasi perhitungan *KNN*

Seperti pada gambar 15 maka keakurasian perhitungan dapat dilihat dari angka pada CA yaitu 0,611 atau 61% dalam persentase.

4.3. Prediksi

Pada tahap prediksi ini merupakan langkah penjelasan untuk dapat mengetahui perkiraan hasil penjualan kain lapis pada penjualan berikutnya. Berikut proses tata caranya. Menambahkan data prediksi ke data *file*



Gambar 16. Data prediksi

Menambahkan *widget file* pada halaman kerja yang berguna untuk menampung data prediksi yang berisi data 1 bulan, data prediksi tersebut hanya berisi variabel hari, penjualan dan jumlah terjual.

The screenshot shows the Orange3 Data Prediksi - Orange window. The File widget is configured to load 'data prediksi.xlsx' from the local file system. The file type is set to 'Automatically detect type'. The Info panel shows 60 instances, 3 features (no missing values), 1 data box (no target variable), and 0 meta attributes. The Columns panel shows a table with 3 rows and 4 columns: Name, Type, Role, and Values.

	Name	Type	Role	Values
1	HARI	numeric	feature	
2	PENJUALAN	categorical	feature	OFF-LINE, ONLINE
3	TERJUAL	numeric	feature	

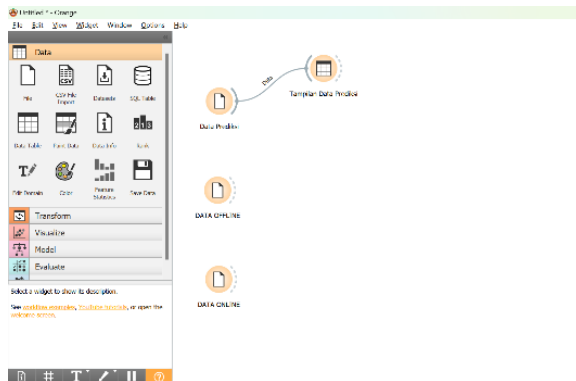
60
Gambar 17. Upload data prediksi

Setelah seperti gambar 17 maka *upload* data prediksi pada bagian *source file*. Setelah kedua *widget* telah terhubung maka akan terlihat isi data prediksi seperti gambar 18 berikut.

INSTANSI	STATUS	STATUS
1	ONLINE	10
2	ONLINE	35
3	ONLINE	40
4	ONLINE	7
5	ONLINE	9
6	ONLINE	50
7	ONLINE	23
8	ONLINE	23
9	ONLINE	75
10	ONLINE	20
11	ONLINE	30
12	ONLINE	27
13	ONLINE	50
14	ONLINE	22
15	ONLINE	75
16	ONLINE	40
17	ONLINE	50
18	ONLINE	25
19	ONLINE	40
20	ONLINE	50

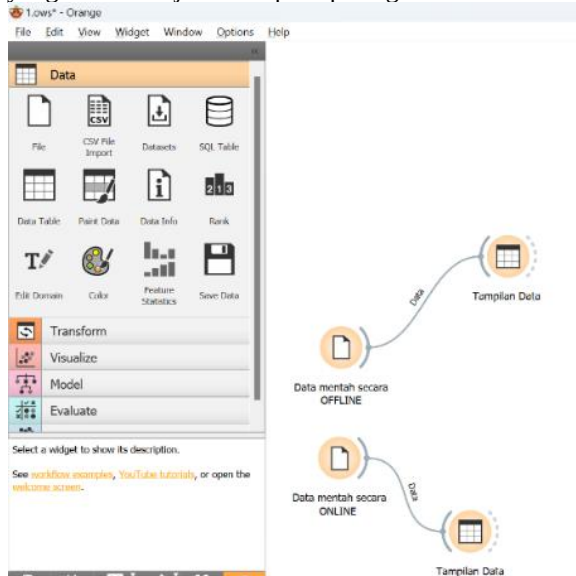
Gambar 18. Tampilan isi data prediksi

Pada tahap selanjutnya, menambahkan *widget* *file* untuk memasukkan data mentah penjualan *offline* dan *online* seperti gambar 19 berikut.



Gambar 19. Data penjualan *offline* online

Setelah memasukkan kedua data penjualan, maka pada tahap selanjutnya tambahkan *widget* data *table* yang akan ditunjukkan seperti pada gambar 20 berikut.



Gambar 20. Data table *offline* online

Pada gambar 20 menampilkan masing masing data yang dimasukkan kedalam *file* penjualan yang akan ditunjukkan pada gambar 21 sebagai berikut.

INSTANSI	STATUS	STATUS
1	ONLINE	10
2	ONLINE	35
3	ONLINE	40
4	ONLINE	7
5	ONLINE	9
6	ONLINE	50
7	ONLINE	23
8	ONLINE	23
9	ONLINE	75
10	ONLINE	20
11	ONLINE	30
12	ONLINE	27
13	ONLINE	50
14	ONLINE	22
15	ONLINE	75
16	ONLINE	40
17	ONLINE	50
18	ONLINE	25
19	ONLINE	40
20	ONLINE	50

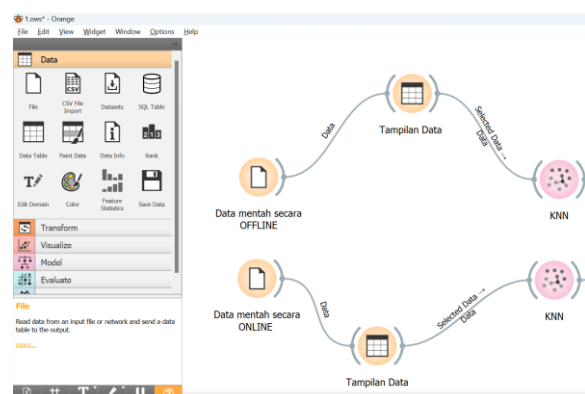
Gambar 21. Tampilan data penjualan secara *offline*

Pada gambar 21 data yang berisi penjualan kain lapis 1 dan kain lapis 2 sebanyak 180 data. Untuk data penjualan secara *online* akan ditampilkan pada gambar 22 sebagai berikut.

INSTANSI	STATUS	STATUS
1	ONLINE	10
2	ONLINE	35
3	ONLINE	40
4	ONLINE	7
5	ONLINE	9
6	ONLINE	50
7	ONLINE	23
8	ONLINE	23
9	ONLINE	75
10	ONLINE	20
11	ONLINE	30
12	ONLINE	27
13	ONLINE	50
14	ONLINE	22
15	ONLINE	75
16	ONLINE	40
17	ONLINE	50
18	ONLINE	25
19	ONLINE	40
20	ONLINE	50

Gambar 22. Tampilan data penjualan secara *online*

Setelah dipastikan bahwa data yang di input dalam *widget* *file* sudah benar, data yang berisi penjualan kain lapis 1 dan kain lapis 2 sebanyak 180 data. Maka tahap selanjutnya adalah dengan menambahkan *widget* metode *KNN* di *layout model* untuk penggunaan metode dalam melakukan perhitungan prediksi sebagai berikut.

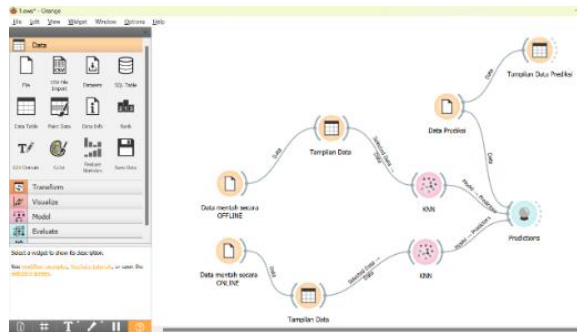


Gambar 23. Menghubungkan data ke metode *KNN*

Setelah terhubungnya data penjualan, maka tahap selanjutnya dapat melakukan prediksi yang ada pada *layout evaluate* dengan *widget predictions*.

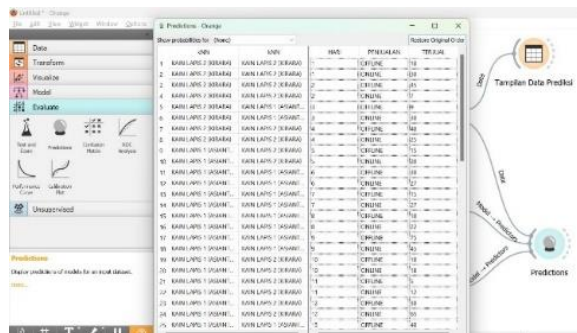
Setelah itu akan memprediksi hasil penjualan untuk beberapa hari kedepan dengan cara

menghubungkan ketiga widget yang kemudian dihubungkan ke widget *predictions* seperti gambar 24 berikut.



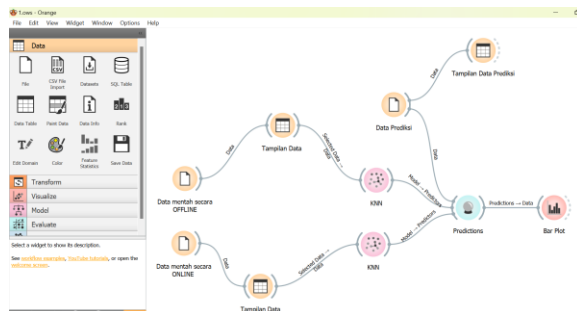
Gambar 24. Tampilan widget prediksi

Setelah *widget predictions* menampilkan hasil dari data prediksi pada gambar 24 dengan cara *double klik* pada widget prediksi maka akan muncul tampilan seperti gambar 25 berikut.



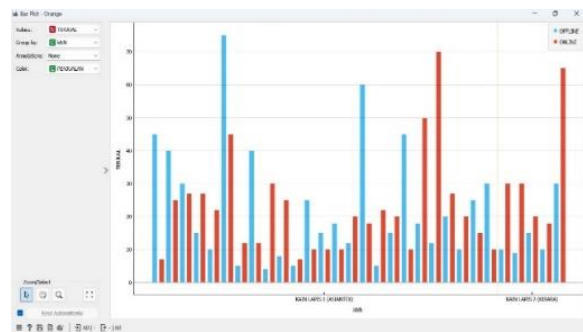
Gambar 25. Hasil prediksi kategori kain lapis

Pada gambar 25 menunjukan hasil prediksi penjualan sebuah kain lapis untuk beberapa hari kedepan, yaitu dengan perbandingan penjualan secara *offline* hari ke 3 untuk data yang di uji yaitu kain lapis 2 dan kain lapis 1 total nya terjual 9 kain. Untuk penjualan secara *online* hari ke 3 dengan data yang di uji yaitu kain lapis 1 dan kain lapis 2 total nya terjual 30 kain, untuk menunjukkan visualisasi data atau analisis pasarnya membutuhkan widget *barplot* yang berguna untuk menampilkan secara jelas perbandingan penjualan antara kain kapis 1 dan kain lapis 2 berdasarkan hasil prediksi, dengan menghubungkan widget prediksi ke widget *bar plot* yang terdapat pada layout *visualize* seperti pada gambar 26 berikut.



Gambar 26. Tampilan widget *barplot*

Untuk melihat pola visualisasi *barplot* prediksi penjualan kain lapis yaitu dengan cara *double klik* pada widget *barplot* tersebut seperti pada gambar 27 berikut.



Gambar 27. Visualisasi *barplot* prediksi penjualan

Pada gambar 27 diatas merupakan penjelasan atau prediksi yang akan banyak terjual untuk kedepannya berdasarkan dari data-data penjualan sebelumnya yang membuat pihak perusahaan mengetahui bahwa kategori kain lapis 1 yang akan banyak dan dominan terjual daripada kain lapis 2 maka dari antara kedua persaingan kain lapis tersebut perusahaan dapat mengetahui untuk mempersiapkan stok selama penjualan kedepan.

5. KESIMPULAN DAN SARAN

Pada penelitian ini penggunaan metode dalam penggunaan pengolahan data (*data mining*) yaitu *KNN* (*K-Nearest Neighbors*) dengan tahap penerapan dengan cara memproses data penjualan sebelumnya dan data prediksi yang dibuat untuk memprediksi data penjualan berikutnya agar dapat dilakukan dengan perhitungan *K-Nearest Neighbour*, yang memiliki tingkat akurasi sebesar 0,611 atau 61% sebagai tingkat persentase yang berarti metode ini berada diatas 0-0,5 memiliki tingkat akurasi yang baik untuk dipergunakan. Pada saat menerapkan data *mining* di aplikasi Orange yaitu harus melewati beberapa tahap dengan cara mempersiapkan data set berupa data penjualan dan data prediksi. Pengujian ini menghasilkan prediksi penjualan beberapa hari kedepan, dimana hasil dari penelitian ini dapat mengetahui penjualan kategori kain lapis 1 yang lebih dominan terjual dari kategori kain lapis 2 sehingga dapat mengetahui minat pasar dan penyediaan stok penjualan. Penelitian ini bermanfaat terhadap kedua pihak yaitu pihak perusahaan yang dapat mengetahui prediksi penjualan selanjutnya dan juga pihak penulis yang dalam berupaya menyelesaikan tugas akhir skripsi yang berjudul "Data Mining metode K-Nearest Neighbour Untuk Regresi Data Penjualan Kain Tekstil". Untuk penelitian selanjutnya agar dapat membandingkan dengan metode prediksi yang lain supaya dapat mengetahui perbedaan tingkat akurasi antar metode dan cara terbaru yang dapat lebih mudah digunakan berdasarkan update sistem ataupun aplikasi lainnya.

DAFTAR PUSTAKA

- [1] W. F and M. W, "Application of Data Mining to Predict Sales of Woven Fabrics Using Linear Regression Case Study," *J. Ilm. Tek. Mesin, Elektro, Dan Komput.*, pp. 27–39, 2022.
- [2] Yuda Irawan, "Penerapan Data Mining Untuk Evaluasi Data Penjualan Menggunakan Metode Clustering Dan Algoritma Hirarki Divisive Di Perusahaan Media World Pekanbaru," *J. Teknol. Inf. Univ. Lambung Mangkurat*, vol. 4, no. 1, pp. 13–20, 2019, doi: 10.20527/jtiulm.v4i1.34.
- [3] A. A. Putri, "Penerapan Data Mining Untuk Memprediksi Penjualan Buah Dan Sayur Menggunakan Metode K-Nearest Neighbor (Studi Kasus : PT . Central Brastagi Utama)," vol. 1, no. 6, pp. 354–361, 2021.
- [4] A. Alfani W.P.R., F. Rozi, and F. Sukmana, "Prediksi Penjualan Produk Unilever Menggunakan Metode K-Nearest Neighbor," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 6, no. 1, pp. 155–160, 2021, doi: 10.29100/jupi.v6i1.1910.
- [5] S. S. Lubis and B. Hendrik, "Implementasi Data Mining Pengelompokan Data Penjualan Berdasarkan Pembelian Dengan Menggunakan Algoritma K-Means Pada UD.Martua," *J. Inf. Sysem Educ. Dev.*, vol. 1, no. 3, pp. 36–41, 2023, [Online]. Available: <https://journal.widyakarya.ac.id/index.php/jusiik-widyakarya/article/view/1531%0Ahttps://journal.widyakarya.ac.id/index.php/jusiik-widyakarya/article/download/1531/1563>
- [6] N. Almumtazah, N. Azizah, Y. L. Putri, and D. C. R. Novitasari, "Prediksi Jumlah Mahasiswa Baru Menggunakan Metode Regresi Linier Sederhana," *J. Ilm. Mat. Dan Terap.*, vol. 18, no. 1, pp. 31–40, 2021, doi: 10.22487/2540766x.2021.v18.i1.15465.
- [7] H. P. Herlambang, F. Saputra, M. H. Prasetyo, D. Puspitasari, and D. Nurlaela, "Perbandingan Klasifikasi Tingkat Penjualan Buah di Supermarket dengan Pendekatan Algoritma Decision Tree, Naive Bayes dan K-Nearest Neighbor," *J. Insa. - J. Inf. Syst. Manag. Innov.*, vol. 3, no. 1, pp. 21–28, 2023, doi: 10.31294/jinsan.v3i1.2097.
- [8] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono, and C. Lauw, "Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Linear Regression," *J. Nas. Teknol. dan Sist. Inf.*, vol. 7, no. 1, pp. 8–17, 2021, doi: 10.25077/teknosi.v7i1.2021.8-17.
- [9] A. F. Boy, "Implementasi data mining dalam memprediksi harga Crude Palm Oil (CPO) pasar domestik menggunakan algoritma regresi linier berganda (Studi kasus dinas perkebunan provinsi Sumatera Utara)," *J. Sci. Soc. Res.*, vol. 4307, no. 2, pp. 78–85, 2020.
- [10] A. D. Alfarizi and A. Andri, "Pemanfaatan Data Mining Dalam Memprediksi Produksi Pada PT Pupuk Sriwidjaja Palembang Menggunakan Algoritma Regresi Linier Berganda," *J. Nas. Ilmu Komput.*, vol. 2, no. 1, pp. 51–63, 2021, doi: 10.47747/jurnalnuk.v2i1.522.
- [11] S. Cumel, David Zamri, Rahmaddeni, "Perbandingan Metode Data Mining untuk Prediksi Banjir dengan Algoritma Naïve Bayes dan KNN," *SENTIMAS Semin. Nas. Penelit. dan ...*, pp. 40–48, 2022, [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas/article/view/353%0Ahttps://journal.irpi.or.id/index.php/sentimas/article/download/353/132>
- [12] S. Mulyati, S. M. Husein, and Ramdhan, "Rancang Bangun Aplikasi Data Mining Prediksi Kelulusan Ujian Nasional Menggunakan Algoritma KNN dan Metode Euclidean Distance," *J. Tek. Inform. Univ. Muhammadiyah Tangerang*, vol. 4, no. 1, pp. 65–73, 2020.
- [13] W. Irmayani, "Visualisasi Data Pada Data Mining Menggunakan Metode Klasifikasi," *J. Khatulistiwa Inform.*, vol. IX, no. I, pp. 68–72, 2021.