

**PENERAPAN *DATA MINING* UNTUK MEMPREDIKSI  
JUMLAH PRODUKSI MENGGUNAKAN METODE  
ALGORITMA *K-NEAREST NEIGHBOR***

**SKRIPSI**



**Oleh:  
Wahyudi  
190210115**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
2023**

**PENERAPAN *DATA MINING* UNTUK MEMPREDIKSI  
JUMLAH PRODUKSI MENGGUNAKAN METODE  
ALGORITMA K-NEAREST**

**SKRIPSI**

**Untuk memenuhi salah satu syarat  
Guna memperoleh gelar sarjana**



**Oleh:  
Wahyudi  
190210115**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
2023**

## SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini saya:

Nama : Wahyudi

NPM : 190210115

Fakultas : Teknik dan Komputer

Program Studi : Teknik Informatika

Menyatakan bahwa "Skripsi" yang penulis buat dengan judul:

**PENERAPAN *DATA MINING* UNTUK MEMREDIKSI JUMLAH PROUKSI  
MENGUNAKAN METODE ALGORITMA K-NEAREST**

Adalah karya sendiri dan bukan "duplikasi" dari karya orang lain. Sepengetahuan saya, didalam naskah skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah di tulis atau di terbitkan oleh orang lain, kecuali yang secara tertulis di kutip di dalam naskah ini dan di sebutkan dalam sumber kutipan dan daftar pustaka. Apabila ternyata di dalam naskah skripsi ini dapat di buktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah skripsi ini digugurkan dan gelar akademik yang saya peroleh di batalkan, serta di proses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun.

Batam, 30 Januari 2023



**Wahyudi**  
**190210115**

**PENERAPAN *DATA MINING* UNTUK MEMPREDIKSI  
JUMLAH PROUKSI MENGGUNAKAN METODE  
ALGORITMA K-NEAREST**

**SKRIPSI**

**Untuk memenuhi salah satu syarat  
Guna memperoleh gelar sarjana**

**Oleh:  
Wahyudi  
190210115**

**Telah disetujui oleh pembimbing pada tanggal  
Seperti tertera dibawah ini  
Batam, 30 Januari 2023**



**Sunarsan Sitohang, S.Kom.,M.TI  
Pembimbing**

## ABSTRAK

Penjualan merupakan faktor utama bagi setiap perusahaan karena dengan adanya penjualan maka suatu perusahaan akan memperoleh keuntungan yang lebih supaya usaha yang di lakukan dapat berjalan dan dapat berkembang. Namun bisa dilihat bahwa persaingan bisnis di era jaman modren sekarang begitu ketat, setiap perusahaan diwajibkan untuk membuat suatu persiapan diri secara profesional dan fleksibel sehingga perusahaan mampu bertahan dan mampu bertumbuh. TAKAdeli menyediakan beragam varian dan rasa terdiri dari brownis, bolu, *cake* lapis dan lainnya. Dalam pemberian informasi mengenai minat konsumen laris, sedang dan tidak laris suatu produk yang di produksi mempengaruhi permintaan konsumen sehingga sering kali perusahaan mengalami kesulitan dalam menentukan produk mana yang laris, sedang dan tidak laris agar dapat di peroduksi lebih banyak. Pada saat ini data penjualan dan produksi yang di peroleh oleh TAKAdeli *Cake Boutique* masih di *input* secara manual kedalam sistem, dengan demikian dibutuhkan suatu prediksi yang tepat untuk memproduksi jumlah produk yang diminati oleh pembeli atau konsumen. Namun dalam menentukan prediksi dibutuhkan sebuah data mining. *Data mining* merupakan proses menemukan sebuah hubungan dengan data yang tidak diketahui oleh pengguna dan menyajikan data dengan cara yang dapat dipahami sehingga hubungan tersebut dapat menjadi dasar pengambilan keputusan. Akan tetapi dalam menentukan produk laris dibutuhkan sebuah algoritma *K-Nearest Neighbor* agar mempermudah perusahaan mengetahui produk – produk mana saja dan varian apa saja yang diminati oleh konsumen menggunakan *microsoft excel* dan *software rapid minner* yang menghasilkan nilai peresentasi nilai akurasi 83,33%..

Kata Kunci: Algoritma *K-Nearest Naighbour*; Data Mining; Prediksi Produksi.

## **ABSTRACT**

*Sales are the main factor for every company because with sales, a company will get more profits so that the business carried out can run and develop. However, it can be seen that business competition in the modern era is so tight, every company is required to make professional and flexible self-preparations so that the company can survive and grow. TAKAdeli provides various variants and flavors consisting of brownies, sponge cakes, layer cakes and others. In providing information about consumer interests that are selling, being and not selling, a product that is produced influences consumer demand so that companies often experience difficulties in determining which products are selling, being and not selling so that they can produce more. At this time the sales and production data obtained by TAKAdeli Cake Boutique are still manually inputted into the system, thus an accurate prediction is needed to produce the number of products that are in demand by buyers or consumers. But in making predictions, data mining is needed. Data mining is the process of finding a relationship with data that is unknown to the user and presenting the data in an understandable way so that the relationship can form the basis of decision making. However, in determining best-selling products, a K-Nearest Neighbor algorithm is needed to make it easier for companies to find out which products and which variants consumers are interested in using Microsoft Excel and Rapid Minner software which produce an accuracy value of 83.33%.*

*Keywords: Data Mining; K-Nearest Naighbour Agorithm; Production Prediction.*

## KATA PENGANTAR

Puji syukur atas berkat rahmat Tuhan yang maha kuasa yang telah melimpahkan segala rahmat dan kuasaNya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Teknik Informatika Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati penulis menyakpaikan ucapan terimakasih kepada:

1. Rektor Universitas Putera Batam
2. Dekan fakultas Teknik Informatika
3. Ketua Program Studi Teknik Informatika
4. Bapak Sunarsan Sitohang, S.Kom.,M.MT Selaku pembimbing Skripsi pada Program Studi Teknik Informatika Universitas Putera Batam
5. Bapak Rahmat fauzi, S.Kom., M.Kom selaku pembimbing akademik Program Studi Teknik Informatika Universitas Putera Batam
6. Dosen dan Staff Universitas Putera Batam
7. Bapak dan Ibu selaku kedua orang tua penulis yang selalu mendoakan dan menyemangati penulis hingga penulisan skripsi ini selesai
8. Keluarga penulis yang selalu mendoakan dan memberikan motivasi kepada penulis agar penelitian ini selesai tepat waktu
9. Teman-teman seperjuangan yang bersedia membagi ilmunya dan sharing pendapat dalam rangka pembuatan skripsi ini
10. Semua pihak yang telah bersedia meluangkan waktu, tenaga dan pikiranya dalam memberikan data/informasi selama penulis membuat skripsi yang tidak dapat penulis sebutkan satu-persatu.

Semoga Tuhan Yang Maha Esa membalas kebaikan dan selalu mencurahkan berkatNya,Amin.

Batam, 30 Januari 2023



Wahyudi

## DAFTAR ISI

Halaman

<b>HALAMAN SAMPUL</b> .....	<b>i</b>
<b>HALAMAN JUDUL</b> .....	<b>ii</b>
<b>SURAT PERNYATAAN ORISINALITAS</b> .....	<b>iii</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>iv</b>
<b>ABSTRAK</b> .....	<b>v</b>
<b>ABSTRACT</b> .....	<b>vi</b>
<b>KATA PENGANTAR</b> .....	Error! Bookmark not defined.
<b>DAFTAR ISI</b> .....	<b>viii</b>
<b>DAFTAR GAMBAR</b> .....	<b>x</b>
<b>DAFTAR TABEL</b> .....	<b>xi</b>
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah.....	4
1.3 Batasan Masalah .....	4
1.4 Rumusan Masalah.....	5
1.5 Tujuan Penelitian .....	5
1.6 Manfaat Penelitian .....	6
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 <i>Knowladge Discoveryin Database (KDD)</i> .....	7
2.1.1 Tahapan proses KDD .....	7
2.2 Data Mining .....	9
2.2.1 Pengertian Data Mining .....	9
2.2.2 Teknik Data Mining .....	9
2.2.3 Keuntungan dan Kelemahan Data mining .....	10
2.2.4 Klasifikasi Data Mining .....	11
2.2.5 Pengelompokan Data Mining.....	11
2.2.6 Operasi Data Mining .....	13
2.3 Produksi .....	13
2.3.1 Faktor-faktor yang mempengaruhi Produksi .....	13
2.3.2 Manfaat Produksi .....	14
2.3.3 Penjualan .....	14
2.4 Prediksi / <i>Forcasting</i> .....	15
2.5 Algoritma <i>K-Nearest Naighbor</i> .....	15
2.5.1 Jarak Euclidean .....	16
2.5.2 Kelebihan dan Kekurangan Algoritma <i>K-Nearest Neighbour</i> .....	17
2.5 <i>Software</i> Pendukung .....	17
2.5.1 <i>Rapid miner</i> .....	18
2.5.2 <i>Microsoft excel</i> .....	19
2.6 Penelitian Terdahulu .....	20
2.7 Kerangka Pemikiran .....	22
<b>BAB III METODE PENELITIAN</b>	



3.1	Desain Penelitian .....	23
3.2	Teknik Pengumpulan Data .....	25
3.3	Operasional Variabel .....	25
3.4	Metode Perancangan Sistem.....	26
3.5	Lokasi dan Jadwal Penelitian.....	26
3.5.1	Lokasi Penelitian.....	26
3.5.2	Jadwal Penelitian.....	27
<b>BAB IV HASIL DAN PEMBAHASAN</b>		
4.1	Hasil Penelitian.....	28
4.1.1	Seleksi Data .....	32
4.1.2	<i>Preprocessing</i> .....	32
4.1.3	<i>Transformation</i> .....	35
4.1.4	<i>K-Nearest Neighbour</i> .....	37
4.2	Hasil Pengujian.....	37
4.3	Implementasi <i>Rapid Miner</i> .....	39
<b>BAB V KESIMPULAN DAN SARAN</b>		
<b>DAFTAR PUSTAKA</b>		
<b>LAMPIRAN</b>		
1.	Pendukung penelitian	
2.	Daftar Riwayat Hidup	
3.	Surat Keterangan Penelitian	

## DAFTAR GAMBAR

<b>Gambar 2. 1</b> Tahapan proses KDD.....	8
<b>Gambar 2. 2</b> Rapid miner .....	19
<b>Gambar 2. 3</b> Kerangka Pemikiran .....	22
<b>Gambar 3. 1</b> Desain Penelitian.....	23
<b>Gambar 3. 2</b> Lokasi Penelitian .....	26
<b>Gambar 4. 1</b> Menghubungkan operator.....	40
<b>Gambar 4. 2</b> Import data.....	41
<b>Gambar 4. 3</b> Step 3 .....	42
<b>Gambar 4. 4</b> Step 4 .....	42
<b>Gambar 4. 5</b> Operator set rule .....	43
<b>Gambar 4. 6</b> Icon run.....	43
<b>Gambar 4. 7</b> Perhitungan Hasil.....	44
<b>Gambar 4. 9</b> Perhitungan presentasi .....	45

## DAFTAR TABEL

<b>Tabel 3. 1</b> Jadwal Penelitian.....	27
<b>Tabel 4. 1</b> Data Set tahun 2020.....	29
<b>Tabel 4. 2</b> Data Set tahun 2021 .....	30
<b>Tabel 4. 3</b> Data Set tahun 2022 .....	31
<b>Tabel 4. 4</b> Seleksi data.....	32
<b>Tabel 4. 5</b> Sample data penjualan brownis .....	33
<b>Tabel 4. 6</b> Sample data penjualan brownis .....	33
<b>Tabel 4. 7</b> Sample data penjualan brownis .....	33
<b>Tabel 4. 8</b> Sample data penjualan chiffon .....	34
<b>Tabel 4. 9</b> Sample data penjualan Chiffon .....	34
<b>Tabel 4. 10</b> Sample data penjualan Chiffon .....	35
<b>Tabel 4. 11</b> Data training produksi kukus coklat cheese .....	35
<b>Tabel 4. 12</b> Data testing/prediksi kukus coklat cheese.....	36
<b>Tabel 4. 13</b> Euclidean Distance .....	38