

**PENERAPAN K-NEAREST NEIGHBOR REGRESSION DALAM
MEMPREDIKSI PENJUALAN PRODUK DI TOKO SHAN BEKASI**

SKRIPSI

Oleh:

Andre Wahyu Saputra

201710225230



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2021**

**PENERAPAN K-NEAREST NEIGHBOR REGRESSION DALAM MEMPREDIKSI
PENJUALAN PRODUK DI TOKO SHAN BEKASI**

SKRIPSI

Oleh:

Andre Wahyu Saputra

201710225230



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2021**



LEMBAR PERSETUJUAN PEMBIMBING

Judul Proposal Skripsi : Penerapan *K-Nearest Neighbor Regression* Dalam
Memprediksi Penjualan Produk Di Toko Shan
Bekasi

Nama Mahasiswa : Andre Wahyu Saputra

Nomor Pokok Mahasiswa : 201710225230

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 16 Juli 2021



NIDN. 0310038006

NIDN. 013077002

LEMBAR PENGESAHAN

Judul Skripsi : Penerapan *K-Nearest Neighbor Regression* Dalam
Memprediksi Penjualan Produk Di Toko Shan
Bekasi

Nama Mahasiswa : Andre Wahyu Saputra

Nomor Pokok Mahasiswa : 201710225230

Program Studi / Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 16 Juli 2021

Bekasi, 22 Juli 2021

Mengesahkan,

Ketua Tim Penguji : Kusdarnowo Hantoro, S.Kom., M.Kom
NIDN. 0329076601

Penguji (I) : Allan Desi Alexander, S.T., M.Kom
NIDN. 0305127404

Penguji (II) : Ratna Salkiawati, S.T., M.Kom
NIDN. 0310038006

Mengetahui,

Ketua Program Studi
Informatika

Dekan
Fakultas Ilmu Komputer

Rakhmat Purnomo, S.Pd., S.Kom., M.Kom.
NIDN. 0322108201

Herlawati, S.Si., M.M., M.Kom.
NIDN. 0311097302

LEMBAR PERNYATAAN BUKAN PLAGIASI



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Andre Wahyu Saputra
NPM : 201710225230
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Penerapan *K-Nearest Neighbor Regression* Dalam
Memprediksi Penjualan Produk Di Toko Shan Bekasi

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Bekasi, 16 Juni 2021
Penulis


10000
METRAI
TEMPER
D430CAJX262778178
Andre Wahyu Saputra

ABSTRAK

Andre Wahyu Saputra. 201710225230. Penerapan *K-Nearest Neighbor Regression* Dalam Memprediksi Penjualan Produk di Toko Shan Bekasi.

Toko Shan Bekasi merupakan toko yang menjual produk peralatan rumah tangga. Permasalahan di Toko Shan Bekasi adalah belum adanya prediksi penjualan akan membuat pemilik toko kesulitan dalam menentukan strategi penjualan maupun dalam penyediaan stok produk barang di masa yang akan datang. Untuk mengetahui prediksi penjualan produk terlaris menggunakan teknik *K-Nearest Neighbor Regression* dengan metode perancangan yang diterapkan pada penelitian ini adalah *Knowledge Discovery in Database (KDD)*. Metode tersebut akan melewati 5 tahap yaitu *data selection, preprocessing, transformation, data mining, interpretation/evaluation*. Hasil penelitian ini adalah nilai $k = 2$ menghasilkan nilai RMSE 0.35810 untuk produk Blender di bulan April, $k = 5$ menghasilkan nilai RMSE 0.48590 untuk produk Mixer di bulan November, $k = 14$ menghasilkan nilai RMSE 0.43311 untuk produk kompor di bulan Januari, $k = 9$ menghasilkan nilai RMSE 0.41592 untuk produk cukur rambut di bulan Desember, $k = 13$ menghasilkan nilai RMSE 0.43405 untuk produk termos pencet di bulan November. Dengan demikian berdasarkan pedoman interpretasi RMSE dapat disimpulkan bahwa seluruh model yang penulis uji memiliki tingkat kesalahan sedang yaitu 0,30 – 0,599.

Kata Kunci: *K-Nearest Neighbor Regression*, Prediksi, RMSE, *Knowledge Discovery in Database*

ABSTRACT

Andre Wahyu Saputra. 201710225230. Application of K-Nearest Neighbor Regression in Predicting Product Sales at Shan S Bekasi.

Toko Shan Bekasi is a store that sells home appliances products. The problem at Toko Shan Bekasi is that the absence of sales predictions will make it difficult for shop owners to determine sales strategies and in providing stock of goods in the future. To find out the prediction of best-selling product sales, the K-Nearest Neighbor Regression technique is used. The design method applied in this research is Knowledge Discovery in Database (KDD). The method will pass through 5 stages, namely data selection, preprocessing, transformation, data mining, interpretation/evaluation. The results of this study are the value of $k = 2$ produces an RMSE value of 0.35810 for Blender products in April, $k = 5$ produces an RMSE value of 0.48590 for Mixer products in November, $k = 14$ produces an RMSE value of 0.43311 for stove products in January, $k = 9$ yielding RMSE value of 0.41592 for shaving products in December, $k = 13$ resulting in RMSE value of 0.43405 for push flask products in November.

Thus, based on the RMSE interpretation guidelines, it can be said that all the models tested by the author have a moderate error rate of 0.30 – 0.599.

Keywords: K-Nearest Neighbor Regression, prediksi, RMSE, K Knowledge Discovery in Database

LEMBAR PERNYATAAN PUBLIKASI

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Andre Wahyu Saputra
NPM : 201710225230
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (Non-Exclusive Royalty-Free Right)**, atas karya ilmiah saya yang berjudul :

Penerapan *K-Nearest Neighbor Regression* Dalam Memprediksi Penjualan Produk Di Toko Shan Bekasi

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi
Pada tanggal : 16 Juni 2021
Yang Menyatakan



Andre Wahyu Saputra

KATA PENGANTAR

Alhamdulillah penulis panjatkan puji dan syukur kepada Allah SWT yang maha pengasih dan maha penyayang sehingga penulis dapat menyelesaikan laporan skripsi ini dengan judul “Penerapan K-Nearest Neighbor Regression Dalam Memprediksi Penjualan Produk di Toko Shan Bekasi”. Yang disusun sebagai syarat untuk mencapai Sarjan S1 Program Studi Informatika Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya.

Penulis mengucapkan terima kasih dan penulis sampaikan kepada pihak yang telah membantu dalam menyelesaikan laporan skripsi ini, terutama kepada:

1. Bapak Irjen Pol. (Purn) Dr. H. Bambang Karsono, Drs., S.H., M.M. Selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Herlawati, S. Si, M. M, M. Kom Selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Rakhmat Purnomo, S. Pd, S. Kom, M. Kom Selaku Kepala Program Studi Informatika Universitas Bhayangkara Jakarta Raya.
4. Ibu Ratna Salkiawati, ST., M.Kom Selaku Dosen Pembimbing Akademik Program Studi Informatika Universitas Bhayangkara Jakarta Raya.
5. Bapak Hendarman Lubis, S. Kom., M. Kom Selaku Dosen Pembimbing Program Studi Informatika Universitas Bhayangkara Jakarta Raya.
6. Kepada kedua orang tua saya Bapak Rudi Hernandi dan Eti Wahyuni yang selalu memberikan dukungan, semangat, dan doa yang tulus.
7. Kepada adik saya Rendi Ardian Saputra yang selalu memberikan semangat untuk tiap harinya.
8. Kepada Ibu Anissa Putri Aprilia Selaku Pemilik Toko Shan Bekasi.
9. Kepada rekan-rekan semuanya yang turut membantu dalam penulisan skripsi ini.
10. Serta seluruh pihak yang membantu penulis selama masa Studi di Universitas Bhayangkara Jakarta Raya yang tidak penulis sebutkan satu per satu.

Penulis juga menyadari bahwa di dalam penyusunan terdapat banyak kekurangan dan kesalahan. Oleh karena itu, penulis mengharapkan kritik dan saran yang bersifat membangun, sehingga selanjutnya dapat menjadi lebih baik di masa yang akan datang.

Akhir kata semoga skripsi ini dapat memberi manfaat bagi pembaca umumnya dan bagi penulis pada khususnya. Terima kasih.

Bekasi, 19 Juni 2021



Andre Wahyu Saputra



DAFTAR ISI

	Halaman
LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERNYATAAN PUBLIKASI	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiv
DAFTAR GAMBAR.....	xvi
DAFTAR RUMUS	xviii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	5
1.4 Batasan Masalah.....	5
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	5
1.7 Tempat dan Waktu Penelitian	6
1.8 Metode Penelitian.....	7
1.8.1 Metode Pengumpulan Data	7
1.8.2 Metode Analisis	7
1.8.3 Metode Perancangan	7
1.8.4 Metode Pengujian.....	8
1.9 Sistematika Penulisan.....	8
BAB II LANDASAN TEORI	10
2.1 Tinjauan Pustaka	10
2.2 Penjualan dan Prediksi (<i>Forecasting</i>)	12

2.2.1	Pengertian Penjualan	12
2.2.2	Pengertian Prediksi (<i>forecasting</i>)	13
2.3	<i>K-Nearest Neighbor</i>	13
2.3.1	Pengertian <i>K-Nearest Neighbor</i>	13
2.3.2	Kelebihan dan Kekurangan <i>K-Nearest Neighbor</i>	15
2.3.2.1	Kelebihan <i>K-Nearest Neighbor</i>	15
2.3.2.2	Kekurangan <i>K-Nearest Neighbor</i>	15
2.3.3	Klasifikasi.....	15
2.3.4	<i>Regression</i>	16
2.4	<i>Data Mining</i>	16
2.4.1	Pengertian <i>Data Mining</i>	16
2.4.2	Pengelompokkan <i>Data Mining</i>	18
2.5	<i>Knowledge Discovery in Database</i>	18
2.5.1	Pengertian <i>Knowledge Discovery in Database</i>	18
2.5.2	Tahapan KDD.....	19
2.6	Min-Max Normalisasi	21
2.7	<i>K-Fold Cross Validation</i>	22
2.8	<i>Root Mean Square Error (RMSE)</i>	23
2.9	<i>Python</i>	25
2.10	Google Colab.....	25
2.11	<i>Flowchart</i>	25
2.11.1	Pengertian <i>Flowchart</i>	25
2.11.2	Simbol <i>Flowchart</i>	26
2.12	<i>Unified Modeling Language (UML)</i>	27
2.12.1	Pengertian <i>Unified Modeling Language (UML)</i>	27
2.12.2	Jenis-Jenis Diagram <i>Unified Modeling Language (UML)</i>	27
BAB III METODOLOGI PENELITIAN		35
3.1	Objek Penelitian	35
3.1.1	Sejarah Toko Shan Bekasi	35
3.1.2	Profil Toko Shan Bekasi	35
3.1.3	Struktur Organisasi Toko Shan Bekasi	36
3.2	Wawancara dan Hasil.....	37

3.2.1	Identitas Narasumber	37
3.2.2	Pertanyaan untuk Narasumber	37
3.2.3	Jawaban Narasumber	37
3.3	Kerangka Penelitian	38
3.3.1	Pengumpulan Data	39
3.3.2	Metode Perancangan	40
3.4	Analisis Sistem Berjalan	41
3.4.1	Sistem Penjualan	41
3.4.2	Sistem Produksi.....	42
3.5	Permasalahan.....	43
3.6	Analisis Usulan Sistem	43
3.6.1	Desain Usulan Sistem	44
3.7	Pengolahan Data.....	45
3.8	Analisis Kebutuhan Sistem	46
BAB IV PERANCANGAN SISTEM DAN IMPLEMENTASI		48
4.1	Perancangan Sistem	48
4.1.1	Pengolahan Data	48
4.1.2	Pembentukan <i>Dataset</i>	50
4.2	Implementasi	58
4.2.1	Normalisasi <i>Dataset</i>	58
4.2.2	Implementasi <i>K-Fold Cross Validation</i>	61
4.2.3	Pembagian <i>Data Training</i> dan <i>Testing</i>	61
4.2.4	Implementasi <i>K-Nearest Neighbor</i>	62
4.3	Proses Pengujian	63
4.3.1	Evaluasi	63
4.3.2	<i>Performance Measure</i>	65
4.3.3	Model Terbaik	66
4.3.4	Prediksi.....	67
4.3.5	Visualisasi	68
BAB V PENUTUP		71
5.1	Kesimpulan	71
5.2	Saran.....	72

DAFTAR PUSTAKA	73
LAMPIRAN.....	76
PLAGIARISME	79
BIODATA MAHASISWA	81
KARTU BIMBINGAN SKRIPSI	82



DAFTAR TABEL

	Halaman
Tabel 1.1 Volume Penjualan Toko Shan Bekasi Pada Tahun 2018 – 2020.....	2
Tabel 1.2 Perbandingan Model.....	3
Tabel 1.3 Waktu Kegiatan Penelitian.....	6
Tabel 2.1 Tinjauan Pustaka.....	10
Tabel 2.2 Pedoman Interpretasi RMSE.....	24
Tabel 2.3 Simbol-simbol pada <i>Flowchart</i>	26
Tabel 2.4 Simbol <i>Use Case Diagram</i>	28
Tabel 2.5 Simbol <i>Activity Diagram</i>	30
Tabel 2.6 Simbol <i>Sequence Diagram</i>	31
Tabel 2.7 Simbol <i>Class Diagram</i>	33
Tabel 2.8 Simbol <i>Deployment Diagram</i>	34
Tabel 3.1 Kuesioner sebelum pembuatan sistem	45
Tabel 3.2 Perangkat keras yang digunakan.....	46
Tabel 3.3 Perangkat lunak yang digunakan	47
Tabel 3.4 Kebutuhan data yang digunakan.....	47
Tabel 4.1 Data Penjualan Produk 2018.....	48
Tabel 4.2 Data Penjualan Produk 2019.....	49
Tabel 4.3 Data Penjualan Produk 2020.....	49
Tabel 4.4 Sampel Penjualan Produk Blender	49
Tabel 4.5 <i>Data Training</i> Blender	50
Tabel 4.6 <i>Data Testing</i> Blender	51
Tabel 4.7 <i>Data Training</i> Mixer.....	52
Tabel 4.8 <i>Data Testing</i> Mixer	53
Tabel 4.9 <i>Data Training</i> Kompor	53
Tabel 4.10 <i>Data Testing</i> Kompor.....	54
Tabel 4.11 <i>Data Training</i> Cukur Rambut.....	55
Tabel 4.12 <i>Data Testing</i> Cukur Rambut	56
Tabel 4.13 <i>Data Training</i> Termos Pencet.....	56
Tabel 4.14 <i>Data Testing</i> Termos Pencet	57

Tabel 4.15 Hasil Nilai RMSE terbaik Setiap Produk.....	66
--	----



DAFTAR GAMBAR

	Halaman
Gambar 1.1 Grafik Data Penjualan Barang Pada Tahun 2018-2020	2
Gambar 2.1 Proses KDD dalam Data Mining.....	19
Gambar 2.2 Diagram 10-fold Cross Validation	23
Gambar 3.1 Struktur Organisasi Toko Shan Bekasi	36
Gambar 3.2 Alur Penelitian di Toko Shan Bekasi	38
Gambar 3.3 Sistem Penjualan Berjalan.....	41
Gambar 3.4 Sistem Produksi Berjalan	42
Gambar 3.5 Desain Data Mining Usulan	44
Gambar 4.1 Script Normalisasi Training	58
Gambar 4.2 Sampel Hasil Normalisasi Data Training Blender	59
Gambar 4.3 Script Pembagian Data Input dan Target Training.....	59
Gambar 4.4 Script Normalisasi Testing	60
Gambar 4.5 Sampel Hasil Normalisasi Penjualan Blender.....	60
Gambar 4.6 Script Pembagian Data Input dan Target Training.....	60
Gambar 4.7 Script K-Fold Cross Validation	61
Gambar 4.8 Hasil Implementasi K-Fold Cross Validation.....	61
Gambar 4.9 Data Training dan Data Testing	61
Gambar 4.10 Hasil Split K-Fold Cross Validation	62
Gambar 4.11 Implementasi K-NN dan Cross Validation	63
Gambar 4.12 Script Evaluasi.....	63
Gambar 4.13 Grafik Nilai RMSE.....	64
Gambar 4.14 Nilai RMSE Per-k	65
Gambar 4.15 Hasil Performance Measure.....	65
Gambar 4.16 Model Terbaik	66
Gambar 4.17 Script Prediksi	67
Gambar 4.18 Hasil Prediksi Penjualan Blender	67
Gambar 4.19 Script Visualisasi.....	68
Gambar 4.20 Visualisasi Hasil Prediksi Penjualan Blender	68
Gambar 4.21 Visualisasi Hasil Prediksi Penjualan Mixer	69

Gambar 4.22 Visualisasi Hasil Prediksi Penjualan Kompor.....	69
Gambar 4.23 Visualisasi Hasil Prediksi Penjualan Cukur Rambut	70
Gambar 4.24 Visualisasi Hasil Prediksi Penjualan Blender	70



DAFTAR RUMUS

	Halaman
$euc = \sqrt{(a_1 - b_1)^2 + \dots + (a_n - b_n)^2}$ (2.1)	14
$euc = \sqrt{(\sum_{i=1}^n (p_i - q_i)^2)}$ (2.2)	14
Normalisasi = $(x - \min) / (\max - \min)$ (2.3)	21
$\sqrt{\frac{1}{n} \sum_i^n (\tilde{y}_i - y_i)^2}$ (2.4)	23



DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Surat Permohonan Izin Penelitian	77
Lampiran 2. Surat Konfirmasi Izin Penelitian	78
Lampiran 3. Source Code.....	79

