

**PENERAPAN ALGORITMA K-NEAREST NEIGHBOR
UNTUK MEMPREDIKSI PENJUALAN SEPEDA MOTOR
TERLARIS PADA PT DAYA ANUGRAH MANDIRI**

SKRIPSI



**Oleh:
Rozimin
160210207**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2022**

**PENERAPAN ALGORITMA K-NEAREST NEIGHBOR
UNTUK MEMPREDIKSI PENJUALAN SEPEDA MOTOR
TERLARIS PADA PT DAYA ANUGRAH MANDIRI**

SKRIPSI

**Untuk memenuhi salah satu
syarat memperoleh gelar Sarjana**



**Oleh:
Rozimin
160210207**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2022**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Rozimin
NPM : 160210207
Fakultas : Teknik dan Komputer
Program Studi : Teknik Informatika

Menyatakan bahwa "Skripsi" yang saya buat dengan judul:

PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK MEMPREDIKSI PENJUALAN SEPEDA MOTOR TERLARIS PADA PT DAYA ANUGRAH MANDIRI

Adalah hasil karya sendiri dan bukan "duplikasi" dari karya orang lain. Sepengetahuan saya, di dalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

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

Batam, 3 Agustus 2022



Rozimin
160210207

**PENERAPAN ALGORITMA K-NEAREST NEIGHBOR
UNTUK MEMPREDIKSI PENJUALAN SEPEDA
MOTOR TERLARIS PADA PT DAYA ANUGRAH
MANDIRI**

SKRIPSI

Untuk memenuhi salah satu
syarat memperoleh gelar Sarjana

Oleh
Rozimin
160210207

Telah disetujui oleh Pembimbing pada
tanggal seperti tertera di bawah ini

Batam, 04 Agustus 2022



Rahmat Fauzi, S.Kom., M.Kom.
Pembimbing

ABSTRAK

Sepeda motor menjadi pilihan semua orang sebagai alat transportasi karena murah dan bisa digunakan dalam waktu yang lama. Kenaikan harga barang-barang konsumsi atau kebutuhan sehari-hari seperti kenaikan harga bahan bakar, bahkan pada hari raya keagamaan (idul fitri), atau sering juga disebut fluktuasi (musiman), data mining merupakan suatu data yang diproses melalui ekstraksi (dari yang sebelumnya belum diketahui, serta tidak bermanfaat) sehingga menjadi ilmu pengetahuan atau informasi dari pola data yang besar jumlahnya. salah satu teknik atau metode yang ada pada data mining adalah teknik K-Nearest Neighbor. algoritma klasifikasi bertujuan untuk memprediksi dataset yang memiliki kelas menjadi dataset kelas yang baru.

Tujuan dari penelitian ini adalah diharapkan penelitian ini dapat membantu perusahaan dalam mengambil keputusan dalam menyediakan saham. Pengolahan data penjualan sepeda motor sebanyak 170 dan 3 atribut yang ada pada pemilihan data menggunakan metode algoritma K-Nearest Neighbor menghasilkan prediksi penjualan sepeda motor Honda dengan tipe metric lebih banyak diminati konsumen dibandingkan dengan tipe sport dan CUB. Penelitian ini menggunakan algoritma K-Nearest Neighbord dengan nilai akurasi sebesar 97,65%. presisi untuk matic 98.00 %, Sport 92.45 % CUB 94.12 % dan *recall* untuk *matic* 97.03%, *Sport* 96.08% dan CUB 88.89%. Pengujian yang dilakukan dengan menggunakan *software Rapidminer* cukup efektif karena hasil akhir yang didapatkan sama dengan perhitungan manual. oleh sebab itu, pengujian dengan *Rapidminer* cukup membantu dalam melakukan pengujian.

Kata kunci: algoritma K-Nearest Neighbord; Penambangan Data; Ramalan

ABSTRACT

Motorcycles are everyone's choice as a means of transportation because they are cheap and can be used for a long time. The increase in the price of consumer goods or daily necessities such as the increase in fuel prices, even on religious holidays (Eid), or often also called fluctuations (seasonal), data mining is data that is processed through extraction (from previously unknown data). known, and not useful) so that it becomes knowledge or information from large data patterns. one of the existing techniques or methods in data mining is the K-Nearest Neighbor technique. Classification algorithm aims to predict datasets that have a class into a new class dataset. The purpose of this study is that it is hoped that this research can assist companies in making decisions in providing shares. Processing 170 motorcycle sales data and 3 attributes in the data selection using the K-Nearest Neighbor algorithm method resulted in predictions of Honda motorcycle sales with the metric type being more attractive to consumers than the sport and CUB types. This study uses the K-Nearest Neighbord algorithm with an accuracy value of 97.65%. precision for matic 98.00%, Sport 92.45% CUB 94.12% and recall for matic 97.03%, Sport 96.08% and CUB 88.89%. Tests carried out using Rapidminer software are quite effective because the final results obtained are the same as manual calculations. therefore, testing with Rapidminer is quite helpful in conducting the test.

Keywords: *K-Nearest Neighbord algorithm; Data Mining; Prediction*

KATA PENGANTAR

Segala puji dan syukur kehadirat Allah SWT yang telah melimpahkan segala rahmat dan karuniaNya, 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 akan 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 menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Putera Batam.
2. Dekan Fakultas Teknik dan Komputer Universitas Putera Batam.
3. Ketua Program Studi Teknik Informatika Universitas Putera Batam.
4. Bapak Rahmat Fauzi, S.Kom., M.Kom. selaku pembimbing Skripsi pada Program Studi Teknik Informatika Universitas Putera Batam.
5. Dosen dan Staff Universitas Putera Batam.
6. Aprillisman selaku koordinator PT Daya Anugrah Mandiri cabang tiban riau bertuahyang telah memberikan izin peneliti untuk melakukan pengumpulan data penelitian.
7. Kepada orang tua, serta anggota keluarga yang telah memberikan semangat dan senantiasa mendoakan.
8. Teman-teman angkatan 2016 kelas tiban Teknik Informatika yang telah berjuang bersama dan memberikan informasi serta dukungan sampai tahap ini

9. Teman-teman yang sangat luar biasa yang telah mendoakan
serta memberikan motivasi.

Semoga Allah SWT membalas kebaikan dan selalu mencurahkan taufik
serta hidayah Nya. Amin.

Batam, 03 Agustus 2022

A handwritten signature in black ink, appearing to be 'Rozimin', enclosed in a light gray rectangular box.

Rozimin

DAFTAR ISI

Halaman

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
HALAMAN PERNYATAAN ORISINILITAS	iii
HALAMAN PENGESAHAN	ivs
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR RUMUS	xiii
BAB I PENDAHULUAN	
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah.....	3
1.3 Batasan Masalah	4
1.4 Perumusan Masalah	4
1.5 Tujuan Penelitian	4
1.6 Manfaat Penelitian	5
1.6.1 Manfaat teoritis.....	5
1.6.2 Manfaat praktis.....	5
BAB II TINJAUAN PUSTAKA	
2.1 Knowledge Discovery in Database (KDD)	6
2.2 Data Mining.....	7
2.2.1 Fungsi <i>Data Mining</i>	8
2.2.2 Kategori <i>Data Mining</i>	9
2.3 Metode <i>Data Mining</i>	9
2.3.1 K-Nearest Neighbor (K-NN).....	9
2.3.2 Algoritma C4.5	10
2.3.3 Algoritma K-Means	11

2.3.4	Algoritma Naïve Bayes.....	11
2.4	Software Pendukung.....	13
2.4.1	Rapidminer.....	13
2.4.2	Microsoft Excel.....	14
2.5	Penelitian Terdahulu.....	14
2.6	Kerangka Pemikiran.....	17

BAB III METODE PENELITIAN

3.1	Desain Penelitian.....	18
3.2	Teknik Pengumpulan Data.....	20
3.3	Operasional Variabel.....	20
3.4	Algoritma Data Mining.....	21
3.4.1	Algoritma K-Nearest Neighbor.....	21
3.4.2	Data Selection.....	21
3.5	Lokasi Dan Jadwal Penelitian.....	27
3.5.1	Lokasi Penelitan.....	27
3.5.2	Jadwal Penelitian.....	28

BAB IV HASIL DAN PEMBAHASAN

4.1	Analisis Data.....	29
4.1.1	Pengumpulan Data.....	29
4.1.2	Preprocessing.....	29
4.1.3	<i>Transformation</i>	32
4.2	Implementasi <i>RapidMiner</i>	48
4.3	Hasil Pengujian Algoritma K-NN.....	53
4.4	Hasil Pengujian.....	56

BAB V KESIMPULAN DAN SARAN

5.1	Kesimpulan.....	57
5.2	Saran.....	58

DAFTAR PUSTAKA.....59

LAMPIRAN

- Lampiran 1. Pendukung Penelitian
- Lampiran 2. Daftar Riwayat Hidup
- Lampiran 3. Surat Keterangan Izin Penelitian

DAFTAR GAMBAR

Halaman

Gambar 2.1 Tahapan <i>Knowledge Discovery in Database</i>	6
Gambar 2.2 Logo <i>Rapidminer</i>	13
Gambar 2.3. <i>Icon Microsoft Excel</i>	14
Gambar 2.4 Kerangka Pemikiran	17
Gambar 3.1 Desain Penelitian	18
Gambar 3.2 Lokasi Penelitian.....	27
Gambar 4.1 Tampilan Awal <i>Rapidminer</i> Versi 99	48
Gambar 4.2 <i>Blank process</i>	49
Gambar 4.3 Halaman Kerja Baru.....	49
Gambar 4.4 <i>Read Excel</i>	50
Gambar 4.5 Tempat Penyimpanan Data	50
Gambar 4.6 <i>Importing data Training</i>	51
Gambar 4.7 Susunan Operator K-NN	51
Gambar 4.8 Model Proses <i>Cross Validation</i>	52
Gambar 4.9 <i>Performance Vektor</i>	52
Gambar 4.10 Prediksi Penjualan Honda Vario	53
Gambar 4.11 Prediksi Penjualan Honda Scoopy	54
Gambar 4.12 Prediksi Penjualan Honda Beat	54
Gambar 4.13 Prediksi Penjualan Honda CB150R	55
Gambar 4.14 Prediksi Penjualan Honda	55

DAFTAR TABEL

Halaman

Tabel 3.1 Data Selection Tahun 2018.	22
Tabel 3.2 Data Selection Tahun 2019.	23
Tabel 3.3 Data Selection Tahun 2020.	25
Tabel 3.4 Jadwal Penelitian.	28
Tabel 4.1 Sampel Data Penjualan Honda Vario.	30
Tabel 4.2 Sampel Data Penjualan Honda Scoopy.....	30
Tabel 4.3 Sampel Data Penjualan Honda Beat.	31
Tabel 4.4 Sampel Data Penjualan Honda CB150R.....	31
Tabel 4.5 Sampel Data Penjualan Honda Sonic.	32
Tabel 4.6 Data <i>Training</i> Penjualan Honda Vario	33
Tabel 4.7 Data <i>Testing</i> Penjualan Honda Vario	34
Tabel 4.8 Data <i>Training</i> Penjualan Honda Scoopy.....	36
Tabel 4.9 Data <i>Testing</i> Penjualan Honda Scoopy.....	37
Tabel 4.10 Data <i>Training</i> Penjualan Honda Beat	39
Tabel 4.11 Data <i>Testing</i> Penjualan Honda Beat	40
Tabel 4.12 Data <i>Training</i> Penjualan Honda CB150R.....	42
Tabel 4.13 Data <i>Testing</i> Penjualan Honda CB150R.....	43
Tabel 4.14 Data <i>Training</i> Penjualan Honda Sonic	45
Tabel 4.15 Data <i>Testing</i> Penjualan Honda Sonic.....	46

DAFTAR RUMUS

Halaman

Rumus 3.4.1. <i>Euclidean distance</i>	10
---	----