

**IMPLEMENTASI DATA MINING MENGGUNAKAN
ALGORITMA REGRESI LINEAR UNTUK
MEMPREDIKSI PENJUALAN PADA SUPERMARKET**

SKRIPSI

Oleh:

Zulfikri

201910225089



PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS BHAYANGKARA JAKARTA RAYA

2023

LEMBAR PERSETUJUAN PEMBIMBING

Judul Tugas Akhir : Implementasi Data Mining Menggunakan
Algoritma Regresi Linear Untuk Memprediksi
Penjualan Pada Supermarket

Nama Mahasiswa : Zulfikri

Nomor Pokok
Mahasiswa : 201910225089

Program Studi / Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian : 04/08/2023



Pembimbing I

Wowon Priatna, S.T., M.Ti

NIDN: 0429118007

Pembimbing II

Dr. Dra. Tyastuti Sri Lestari, M.M

NIDN: 0327036701

LEMBAR PENGESAHAN

Judul Tugas Akhir : IMPLEMENTASI DATA MINING
MENGUNAKAN ALGORITMA REGRESI
LINEAR UNTUK MEMPREDIKSI
PENJUALAN PADA SUPERMARKET

Nama Mahasiswa : Zulfikri

Nomor Pokok Mahasiswa : 201910225089

Program Studi / Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Tugas Akhir : 04 Agustus 2023

Jakarta, 11 Agustus 2023

MENGESAHKAN,

Ketua Tim Penguji : Prima Dina Atika, S.Kom., M.Kom
NIDN. 0311037107

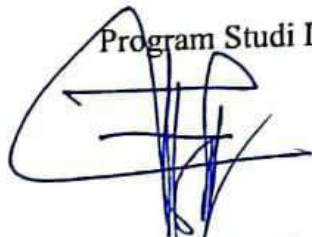
Penguji I : Asep Ramdhani Mahbub, S.Kom., M.Kom
NIDN. 0329087703

Penguji II : Wowon Priatna, S.T., M.Ti
NIDN. 0429118007

MENGETAHUI,

Ketua

Program Studi Informatika



Ahmad Fahurrozi, S.E., M.M.S.I

NIP. 2012486

Dekan

Fakultas Ilmu Komputer



Dr. Dra. Tyastuti Sri Lestari, M.M

NIP. 1408206



LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Zulfikri
NPM : 201910225089
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Implementasi data mining menggunakan algoritma regresi linear untuk memprediksi data penjualan pada supermarket.

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.

Jakarta, 11 Agustus 2023

Penulis


E3AKX543907385

Zulfikri

ABSTRAK

Zulfikri. 201910225089. Implementasi data mining menggunakan algoritma regresi linear untuk memprediksi data penjualan pada supermarket. Bekasi: Universitas Bhayangkara Jakarta Raya. 2023

Penumpukan stok di gudang maupun kekurangan stock pada saat permintaan sedang tinggi merupakan hal yang sering dialami oleh kebanyakan supermarket, oleh karna itu mengetahui atau memprediksi penjualan dimasa depan sangat dibutuhkan untuk mengurangi kerugian tersebut. Pada penelitian ini data penjualan 2020 sampai 2021 digunakan untuk memprediksi penjualan dimasa mendatang dengan menggunakan algoritma regresi linear dan alur proses knowledge discovery in database (KDD). Hasil evaluasi yang didapat dari hasil prediksi paling baik ada pada kategori Fashion Accessories dengan RMSE/MAPE 1,019/12% dan paling rendah pada kategori Home and Lifestyle dengan RMSE/MAPE 1,255/15%. Hasil tersebut dapat dikatakan baik untuk digunakan dengan hasil dari evaluasi MAPE.

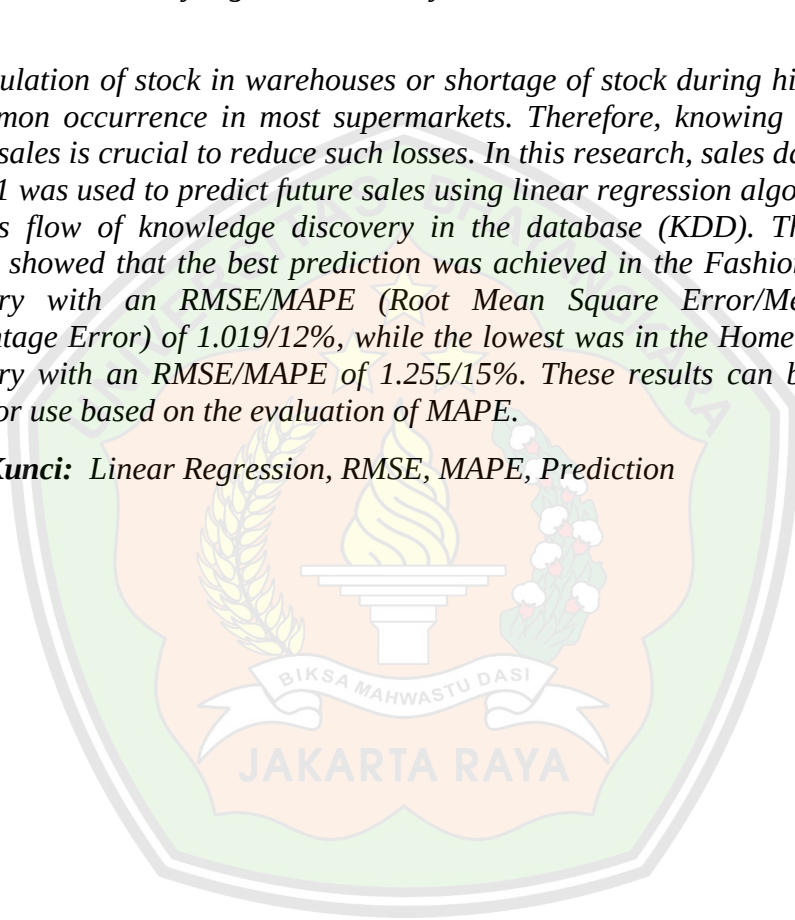
Kata Kunci : *Regresi linear; RMSE; MAPE; Prediksi*

ABSTRACT

Zulfikri. 201910225089. *Implementation of data mining using linear regression algorithm to predict sales data in a supermarket. Bekasi: Faculty of Computer Science. Jakarta Bhayangkara University. 2023*

Accumulation of stock in warehouses or shortage of stock during high demand is a common occurrence in most supermarkets. Therefore, knowing or predicting future sales is crucial to reduce such losses. In this research, sales data from 2020 to 2021 was used to predict future sales using linear regression algorithm and the process flow of knowledge discovery in the database (KDD). The evaluation results showed that the best prediction was achieved in the Fashion Accessories category with an RMSE/MAPE (Root Mean Square Error/Mean Absolute Percentage Error) of 1.019/12%, while the lowest was in the Home and Lifestyle category with an RMSE/MAPE of 1.255/15%. These results can be considered good for use based on the evaluation of MAPE.

Kata Kunci: Linear Regression, RMSE, MAPE, Prediction



**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK**

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

Nama : Zulfikri
NPM : 201910225089
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 :

**IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA
REGRESI LINEAR UNTUK MEMPREDIKSI PENJUALAN PADA
SUPERMARKET**

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 ijin 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 : 11 Agustus 2023
Yang Menyatakan


Zulfikri

KATA PENGANTAR

Segala puji bagi Allah SWT, yang telah melimpahkan Rahmat dan hidayahnya sehingga saya dapat menyelesaikan penelitian skripsi yang berjudul **“Implementasi Data Mining Menggunakan Algoritma Regresi Linear Untuk Memprediksi Penjualan di Supermarket”**. Namun masih ada kewajiban yang harus saya kerjakan yaitu pembuatan laporan tertulis Penelitian Skripsi. Dengan adanya Penelitian Skripsi, saya mendapatkan banyak sekali wawasan, dan ilmu yang bermanfaat. Skripsi ini saya kerjakan sebagai tanggung jawab dan bukti tertulis akan syarat lulus pada mata kuliah skripsi yang saya ambil pada semester 8 ini.

Dalam penyusunan penelitian skripsi ini saya banyak sekali mendapatkan bantuan dari berbagai pihak, oleh sebab itu saya ingin mengungkapkan rasa terima kasih kepada :

1. Bapak Irjen Pol (Purn) Prof. Dr. Drs. Bambang Karsono, S.H., M.M. Selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Dr. Dra. Tyastuti Sri Lestari, M.M. selaku Dekan Fakultas Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya dan selaku Dosen Pembimbing II dalam penelitian proposal skripsi Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I. selaku Kepala Program Studi sejak awal berdirinya Fakultas Ilmu Komputer Universitas Bhayangkara

Jakarta Raya.

4. Bapak Wowon Priatna, S.T, M.T.I selaku Dosen Pembimbing I dalam penelitian proposal skripsi Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
5. Ibu Rafika Sari, S.SI.,M.SI. selaku dosen Pembimbing Akademik.
6. Kedua orang tua dan adik kandung saya yang selalu memberikan dukungan terhadap penulis supaya menyelesaikan skripsi.
7. Kepada Mas Tito dan Mbak Risty yang telah membiayai kuliah penulis.
8. Kepada Risyia Aulia selaku teman yang menemani saya saat membuat skripsi ini.
9. Serta teman-teman kelas A2 angkatan 2019 yang selalu membantu dari semester 1 sampai semester 8.

Penelitian Skripsi yang saya susun ini, mungkin masih terdapat banyak kekurangan dan masih jauh dari sempurna, oleh karena itu saya dengan hati terbuka menerima masukan baik berupa kritik, atau saran-saran yang dapat dibawah untuk mengikuti kekurangan-kekurangan yang ada. Saya harap tulisan ini bermanfaat, baik bagi saya pribadi maupun bagi siapapun yang membacanya.

Jakarta, 11 Agustus 2023

Hormat saya



Zulfikri

DAFTAR ISI

	Halaman
LEMBAR PERSETUJUAN	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
ABSTRACT	v
LEMBAR PERNYATAAN PUBLIKASI	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Rumusan Masalah	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Batasan Masalah.....	4
1.7 Sistematika Tugas Akhir	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terdahulu.....	6
2.2 Data Sains.....	9
2.3 <i>Data Mining</i>	10
2.4 <i>Knowledge Data Discovery (KDD)</i>	10
2.5 <i>Feature Selection</i>	14
2.6 Regresi Linear	15
2.7 Supermarket.....	17
2.8 Metode Evaluasi	17
2.8.1 <i>Root Mean Squared Error (RMSE)</i>	17
2.8.2 <i>Mean Absolute Percentage Error (MAPE)</i>	18

BAB III METODOLOGI PENELITIAN	20
3.1 Kerangka Penelitian	20
3.2 <i>Data Selection</i>	21
3.3 <i>Data Pre-processing</i>	21
3.4 <i>Data Transformation</i>	21
3.5 <i>Data Mining</i>	21
3.6 <i>Evaluation</i>	22
3.7 Prediksi Manual.....	23
3.8 Pembahasan	23
BAB IV HASIL DAN PEMBAHASAN	25
4.1 Hasil Dengan <i>Script Python</i>	25
4.1.1 <i>Data Selection</i>	25
4.1.2 <i>Data Pre-processing</i>	28
4.1.3 <i>Data Transformation</i>	29
4.1.4 <i>Data Mining</i>	30
4.1.5 <i>Evaluation</i>	34
4.2 Hasil Perhitungan Manual	37
4.2.1 Perhitungan Manual <i>Electronic Accessories</i>	37
4.2.2 Perhitungan Manual <i>Health and Beauty</i>	39
4.2.3 Perhitungan Manual <i>Sport and Travel</i>	41
4.2.4 Perhitungan Manual <i>Food and Beverage</i>	44
4.2.5 Perhitungan Manual <i>Fashion Accessories</i>	46
4.2.6 Perhitungan Manual <i>Home and Lifestyle</i>	48
4.3 Pembahasan	53
BAB V PENUTUP	56
5.1 Kesimpulan.....	56
5.2 Saran	57
DAFTAR PUSTAKA	58
LAMPIRAN.....	61

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	8
Tabel 2. 2 Tabel Evaluasi RMSE	18
Tabel 2. 3 Tabel Evaluasi MAPE.....	19
Tabel 4. 1 Keterangan Atribut.....	26
Tabel 4. 2 Pembagian data	30
Tabel 4. 3 Koefisien dan Intercept	31
Tabel 4. 4 Hasil evaluasi	35
Tabel 4. 5 Sampel <i>Electronic Accessories</i>	37
Tabel 4. 6 Hasil evaluasi <i>Electronic Accessories</i>	39
Tabel 4. 7 Sampel <i>Health and Beauty</i>	39
Tabel 4. 8 Hasil evaluasi <i>Health and Beauty</i>	41
Tabel 4. 9 Sampel <i>Sport and Travel</i>	41
Tabel 4. 10 Evaluasi hasil prediksi <i>Health and Beauty</i>	43
Tabel 4. 11 Sampel <i>Food and Beverage</i>	44
Tabel 4. 12 Hasil Evaluasi <i>Food and Beverage</i>	45
Tabel 4. 13 Sampel <i>Fashion Accessories</i>	46
Tabel 4. 14 Hasil Evaluasi <i>Fashion Accessories</i>	47
Tabel 4. 15 Sampel <i>Home and Lifestyle</i>	48
Tabel 4. 16 Hasil Evaluasi <i>Home and Lifestyle</i>	50
Tabel 4. 17 Tabel Evaluasi Perhitungan Manual.....	50
Tabel 4. 18 Evaluasi Keseluruhan Data Perhitungan Manual	52
Tabel 4. 19 Perbandingan Hasil Evaluasi Python dan Perhitungan Manual	54

DAFTAR GAMBAR

Gambar 2. 1 Alur Pengolahan Data Sains	10
Gambar 2. 2 Alur KDD	11
Gambar 3. 1 Kerangka Penelitian.....	20
Gambar 3. 2 Alur Proses <i>Data Mining</i>	20
Gambar 4. 1 Data Yang Digunakan	25
Gambar 4. 2 Data 2020.....	27
Gambar 4. 3 Data 2021.....	27
Gambar 4. 4 <i>Null Data</i>	28
Gambar 4. 5 Pemisahan kolom datetime.....	29
Gambar 4. 6 Variable akhir yang dipakai.....	29
Gambar 4. 7 Hasil prediksi <i>Electronic Accessories</i>	31
Gambar 4. 8 Hasil prediksi <i>Health and Beauty</i>	32
Gambar 4. 9 Hasil prediksi <i>Sport and Travel</i>	32
Gambar 4. 10 Hasil prediksi <i>Food and Beverages</i>	33
Gambar 4. 11 Hasil prediksi <i>Fashion Accessories</i>	33
Gambar 4. 12 Hasil prediksi <i>Home and Lifestyle</i>	34
Gambar 4. 13 Evaluasi RMSE <i>Python</i>	36
Gambar 4. 14 Evaluasi MAPE <i>Python</i>	36
Gambar 4. 15 Evaluasi RMSE Sample Perhitungan Manual	51
Gambar 4. 16 Evaluasi MAPE Sample Perhitungan Manual	51
Gambar 4. 17 Evaluasi RMSE Perhitungan Manual Seluruh Data Penjualan	52
Gambar 4. 18 Evaluasi MAPE Perhitungan Manual Seluruh Data Penjualan	53

DAFTAR LAMPIRAN

Lampiran 1. Lembar Orisinalitas	62
Lampiran 2. Biodata Mahasiswa.....	63
Lampiran 3. Kartu Bimbingan Tugas Akhir	64
Lampiran 4. Keseluruhan Data.....	65

