

**ESTIMASI WAKTU TEMPUH PADA DATASET PENGIRIMAN JERUK
MENGUNAKAN APLIKASI RAPID MINER**



**UNIVERSITAS
TEKNOLOGI
SUMBAWA**

Oleh :

Zainul Hasan (20.01.013.080)
Reinaldy Adiwijaya (20.01.013.062)
Abdullah Azzam (20.01.013.075)
Muhammad Umar Alfajri (20.01.013.076)

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS REKAYASA SISTEM UNIVERSITAS
TEKNOLOGI SUMBAWA
TAHUN AJARAN
2020/2021**

ABSTRAK

ESTIMASI WAKTU TEMPUH PADA DATASET PENGIRIMAN JERUK MENGUNAKAN APLIKASI RAPID MINER

Perusahaan Jeruk Xyz merupakan perusahaan yang bergerak dibidang penjualan jeruk di Kota xxx. perusahaan Jeruk Xyz mencatat data pengiriman jeruk terhadap customernya. dalam dataset yang tercatat, tersedia berbagai data berguna yang bisa dianalisis untuk diolah sebagai kebutuhan strategi dari perusahaan jeruk XYZ. Penerapan estimasi data mining untuk menganalisis hasil estimasi dari dataset pengirim jeruk yang nantinya bisa digunakan untuk kebutuhan strategi penjualan yang lebih efektif dan efisien pada perusahaan jeruk XYZ.

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Allah Ta'ala karena karunianya penulis dapat menyelesaikan Laporan tentang estimasi ini sebagai pertanggungjawaban dari tugas akhir. Laporan yang berjudul “estimasi waktu tempuh pada dataset pengiriman jeruk menggunakan aplikasi rapid miner “ ini diajukan sebagai bukti telah mengerjakan tugas akhir.

Pemilihan tema ini didasari atas karakteristik data mining yang dirasa penulis sangat unik dalam penerapan bahasa jurnalistik. Data penerapan bahasa jurnalistik tersebut dapat direpresentasikan dalam tulisan berupa artikel hasil liputan yang ditulis oleh penulis. Semoga dengan adanya laporan ini dapat menambah wawasan terhadap data mining estimasi.

Penulis menyadari bahwa masih banyak kesalahan dalam penyusunan laporan ini, baik dari segi EBI, kosakata, tata bahasa, etika maupun isi. Maka dari itu penulis sangat mengharapkan kritik dan saran seluas-luasnya dari pembaca yang kemudian akan penulis jadikan sebagai evaluasi.

Demikian semoga laporan ini bisa diterima sebagai ide atau gagasan yang menambah kekayaan intelektual dalam bidang kajian data mining. Semoga laporan saya ini dapat bermanfaat bagi pembaca dan juga untuk penulis sendiri.

Semarang, 3 Desember 2019

Penulis

DAFTAR ISI

HALAMAN DEPAN.....	Error! Bookmark not defined.
ABSTRAK.....	ii
KATA PENGANTAR.....	iii
DAFTAR ISI	iv
BAB I PENDAHULUAN.....	5
1.1. Latar Belakang	5
1.2. Rumusan Masalah	5
1.3. Tujuan Penilitan	5
BAB II LANDASAN TEORI.....	6
2.1. Tinjauan Pustaka	6
2.7. Dasar Teori.....	6
BAB III METODOLOGI PENELITIAN	7
3.1 Metode Penelitian.....	7
3.2 Alat dan Bahan Percobaan	7
BAB IV PEMBAHASAN	8
4.1. Atribut Dataset	8
4.2. Seleksi Data.....	9
4.3. Menerapkan Clustering dan Evaluasi.....	9
4.3.1. Penerapan Clustering	10
4.3.2. Final Clustering.....	11
BAB V PENUTUP	13
5.1. Kesimpulan	13
5.2. Saran.....	13
DAFTAR PUSTAKA.....	14

BAB I

PENDAHULUAN

1.1. Latar Belakang

Dalam berbagai kegiatan, sebuah perencanaan dan perkiraan memang perlu dilakukan. Terlebih lagi kegiatan yg membutuhkan anggaran dana secara khusus maupun proyek pekerjaan yg dilakukan dengan batasan waktu. Dalam hal ini, rencana perkiraan atau estimasi, bisa menjadi tolok ukur dan gambaran umum dana atau waktu yang dibutuhkan untuk menyelesaikan pekerjaan.

Dari hal tersebut, dapat dipahami bahwa estimasi adalah metode untuk memperkirakan sesuatu agar bisa mendapat gambaran umum. Ini menjadi salah satu metode yang perlu dipahami oleh berbagai organisasi atau lembaga dalam menerapkan setiap kegiatan kerjanya.

Bukan hanya untuk mengetahui gambaran umum, estimasi juga penting dilakukan untuk menyusun berbagai rencana guna mengantisipasi situasi masalah tertentu. Sehingga, jika terjadi suatu kondisi yang menjadi kendala jalannya kegiatan, dapat diantisipasi dengan baik. Perencanaan estimasi dalam sebuah kegiatan juga penting dilakukan agar pekerjaan yang akan dilaksanakan efektif dan sesuai dengan target.

Selain itu, juga terdapat berbagai manfaat yang bisa didapatkan dari penerapan estimasi yang baik. Oleh karena itu, penting untuk mengetahui apa yang dimaksud dengan estimasi, apa saja manfaat yang bisa didapatkan, dan berbagai contoh penerapan estimasi di bidang yang berbeda-beda. Waktu tempuh adalah lamanya waktu yang terpakai dalam perjalanan untuk menempuh suatu jarak tertentu. Ukurannya adalah ukuran waktu detik, menit, jam, hari, pekan dan seterusnya. untuk mengestimasi waktu tempuh tersbeut penulis menggunakan sebuah aplikasi rapidminer.

RapidMiner adalah platform perangkat lunak ilmu data yang dikembangkan oleh perusahaan bernama sama dengan yang menyediakan lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran dalam, penambangan teks, dan analisis prediktif. Hal ini digunakan untuk bisnis dan komersial, juga untuk penelitian, pendidikan, pelatihan, rapid prototyping, dan pengembangan aplikasi serta mendukung semua langkah dalam proses pembelajaran mesin termasuk persiapan data, hasil visualisasi, validasi model, dan optimasi.

1.2. Rumusan Masalah

Berdasarkan permasalahan yang ada, rumusan masalah yang dapat disimpulkan yaitu : Bagaimana cara menganalisis data yang ada untuk mencari estimasi waktu tempuh pengiriman pada dataset perusahaan jeruk XYZ?

1.3. Tujuan Penelitian

untuk mengetahui cara mengestimasi dataset sebagai bahan pembelajaran data mining pada sebuah dataset sebagai informasi yang berguna bagi perusahaan jeruk xyz.

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Hal yang perlu dipahami mengenai estimasi yang pertama adalah definisi atau pengertiannya terlebih dahulu. Dalam Kamus Besar Bahasa Indonesia, estimasi memiliki arti perkiraan, penilaian, atau pendapat. Sedangkan berdasarkan istilah, terdapat ahli yang menyebutkan bahwa estimasi adalah suatu pengukuran yang didasarkan pada hasil kuantitatif atau hasil yang akurasi dapat diukur dengan angka. Dalam hal ini, estimasi biasanya digunakan dengan menggunakan perhitungan angka sebagai tolok ukur yang jelas dan pasti. Misalnya seperti estimasi statistik pada proyek sensus penduduk, estimasi biaya pengerjaan proyek, estimasi biaya untuk menggelar suatu acara dan lain sebagainya.

Estimasi pengiriman seringkali ditemukan pada saat masyarakat melakukan pengiriman barang dengan menggunakan jasa pengiriman. Biasanya pada saat pelanggan e commerce membeli suatu barang pada sebuah marketplace, maka pelanggan tersebut mendapatkan sebuah estimasi pengiriman barang. Jadi apakah sebenarnya estimasi pengiriman barang?

Estimasi pengiriman barang adalah perkiraan waktu pengiriman dengan perhitungan waktu dan jarak sampai kepada penerima barang. Biasanya estimasi waktu pengiriman diberikan kepada pelanggannya berdasarkan dengan berapa jauh barang dikirim dan dengan menggunakan produk apa barang dikirimkan. Setiap produk pada perusahaan jasa pengiriman memang memiliki estimasi pengiriman yang berbeda-beda, ada yang same day, ada yang 3 hari sampai ada juga yang besok sampai.

Dengan mengetahui perkiraan barang sampai maka kamu sebagai penjual dapat memilih produk jasa sesuai dengan kualitas barang yang ingin kamu kirimkan. Seperti contohnya makanan tentu harus menggunakan produk same day, ataupun next day agar barang tidak keburu kadaluarsa di jalan. Sebagai acuan perkiraan penulis menggunakan metode estimasi linier regresi.

Regresi linear merupakan suatu pendekatan untuk memantapkan hubungan antara satu atau lebih variabel dependen (regresi linear sederhana) dan juga variabel independen (regresi linier banyak). Salah satu aplikasi dari regresi linear adalah untuk melakukan prediksi berdasarkan data-data yang telah dimiliki sebelumnya. Dengan asumsi hubungan di antara variabel-variabel tersebut, dapat didekati oleh suatu persamaan garis lurus, maka model yang mendekati hubungan antar variabel di data tersebut disebut sebagai pemantapan 1 regresi linear.

2.2 Dasar Teori

Alur proses penelitian mengadopsi model CRISP-DM (Cross Standard Industries Process for Data Mining). Alur Proses laporan ini dapat dijelaskan sebagai berikut :

- a. Business Understanding : Tahap ini memahami tujuan dan kebutuhan dari sudut pandang bisnis, kemudian menterjemahkan pengetahuan ini ke dalam pendefinisian masalah pada data mining. Selanjutnya akan ditentukan rencana dan strategi untuk mencapai tujuan tersebut.
- b. Data Understanding : Tahap ini dimulai dengan pengumpulan data yang kemudian akan dilanjutkan dengan proses untuk mendapatkan pemahaman yang mendalam tentang data, mengidentifikasi masalah kualitas data, atau untuk mendeteksi adanya bagian yang menarik dari data yang dapat digunakan untuk hipotesa untuk informasi yang tersembunyi.
- c. Data Preparation : Tahap ini meliputi semua kegiatan untuk membangun dataset akhir (data yang akan diproses pada tahap pemodelan) dari data mentah. Tahap ini dapat diulang beberapa kali. Pada tahap ini juga mencakup pemilihan tabel, record, dan atribut-atribut data, termasuk proses pembersihan dan transformasi data untuk kemudian dijadikan masukan dalam tahap pemodelan.

BAB III METODOLOGI PENELITIAN

3.1 Metode Penelitian

Setiap penelitian penting dalam menentukan metode yang akan digunakan. Metode penelitian merupakan langkah-langkah untuk meneliti suatu objek, yang akan memandu peneliti tentang urutan-urutan penelitian yang akan dilakukan dan meliputi teknik serta prosedur yang akan digunakan dalam penelitian.

Perubahan nilai suatu variabel tidak selalu terjadi dengan sendirinya, namun perubahan nilai variabel itu dapat pula disebabkan oleh berubahnya variabel lain yang berhubungan dengan variabel tersebut. Untuk mengetahui pola nilai suatu variabel yang disebabkan oleh variabel lain diperlukan alat analisis yang memungkinkan kita untuk membuat perkiraan nilai variabel tersebut pada nilai tertentu variabel yang mempengaruhinya.

Teknik yang umum digunakan untuk menganalisis hubungan antara dua atau lebih variabel dalam ilmu statistik adalah analisis regresi. Analisis regresi adalah teknik statistik yang berguna untuk memeriksa dan memodelkan hubungan diantara variabel-variabel. Analisis regresi berguna dalam menelaah hubungan dua variabel atau lebih dan terutama untuk menelusuri pola hubungan yang modelnya belum diketahui dengan sempurna, sehingga dalam penerapannya lebih bersifat eksploratif dan mudah untuk dipahami.

Persamaan regresi yang digunakan untuk membuat taksiran mengenai nilai variabel terikat disebut persamaan regresi estimasi, yaitu suatu formula yang matematis yang menunjukkan hubungan keterkaitan antara satu atau beberapa variabel yang nilainya sudah diketahui dengan satu variabel yang nilainya belum diketahui.

3.2 Alat dan Bahan Percobaan

Seluruh bahan percobaan yang telah dilakukan pada laporan ini menggunakan:

- a. Rapid Miner : Sebagai aplikasi penunjang visualisasi data
- b. Dataset pengiriman jeruk : Adalah sebuah dataset open-source yang telah diunduh dari github seseorang yakni penyedia dataset publik
- c. Linier Regresion : Merupakan algoritma yang diterapkan untuk mengestimasi data

BAB IV PEMBAHASAN

4.1 Atribut Dataset

Row No.	Kd_Pesanan	Jmlh_Pesanan (kg)	Jmlh_Tikungan	jarak (km)	Waktu_Tempuh (km/jam)
1	1	5	15	4	15
2	2	10	12	3.100	10
3	3	6	13	3.500	12
4	4	2	15	4	15
5	5	4	14	3.700	9
6	6	7	12	3.100	11
7	7	3	11	3	12
8	8	5	13	3.500	10
9	9	4	8	2	9
10	10	3	9	2.400	10
11	11	5	13	3.500	12
12	12	3	15	4	13
13	13	2	14	3.700	9
14	14	1	8	2	10
15	15	3	9	2.400	9

Gambar 1. Dataset Awal Pengiriman Jeruk

Berikut adalah penjelasan terkait atribut yang ada pada dataset pengiriman jeruk perusahaan xyz, yang memiliki 5 kolom tabel sebagai berikut :

Kd_pesanan = Merupakan kode transaksi pemesanan jeruk pada perusahaan xyz

Jmlh_pesanan (kg) = jumlah pesanan jeruk dalam satuan kg

Jmlh_Tikungan = Merupakan jumlah setiap tikungan yang ada pada jalur pengiriman jeruk ke customer.

Jarak (Km) = Merupakan Jarak tempuh pengiriman jeruk ke customers dalam satuan kilometer (Km)

Waktu_Tempuh(Km/Jam) = Merupakan waktu tempuh dari pengiriman jeruk ke customers dalam satuan km/jam.

4.2 Seleksi Data

Bab ini membahas proses pemilihan data. Telah diketahui sedari awal bahwa clustering hanya akan berjalan jika data yang dihimpun merupakan atribut integer.

Dapat dilihat pada *gambar 1* yang merupakan dataset bawaan, adalah dataset tanpa adanya perubahan data. Pengolahan data dilakukan agar penerapan estimasi pada dataset pengiriman jeruk dapat dijalankan. Beberapa atribut pada dataset tersebut tidaklah berupa angka, yang mana harus dimodifikasi terlebih dahulu. Dapat dilihat di bawah ini, yang merupakan data pengiriman jeruk awal yang belum dimodifikasi.

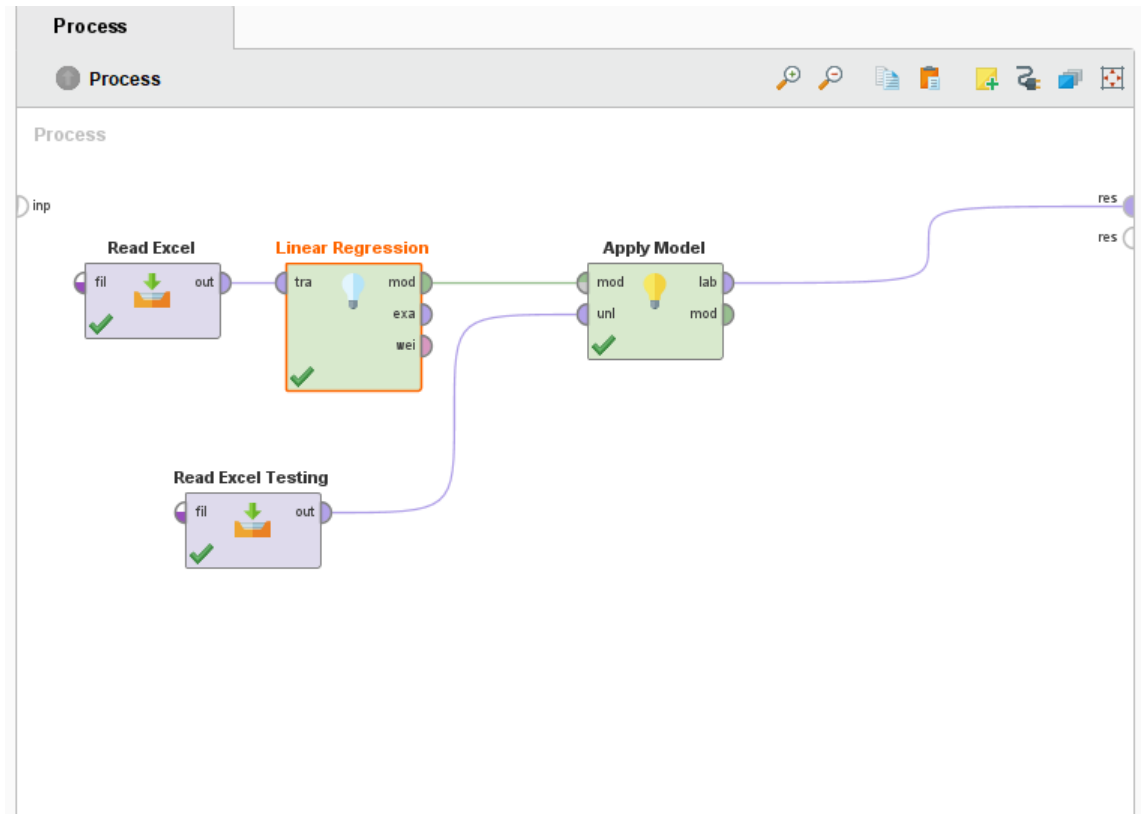
Pada tahapan final penyeleksian data, haruslah memprioritaskan atribut berupa integer. Namun, ada pengecualian dalam pemrosesan data ini. Tidak semua atribut berupa integer layak dijadikan sebagai acuan.

Row No.	Kd_Pesanan	Jmlh_Pesanan (kg)	Jmlh_Tikungan	jarak (km)	Waktu_Tempuh (km/jam)
1	1	5	15	4	15
2	2	10	12	3.100	10
3	3	6	13	3.500	12
4	4	2	15	4	15
5	5	4	14	3.700	9
6	6	7	12	3.100	11
7	7	3	11	3	12
8	8	5	13	3.500	10
9	9	4	8	2	9
10	10	3	9	2.400	10
11	11	5	13	3.500	12
12	12	3	15	4	13
13	13	2	14	3.700	9
14	14	1	8	2	10
15	15	3	9	2.400	9

Gambar 1. 2

4.3 Menerapkan Estimasi Linier Regresion pada DataSet

Telah kami dapati bahwa, bilamana jumlah pengelompokkan dari linier regresion berbeda, maka hasil estimasi akan juga berbeda. Juga telah dibandingkan melalui tiap estimasi.



Gambar 2. Mencari jumlah linier regresion yang sesuai dengan estimasi

Dan hasil final proses estimasi seperti gambar berikut.

Row No.	Waktu_Tempuh (km/jam)	prediction(Waktu_Tempuh (km/jam))	Jmlh_Pesan...	Jmlh_Tikun...	jarak (km)
1	15	21.602	5	15	4
2	10	21.602	10	12	3.100
3	12	21.602	6	13	3.500
4	15	21.602	2	15	4
5	9	21.602	4	14	3.700
6	11	21.602	7	12	3.100
7	12	21.602	3	11	3
8	10	21.602	5	13	3.500
9	9	21.602	4	8	2
10	10	21.602	3	9	2.400
11	12	21.602	5	13	3.500
12	13	21.602	3	15	4
13	9	21.602	2	14	3.700
14	10	21.602	1	8	2

BAB V

PENUTUP

5.1. Kesimpulan

Dari uraian yang telah dijelaskan pada bab sebelumnya, maka dapat disimpulkan hal-hal sebagai berikut :

1. Dapat ditarik kesimpulan bahwa dalam melakukan estimasi menggunakan metode linear regression diperlukan dua dataset yang digunakan menjadi data training dan data testing.
2. Penerapan menggunakan metode estimasi hanya akan berjalan jika atribut dataset bernilai *integer*.

5.2. Saran

Diharapkan penelitian selanjutnya mampu mengimplementasikan metode yang digunakan lebih baik dan akurat.

DAFTAR PUSTAKA

<https://www.merdeka.com/jateng/estimasi-adalah-metode-untuk-memperkirakan-sesuatu-ketahui-manfaatnya-kln.html>

<https://kargo.tech/blog/estimasi-pengiriman/>

https://id.wikipedia.org/wiki/Regresi_linear