

PENERAPAN DATA MINING UNTUK PREDIKSI UJUK KERJA OPERASIONAL PENAMBANGAN BATUBARA

TEMU PROFESI TAHUNAN (TPT) XXXI PERHAPI 2022

¹Sumarjono dan ²M. Arif Saputra

¹Senior Engineer Mining PT. Kaltim Prima Coal

² Mining Engineer PT. Kaltim Prima Coal

Abstract

Dalam operasional penambangan batubara 24 jam sehari dan 7 hari seminggu yang menerapkan Fleet Management System (FMS) menghasilkan data ujuk kerja operasional seperti spotting time, loading time, payload, Physical Availability, digrate & productivity yang menjadi perhatian semua pemangku kepentingan dalam operasional tambang. Data tersebut secara individu dianalisis sesuai kebutuhan masing-masing pihak. Perkembangan teknologi dan sistem informasi menggunakan data mining sebagai metode pengambilan dan pemrosesan data yang berkembang. Data yang menumpuk dari waktu ke waktu, tidak akan memiliki nilai yang berharga jika tidak dilakukan analisis. Untuk menganalisis data historis, maka perlu digunakan metode yang tepat. Metoda Data Mining merupakan solusi yang tepat untuk menganalisis data historis tersebut. Tujuan dari penelitian ini adalah untuk mendapatkan pola dari data operasional dan dapat melakukan prediksi terhadap ujuk kerja yang akan dicapai dalam jangka waktu tertentu. Dalam penelitian ini, penulis menggunakan teknik pengayaan data mining dan software yang digunakan adalah Rapidminer. Pengujian menggunakan Rapidminer didapatkan model dari suatu proses operasional berdasarkan data historical dan model tersebut akan diimplementasikan sebagai prediksi terhadap capaian ujuk kerja operasional. Penelitian ini akan dilakukan untuk mengukur keakuratan data hasil prediksi menggunakan Regresi Linier dengan tujuan untuk mendapatkan hasil tingkat akurasi yang ditawarkan oleh metode tersebut.

Kata Kunci: Prediksi, Ujuk Kerja Operasional Penambangan, Data Mining, Regresi Linier, Rapidminer

Abstract

In coal mining operations, 24 hours a day and 7 days a week that apply the Fleet Management System (FMS) it produces operational performance data such as spotting time, loading time, payload, Physical Availability, digging rate & productivity which is the concern of all stakeholders in mining operations. The data is analyzed according to the needs of each party. The development of technology and information systems using data mining as a method of data retrieval and processing is growing. The data accumulates over time will have no value if it is not analyzed. To analyze historical data, it is necessary to use the right method. The Data Mining method is the right solution to analyze the historical data. The purpose of this study is to obtain patterns from operational data and be able to predict the performance to be achieved within a certain period of time. In this study, the author uses data mining enrichment techniques and the software used is Rapidminer. Testing using Rapidminer obtained a model of an operational process based on historical data and the model will be implemented as a prediction of operational performance. This research will be conducted to measure the accuracy of the predicted data using Linear Regression with the aim of getting the results of the level of accuracy offered by the method.

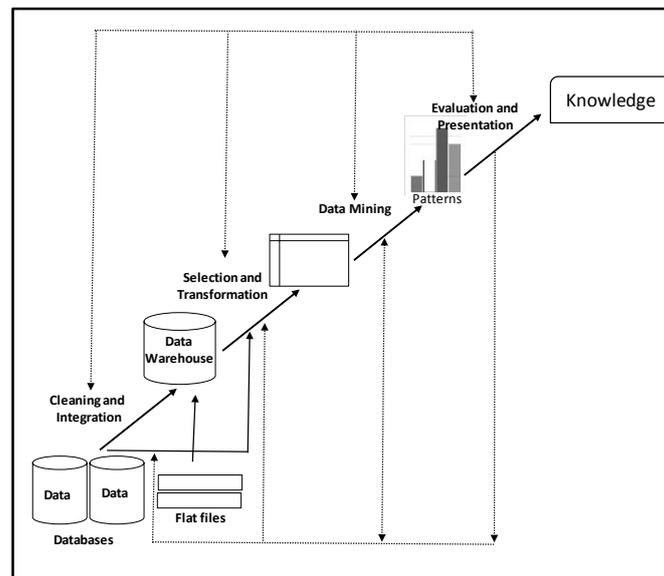
Keywords: Prediction, Apply for Mining Operations, Data Mining, Regression Linier, Rapidminer

1 PENDAHULUAN

1.1 Latar Belakang

Dalam operasional penambangan batubara 24 jam sehari dan 7 hari seminggu yang menerapkan Fleet Management System (FMS) menghasilkan data ujuk kerja operasional seperti spotting time, loading time, payload, Physical Availability, digrate & productivity yang menjadi perhatian semua pemangku kepentingan dalam operasional tambang. Data tersebut secara individu dianalisis sesuai kebutuhan masing-masing pihak.

Perkembangan teknologi dan sistem informasi menggunakan data mining sebagai metode pengambilan dan pemrosesan data yang berkembang. Data yang menumpuk dari waktu ke waktu, tidak akan memiliki nilai yang berharga jika tidak dilakukan analisis. Untuk menganalisis data historis, maka perlu digunakan metode yang tepat. Dalam pengertian Data Mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar (Davies, 2004). Data Mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data (Pramudiono, 2007). Data mining, sering juga disebut sebagai knowledge discovery in database (KDD).



Gambar 1.1 Tahapan data mining Han: 2006)

Prediksi atau biasa disebut peramalan adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil (Herdianto, 2013). Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi. Ketepatan prediksi tergantung dengan pengolahan data disejak awal.

Untuk melihat seberapa jauh kemajuan yang sudah tercapai dalam suatu operasional dalam hal ini adalah operasional penambangan batubara, perlu adanya suatu sarana monitoring bersama ujuk kerja operasional (KPI atau Key Performance Indicators). Dalam tulisan ini penulis membuat prediksi KPI operasional penambangan yaitu productifitas alat muat berguna sebagai sebuah alat ukur yang bisa memberikan informasi yang valid akan keberhasilan yang sudah dicapai dalam waktu ke waktu.

Linear regression pada umumnya akan membuat prediksi berdasarkan nilai yang sudah ada sebelumnya. Metode ini membangun fungsi persamaan linear, yang dibentuk oleh konstanta dan

variabel yang mewakili garis lurus, jika dikonversi menjadi bentuk grafik. Setelah mendapatkan persamaan linier, prediksi dapat dibuat dengan mengganti nilai variabel uji pada persamaan linier. Setelah itu, nilai prediksi dari variabel yang diharapkan akan diperoleh [4]. Metode linear regression sudah sangat sering digunakan pada penelitian yang berhubungan dengan prediksi atau peramalan, karena banyak library pada berbagai macam bahasa pemrograman dan aplikasi sudah menerapkan metode tersebut, salah satunya adalah aplikasi RapidMiner.

Penulis dengan memanfaatkan aplikasi RapidMiner sebelumnya dilakukan oleh Vincentius Riandaru Prasetyo dkk [2]. Penelitian tersebut mencoba menerapkan aplikasi RapidMiner untuk melakukan prediksi nilai tukar rupiah terhadap USD dengan menggunakan metoda Linier Regresion. Penelitian lainnya yang menerapkan RapidMiner melakukan prediksi factor kehadiran karyawan dengan menggunakan algoritma C4.5, Random tree & random forest oleh Riza fahlevi dkk (5). Dan tentunya masih banyak penelitian lainnya yang menggunakan aplikasi RapidMiner untuk melakukan prediksi. Berbeda dengan penelitian sebelumnya,

1.2 Tujuan

saat ini penulis mencoba menerapkan atau menggunakan aplikasi rapidMiner ini untuk melakukan prediksi data KPI operasional penambangan yang basis datanya realtime yang diperoleh dari Fleet Management System.

2 METODE

2.1. Pengumpulan Data

Pada uji prediksi ini, data dikumpulkan dengan cara mengunduh historical-data, di dispatch system melalui minVu sebagai data warehouse yang dimiliki perusahaan. Banyaknya data yang diunduh yaitu data harian unjuk kerja operasional yang terdiri dari Tanggal, Spoting time, Loading Time & produktivitas alat gali/muat. Beberapa atribut yang terdapat pada data yang akan diprediksi ditunjukkan pada table 1 berikut.

Table 1 Atribut Data unjuk kerja operasional penambangan

Kolom	Jenis	Keterangan
Tanggal	Date	Tanggal Operasi
Spoting Time (min)	Numerik	Waktu alat gali/muat melakukan manuver
Loading Time (min)	Numerik	Waktu alat gali/muat melakukan proses pengisian
Productivity (bcm/hr)	Numerik	Production rate dalam satuan bcm/hr

Data yang didapatkan akan dibagi menjadi 2 bagian yaitu sebagai data training dan data testing. Data training tersebut akan digunakan pada metode linear regression untuk mendapatkan persamaan atau model prediksi. Bagian kedua yaitu data testing yang terdiri dari 10% data unjuk kerja operasional. Proses pengolahan data training dan testing akan menggunakan aplikasi RapidMiner. Pada penelitian ini, atribut yang digunakan untuk mendapatkan model prediksi dengan metode linear regression adalah Spoting Time, Loading Time, dan Productivity. Atribut tanggal digunakan untuk menunjukkan time series dari model prediksi.

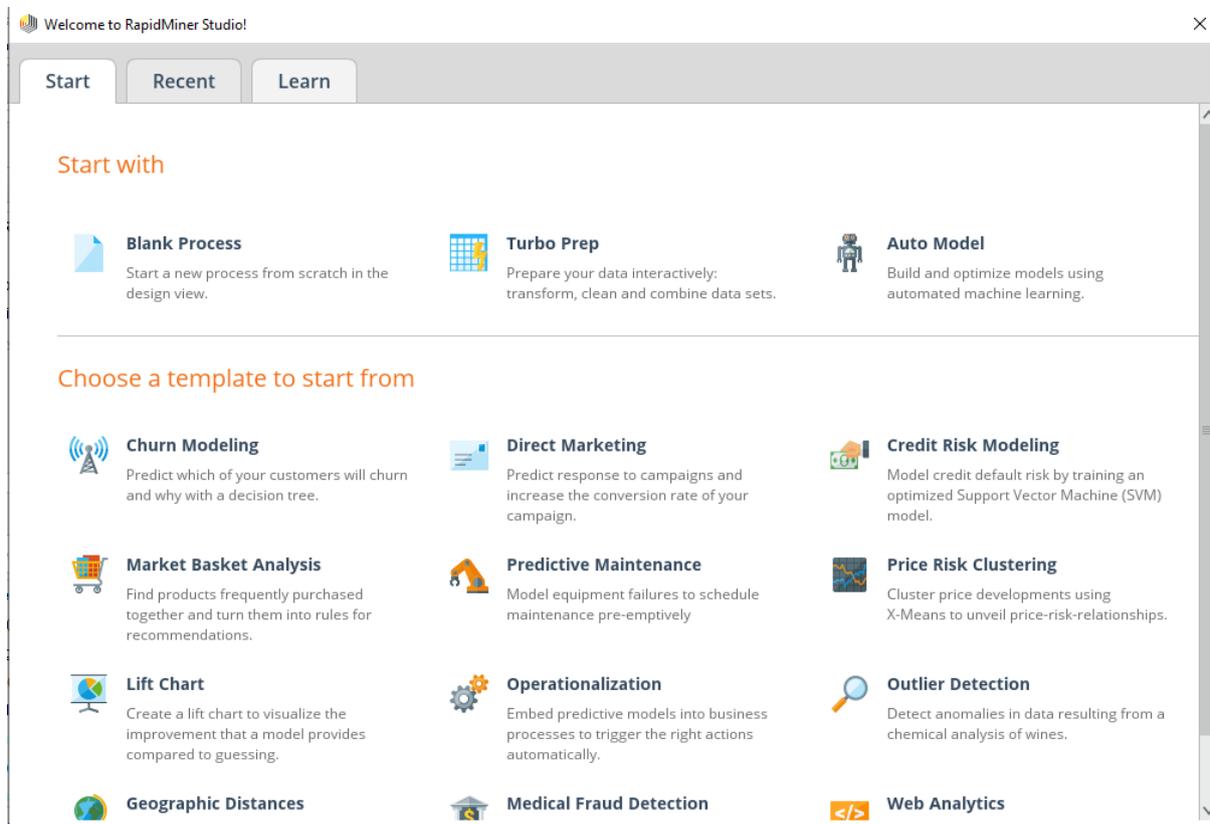
2.2. Aplikasi RapidMiner

RapidMiner adalah aplikasi atau perangkat lunak yang berfungsi sebagai alat pembelajaran dalam ilmu data mining. Platfrom dikembangkan oleh perusahaan yang didedikasikan untuk semua

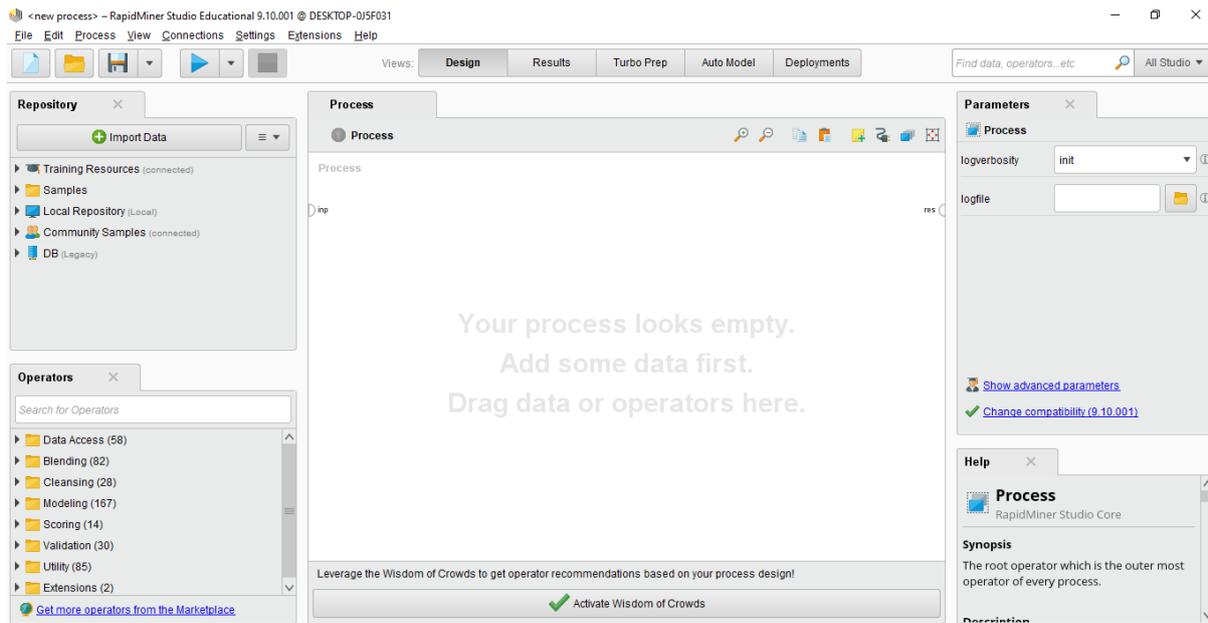
langkah yang melibatkan sejumlah besar data dalam bisnis komersial, penelitian, pendidikan, pelatihan, dan pembelajaran. RapidMiner memiliki sekitar 100 solusi pembelajaran untuk pengelompokan, klasifikasi dan analisis regresi [3]. RapidMiner juga mendukung sekitar 22 format file, seperti .xls, .csv, dan sebagainya. RapidMiner membawa kecerdasan buatan kepada perusahaan melalui platform ilmu data yang terbuka dan dapat diskalakan. RapidMiner dibangun untuk tim analisis, mengintegrasikan seluruh siklus ilmu data, dari persiapan data hingga pembelajaran mesin hingga penyebaran model prediksi.

RapidMiner adalah perangkat lunak independen yang digunakan untuk menganalisa data dan mesin penambangan data, yang dapat diintegrasikan dengan berbagai bahasa pemrograman secara mudah. RapidMiner ditulis dengan bahasa pemrograman Java, sehingga dapat bekerja pada banyak sistem operasi. RapidMiner menyediakan UI untuk mendesain pipa analisis, di mana akan menghasilkan file XML yang dapat menjelaskan proses analisis yang ingin diterapkan oleh pengguna ke data. RapidMiner akan membaca file ini untuk menjalankan analisa secara otomatis [3].

RapidMiner menyediakan tampilan (UI) yang ramah pengguna, sehingga memudahkan pengguna saat menggunakannya. Tampilan yang terdapat pada RapidMiner disebut Perspective. Terdapat 3 Perspective, yaitu Welcome Perspective, Design Perspective dan Result Perspective. Saat user membuka aplikasi RapidMiner, user akan melihat Welcome Perspective, seperti yang ditunjukkan pada Gambar 2.1. Design Perspective merupakan tampilan di mana user akan bekerja di RapidMiner. Design Perspective adalah perspective utama pada RapidMiner, yang digunakan sebagai tempat untuk merancang dan mengelola proses analisis. Seperti yang ditunjukkan pada Gambar 2.2, terdapat beberapa view pada perspective ini dengan fungsinya masing-masing. View dapat mendukung proses analisis data. Result Perspective adalah tampilan di mana hasil analisis akan ditampilkan. Tampilan yang dihasilkan bisa bermacam-macam, seperti grafik, teks, tree, tabel, dan sebagainya, tergantung dari data dan metode yang dipakai untuk melakukan analisis.

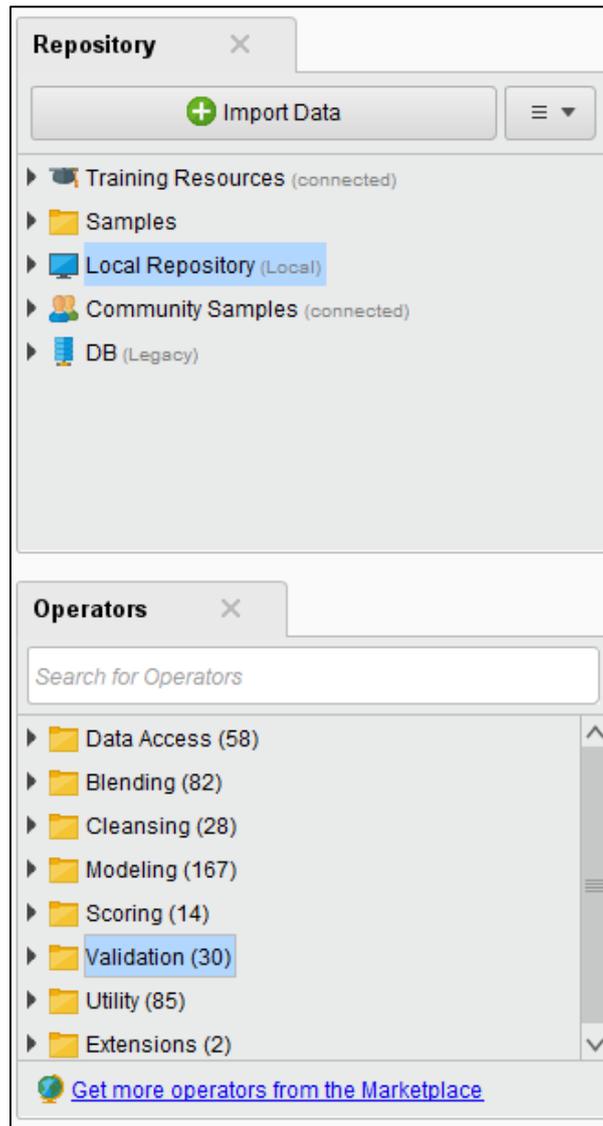


Gambar 2.1 Welcome Perspective



Gambar 2.2 Design Perspective

Sebagai tampilan untuk lingkungan kerja, Design Perspective memiliki beberapa view. Operator View adalah view yang paling utama dan penting. Semua operator atau langkah-langkah kerja dalam RapidMiner ditampilkan dalam operator view dalam bentuk grup hierarkis, sehingga operator ini dapat digunakan dalam proses analisis, seperti yang ditunjukkan pada Gambar 2.3. Berdasarkan Gambar 3 di bawah, operator view dibagi menjadi beberapa kelompok operator, yang pertama adalah Process Control. Operator ini berisi logika looping dan percabangan yang mengatur aliran proses analisis data. Yang kedua adalah Utility yang berfungsi sebagai operator bantuan seperti macros, login, subproses dan lain-lain. Ketiga yaitu Repository Access, di mana operator ini berfungsi untuk membaca atau menulis akses pada repository. Operator yang keempat adalah Import yang digunakan untuk membaca data dan objek dari file, database, dan sebagainya. Operator kelima yaitu Export yang mempunyai fungsi yang berkebalikan dari Import, di mana operator ini digunakan untuk menulis data ke dalam format tertentu. Selanjutnya ada operator Data Transformation yang berguna untuk transformasi data dan metadata. Kemudian operator Modelling yang berisi segala macam metode dan teknik data mining untuk pengolahan data. Operator terakhir adalah Evaluation yang berfungsi untuk mengevaluasi kualitas output yang dihasilkan.



Gambar 2.3 Operator View

Selain operator view, repository view juga merupakan komponen yang cukup penting dalam Design Perspective. Pada view ini, user dapat mengelola dan mengatur proses analisis data menjadi sebuah proyek, di mana pada saat yang sama view tersebut berperan sebagai sumber data dan hal-hal yang terkait dengan metadata. Process view juga berfungsi untuk menunjukkan tahap-tahap tertentu dalam proses analisis dan juga berfungsi sebagai tautan ke langkah-langkah berikutnya.

2.3. Linear Regression

Regresi adalah teknik membangun model yang digunakan untuk memprediksi nilai input data yang diberikan. Regresi adalah suatu ukuran statistik yang digunakan untuk menentukan kekuatan hubungan antara variabel terikat (dependen) dengan variabel bebas (independen). Metode utama untuk peramalan adalah membangun model regresi dengan mencari hubungan antara satu atau lebih variabel bebas atau prediktor (X) dan variabel terikat atau respons (Y). Regresi linier memodelkan hubungan antara variabel skalar dan satu atau lebih variabel penjelas. Metode regresi linier didasarkan pada pola hubungan data pada masa lalu. Pada penelitian ini dilakukan pembentukan model dari data training yang telah dikumpulkan sebelumnya. Model terbentuk dari persamaan linear regression yang secara umum ditulis seperti persamaan (1).

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n \quad (1)$$

Pada persamaan (1) di atas, Y adalah variabel dependen, yang tergantung pada nilai X (variabel independen). Nilai a adalah suatu konstanta dan b adalah suatu koefisien regresi dari variabel X. Untuk mendapatkan nilai a dan b terhadap nilai variabel X, maka dapat ditulis dalam persamaan (2) dan persamaan (3).

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} \quad (2)$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \quad (3)$$

Penulis memanfaatkan aplikasi RapidMiner untuk mendapatkan model persamaan dari data training yang telah dikumpulkan.

2.4. Evaluasi Root Mean Square Error

Evaluasi dilakukan untuk melihat hasil prediksi yang dihasilkan oleh RapidMiner. Hasil yang RMSE atau Root Mean Square Error adalah akar kuadrat dari kuadrat kesalahan rata-rata yang dihasilkan dari perhitungan suatu metode. RMSE digunakan untuk membandingkan nilai yang diprediksi oleh model hipotetis dengan nilai dari hasil pengamatan. Dengan kata lain, RMSE mengukur kualitas kesesuaian antara data aktual dan model prediksi. RMSE adalah salah satu ukuran yang paling umum digunakan untuk model regresi. RMSE dapat dihitung dengan menggunakan persamaan (4).

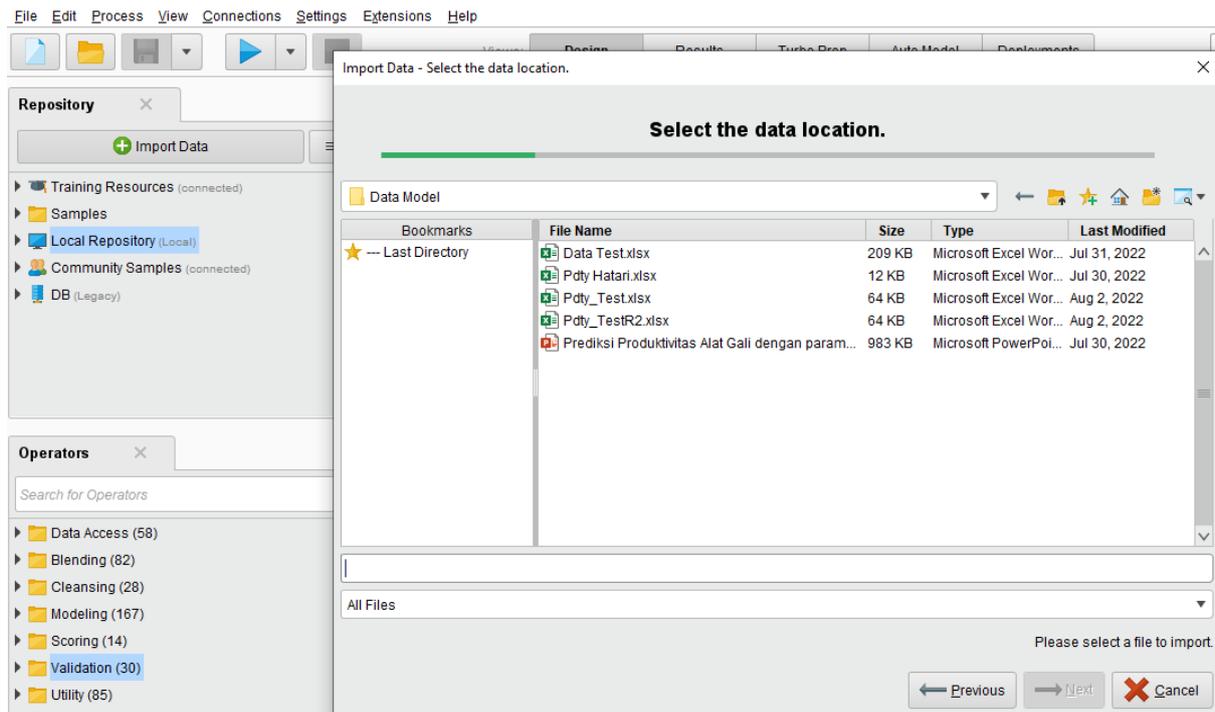
$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Y_t - Y'_t)^2}{n}} \quad (4)$$

Di mana Y_t adalah nilai aktual pada periode t, dan Y'_t adalah nilai hasil prediksi pada periode t. Untuk jumlah prediksi diwakili oleh nilai n. RapidMiner juga menyediakan operator untuk mengevaluasi kinerja regresi linier yaitu menghitung nilai root mean square error. Dalam ukuran ini model prediksi dikatakan paling baik RMSE adalah 0 (nol).

3 HASIL DAN PEMBAHASAN

3.1. Desain Model Prediksi

Pemodelan dilakukan dengan memanfaatkan aplikasi RapidMiner. Proses diawali dengan mengambil data training yang telah disediakan sebelumnya dengan cara menekan klik pada tab Import Data, dan pilih pada local folder yang menyimpan data tersebut, karena data training yang dikumpulkan berupa file excel. Proses tersebut juga dapat dilihat pada Gambar 3.1.



Gambar 3.1 Proses Import Data

Langkah berikutnya adalah memilih file excel data training yang akan digunakan. Klik tombol Import Configuration Wizard, seperti yang diperlihatkan pada Gambar 3.1, kemudian pilih file excel yang digunakan. Ada beberapa step setelah memilih file yaitu memilih data, anotasi, dan atribut data. Dikarenakan data yang digunakan sudah sesuai kebutuhan, maka tidak perlu pengaturan khusus pada step-step tersebut.

Import Data - Format your columns. X

Format your columns.

Replace errors with missing values ⓘ

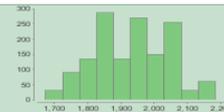
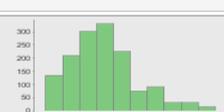
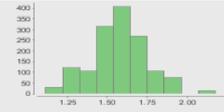
	Date <i>date</i>	Time <i>time</i>	Spot Time <i>real</i>	Load Time <i>real</i>	Productivity <i>real label</i>
1	Jan 1, 2022	7:30:00 AM SGT	0.995	1.443	1917.680
2	Jan 1, 2022	7:45:00 AM SGT	1.048	1.481	2119.285
3	Jan 1, 2022	8:00:00 AM SGT	1.034	1.651	1971.555
4	Jan 1, 2022	8:15:00 AM SGT	1.210	1.494	1840.570
5	Jan 1, 2022	8:30:00 AM SGT	1.005	1.626	1833.908
6	Jan 1, 2022	8:45:00 AM SGT	1.085	1.543	1861.424
7	Jan 1, 2022	9:00:00 AM SGT	1.292	1.464	1749.571
8	Jan 1, 2022	9:15:00 AM SGT	1.122	1.695	1780.299
9	Jan 1, 2022	9:30:00 AM SGT	1.115	1.751	1797.424
10	Jan 1, 2022	9:45:00 AM SGT	1.108	1.696	1790.831
11	Jan 1, 2022	10:00:00 AM SGT	1.151	1.299	1815.808
12	Jan 1, 2022	10:15:00 AM SGT	1.057	1.286	2055.125

✔ no problems.

← Previous
Next →
✗ Cancel

Gambar 3.2 Pemilihan File Data

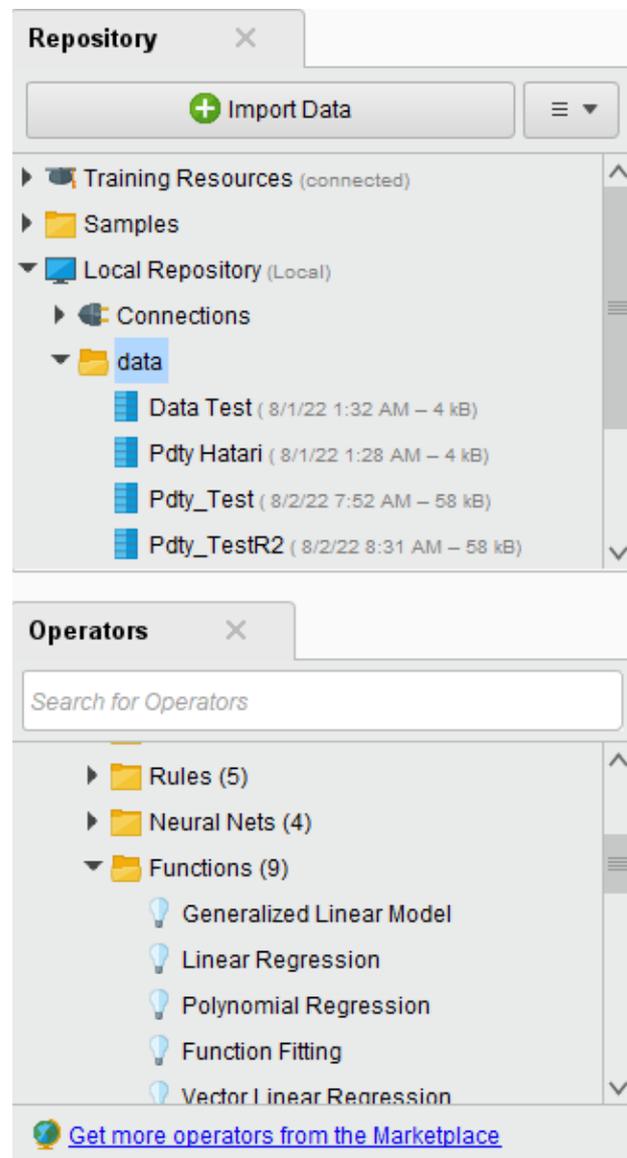
Result History ExampleSet (/Local Repository/data/Pdty_TestR3) X

Name	Type	Missing	Statistics	Filter (5 / 5 attributes):
Productivity	Real	0	 Min: 1674.135 Max: 2182.809	Search for Attributes
Date	Date-time	0	Earliest date: Jan 1, 2022 Latest date: Jan 15, 2022 Duration: 14 days	
Time	Time	0	Min: 12:00:00 AM Max: 11:45:00 PM Average: 11:52:19.0700	
Spot Time	Real	0	 Min: 0.894 Max: 1.376	
Load Time	Real	0	 Min: 1.117 Max: 2.173	

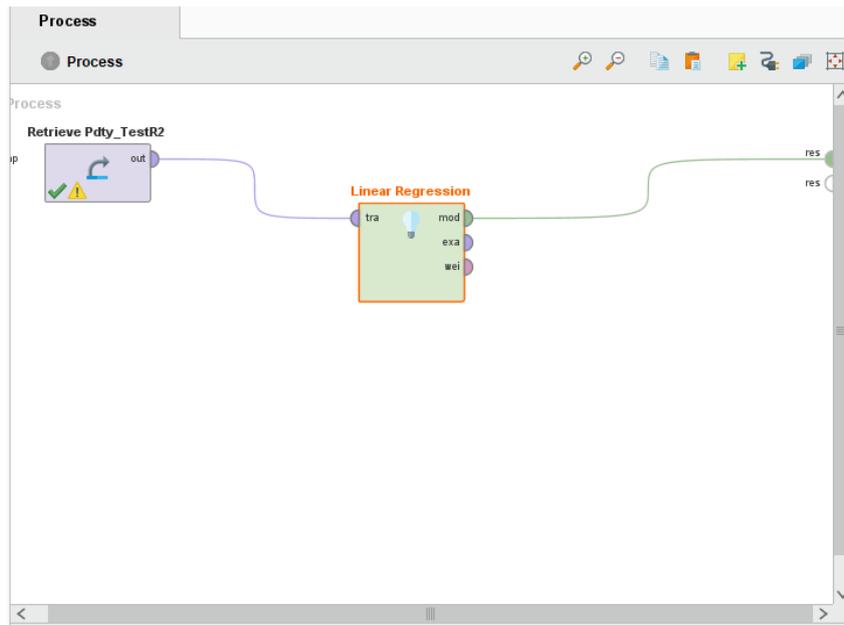
Showing attributes 1 - 5 Examples: 1,441 | Special Attributes: 1 | Regular Attributes: 4

Gambar 3.3 Visualisasi Data

Langkah berikutnya setelah selesai melakukan Import Data adalah menambahkan operator Linear Regression. Operator tersebut termasuk dalam kategori Function Fitting, seperti yang ditunjukkan pada Gambar 3.4. Setelah menambahkan operator Linear Regression, langkah selanjutnya adalah menghubungkannya ke socket “tra” pada operator Linear Regression dengan socket “out”. Selanjutnya, menghubungkan socket “mod” pada operator Linear Regression dengan socket “res” untuk menampilkan model yang diperoleh, seperti yang ditunjukkan pada Gambar 3.5.



Gambar 3.4 Lokasi Operator Linear Regression



Gambar 3.5 Pembentukan Model Linear Regression

Operator Linear Regression menggunakan kriteria Akaike untuk pemilihan model. Akaike merupakan ukuran kebaikan relatif dari kecocokan suatu model statistik. Ukuran tersebut pada dasarnya merupakan konsep entropi informasi, di mana konsep tersebut menawarkan ukuran relatif dari informasi yang hilang ketika model digunakan untuk menggambarkan realitas yang ada. Selain itu, dapat juga untuk menggambarkan tradeoff antara bias dan variansi atau secara singkat berbicara antara kompleksitas model dan akurasi. Apabila proses pada Gambar 3.5 di atas dijalankan, maka akan ditampilkan model dari data training tersebut. Model linear regression yang dihasilkan, dapat dilihat pada Gambar 3.6.

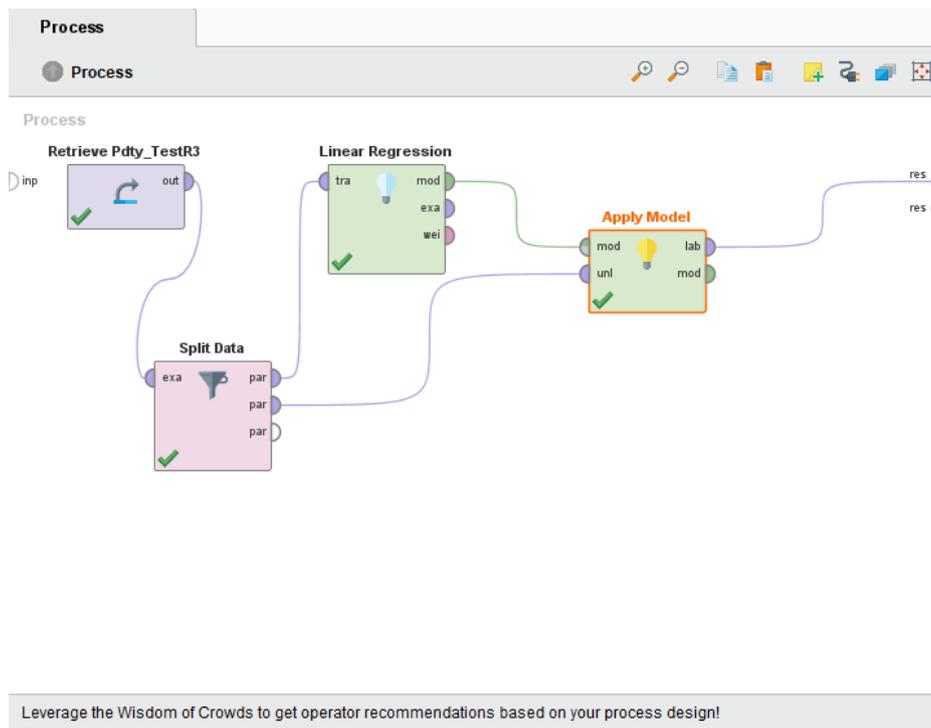
LinearRegression

- 791.443 * Spot Time
 - 155.702 * Load Time
 + 3012.383

Gambar 3.6 Model Linear Regression

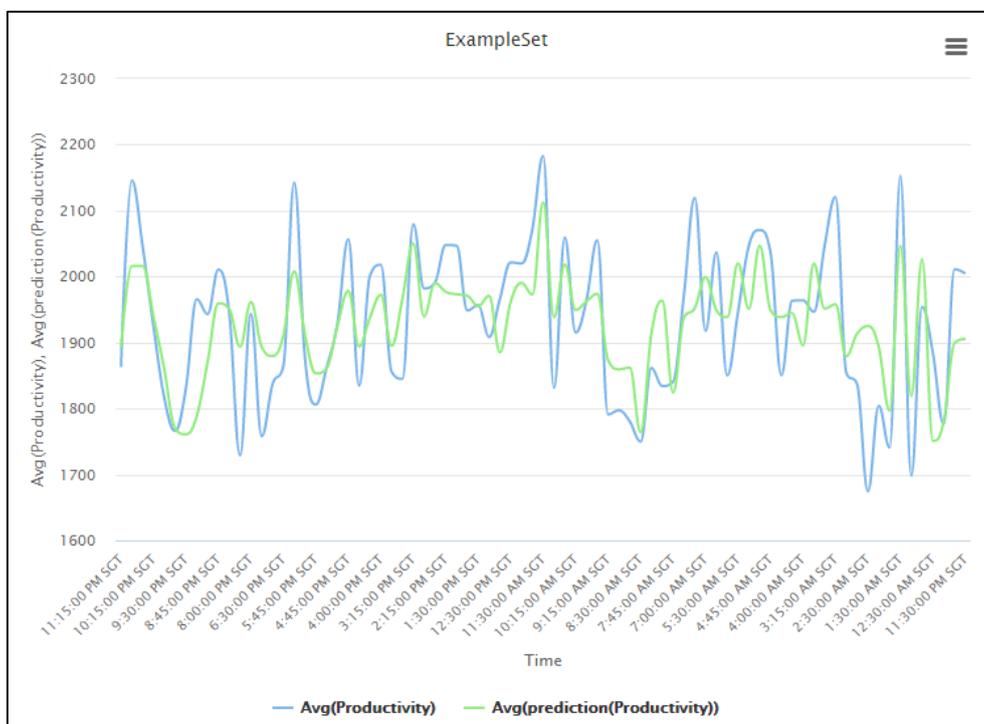
3.2. Proses Uji Coba Dan Evaluasi

Proses uji coba dijalankan dengan cara melakukan import data testing yang ada dengan cara yang sama dengan import data training, kemudian diaplikasikan dengan operator Apply Model, di mana model linear regression telah didapatkan sebelumnya. Operator Apply Model termasuk ke dalam kategori Confidences. Langkah berikutnya adalah menghubungkan socket-socket yang ada seperti yang ditunjukkan pada Gambar 3.7.



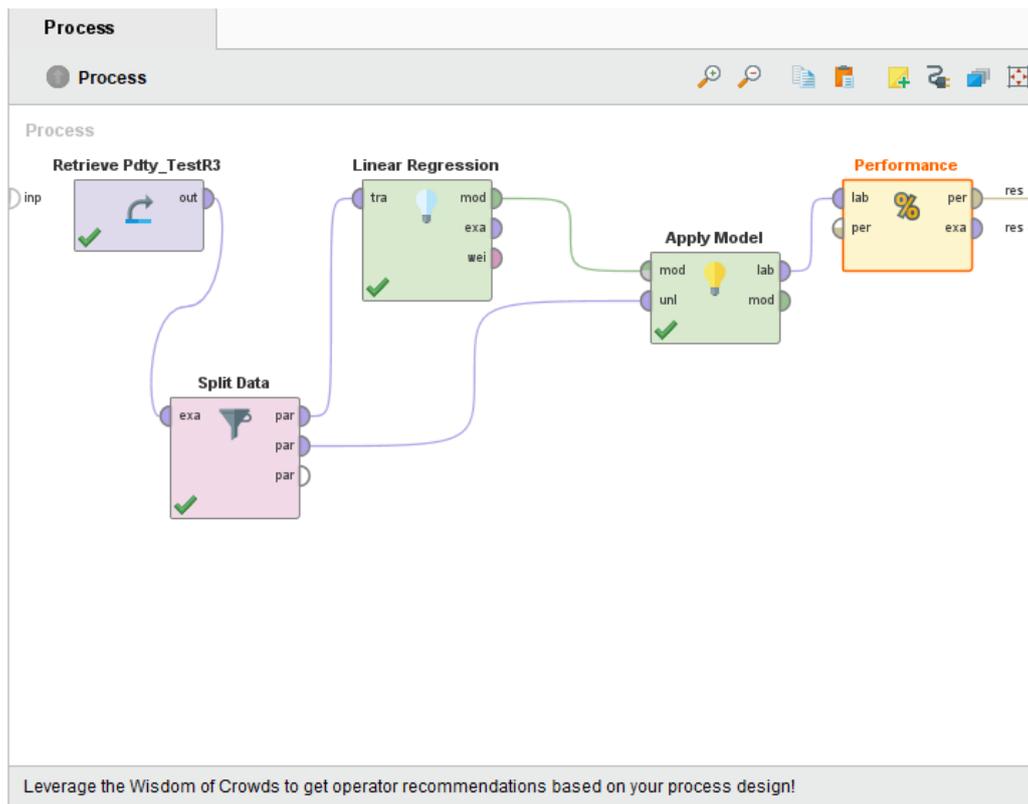
Gambar 3.7 Uji Coba Data Testing Pada Model Regresi

Hasil prediksi yang dilakukan oleh aplikasi RapidMiner terhadap data testing dengan model linear regression yang ada, dapat dilihat pada Gambar 3.8. Berdasarkan Gambar 3.9, dapat dilihat bahwa hasil prediksi dapat dilihat



Gambar 3.8 Hasil Prediksi

Hasil prediksi diatas dilakukan uji performance untuk mengevaluasi terhadap data yang diprediksi (Data Training & Data Testing).



Gambar 3.9 Proses Uji Performa Model Linear Regression

Hasil dari proses uji performa model linear regression diperlihatkan pada Gambar 3.10. Hasil dari perhitungan RMSE yang didapatkan tergolong cukup besar yaitu 87,518. Dikatakan cukup besar karena RMSE yang didapat masih jauh mendekati kisaran 0 sampai 1.

```

root_mean_squared_error
root_mean_squared_error: 87.518 +/- 0.000

```

Gambar 3.10 Hasil Uji Performance Prediksi

Ada beberapa faktor yang berkontribusi terhadap nilai RMSE yang besar yaitu data training yang terlalu sedikit, kesalahan asumsi bahwa data memiliki hubungan yang linear. Selain itu, fitur yang digunakan mungkin tidak memiliki korelasi yang cukup tinggi dengan nilai yang kita coba prediksi,

Nilai RMSE yang besar ini dapat disebabkan karena hanya menggunakan satu model atau fungsi saat melakukan prediksi. Hal ini dapat mengakibatkan hasil prediksi yang kurang tepat, sehingga data yang ada perlu dilatih ulang untuk menghasilkan model prediksi yang baru. Untuk meningkatkan keakuratan nilai prediksi, dapat ditambahkan beberapa faktor / variabel lain yang mempengaruhi productivity seperti: ready hour, Produksi OB, Match Factor dan sebagainya.

4 KESIMPULAN

Berdasarkan hasil percobaan yang dilakukan sebelumnya, maka dapat diambil kesimpulan bahwa metode linear regression cukup baik diterapkan pada kasus prediksi. Untuk performa model linear regression yang dihitung menggunakan root mean square error dihasilkan nilai sebesar 87,518. Hal tersebut menandakan performa model yang dihasilkan kurang baik.

Kekurangan dari metode ini adalah penggunaan satu model atau fungsi saja dalam perhitungan prediksi. Hal ini akan mengakibatkan ketidakakuratan prediksi apabila suatu ketika terjadi lonjakan atau penurunan yang cukup besar pada spotting time & Loading Time, sehingga perlu dilakukan proses training ulang untuk membuat model yang baru. Untuk meningkatkan akurasi nilai prediksi, dapat ditambahkan beberapa faktor / variabel lain yang mempengaruhi Productivitas Alat Gali, seperti: ready hour, Produksi OB, Match Factor dan sebagainya.

DAFTAR PUSTAKA

1. Fitri Marisa, "Edukasi Data Mining (Konsep Dan Penerapan)," Jurnal Teknologi Informasi Vol. 4 No. 2, Universitas Widyagama Malang.
2. Vincentius Riandaru Prasetyo, Hamzah Lazuardi, Aldo Adhi Mulyono, Christian Lauw "Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Regresi Linier," Jurnal Nasional Teknologi dan Sistem Informasi VOL. 07 NO. 01 (2021) 008-01.
3. D. Aprillia, D. A. Baskoro, L. Ambarwati dan I. W. S. Wicaksana, Belajar Data Mining dengan RapidMiner, Jakarta, 2013.
4. Herdianto, "Prediksi Kerusakan Motor Induksi Menggunakan Metoda Jaringan Saraf Tiruan" Universitas Sumatra Utara, 2013.
5. Riza Fahlevi, Antonius Yadi Kuntoro, Lasman Efendi dan Ridatu Oca Nitra "Prediction of Employee Attendance Factors Using C4.5 Algorithm, Random Tree, Random Forest". R. Fahlevi, et al./Semesta Teknika, Vol. 23, No. 1, 39-53, Mei 2020