



PENERAPAN CRISP-DM MENGGUNAKAN MLR K-FOLD PADA DATA SAHAM PT. TELKOM INDONESIA (PERSERO) TBK (TLKM) (STUDI KASUS: BURSA EFEK INDONESIA TAHUN 2015-2022)

Agung Pambudi¹⁾, Zaenal Abidin²⁾, Permata³⁾

Sistem Informasi/Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia
Jl.ZA. Pagar Alam No. 9-11, Labuhan Ratu, Kec. Kedaton, Bandar Lampung

Email: ¹agung_pambudi@teknokrat.ac.id, ²zabin@teknokrat.ac.id, ³permata@teknokrat.ac.id

Abstract

Data mining techniques have been widely adopted over the last few decades, especially in the business and finance domains. To achieve sustainable benefits from this technique, an organization or company must adopt a standardized process for managing data mining projects, mostly using CRISP-DM. Research has shown that this standard process is often not used as it has been defined to meet the various needs of data mining projects. Understanding company goals and converting them into data mining goals is the primary emphasis of the first phase of CRISP-DM, which is referred to as the Business Understanding (BU) phase. This helps decide the design strategy and the resources that are necessary. After that, the Data Understanding (DU) phase will collect and investigate the preliminary data in order to acquire a better understanding of the data quality. After this, different activities in the Data Preparation (DP) stage are carried out in order to build the final dataset from the raw dataset. These activities may include record and feature selection, data transformation and cleaning for modeling tools, or other similar tasks. The Modeling (MO) phase includes the selection and use of a number of different modeling techniques to the dataset that has been prepared. During the Evaluation (EV) phase, the performance of the model is assessed and reviewed before being placed within the framework of the business objectives. The process of deploying the model in an end-user context is then described in the Deployment (DE) phase of the project. To increase understanding of how the data mining process is used as a reference standard, expanded and adapted in practice, this research reports a case study in finance, especially stock trading. The case study was conducted based on documentation from a portfolio of data mining projects, supplemented by data sourced directly from the Indonesia Stock Exchange. The results reveal that CRISP-DM has impacts and the mechanisms used to predict the stock of PT. Telkom Indonesia (Persero) Tbk (TLKM). This study provides an R2 value of 100% with a MAPE value of 0,0013% when implementing CRISP-DM or similar processes in business and finance.

Keyword: CRISP-DM, Data Mining, Indonesia Stock Exchange, Business, Stock Trading

Abstrak

Teknik data mining telah diadopsi secara luas selama beberapa dekade terakhir, terutama dalam domain bisnis dan keuangan. Untuk mencapai manfaat berkelanjutan dari teknik ini, sebuah organisasi atau perusahaan harus mengadopsi proses standar untuk mengelola proyek penambangan data, sebagian besar menggunakan CRISP-DM. Penelitian telah menunjukkan bahwa proses standar ini sering tidak digunakan seperti yang telah ditentukan untuk memenuhi berbagai kebutuhan proyek data mining. Fase awal CRISP-DM, Business Understanding (BU), berfokus pada pemahaman tujuan bisnis dan ترجمahan ke dalam tujuan data mining untuk menentukan rencana desain dan sumber daya yang diperlukan. Fase Data Understanding (DU) kemudian mengumpulkan dan mengeksplorasi data awal untuk mendapatkan wawasan tentang kualitas data. Dataset akhir kemudian dibuat dari dataset mentah melalui berbagai kegiatan dalam tahap Data Preparation (DP), seperti pemilihan catatan dan fitur atau transformasi dan pembersihan data untuk alat pemodelan. Beberapa teknik pemodelan kemudian dipilih pada fase Modeling (MO) dan diterapkan pada dataset yang telah disiapkan. Kinerja model dievaluasi dalam fase Evaluasi (EV) dan dimasukkan ke dalam konteks tujuan bisnis. Fase Deployment (DE) kemudian menjelaskan proses penerapan model dalam konteks pengguna akhir. Untuk meningkatkan pemahaman tentang bagaimana proses data mining digunakan sebagai standar acuan, diperluas dan diadaptasi dalam praktik, penelitian ini melaporkan studi kasus di bidang keuangan khususnya perdagangan saham. Studi kasus dilakukan berdasarkan dokumentasi dari portofolio proyek data mining, dilengkapi dengan data yang bersumber langsung dari Bursa Efek Indonesia. Hasilnya mengungkapkan dalam CRISP-DM memiliki dampak dan mekanisme yang digunakan untuk mengatasi prediksi saham pada PT. Telkom Indonesia (Persero) Tbk (TLKM). Studi ini memberikan nilai R^2 sebesar 100% dengan nilai MAPE sebesar 0,0013% dengan menerapkan CRISP-DM di bidang bisnis dan keuangan.

Kata Kunci: *CRISP-DM, Data Mining, Indonesia Stock Exchange, Business, Stock Trading*

1. PENDAHULUAN

Saham adalah sebuah investasi berbentuk surat berharga yang merupakan bukti sah dibandingkan jenis investasi lainnya. Orang tertarik pada saham karena "pertumbuhan modal jangka panjang (*capital gain*), *dividen*, dan perlindungan nilai tukar mata uang terhadap inflasi". Di sektor telekomunikasi, PT. Telkom Indonesia (Persero) Tbk (TLKM) masih menjadi emiten yang paling menonjol. Ada 45 saham paling likuid di Bursa Efek Indonesia yang masuk ke daftar LQ45, termasuk saham PT. Telkom Indonesia (Persero) Tbk (TLKM) yang juga merupakan saham *grade A*. Beberapa *tools* atau alat baru telah dirancang oleh para investor dengan bantuan software RapidMiner menggunakan total data yang terkumpul sebanyak 1.415 *record*. Meskipun penggunaan indikator ini memberi mereka beberapa informasi tentang harga penutupan saham, tetapi masih sulit untuk memiliki prediksi akurat dari tren harian hingga mingguan. Indikator teknikal tersebut dapat mengandung apakah suatu tren akan berlanjut. Sebuah penelitian oleh Shakhla, S., Shah, B., Shah, N., Unadkat, V., & Kanani, P. (2018) menggunakan model *multiple linear regression* (regresi linier berganda) yang dapat memprediksi harga saham pada waktu yang akan datang menggunakan data harga saham yang diperoleh dari *Yahoo! Finance*, diperoleh hasil akurasi menggunakan R^2 sebesar 0,914.

Penelitian terdahulu belum banyak menggunakan dan mengoptimalkan model ini secara

efektif. Untuk itu penelitian ini menggunakan *CRISP-DM* dan *K-Fold* sebagai metodologi yang direkomendasikan dan digunakan secara luas dengan melakukan studi kasus pada Bursa Efek Indonesia di mana peneliti mempelajari koleksi proyek penambangan data berdasarkan dokumentasi dan wawancara semi-terstruktur dengan para pemangku kepentingan proyek.

2. METODE PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menggunakan metodologi data mining *CRISP-DM* (*Cross Industry Standard Process for Data Mining*) pertama kali didirikan pada akhir 1990-an oleh empat perusahaan: *Integral Solutions Ltd.* (penyedia solusi data mining komersial), *NCR* (penyedia database), *DaimlerChrysler* (penyedia mobil produsen), dan *OHRA* (perusahaan asuransi) dan ditetapkan sebagai proses standar untuk strategi pemecahan masalah data mining. Di Indonesia, *CRISP-DM* dijadikan acuan Standar Kompetensi Kerja Nasional: KepMen Ketenagakerjaan No 299 tahun 2020.



MENTERI KETENAGAKERJAAN
REPUBLIK INDONESIA

KEPUTUSAN MENTERI KETENAGAKERJAAN

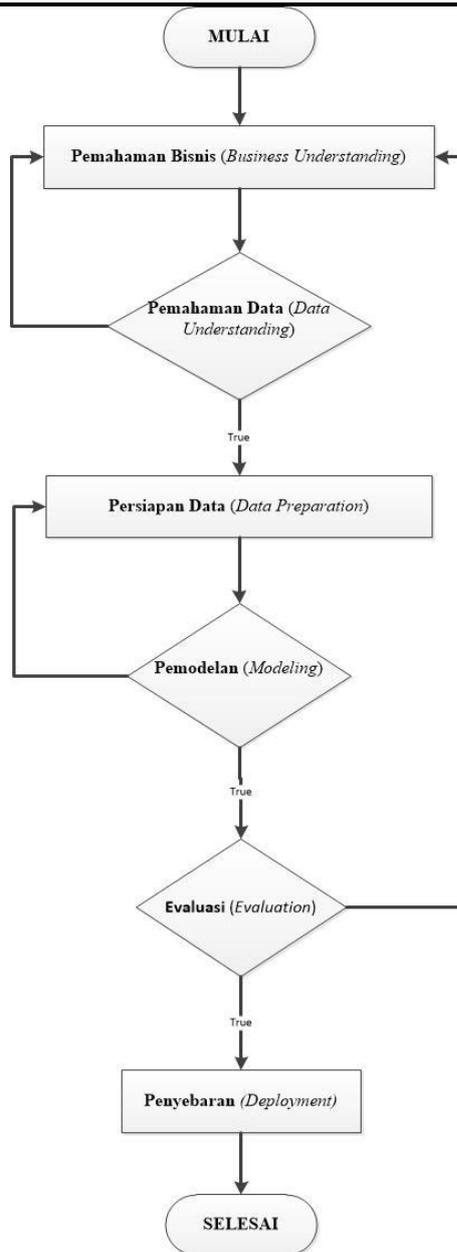
REPUBLIK INDONESIA

NOMOR 299 TAHUN 2020

TENTANG

PENETAPAN STANDAR KOMPETENSI KERJA NASIONAL INDONESIA
KATEGORI INFORMASI DAN KOMUNIKASI GOLONGAN POKOK AKTIVITAS
PEMROGRAMAN, KONSULTASI KOMPUTER DAN KEGIATAN YANG
BERHUBUNGAN DENGAN ITU (YBDI) BIDANG KEAHLIAN *ARTIFICIAL
INTELLIGENCE* SUBBIDANG *DATA SCIENCE*

Gambar 1. SKKNI *Data Science*.



Gambar 2. Bagan alur CRISP-DM.

1. Pemahaman Bisnis (*Business Understanding*)

Selama tahap awal ini, fokus utama adalah memperoleh pemahaman tentang tujuan dan persyaratan bisnis. Setelah itu informasi yang diperoleh digunakan untuk merumuskan definisi masalah data mining dan rencana awal yang ditujukan untuk mencapai tujuan bisnis. Pemahaman bisnis membutuhkan komunikasi dan kejelasan. Kesulitannya di sini adalah bahwa para pemangku kepentingan memiliki tujuan, bias, dan modalitas yang berbeda dalam menghubungkan informasi. Tanpa perspektif yang jelas, ringkas, dan lengkap tentang apa tujuan proyek sehingga sumber daya akan terbuang sia-sia.

2. Pemahaman Terhadap Data (*Data Understanding*)

Pemahaman data meliputi kegiatan untuk mempersiapkan, menilai kebutuhan data, dan mencakup pengumpulan data. Ada hubungan erat antara pemahaman bisnis dan pemahaman data yaitu

merumuskan masalah data mining dan pemahaman tentang data yang tersedia.

Data dikumpulkan pada tahap proses ini. Pemahaman tentang apa yang diinginkan dan dibutuhkan bisnis akan menentukan data apa yang dikumpulkan, dari sumber apa dan dengan metode apa.

3. Persiapan Data (*Data Preparation*)

Persiapan data berlangsung setelah data dikumpulkan dan perlu diidentifikasi, dipilih, dibersihkan dan kemudian disusun ke dalam bentuk ataupun format yang diinginkan. Setelah dataset dipilih, maka harus diperiksa untuk kasus yang meragukan, hilang, atau ambigu.

Persiapan data kemungkinan akan dilakukan beberapa kali, dan tidak dalam urutan yang ditentukan.

4. Pemodelan (*Modeling*)

Pemodelan merupakan penerapan algoritma untuk menemukan, mengidentifikasi dan menampilkan pola. Setelah siap untuk digunakan, data harus diekspresikan melalui model apa pun yang sesuai, memberikan wawasan yang berarti, dan semoga pengetahuan baru. Penggunaan model mengungkapkan pola dan struktur dalam data yang memberikan wawasan tentang fitur yang menarik. Pemilihan model didasarkan pada tipe data, dan beberapa model umumnya digunakan untuk memperkirakan, memprediksi, mengklasifikasikan, mengelompokkan, atau melihat asosiasi. Adapun penggunaan model pada penelitian ini adalah regresi linier berganda (*multiple linear regression*) yang dirancang untuk melakukan prediksi.

5. Evaluasi (*Evaluation*)

Model yang dipilih harus diuji. Langkah evaluasi dilakukan untuk mengetahui performansi model yang dihasilkan, model yang dihasilkan pada tahap sebelumnya diuji dengan menggunakan data testing sebagai data masukan. Ada banyak metode untuk mengevaluasi model, seringkali tergantung pada jenis tugas yang dilakukan dan informasi yang ingin diketahui. Pada penelitian ini akan menggunakan metrik paling umum untuk mengevaluasi prediksi:

- *R²* atau R Kuadrat (*R Squared*)
- *ROC-AUC*
- *Mean Absolute Error (MAE)*
- *Mean Squared Error (MSE)*
- *Root Mean Squared Error (RMSE)*
- *Mean Absolute Percentage Error (MAPE)*

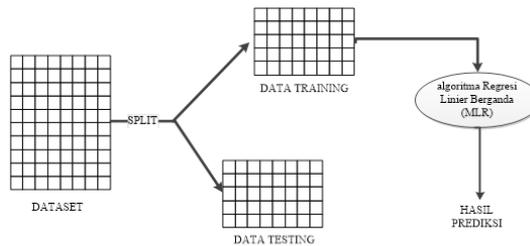
6. Penyebaran (*Deployment*)

Pada langkah penyebaran (*deployment*), pengetahuan atau informasi yang diperoleh disajikan dalam bentuk khusus untuk digunakan pengguna akhir. Fase penyebaran dapat berupa pembuatan aplikasi atau laporan sederhana.

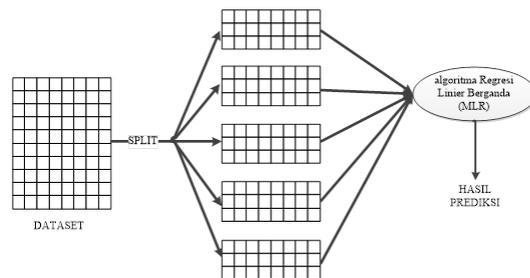
2.2 Penentuan Data Latih dan Data Uji

Tujuan dari pengujian model adalah untuk mengetahui apakah solusi yang dihasilkan oleh model tersebut valid. Membentuk data training sebagai pembentuk model dan data testing untuk pengujian model. Dataset dibagi menjadi dua bagian, data latih dan data uji. Pengujian dilakukan dengan jumlah data latih dan data uji yang berbeda. dijelaskan sebagai berikut: 80% data pelatihan dan 20% data uji (Fadilah, W. R. U., Agfiannisa, D., & Azhar, Y 2020). Pengolahan dataset yang dibagi menjadi dataset training dan testing 80%:20% dengan jumlah 1834 record yang terdiri dari beberapa

atribut. Data tersebut untuk melihat lebih jelasnya bisa dilihat di <https://doi.org/10.34740/KAGGLE/DSV/4116840>.



Gambar 3. Membagi *dataset* menjadi *data training* dan *data test*



Gambar 4. Membagi dataset menggunakan *K-Fold Cross-Validation*

3. HASIL DAN PEMBAHASAN

3.1 Business Understanding

Investor perlu mengambil tindakan yang tepat untuk mengurangi risiko yang ada dengan memahami naik turunnya saham PT. Telkom Indonesia (Persero) Tbk (TLKM) dimasa depan. Selain itu, investor perlu mengantisipasi langkah-langkah kebijakan yang tepat untuk mengambil keputusan membeli ataupun menjual saham PT. Telkom Indonesia (Persero) Tbk (TLKM) dikarenakan harga saham yang fluktuatif.

3.2 Data Understanding

Pemahaman data bergantung pada pemahaman bisnis. Data dikumpulkan pada tahap proses ini. Pemahaman tentang apa yang diinginkan dan dibutuhkan bisnis akan menentukan data apa yang dikumpulkan, dari sumber apa dan dengan metode apa. *CRISP-DM* menggabungkan tahapan persyaratan data dan pengumpulan data.

Data yang diperoleh sesuai dengan kebutuhan penelitian dapat diunduh pada alamat <https://doi.org/10.34740/KAGGLE/DSV/4116840> dan sudah diolah kebentuk CSV.

Pada penelitian ini fitur yang akan di fokuskan yaitu *previous*, *open_price*, *first_trade*, *high*, *low*,

close, index_individual, offer, bid. Masing-masing fitur tersebut dijelaskan pada tabel dibawah ini.

Tabel 1. Deskripsi masing-masing fitur

Fitur	Indikator	Ukuran	Tipe Data
<i>close</i>	Harga pada saat jam kerja bursa saham telah berakhir tiap hari.	Tingkat harga	Float
<i>previous</i>	Harga penutupan hari bursa sebelumnya	Tingkat harga	Float
<i>open_price</i>	Harga saat awal dimulainya bursa saham tiap hari.	Tingkat harga	Float
<i>first_trade</i>	Harga saham yang dijual pada awal perdagangan saham sebuah emiten	Tingkat harga	Float
<i>high</i>	Harga tertinggi yang diperoleh emiten di bursa saham per hari.	Tingkat harga	Float
<i>low</i>	Harga terendah yang diperoleh emiten di bursa saham per hari.	Tingkat harga	Float
<i>index_individual</i>	Pergerakan harga masing-masing saham terhadap harga dasarnya atas masing-masing saham suatu emiten	Tingkat harga	Float
<i>offer</i>	Nilai penawaran harga jual pada hari tersebut	Tingkat harga	Float
<i>bid</i>	Nilai penawaran harga beli pada hari tersebut	Tingkat harga	Float

3.3 Data Preparation

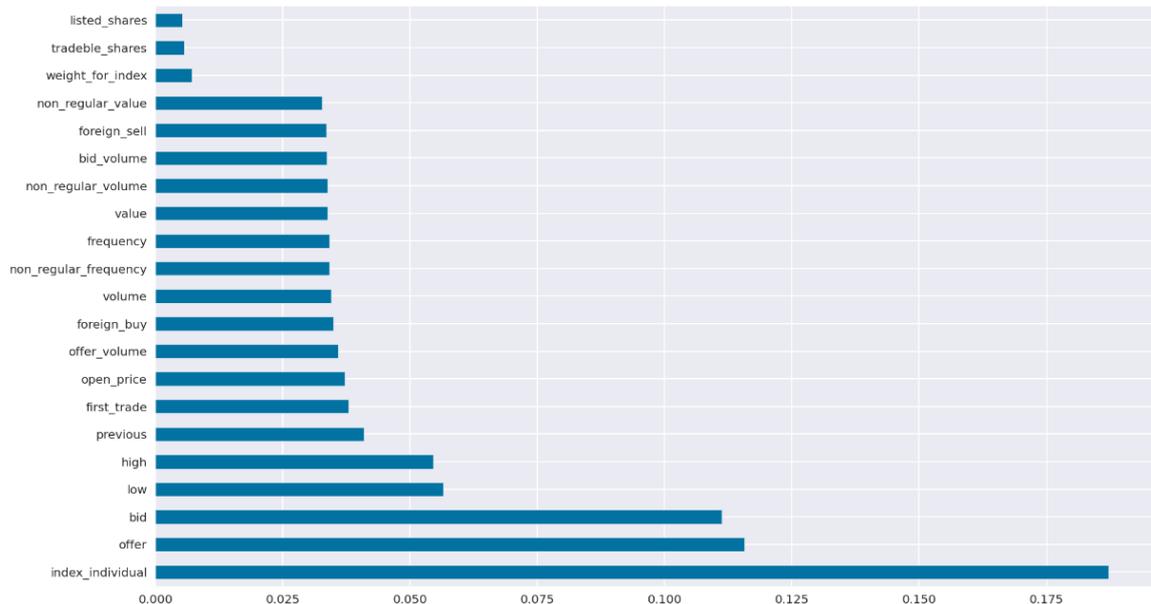
Feature Importance (pentingnya fitur) digunakan untuk menilai setiap fitur data, semakin tinggi skornya, semakin penting atau relevan fitur tersebut terhadap variabel target. *Feature Importance* adalah *class* bawaan yang dilengkapi dengan *classifier* berbasis pohon, dan peneliti akan menggunakan *classifier* berbasis pohon untuk mengekstrak 25 fitur dari dataset PT. Telkom Indonesia (Persero) Tbk (TLKM).

```
from sklearn.ensemble import ExtraTreesClassifier
import matplotlib.pyplot as plt
model = ExtraTreesClassifier()
model.fit(X,y)

print(model.feature_importances_)

feat_importances = pd.Series(model.feature_importances_, index=X.columns)
feat_importances.nlargest(25).plot(kind='barh')
plt.show()
```

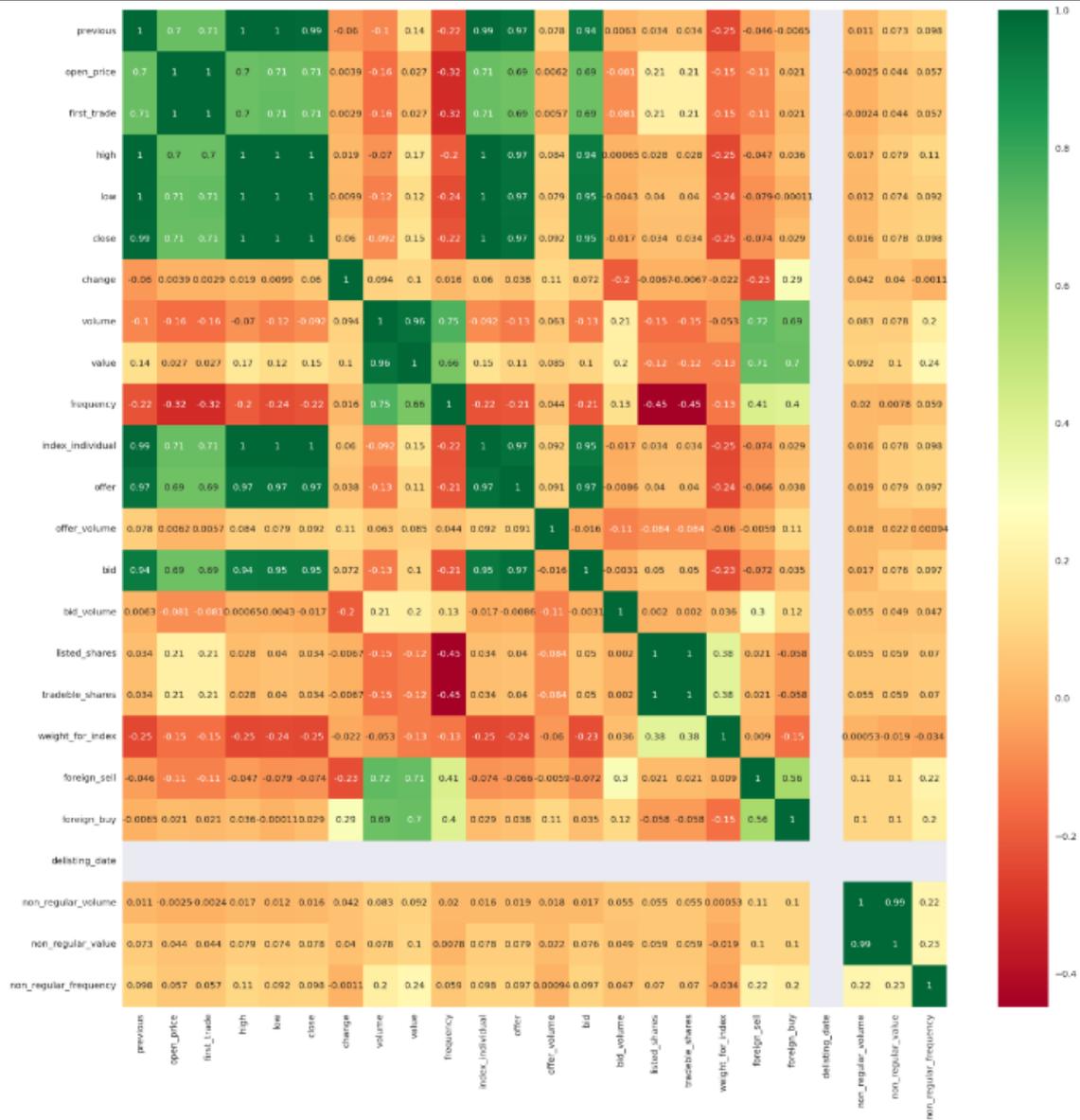
```
[0.04093328 0.0371906 0.03782011 0.05451715 0.0565882 0.03445955
0.03378764 0.03409028 0.18709152 0.11562202 0.0358991 0.11120055
0.03364001 0.00521921 0.0056441 0.00711307 0.03352343 0.03498758
0.0337608 0.03276207 0.03414972]
```



Gambar 5. Pemilihan *feature* menggunakan *classifier* berbasis pohon

Terlihat pada gambar diagram diatas, fitur `index_individual` memiliki skor paling tinggi yaitu 0,18709152 diantara fitur yang lain.

Korelasi menggambarkan bagaimana fitur berhubungan satu sama lain dengan variabel target. Korelasi bisa positif atau negatif. Peneliti menggunakan *python library* yaitu *seaborn* untuk menggambarkan peta panas dari fitur yang berkorelasi.



Gambar 6. Mencari korelasi menggunakan Matriks Korelasi dan ditampilkan dalam bentuk *Heatmap*.

3.4 Modeling

Modeling melibatkan penggunaan model pembelajaran mesin (*machine learning*) untuk memahami pola data dan memperoleh wawasan. Wawasan bisa berupa visualisasi pola atau prediksi nilai masa depan. Tahapan ini merupakan bagian paling menarik dari proyek ilmu data (*data science*) karena melibatkan pembelajaran mesin sebagai salah satu komponen terpentingnya.



Gambar 7. Perbandingan *data testing* pada harga aktual dan harga prediksi

Tabel 2. Hasil perhitungan MAE dan MSE pada *data testing*

Metode	Test MAE	Test MSE
Multiple Linear Regression	0,04732	0,00300

3.5 Evaluation

Tahap evaluasi akan memeriksa apakah pengetahuan yang telah ditemukan tersebut baru dan menarik. Hanya model yang disetujui yang dipertahankan, dan seluruh proses ditinjau kembali untuk mengidentifikasi tindakan alternatif mana yang dapat diambil untuk meningkatkan hasil yang terbaik.

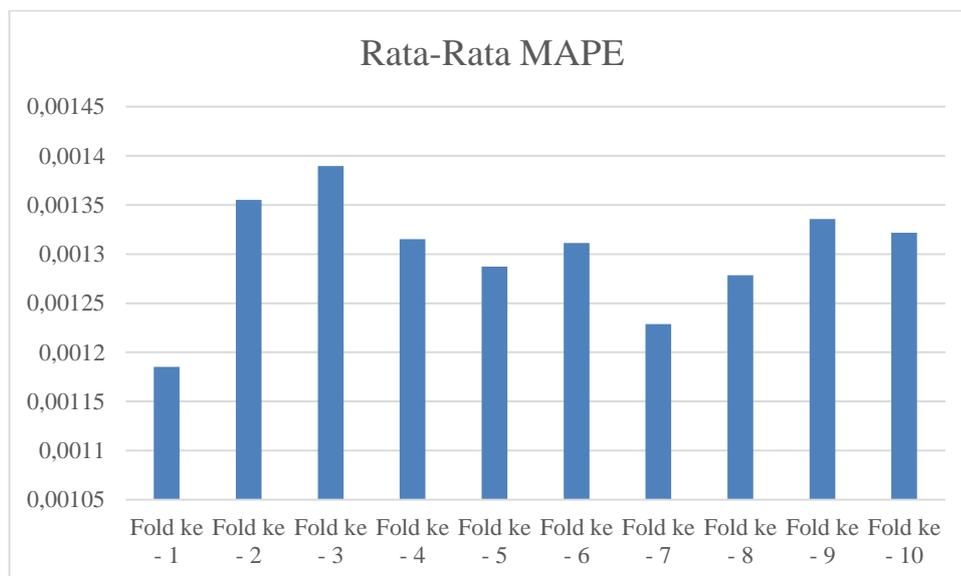
Tabel 3. Hasil perhitungan akurasi pada algoritma MLR

R2	ROC-AUC	MAE	MSE	RMSE	MAPE
1,000	1,000	0,0473	0,003	0,0548	0,0013

Tabel 4. Hasil perhitungan akurasi pada algoritma MLR + KFold

R2	ROC-AUC	MAE	MSE	RMSE	MAPE
----	---------	-----	-----	------	------

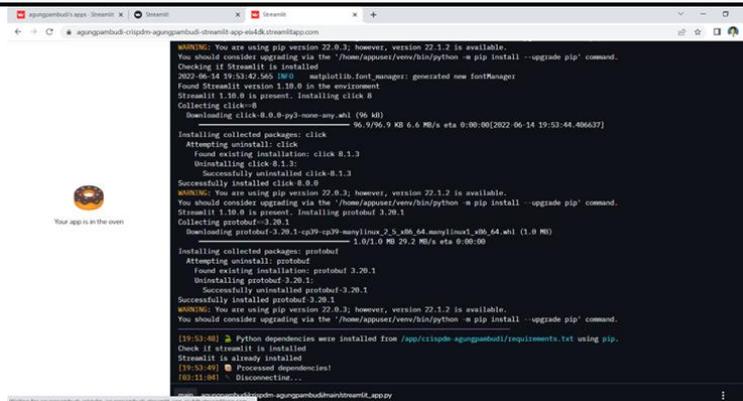
	R2	ROC-AUC	MAE	MSE	RMSE	MAPE
Fold ke - 1	1,000	1,000	0,042487	0,00251	0,05011	0,001185
Fold ke - 2	1,000	1,000	0,049366	0,00322	0,05672	0,001355
Fold ke - 3	1,000	1,000	0,049461	0,00317	0,05629	0,00139
Fold ke - 4	1,000	1,000	0,047794	0,0031	0,05564	0,001315
Fold ke - 5	1,000	1,000	0,045573	0,00284	0,05331	0,001287
Fold ke - 6	1,000	1,000	0,046041	0,00281	0,05303	0,001311
Fold ke - 7	1,000	1,000	0,044149	0,00268	0,05174	0,001229
Fold ke - 8	1,000	1,000	0,046132	0,00285	0,05337	0,001278
Fold ke - 9	1,000	1,000	0,047901	0,00297	0,05452	0,001336
Fold ke - 10	1,000	1,000	0,047148	0,00306	0,05534	0,001322



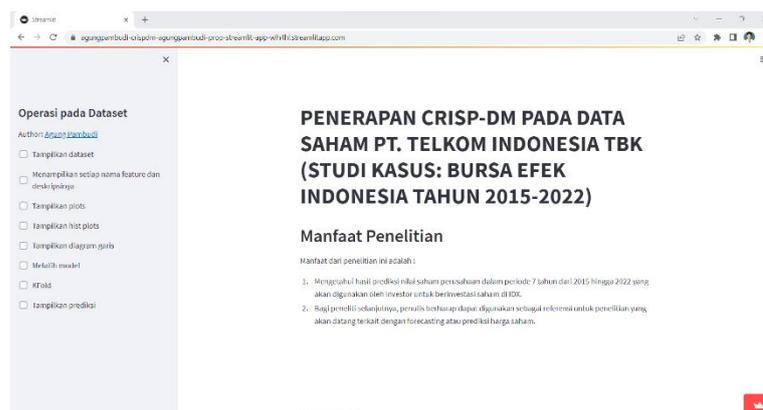
Gambar 8. Hasil rata-rata perhitungan MAPE pada MLR + Kfold

3.6 Deployment

Deployment model ke dalam sistem sering kali memerlukan pengodean ulang model untuk membuatnya lebih cepat atau lebih kompatibel dengan sistem yang sudah ada di lingkungan produksi.



Gambar 9. Proses *deployment* menggunakan *Streamlit*



Gambar 10. Hasil *deployment* ke *Streamlit Cloud* (<https://agungpambudi.streamlit.app>).

4. KESIMPULAN

Penerapan CRISP-DM merupakan langkah yang tepat untuk dipilih pada penelitian ini, karena dapat menyelesaikan sebuah masalah secara lebih terstruktur dan menghasilkan tujuan yang tepat bagi investor. Hasil pemodelan menggunakan *Multiple Linear Regression* dan *K-Fold* didapatkan nilai R^2 sebesar 100% dengan nilai *MAPE* sebesar 0,0013%. Model yang diperoleh memiliki nilai *ROC-AUC* sebesar 1,000 dan nilai *RMSE* sebesar 0,0548 yang tergolong regresi luar biasa (*Excellent Regression*). Nilai maksimum yang diperoleh dari fitur *index_individual* adalah 0,18388841, sehingga menjadi fitur yang memiliki pengaruh besar pada proses regresi, diikuti dengan fitur *offer* memiliki skor 0.11562202, fitur *bid* memiliki skor 0.11120055, fitur *low* memiliki skor 0.0565882, fitur *high* memiliki skor 0.05451715, fitur *previous* memiliki skor 0.04093328, fitur *first_trade* memiliki skor 0.03782011, fitur *open_price* memiliki skor 0.0371906, fitur *offer_volume* memiliki skor 0.0358991, fitur *foreign_buy* memiliki skor 0.03498758, fitur *volume* memiliki skor 0.03445955, fitur *non_regular_frequency* memiliki skor 0.03414972, fitur *frequency* memiliki skor 0.03409028, fitur *value* memiliki skor 0.03378764, fitur *non_regular_volume* memiliki skor 0.0337608, fitur *bid_volume* memiliki skor 0.03364001, fitur *foreign_sell* memiliki skor 0.03352343, fitur *non_regular_value* memiliki skor 0.03276207, fitur *weight_for_index* memiliki skor 0.00711307, fitur *tradeable_shares* memiliki skor 0.0056441, dan skor yang paling kecil yaitu pada fitur *listed_shares* memiliki skor 0.00521921 dari 21 fitur yang telah diseleksi. Prediksi saham dalam investasi diperlukan untuk meminimalkan kerugian yang akan terjadi sehingga optimalisasi keuntungan dapat dicapai.

REFERENCES

- [1] Rakhmawati, D. (2017). Analisis Regresi Dummy Untuk Mengetahui Pengaruh Sektor Bursa Efek Indonesia (Bei) Terhadap Indeks Harga Saham Lq45. *Jurnal Pro Bisnis*, 10(2), 1–15.
- [2] Wahid, A. (2018). Analisis Transaksi Perdagangan Saham untuk Mendeteksi Manipulasi Harga Saham dengan Outlierr Detection. 1–115.
- [3] Ariesanto Akhmad, E. P. (2020). Data Mining Menggunakan Regresi Linear untuk Prediksi Harga Saham Perusahaan Pelayaran. *Jurnal Aplikasi Pelayaran Dan Kepelabuhanan*, 10(2), 120. <https://doi.org/10.30649/japk.v10i2.83>
- [4] Priyadi, I., Santony, J., & Na'am, J. (2019). Data Mining Predictive Modeling for Prediction of Gold Prices Based on Dollar Exchange Rates, Bi Rates and World Crude Oil Prices. *Indonesian Journal of Artificial Intelligence and Data Mining*, 2(2), 93. <https://doi.org/10.24014/ijaidm.v2i2.6864>
- [5] Sunaryo. (2020). Pengaruh Volume Penawaran Saham Terhadap Risiko Sistematis dan Implikasinya Terhadap Harga Saham. *Mediastima*, 1(2), 105–112.
- [6] Nababan, L. U. (2019). Penerapan Model Regresi Data Panel Pada Analisis Harga Saham Perusahaan Batubara. *Akuntabel*, 16(1), 81–97.
- [7] Cokrodiharjo, V. R., & Chalid, D. A. (2021). Peramalan Harga Saham Menggunakan Metode Support Vector Machine (SVM): Studi Kasus Saham di Bursa Efek Indonesia (BEI). *Jpmb* 61, 3(1), 61–74.
- [8] Widodo, E., & Irmayanti, N. A. (2019). Comparison of Truncated Spline Regression with Simple Linear Regression metode on The Stock Price of Mining Company in Indonesia (Case study: ADRO . JK, ITMG . JK, and ANTM . JK) Perbandingan Metode Regresi Spline Truncated dengan Regresi Linear Sede. *Jurnal Ilmu-Ilmu MIPA*, 19. <https://doi.org/10.20885/eksakta.vol19.iss2.art5>
- [9] Adhiva, J., Putri, S. A., & Setyorini, S. G. (2020). Prediksi Hasil Produksi Kelapa Sawit Menggunakan Model Regresi Pada PT . Perkebunan Nusantara V. *Seminar Nasional Teknologi Informasi, Komunikasi Dan Industri*, 155–162.
- [10] Izzah, A., & Widyastuti, R. (2017). Prediksi Harga Saham Menggunakan Improved Multiple Linear Regression untuk Pencegahan Data Outlier. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 2(3), 141–150. <https://doi.org/10.22219/kinetik.v2i3.268>
- [11] Savitri, U. S., Saepudin, D., & Umbara, R. F. (2019). Prediksi Harga Saham serta Pemberian Keputusan Jual Beli Menggunakan Metode Regression Tree dan Self Organizing Map (SOM). *E-Proceeding of Engineering* :, 6(1), 2527–2538.
- [12] Khaira, U., Utomo, P. E. P., Suratno, T., & Gulo, P. C. S. (2021). Prediksi Indeks Harga Saham Gabungan (IHSG) Menggunakan Algoritma Autoregressive Integrated Moving Average (ARIMA). *JUSS (Jurnal Sains Dan Sistem Informasi)*, 2(2), 11–17. <https://doi.org/10.22437/juss.v2i2.8449>
- [13] Puspitaningtyas, Z. (2017). Prediksi pergerakan harga saham menggunakan analisis profitabilitas. *Performance: Jurnal Personalia, Financial ...*, June 2016, 29–35. <https://doi.org/10.20884/1.performance.2017.24.2.697>
- [14] Laksamana, M. A. H., Amroni, & Toscani, A. N. (2021). Penerapan Data Mining untuk Memprediksi Jumlah Total Produksi Hcl Pada Perusahaan PT . Lontar Papyrus Menggunakan Algoritma Regresi Linier Berganda. *Jurnal Ilmiah Mahasiswa Teknik Informatika*, 3(2), 187–198.
- [15] Purwadi, P., Ramadhan, P. S., & Safitri, N. (2019). Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Deli Serdang. *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika Dan Komputer)*, 18(1), 55. <https://doi.org/10.53513/jis.v18i1.104>
- [16] Suryanto, A. A. (2019). Penerapan Metode Mean Absolute Error (Mea) Dalam Algoritma Regresi Linear Untuk Prediksi Produksi Padi. *Saintekbu*, 11(1), 78–83. <https://doi.org/10.32764/saintekbu.v11i1.298>
- [17] Fahlepi, M., & Widjaja, A. (2019). Penerapan Metode *Multiple Linear Regression* Untuk Prediksi Harga Sewa Kamar Kost. *Jurnal Strategi*, 1(November), 615–629.
- [18] Ayuni, G. N., & Fitriana, D. (2019). Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ. *Jurnal Telematika*, 14(2), 79–86. <https://journal.ithb.ac.id/telematika/article/view/321>



- [19] Merfin, M., & Oetama, R. S. (2019). Prediksi Harga Saham Perusahaan Perbankan Menggunakan Regresi Linear Studi Kasus Bank BCA Tahun 2015-2017. *Ultimatics: Jurnal Teknik Informatika*, 11(1), 11–15. <https://doi.org/10.31937/ti.v11i1.1239>
- [20] Solihat, I., Jasmi, J., & Fachruddin, F. (2020). Prediksi Jumlah Produksi Kelapa Sawit Dengan Regresi. *Snastikom*, 1(1), 17–26.
- [21] Pangestu, A. A., Info, A., Regression, M. L., & Volume, S. (2021). Prediksi Volume Penjualan Gas Pt Pgn (Persero) Menggunakan Regresi Linear Berganda. 1(1), 14–18.
- [22] Agustia, D., Muhammad, N. P. A., & Permatasari, Y. (2020). Earnings management, business strategy, and bankruptcy risk: evidence from Indonesia. *Heliyon*, 6(2), e03317. <https://doi.org/10.1016/j.heliyon.2020.e03317>
- [23] Kumar, R., Kumar, P., & Kumar, Y. (2020). Time Series Data Prediction using IoT and Machine Learning Technique. *Procedia Computer Science*, 167(2019), 373–381. <https://doi.org/10.1016/j.procs.2020.03.240>
- [24] Sharma, A., Modak, S., & Sridhar, E. (2019). *Data Visualization and Stock Market and Prediction*. 2037–2040.
- [25] Lubis, A. A., Wong, N. P., & Sinaga, F. M. (2020). Prediksi Akurasi Perusahaan Saham Menggunakan SVM dan K-Fold Cross Validation. *Jurnal SIFO Mikroskil*, 21(1), 11–19. <https://mikroskil.ac.id/ejurnal/index.php/jsm/article/view/718>
- [26] Shakhla, S., Shah, B., Shah, N., Unadkat, V., & Kanani, P. (2018). Stock price trend prediction using multiple linear regression. *International Journal of Engineering Science Invention (IJESI)*, 7(10), 29–33. www.ijesi.org
- [27] Chen, S., & Chen, S. (2020). Forecasting Daily Stock Market Return with Multiple Linear Regression Forecasting Daily Stock Market Return with Multiple Linear Regression. 1–10.