Analisis Saham Telkom menggunakan Algorithm K-Nearest Neighbors dan Optimasi Grid Search

First Author #1, Second Author \*2, Third Author #3

# *First-Third Department, First-Third University*

*Address Including Country Name*

1 first.author@first-third.edu

3 third.author@first-third.edu

\**Second Company*

*Address Including Country Name*

2 second.author@second.com

Received on dd-mm-yyyy, revised on dd-mm-yyyy, accepted on dd-mm-yyyy

**Abstract**

An abstract of no more than 200 words should state the purposes of the study, procedures, ﬁndings and conclusions. It should be written in Times New Roman with font size of 9. Abstract should be followed by four to eight key words to assist cross-indexing. The key words should be ordered alphabetically and separated by commas. For your paper to be published in the journal, you must use this document as both an instruction set and as a template into which you can type your own text. If your paper does not conform to the required format, you will be asked to ﬁx it. In the paper, authors are suggested to present their articles in the sections structure: Introduction - Previous Study/Literature Review - The Proposed Method/Algorithm/Procedure speciﬁcally designed (optional) - Research Method Results and Discussion - Conclusion. Authors may present complex proofs of theorems or non-obvious proofs of correctness of algorithms after introduction section (obvious theorems and straightforward proofs of existing theorems are not needed).

**Keywords:** INISTA, journal, MS.Word template

1. Introduction

**M**emiliki peran yang esensial, saham menjadi salah satu kunci menciptakan deviden terhadap suatu perusahaan. Berkembangnya instrumen perdagangan menghadirkan pasar saham sebagai salah satu langkah transaksi yang hadir didalamnya. Persepsi keuntungan merupakan faktor utama yang mempengaruhi minat berinvestasi di saham [1]. Pada kenyataannya terdapat banyaknya faktor yang dapat mempengaruhi fluktuasi nilai didalamnya. Misalnya kemunculnya pandemi covid-19 yang blablablabla menjadi salah satu faktor yang dapat mempengaruhi tingkat fluktuasi yang terjadi [?]. Oleh karenanya diperlukan instrumen yang dapat membantu dalam monitoring arah pergerakan nilainya.

Perkembangan teknologi yang memadai memunculkan menghadirkan bantuan terhadap berbagai masalah. Salah satunya melalui algoritma K-Nearest Neighbors yang memiliki potensi besar untuk meningkatkan efisiensi, efektivitas, dan keamanan sistem keuangan [?]. Algoritma tersebut kerap digunakan sebagai salah satu langkah yang dapat digunakan.

Berdasarkan penelitian sebelumnya, algoritma tersebut mampu … [?]. Penelitian lainnya mengungkapkan bahwa algoritma K-Nearest neighbors dapat …[?]. Penelitian yang dilakukan oleh …, algoritma K-Nearest neighbors memiliki kelebihan … [?]. Dibandingkan dengan algoritma … , algoritma K-Nearest neigbors memiliki keunggulan … [?]. Dengan demikian, algoritma K-Nearest Neighbors dengan metode optimasi *Grid Search,* dapat digunakan sebagai langkah prediksi saham.

Tujuan dari penelitian ini ialah agar dapat menemukan salah satu metode optimasi prediksi dari algoritma K-Nearest Neighbors. Melalui penelitian ini diharapkan dapat membandingkan ataupun menemukan salah satu langkah terbaik dari berbagai metode *machine learning yakni unsupervised learning*.

II. Literature Review

* + - 1. Algoritma K-Nearest Neighbors

Algoritma K-Nearest Neighbors sering kali digunakan sebagai salah satu algoritma *unsupervised learning* dalam *machine learning* dalam memprediksi berbagai kasus berdasarkan .. [?]. Pada penelitian … , algoritma ini merupakan … [?]. Adapun perhitungan algoritma K-Nearest Neighbors ialah sebagai berikut :

Sebagaimana jks merupakan sakjsak dan kejw merupakan kejewkejw

* + - 1. Metode Optimasi *Search Grid*

Algoritma K-Nearest Neighbors sering kali digunakan sebagai salah satu algoritma *unsupervised learning* dalam *machine learning* dalam memprediksi berbagai kasus berdasarkan .. [?]. Pada penelitian … , algoritma ini merupakan … [?]. Adapun perhitungan algoritma K-Nearest Neighbors ialah sebagai berikut :

Sebagaimana jks merupakan sakjsak dan kejw merupakan kejewkejw

III. Research Method

Penelitian yang dilakukan menggunakan metode *mix method*. Dimana pendekatakan penelitian menggunakan metode campuran kualitatif dan kuantitatif sebagai studi literatur yang mendukung penelitian terhadap data yang diteliti [?]. Pada gambar 1 menunjukan tahapan penelitian yang akan dilakukan.

A diagram of a model

Description automatically generated

Gambar. 1. Tahapan penelitian

* + - 1. Pengambilan Data

Penelitian yang dilakukan menggunakan data PT. Telkom … dalam rentang tahun 2021 sampai dengan April 2024 yang didapatkan melalui situs kaggle [?]. Secara keseluruhan, data yang didapatkan yakni sebanyak 805 buah sampel data mentah. Masing – masing sampel terdiri dari berbagai atribut yakni, tanggal, harga buka, harga tertinggi, harga terendah, harga tutup dan volume. Melalui sampel data yang telah didapatkan, diharapkan penelitian dapat memprediksi harga saham dengan mempertimbangkan korelasi antara data dan atribut dari masing - masing sampel yang terdapat didalamnya.

* + - 1. Eksplorasi dan Analisa Data

Eksplorasi terhadap sampel data, digunakan sebagai acuan validalitas terhadap sampel data penelitian yang telah didapatkan. Eksplorasi data digunakan untuk memonitoring kelengkapan nilai – nilai atribut dari masing – masing sampel data penelitian. Metode korelasi *pearson* digunakan untuk menilai seberapa erat hubungan antar variabel dalam sebuah penelitian dengan melihat seberapa besar atau kecilnya nilai koefisien (r) yang dihasilkan dari analisis data [?]. Metode korelasi *pearson* dapat dituliskan sebagai berikut [?] : . Selain dengan metode korelasi *pearson*, teknik *engineer feature*  dan *feature selection* digunakan untuk menilai tingkat hubungan antara fitur dan target. Adapun teknik *engineer feature* dan *feature selection* dapat dituliskan sebagai berikut :

* + - 1. Pengembangan Model

Langkah selanjutnya penelitian mengacu pada pembuatan model terhadap data yang telah di eksplorasi sebelumnya. Pada langkah ini penelitian membagi 80% dari total data yang telah didapatkan menjadi data *training* dan 20% dari data yang tersisa menjadi data *testing* terhadap model. Nilai *grid* yang menjadi parameter model prediksi terbaik diambil melalui metode *Grid Search* sebagai optimasi pencarian *hyperparameter* terbaik *grid* yang digunakan pada model prediksi.

* + - 1. Pengujian Akurasi Prediksi

Berdasarkan model prediksi yang telah dibentuk sebelumnya, model diuji menggunakan 20% dari total data *testing* yang telah ditetapkan sebelumnya. Pengujian model melibatkan metode *Mean Squared Eror* (MSE), *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan *R-squared* sebagai metode uji dari prediksi yang telah dihasilkan model sebelumnya. Adapun perhitungan berdasarkan metode – metode tersebut sebagai berikut :

MSE= x (?)

RMSE = (?)

MAE = x (?)

*R-sequared*= (2) (?)

Dimana n meripakan jumlah observasi dan Σ sebagai notasi penjumlahan dan | sebagai notasi nilai absolut.

IV. Results And Discussion

Dari hasil penelitian yang dilakukan, dilakukan pensortiran data. Data yang diambil ialah data yang berada dalam rentang tahun 2021 hingga 2024. Dari 4732 buah data, didapatkan sebanyak 805 data berdasarkan tahun yang diinginkan. Melalui visualisasi data, dalam rentang 2021 hingga 2024 ditemukan bahwa jumlah harga tertinggi terdapat pada tanggal 23 Agustus 2022 dengan harga akhir mencapai 4558 dari harga buka sebesar 4471.



Gambar 2. Visualisasi data

Eksplorasi data yang dilakukan digunakan untuk mengetahui jumlah besaran aktivitas pergerakan dari data yang telah diambil sebelumnya. Dari hasil eskplorasi data yang telah dilakukan pada penelitian, ditemukan hasil deskripsi statik yang ditunjukan pada tabel 1.

Tabel 1

Deskripsi statik

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Open** | **High** | **Low** | **Close** | **Volume** |
| *Count* | 805 | 805 | 805 | 805 | 805 |
| *Mean* | 3683.178882 | 3719.972671 | 3643.485714 | 3681.972671 | 1.088572 |
| Deviasi Standar | 439.724117 | 440.637165 | 435.799147 | 439.315061 | 7.695868 |
| Nilai Terendah | 2758.000000 | 2793.000000 | 2679.000000 | 2740.000000 | 7.959900 |
| Kuartil Pertama | 3397.000000 | 3434.000000 | 3351.000000 | 3397.000000 | 6.516750 |
| Kuartil Kedua | 3804.000000 | 3841.000000 | 3767.000000 | 3800.000000 | 8.974280 |
| Kuartil Ketiga | 3980.000000 | 4010.000000 | 3944.000000 | 3980.000000 | 1.258177 |
| Nilai Tertinggi | 4510.000000 | 4577.000000 | 4462.000000 | 4558.000000 | 1.155861 |

Dimana berdasarkan deskripsi statik yang telah diberikan merupakan ringkasan harga data finansial saham berdasarkan atribut – atribut yang ada yakni *open, high, low close, dan volume*. Berdasarkan data statik yang ditemukan dalam penelitian, menjelaskan bahwa rata – rata harga penutupan saham akan sedikit lebih rendah dibandingkan harga buka awal saham. Berdasarkan deskripsi statik tersebut juga menunjukan perkembangan harga saham. Hal ini dibuktikan dengan miningkatnya harga kuartil pertama sampai dengan ketiga yang kian meningkat. Nilai yang terjadi dalam distribusi volume terhadap frekuensi data juga dapat ditunjukan pada gambar 2.

A graph of a number of blue and black bars

Description automatically generated with medium confidence

Gambar 3. Visualisasi data

Berdasarkan hasil perhitungan korelasi yang telah dilakukan, masing – masing atribut memiliki keterikatan yang kuat dengan nilai positif yang ditampilkan pada gambar 4.

A screenshot of a graph

Description automatically generated

Gambar 4. Visualisasi data

Pada penelitian yang dilakukan ini menggunakan nilai atribut pada variabel *close* sebagai data target prediksi dari penelitian yang dilakukan. Dari perhitungan korelasi matriks yang dilakukan, menunjukan bahwa atribut nilai *close* memiliki respon yang cukup positif pada beberapa *feature*. Nilai korelasi matriks ditampilkan pada table 2.

Berdasarkan pada perhitungan dengan menggunakan algoritma K-Nearest Neighbors menghasilkan

V. Conclusion

Summarize sentences the primary outcomes of the study in a paragraph. Explain if the claims in this section supported by the results, and if they seem reasonable. Also, describe whether the result support or contradict previous theories, and explain how the research has moved the body of scientiﬁc knowledge forward.

Acknowledgment

The authors would like to thank...

References

1. H. Kopka and P. W. Daly, “A Guide to LATEX”, 3rd ed. Harlow, England: Addison-Wesley, (1999).
2. J. K. Author, “Title of chapter in the book”, in Title of His Published Book, xth ed. City of Publisher, Country if not
3. USA: Abbrev. of Publisher, ch. x.sec. x, (year): pp-pp.
4. Author. (year, month). Title. Journal. volume (issue), pages. Available: site/path/ﬁle
5. J. K. Author, “Title of paper”, in Abbreviated Name of Conf., City of Conf., Abbrev. State (if given), (year): pp-pp.
6. J. K. Author, “Title of dissertation”, Ph.D. dissertation, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, (year).
7. Ramadhani R. D., Priyanto, A., Sidiq, M. F. 2018., Enhancement of K-Parameter Using Hybrid Stratified Sampling and Genetic Algorithm, INFOTEL, 10 (1), pp.22-27.
8. Iskandar A. R., 2018., Transition Strategies of Change Management For the Succesful Implementation of Data Warehouse of Higher Education in Indonesia, Journal of Data Science and Its Applications (JDSA), 1(1), pp.28-38.
9. Amalia A. E., Airlangga, G., Thohari, A. N. A., 2018., Breast Cancer Image Segmentation Using K-Means Clustering Based on GPU Cuda Parallel Computing, INFOTEL, 10 (1), pp.33-38.
10. Usada E., 2018., Penggabungan Metode Inferensi Fuzzy Dengan Operator Prewitt Untuk Deteksi Tepi, INFOTEL, 9 (3), pp.271-277.