

**PERBANDINGAN KINERJA MODEL PREDIKSI  
ALGORITMA REGRESI LINEAR DAN *K-NEAREST  
NEIGHBORS* DALAM KASUS JUMLAH PELANGGAN  
TELEKOMUNIKASI**

**DRAFT SKRIPSI**

Diajukan untuk memenuhi sebagian syarat untuk memperoleh gelar Sarjana  
Teknik Program Studi Sistem Telekomunikasi



Oleh

Aldewo Dillon Perkasa

NIM. 1905342

**PROGRAM STUDI SISTEM TELEKOMUNIKASI**

**KAMPUS UPI DI PURWAKARTA**

**UNIVERSITAS PENDIDIKAN INDONESIA**

**2023**

**PERBANDINGAN KINERJA MODEL PREDIKSI  
ALGORITMA REGRESI LINEAR DAN  
*K-NEAREST NEIGHBORS* DALAM KASUS  
JUMLAH PELANGGAN TELEKOMUNIKASI**

Oleh

Aldewo Dillon Perkasa

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat memperoleh gelar Sarjana Teknik pada Program Studi Telekomunikasi

© Aldewo Dillon Perkasa 2023  
Universitas Pendidikan Indonesia  
Agustus 2023

Hak Cipta dilindungi undang-undang.  
Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian,  
dengan dicetak ulang, difoto kopi, atau cara lainnya tanpa ijin dari penulis.

**LEMBAR PENGESAHAN**

**ALDEWO DILLON PERKASA**

**PERBANDINGAN KINERJA MODEL PREDIKSI ALGORITMA  
REGRESI LINEAR DAN *K-NEAREST NEIGHBORS* DALAM KASUS  
JUMLAH PELANGGAN TELEKOMUNIKASI**

Disetujui dan disahkan oleh pembimbing :

**Pembimbing I**



**Ichwan Nur Ichwan, S. E., M. T.**

**NIP. 920200119900330101**

**Pembimbing II**

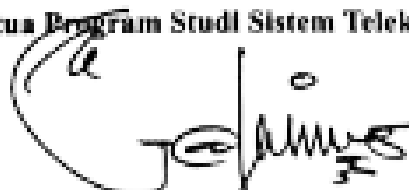


**Hafiyah Putra Prutama, S. ST, M. T.**

**NIP. 920190219921224101**

Mengetahui

**Ketua Program Studi Sistem Telekomunikasi**



**Galura Muhammad Suranegara, M. T.**

**NIP.920190219920111101**

## ABSTRAK

Dalam ilmu komputer dan *data science*, perbandingan antara algoritma - algoritma memiliki peran penting. Penelitian ini membandingkan performa dua algoritma untuk prediksi atau *forecasting*, dengan data dari Badan Pusat Statistik mengenai Jumlah Pelanggan *Wired* dan *Wireless* Telekomunikasi. Evaluasi dilakukan dengan metrik seperti *Mean Absolute Deviation* (MAD) untuk regresi linear dan *k-nearest neighbors*. Hasilnya, MAD regresi linear: 1.520 (*wired*) dan 46.140 (*wireless*), serta *k - nearest neighbors*: 827 (*wired*) dan 35.723 (*wireless*). *Mean Square Error* (MSE) regresi linear: 27.850 (*wired*) dan 24.907.799 (*wireless*), *k-nearest neighbors*: 10.282 (*wired*) dan 21.165.042 (*wireless*). *Root Mean Square Error* (RMSE) regresi linear: 1.668 (*wired*) dan 49.907 (*wireless*), *k-nearest neighbors*: 1.014 (*wired*) dan 46.005 (*wireless*). *Mean Absolute Percentage Error* (MAPE) regresi linear: 15,10% (*wired*) dan 13,48% (*wireless*), *k-nearest neighbors*: 8,57% (*wired*) dan 9,87% (*wireless*). Hasil menunjukkan *k-nearest neighbors* lebih unggul daripada regresi linear dalam prediksi. Algoritma ini, yang utamanya berfungsi sebagai klasifikasi, dapat mengungguli regresi linear dalam tugas prediksi atau *forecasting*.

## **ABSTRACT**

*In computer science and data science, the comparison between algorithms plays a crucial role. This research compares the performance of two algorithms for prediction or forecasting using data from the Central Statistics Agency (Badan Pusat Statistik) regarding the number of Wired and Wireless Telecommunication Customers. Evaluation is conducted using metrics such as Mean Absolute Deviation (MAD) for linear regression and k-nearest neighbors. The results show that for linear regression, MAD is 1,520 (wired) and 46,140 (wireless), while for k-nearest neighbors, it is 827 (wired) and 35,723 (wireless). Mean Square Error (MSE) for linear regression is 27,850 (wired) and 24,907,799 (wireless), and for k-nearest neighbors, it is 10,282 (wired) and 21,165,042 (wireless). Root Mean Square Error (RMSE) for linear regression is 1,668 (wired) and 49,907 (wireless), and for k-nearest neighbors, it is 1,014 (wired) and 46,005 (wireless). Mean Absolute Percentage Error (MAPE) for linear regression is 15.10% (wired) and 13.48% (wireless), while for k-nearest neighbors, it is 8.57% (wired) and 9.87% (wireless). The results indicate that k-nearest neighbors outperform linear regression in prediction. This algorithm, primarily functioning as classification, can surpass linear regression in prediction or forecasting tasks.*

## DAFTAR ISI

LEMBAR PENGESAHAN.....	i
PERNYATAAN KEASLIAN SKRIPSI DAN BEBAS PLAGIARISME.....	ii
KATA PENGANTAR.....	iii
UCAPAN TERIMA KASIH.....	iv
ABSTRAK.....	vi
ABSTRACT.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMPIRAN.....	xi
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian.....	4
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian.....	5
1.5.1 Manfaat Teoritis.....	5
1.5.2 Manfaat Praktis.....	5
1.6 Struktur Organisasi Skripsi.....	6
BAB II.....	8
KAJIAN PUSTAKA.....	8
2.1 Forecasting.....	8
2.1.1 Quantitative Forecasting.....	9
2.1.2 Qualitative Forecasting.....	10
2.2 Machine Learning.....	11
2.2.1 Supervised Learning.....	12
2.2.2 Unsupervised Learning.....	13
2.2.3 Reinforcement Learning.....	14
2.3 Python Programming Language.....	14
2.4 Algoritma Regresi Linear.....	16
2.5 Algoritma K-Nearest Neighbors (KNN).....	18
2.6 Penelitian Terdahulu.....	21
BAB III.....	25
METODE PENELITIAN.....	25
3.1 Jenis Penelitian.....	25
3.2 Prosedur Penelitian.....	25
3.2.1 Hipotesis.....	25
3.2.2 Alur Pengolahan Data.....	26

3.2.3 Alur Evaluasi Data.....	28
3.3 Metode Pengumpulan Data.....	29
3.4 Data Penelitian.....	30
3.5 Analisis Data.....	31
3.5.1 Mean Absolute Deviation.....	31
3.5.2 Mean Square Error.....	32
3.5.3 Root Mean Square Error.....	32
3.5.4 Mean Absolute Percentage Error.....	33
3.6 Alat Penelitian.....	33
3.6.1 Perangkat Lunak.....	33
3.6.2 Perangkat Keras.....	34
BAB IV.....	35
TEMUAN DAN PEMBAHASAN.....	35
4.1 Rancangan Model Prediksi.....	35
4.1.1 Rancangan Model Prediksi Algoritma Regresi Linear.....	35
4.1.2 Rancangan Model Prediksi Algoritma K - Nearest Neighbors.....	38
4.2 Pembahasan Perubahan Jumlah Pelanggan.....	40
4.2.1 Perubahan Jumlah Pelanggan Algoritma Regresi Linear.....	40
4.2.2 Perubahan Jumlah Pelanggan Algoritma K-Nearest Neighbors.....	42
4.2.3 Perbandingan Perubahan Jumlah Pelanggan.....	44
4.3 Pembahasan Hasil Evaluasi Model Prediksi.....	45
4.3.1 Hasil Perhitungan Evaluasi Matriks.....	45
4.3.2 Perbandingan Hasil Evaluasi Matriks.....	46
BAB V.....	52
SIMPULAN, IMPLIKASI DAN REKOMENDASI.....	52
5.1 Kesimpulan.....	52
5.2 Implikasi.....	53
5.2.1 Implikasi Teoritis.....	53
5.2.2 Implikasi Praktis.....	53
5.2.3 Implikasi untuk penelitian masa depan.....	53
5.3 Rekomendasi.....	54
DAFTAR PUSTAKA.....	55
LAMPIRAN.....	58
Biografi Penulis.....	58
Syntax Program.....	60
Algoritma Regresi Linear.....	60
Algoritma K - Nearest Neighbors.....	63

## DAFTAR PUSTAKA

- Almumtazah, N., Azizah, N., Putri, Y. L., & Novitasari, D. C. R. (2021). Prediksi Jumlah Mahasiswa Baru Menggunakan Metode Regresi Linier Sederhana. *Jurnal Ilmiah Matematika Dan Terapan*, 18(1), 31–40.  
<https://doi.org/10.22487/2540766X.2021.v18.i1.15465>
- Analysis and design of algorithms. A critical comparison of different works on algorithms* (1. Auflage, digitale Originalausgabe). (2019). GRIN Verlag.
- Chandra, W., Suprihatin, B., & Resti, Y. (2023). Median-KNN Regressor-SMOTE-Tomek Links for Handling Missing and Imbalanced Data in Air Quality Prediction. *Symmetry*, 15(4), 887.  
<https://doi.org/10.3390/sym15040887>
- Chatfield, C. (2001). *Time-series forecasting*. Chapman & Hall/CRC.
- Derisma, D. (2020). Perbandingan Kinerja Algoritma untuk Prediksi Penyakit Jantung dengan Teknik Data Mining. *Journal of Applied Informatics and Computing*, 4(1), 84–88. <https://doi.org/10.30871/jaic.v4i1.2152>
- Desiani, A. (2022). Perbandingan Implementasi Algoritma Naïve Bayes dan K-Nearest Neighbor Pada Klasifikasi Penyakit Hati. *SIMKOM*, 7(2), 104–110. <https://doi.org/10.51717/simkom.v7i2.96>
- Diansyah, S. (2022). Klasifikasi Tingkat Kepuasan Pengguna dengan Menggunakan Metode K-Nearest Neighbour (KNN). *Jurnal Sistim Informasi Dan Teknologi*, 7–12. <https://doi.org/10.37034/jsisfotek.v4i1.114>
- Dwi Selvy Wisdayani. (2019). *Perbandingan Algoritma K-Nearest Neighbor dan Naive Bayes untuk Klasifikasi Tingkat Keparahan Korban Kecelakaan*



- Lalu Lintas di Kabupaten Pati Jawa Tengah* [Universitas Muhammadiyah Semarang]. <http://repository.unimus.ac.id/id/eprint/3832>
- Elsayed, K. M. T. (2015). Mean Absolute Deviation: Analysis and Applications. *International Journal of Business and Statistical Analysis*, 2(2), 63–74. <https://doi.org/10.12785/ijbsa/020201>
- Erickson, J. (2019). *Algorithms* (1st paperback edition). Selbstverlag.
- Fachid, S., & Triayudi, A. (2022). Perbandingan Algoritma Regresi Linier dan Regresi Random Forest Dalam Memprediksi Kasus Positif Covid-19. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 6(1), 68. <https://doi.org/10.30865/mib.v6i1.3492>
- Ford, M. (2018). *Architects of intelligence: The truth about AI from the people building it* (First published: November 2018). Packt Publishing.
- G. N. Ayuni & D. Fitriana. (2019). Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ. *Jurnal Telematika*, 14(2), 79–86.
- Hodson, T. O., Over, T. M., & Foks, S. S. (2021). Mean Squared Error, Deconstructed. *Journal of Advances in Modeling Earth Systems*, 13(12). <https://doi.org/10.1029/2021MS002681>
- Ibnu Daqiqil Id. (2021). *MACHINE LEARNING: Teori, Studi Kasus dan Implementasi Menggunakan Python*. UR PRESS.
- Izzan Arimi, Ratna Purwaningsih, & Zainal Fanani Rosyada. (2022). Metode K-Nearest Neighbor untuk Memprediksi Penjualan Produk pada UMKM Pengolahan Ikan Maju Jaya. *Industrial Engineering Online Journal*, 12(1). <https://ejournal3.undip.ac.id/index.php/ieoj/article/view/37165>

- Khairiati, A., Budiarti, R., & Purnaba, I. G. P. (2022). Perbandingan Analisis Regresi Linear dengan Analisis Regresi Copula pada Data Keuangan. *Jambura Journal of Mathematics*, 4(2), 209–219.  
<https://doi.org/10.34312/jjom.v4i2.13829>
- Khazae Poul, A., Shourian, M., & Ebrahimi, H. (2019). A Comparative Study of MLR, KNN, ANN and ANFIS Models with Wavelet Transform in Monthly Stream Flow Prediction. *Water Resources Management*, 33(8), 2907–2923. <https://doi.org/10.1007/s11269-019-02273-0>
- Kleinberg, J., & Tardos, É. (2006). *Algorithm design*. Pearson/Addison-Wesley.
- Lazzeri, F. (2020). *Machine Learning for Time Series Forecasting With Python*. (1st ed.). John Wiley and Sons.
- Lee, W.-M. (2019). *Python machine learning*. John Wiley and Sons.
- Levitin, A. (2012). *Introduction to the design & analysis of algorithms* (3rd ed). Pearson.
- M. A. Moon., Chapter 5. Quantitative Forecasting Techniques. (2018). , *Demand and Supply Integration* (pp. 87–112). De Gruyter.  
<https://doi.org/10.1515/9781501506024-005>
- Mariko, M. (2021). Perbandingan Algoritma Apriori Dan Algoritma Fp-Growth Untuk Rekomendasi Item Paket Pada Konten Promosi. *EXPLORE*, 11(2), 24. <https://doi.org/10.35200/explore.v11i2.438>
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed). Wiley.
- Sarker, I. H. (2021). *Machine Learning: Algorithms, Real-World Applications and*

Research Directions. *SN Computer Science*, 2(3), 160.

<https://doi.org/10.1007/s42979-021-00592-x>

Swamidass, P. M. (Ed.). (2006). *Encyclopedia of production and manufacturing management*.

Yan, X., & Su, X. G. (2009). *Linear Regression Analysis: Theory and Computing*.

WORLD SCIENTIFIC. <https://doi.org/10.1142/6986>