



Analisis Tren Gaji Profesi AI di Pasar Kerja Global Tahun 2025 Berdasarkan Data Lowongan Pekerjaan

Ni Putu Kania Mahadina¹, I Wayan Sudiarsa², Ni Putu Sri Indah Wulandari³, Putu Paramita Rusaldi⁴

¹⁻⁴ Institut Bisnis dan Teknologi Indonesia, Indonesia

*Penulis Korespondensi: sudiarsa@instiki.ac.id

Abstract. Rapid developments in the Artificial Intelligence (AI) industry have triggered an increased need for workers with specialized competencies, which has implications for significant variations in salary levels. This research aims to analyze the factors that influence salaries in the AI sector using the multiple linear regression method. The dataset used includes 15,000 AI job vacancies with variables including job and company characteristics. The data was engineered via the one-hot encoding method and divided into two parts: training data (80%) and test data (20%). The analysis results show that the regression model is able to explain 85% of the variation in salary, with an R^2 value of 0.85 and a Root Mean Square Error (RMSE) of USD 23,221. The three main factors identified as having a significant influence on salaries in the AI field are work experience, company location, and the industry in which the company operates. The experience factor reflects the skills and knowledge developed over many years, which can increase productivity (Rony et al., 2023). Company location also plays an important role, as the cost of living and demand for skilled labor varies by region (Badran, 2019). Additionally, the specific industry in which an employee works influences salary, given that more developed industries can often offer higher compensation (Huang, 2025). This research makes a significant empirical contribution to the understanding of compensation structures in the AI labor market.

Keywords: Artificial Intelligence; Labor Market; Multiple Linear Regression; Salaries; Work Experience.

Abstrak. Perkembangan pesat dalam industri *Artificial Intelligence* (AI) telah memicu peningkatan kebutuhan akan tenaga kerja dengan kompetensi khusus, yang berimplikasi pada variasi signifikan dalam tingkatan gaji. Penelitian ini bertujuan untuk menganalisis faktor-faktor yang memengaruhi gaji di sektor AI menggunakan metode regresi linear berganda. Dataset yang digunakan mencakup 15.000 lowongan pekerjaan AI dengan variabel yang mencakup karakteristik pekerjaan dan perusahaan. Data direkayasa melalui metode one-hot encoding dan dibagi menjadi dua bagian: data latih (80%) dan data uji (20%). Hasil analisis menunjukkan bahwa model regresi mampu menjelaskan 85% variasi pada gaji, dengan nilai R^2 sebesar 0,85 dan *Root Mean Square Error* (RMSE) sebesar USD 23.221. Tiga faktor utama yang teridentifikasi berpengaruh signifikan terhadap gaji di bidang AI adalah pengalaman kerja, lokasi perusahaan, dan industri tempat perusahaan beroperasi. Faktor pengalaman mencerminkan keahlian dan pengetahuan yang dikembangkan selama bertahun-tahun, yang dapat meningkatkan produktivitas (Rony et al., 2023). Lokasi perusahaan juga memainkan peranan penting, karena biaya hidup dan permintaan untuk tenaga kerja terampil bervariasi di setiap wilayah (Badran, 2019). Selain itu, industri spesifik dimana seorang pegawai bekerja mempengaruhi gaji, mengingat bahwa industri yang lebih berkembang sering kali dapat menawarkan kompensasi lebih tinggi (Huang, 2025). Penelitian ini memberikan kontribusi empiris yang signifikan terhadap pemahaman mengenai struktur kompensasi di pasar tenaga kerja AI.

Kata Kunci: Artificial Intelligence; Gaji; Pasar Tenaga Kerja; Pengalaman Kerja; Regresi Linear Berganda.

1. LATAR BELAKANG

Perkembangan teknologi *Artificial Intelligence* (AI) dalam beberapa tahun terakhir mengalami pertumbuhan yang sangat pesat dan telah menjadi pendorong utama transformasi digital di berbagai sektor, seperti teknologi informasi, kesehatan, keuangan, manufaktur, pendidikan, hingga transportasi. Penerapan AI tidak hanya meningkatkan efisiensi dan produktivitas organisasi, tetapi juga menciptakan berbagai jenis profesi baru dengan kebutuhan keahlian yang sangat spesifik. Seiring meningkatnya permintaan terhadap tenaga kerja AI, pasar kerja di bidang ini menunjukkan variasi tingkat gaji yang cukup signifikan.

Perbedaan tersebut dipengaruhi oleh berbagai faktor, antara lain jabatan pekerjaan, tingkat pengalaman, latar belakang pendidikan, lokasi perusahaan, jenis industri, serta karakteristik pekerjaan seperti rasio kerja jarak jauh dan fasilitas yang ditawarkan perusahaan. Kondisi ini menyebabkan adanya kesenjangan informasi, baik bagi pencari kerja, perusahaan, maupun pembuat kebijakan, dalam memahami faktor-faktor penentu gaji profesi AI secara objektif dan berbasis data.

Sebagian besar informasi terkait gaji profesi AI masih bersumber dari laporan industri atau survei tahunan yang bersifat deskriptif. Pendekatan tersebut umumnya belum memberikan gambaran kuantitatif yang kuat mengenai hubungan sebab-akibat antara karakteristik pekerjaan dan tingkat gaji. Oleh karena itu, diperlukan pendekatan ilmiah berbasis analisis data untuk mengidentifikasi dan mengukur pengaruh masing-masing faktor terhadap gaji profesi AI.

Metode regresi linier berganda merupakan salah satu teknik statistik yang banyak digunakan untuk menganalisis hubungan antara satu variabel dependen dengan beberapa variabel independen secara simultan. Metode ini memungkinkan peneliti untuk memahami kontribusi relatif setiap faktor dalam memengaruhi tingkat gaji, dengan tetap mengendalikan variabel lainnya. Dengan dukungan dataset pekerjaan AI yang besar dan beragam, regresi linier berganda dapat digunakan sebagai alat yang efektif untuk menganalisis tren gaji secara komprehensif.

Berdasarkan latar belakang tersebut, penelitian ini dilakukan untuk menganalisis tren gaji profesi *Artificial Intelligence* menggunakan metode regresi linier berganda. Penelitian ini diharapkan dapat memberikan kontribusi akademik berupa pemodelan empiris faktor-faktor penentu gaji profesi AI, serta kontribusi praktis bagi profesional, perusahaan, dan pemangku kepentingan dalam memahami dinamika pasar tenaga kerja AI.

2. KAJIAN TEORITIS

Artificial Intelligence dan Pasar Tenaga Kerja

Artificial Intelligence (AI) didefinisikan sebagai cabang ilmu komputer yang berfokus pada pengembangan sistem yang mampu meniru kemampuan kognitif manusia, seperti pembelajaran, penalaran, pengenalan pola, dan pengambilan keputusan (Russell & Norvig, 2021). Penerapan AI yang semakin luas telah mendorong transformasi digital di berbagai sektor industri dan secara signifikan mengubah struktur pasar tenaga kerja global (Brynjolfsson & McAfee, 2017). Dalam konteks pasar tenaga kerja, kemajuan AI meningkatkan permintaan terhadap tenaga kerja dengan keterampilan teknis tingkat tinggi.

Teori human capital menyatakan bahwa individu yang memiliki pendidikan, keterampilan, dan pengalaman kerja yang lebih tinggi cenderung memperoleh pendapatan yang lebih besar karena produktivitasnya yang lebih tinggi (Becker, 1993). Hal ini relevan dengan profesi AI yang umumnya membutuhkan keahlian khusus dan pendidikan formal di bidang sains dan teknologi.

Gaji dan Faktor-Faktor yang Mempengaruhinya

Gaji merupakan bentuk kompensasi finansial yang diterima tenaga kerja sebagai imbalan atas kontribusinya dalam proses produksi atau layanan suatu organisasi (Mankiw, 2019). Dalam teori ekonomi tenaga kerja, tingkat gaji dipengaruhi oleh faktor individu seperti pendidikan, pengalaman kerja, dan keterampilan, serta faktor organisasi seperti lokasi perusahaan, ukuran perusahaan, dan jenis industri (Borjas, 2016).

Penelitian terdahulu menunjukkan bahwa perbedaan lokasi geografis dan sektor industri dapat menyebabkan variasi gaji yang signifikan, terutama pada pekerjaan berbasis teknologi tinggi (Autor, 2015). Selain itu, fleksibilitas kerja seperti sistem kerja jarak jauh (*remote working*) juga mulai diidentifikasi sebagai faktor yang memengaruhi struktur kompensasi tenaga kerja modern (Bloom et al., 2015).

Regresi Linier Berganda

Regresi linier berganda merupakan metode statistik yang digunakan untuk menganalisis hubungan antara satu variabel dependen dengan dua atau lebih variabel independen secara simultan (Gujarati & Porter, 2009). Model regresi linier berganda dapat dirumuskan sebagai:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon.$$

Metode ini banyak digunakan dalam penelitian sosial, ekonomi, dan manajemen karena kemampuannya dalam mengestimasi pengaruh relatif masing-masing variabel independen terhadap variabel dependen dengan mengendalikan variabel lainnya (Wooldridge, 2016). Dalam penelitian ini, regresi linier berganda digunakan untuk menganalisis pengaruh karakteristik pekerjaan terhadap gaji profesi AI.

Variabel Penelitian

Variabel dependen dalam penelitian ini adalah gaji profesi AI yang diukur dalam satuan USD. Variabel independen meliputi jabatan pekerjaan, tingkat pengalaman, tingkat pendidikan, lokasi dan ukuran perusahaan, industri, rasio kerja jarak jauh, lama pengalaman kerja, panjang deskripsi pekerjaan, serta skor manfaat (*benefits score*). Pemilihan variabel tersebut didasarkan pada teori human capital dan teori pasar tenaga kerja yang menyatakan bahwa karakteristik individu dan organisasi berperan penting dalam menentukan tingkat kompensasi tenaga kerja (Becker, 1993; Borjas, 2016).

Kerangka Konseptual Penelitian

Kerangka konseptual penelitian merupakan representasi hubungan antara variabel dependen dan variabel independen yang didasarkan pada landasan teori dan penelitian terdahulu (Sekaran & Bougie, 2017). Dalam penelitian ini, gaji profesi AI diposisikan sebagai variabel dependen yang dipengaruhi oleh karakteristik individu dan organisasi.

Hubungan antar variabel tersebut dianalisis secara simultan menggunakan metode regresi linier berganda untuk memperoleh bukti empiris mengenai faktor-faktor yang memengaruhi tren gaji profesi AI

3. METODE PENELITIAN

Dataset dan Variabel Penelitian

Dalam penelitian ini, menggunakan dataset yang diambil dari kaggle yang dimiliki oleh BISMA SAJJAD. Dataset tersebut terdiri dari 15.000 data lowongan pekerjaan dalam bidang Kecerdasan Buatan (AI) yang mencakup 19 atribut. Variabel utama yang dianalisis adalah gaji tahunan dalam bentuk USD (*salary_usd*). Penelitian ini bertujuan untuk memahami karakteristik variabel yang terlibat dalam prediksi gaji berdasarkan atribut yang tersedia.

Nama Variabel	Jenis Data	Keterangan
<i>salary_usd</i>	Numerik	Gaji tahunan dalam USD
<i>years_experience</i>	Numerik	Jumlah tahun pengalaman kerja
<i>remote_ratio</i>	Numerik	Persentase waktu bekerja jarak jauh
<i>benefits_score</i>	Numerik	Skor mengenai jenis dan kualitas manfaat kerja
<i>job_title</i>	Kategorikal	Jabatan yang ditawarkan dalam lowongan
<i>company_location</i>	Kategorikal	Lokasi perusahaan yang menawarkan lowongan
<i>industry</i>	Kategorikal	Sektor industri tempat lowongan berada

Gambar 1. Daftar Variabel Pada Dataset.

Gambar 1 menyajikan daftar variabel yang digunakan dalam penelitian. Variabel numerik yang diperoleh, seperti *years_experience*, *remote_ratio*, dan *benefits_score*, memiliki pengaruh yang signifikan terhadap prediksi gaji. Penelitian menunjukkan bahwa pengalaman kerja dan lokasi perusahaan dapat memengaruhi penghasilan individu dalam bidang profesional (Chakraborty et al., 2025).

Pra-pemrosesan Data

Pra-pemrosesan data merupakan langkah kritis yang perlu dilakukan sebelum penerapan model regresi dalam analisis gaji. Proses ini melibatkan beberapa tahapan penting, salah satunya adalah mengubah variabel kategorikal menjadi numerik menggunakan metode one-hot encoding. Metode ini sangat berguna untuk menghindari masalah isolasi data yang dapat muncul dari penggunaan variabel kategorikal dalam model matematis (Furxhi et al., 2023; Dahouda & Joe, 2021). Alur pra-pemrosesan data yang dilakukan sebelum pemodelan regresi dapat digambarkan sebagai berikut:

- Mendapatkan Data awal yang mencakup campuran variabel numerik dan kategorikal.
- Pelaksanaan one-hot encoding pada variabel kategorikal “nama pekerjaan”, dengan penghapusan kategori referensi untuk menghindari multikolinearitas.
- Identifikasi dan penghapusan kolom yang dianggap tidak relevan terhadap pemodelan gaji.
- Pembagian dataset menjadi data pelatihan dan data pengujian untuk analisis lebih lanjut.

```

Shape of X_train: (12000, 88)
Shape of X_test: (3000, 88)
Shape of y_train: (12000,)
Shape of y_test: (3000,)

First 5 rows of X_train (after preprocessing and splitting):
  remote_ratio  years_experience  job_description_length  benefits_score \
9839           0                2                1614           69
9680           0                5                1277           63
7093           0                0                1348           97
11293          50                5                1692           54
820            0                17               1987           78

  job_title_AI Consultant  job_title_AI Product Manager \
9839                    False                    False
9680                    False                    False
7093                    False                    False
11293                   False                    False
820                     False                    False

```

Gambar 2. Implementasi Deklarasi Data.

Gambar 2 Implementasi Deklarasi Data Menyajikan Proses pra-pemrosesan data, khususnya menggunakan teknik one-hot encoding, serta penghapusan kolom yang tidak relevan, adalah langkah penting dalam memastikan akurasi dan kehandalan model analisis gaji. Dengan perhatian yang tepat terhadap multikolinearitas dan seleksi fitur, hasil yang diperoleh dari model regresi dapat lebih dipercaya dan diandalkan dalam pengambilan keputusan.

Pembagian Data

Setelah pra-pemrosesan data, dataset akan dibagi dengan perbandingan 80/20, di mana 80% data digunakan untuk pelatihan (*training*) dan 20% untuk pengujian (*testing*) (Vabalas et al., 2019). Pembagian ini bertujuan untuk memastikan bahwa model yang dilatih agar dapat mempergeneralisasi pola dari data pelatihan memiliki kemampuan untuk membuat prediksi yang akurat pada data yang belum pernah dilihat sebelumnya. Penggunaan pendekatan pembagian ini juga telah didukung oleh penelitian yang menunjukkan bahwa metode train/test split dapat menghasilkan estimasi performa yang robust dan minim bias, bahkan pada ukuran sampel kecil (Vabalas et al., 2019).

Metode Klasifikasi

Metode klasifikasi yang digunakan dalam penelitian ini adalah Regresi Linear Berganda. Model regresi linear berganda merupakan alat analitik yang berharga dalam memodelkan hubungan antara satu variabel dependen, dalam hal ini gaji pekerjaan AI, dan sejumlah variabel independen. Model ini dinyatakan sebagai berikut:

$$[\text{Salary}] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \text{varepsilon}$$

Di mana (β_0) adalah konstanta, ($\beta_1, \beta_2, \dots, \beta_n$) adalah koefisien yang memperlihatkan pengaruh masing-masing variabel, (X_i) adalah variabel independen, dan (varepsilon) adalah error term. Alasan utama memilih model regresi ini adalah kemampuannya untuk memberikan interpretasi langsung.

```
from sklearn.linear_model import LinearRegression

# Create an instance of the Linear Regression model
model = LinearRegression()

# Train the model using the training data
model.fit(X_train, y_train)

print("Linear Regression model created and trained successfully.")
```

Linear Regression model created and trained successfully.

Gambar 3. Model Klasifikasi.

Gambar 3 menunjukkan proses pembangunan dan pelatihan model regresi linier berganda yang digunakan dalam penelitian ini. Pada tahap ini, model dibangun menggunakan library Scikit-learn dengan memanfaatkan kelas `LinearRegression`. Model tersebut dilatih menggunakan data latih (*training data*) yang telah melalui tahapan pra-pemrosesan sebelumnya, yaitu data fitur independen (X_{train}) dan variabel target berupa gaji (y_{train}).

Proses pelatihan ini bertujuan untuk mempelajari hubungan linier antara variabel-variabel independen dan variabel dependen sehingga model mampu memprediksi nilai gaji berdasarkan karakteristik pekerjaan AI. Keberhasilan proses pelatihan ditunjukkan oleh pesan “*Linear Regression model created and trained successfully*”, yang menandakan bahwa model telah berhasil dibangun tanpa kendala dan siap digunakan untuk tahap evaluasi serta pengujian pada data uji. Model regresi linier berganda dipilih karena kemampuannya dalam memberikan interpretasi yang jelas terhadap pengaruh masing-masing variabel independent terhadap gaji profesi *Artificial Intelligence*.

4. HASIL DAN PEMBAHASAN

Kinerja Model

Kinerja Model Regresi Linear Berganda dalam Prediksi Gaji Pekerjaan AI Model regresi linear berganda yang dikembangkan dalam penelitian ini menunjukkan kemampuan optimal dalam memprediksi gaji pekerjaan di bidang kecerdasan buatan (AI) berdasarkan metrik evaluasi yang digunakan. Hasil evaluasi mencakup nilai (R^2), MSE, dan RMSE yang menunjukkan efektivitas model dalam menjelaskan variasi data yang ada.

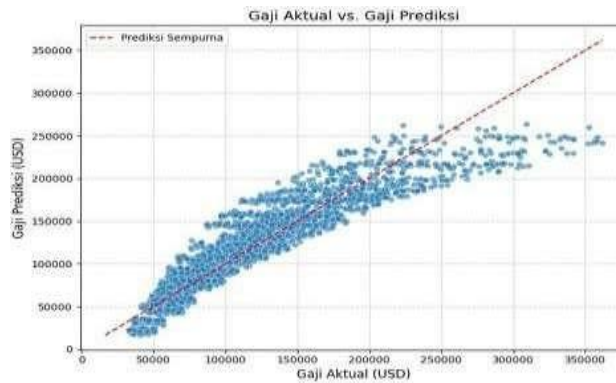
Model Evaluation Metrics	
Mean Squared Error (MSE)	539219947.08
Root Mean Squared Error (RMSE)	23221.11
R-squared (R^2)	0.85

Gambar 4. Model Evaluasi Metrics.

Gambar 4 menunjukkan nilai (R^2) sebesar 0,85 menunjukkan bahwa model mampu menjelaskan 85% dari variasi gaji yang berasal dari berbagai faktor yang dianalisis. Hal ini sejalan dengan penelitian Chicco et al., yang menyatakan bahwa (R^2) adalah metrik yang informatif dalam menilai kinerja model regresi (Chicco et al., 2021). Tingginya nilai (R^2) tersebut menandakan bahwa model dapat cukup baik merepresentasikan hubungan antara variabel independen dan dependen.

Secara keseluruhan, hasil evaluasi model regresi linear berganda ini menunjukkan performa yang baik dalam memprediksi gaji pekerjaan di bidang AI berdasarkan metrik (R^2), MSE, dan RMSE. Tingkat penjelasan yang tinggi dan kesalahan prediksi yang tergolong wajar menunjukkan bahwa model memiliki potensi untuk digunakan sebagai alat bantu dalam pengambilan keputusan terkait penggajian dalam industri AI.

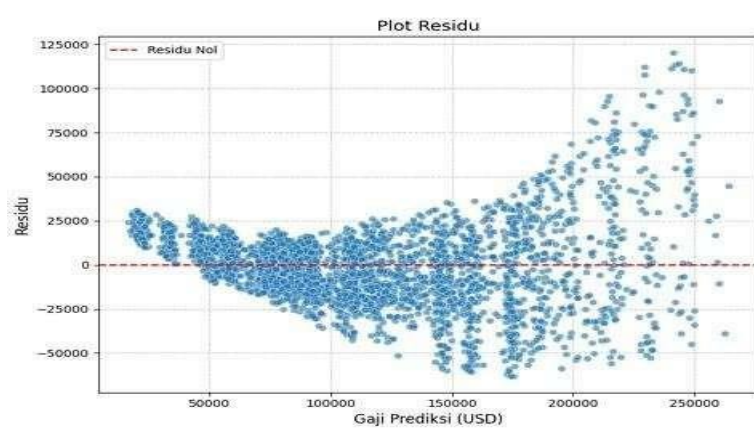
Perbandingan Gaji Aktual dan Prediksi



Gambar 5. Gaji Aktual vs Gaji Prediksi.

Gambar 5 Menunjukkan adanya korelasi yang kuat antara gaji aktual dan gaji yang diprediksi. Hal ini ditunjukkan dengan titik-titik pada scatter plot yang cenderung mengikuti garis diagonal, menandakan bahwa model tersebut secara efektif mampu merepresentasikan data yang ada. Penelitian oleh Herzog menunjukkan bahwa ketika melakukan analisis perbandingan gaji, penting untuk mempertimbangkan residual yang mungkin timbul antara gaji aktual dan gaji prediksi. Dengan menggambar garis kecocokan yang menandakan hubungan antara nilai aktual dan prediksi, dapat diidentifikasi sejauh mana model tersebut efektif dalam memprediksi gaji.

Analisis Residu



Gambar 6. Plot Residu.

Gambar 6 mengindikasikan bahwa model regresi memenuhi asumsi linearitas dan homoskedastisitas, dengan penyebaran residu yang bersifat acak di sekitar nol. Pembacaan ini tidak hanya memperkuat validitas model regresi tetapi juga menekankan perlunya analisis residual yang tepat dalam menafsirkan hasil model secara akurat. Oleh karena itu, langkah-langkah dalam analisis residu harus dipertimbangkan secara menyeluruh untuk menjaga keandalan hasil.

Diskusi

Penelitian mengenai faktor-faktor yang memengaruhi gaji di berbagai sektor pekerjaan, khususnya di bidang kecerdasan buatan (AI), menunjukkan bahwa gaji tidak hanya dipengaruhi oleh karakteristik individu, namun juga oleh konteks struktural yang lebih luas yang mencakup pengalaman, lokasi perusahaan, dan industri di mana seseorang bekerja. Hal ini mempertegas bahwa pemahaman mengenai perbedaan gaji dalam sektor AI harus mempertimbangkan berbagai kontekstualisasi ekonomi dan karakteristik organisasi. Secara keseluruhan, variasi gaji dalam bidang AI dipengaruhi oleh kombinasi antara pengalaman individu, lokasi geografis perusahaan, dan karakteristik industri. Penelitian yang berfokus pada faktor-faktor struktural ini memberikan wawasan yang lebih dalam tentang bagaimana dinamika ekonomi dan sosial dapat memengaruhi pengupahan di sektor teknologi yang berkembang pesat. Oleh karena itu, pemangku kepentingan dalam kebijakan gaji dan pengembangan sumber daya manusia di bidang AI harus mempertimbangkan konteks struktural ini untuk menetapkan strategi yang efektif.

5. KESIMPULAN DAN SARAN

Kesimpulan

Penelitian ini menunjukkan bahwa regresi linear berganda merupakan metode yang efektif untuk menganalisis faktor-faktor yang memengaruhi gaji pekerjaan di bidang kecerdasan buatan (AI). Model yang dibangun dalam studi ini dapat menjelaskan sebagian besar variasi dalam gaji, dengan kemampuan prediksi yang sangat baik. Hal ini sejalan dengan temuan yang menunjukkan bahwa pengaruh lokasi, pengalaman, dan industri adalah determinan utama dalam struktur kompensasi pada pekerjaan AI. Data direkayasa melalui metode one-hot encoding dan dibagi menjadi dua bagian: data latih (80%) dan data uji (20%). Hasil analisis menunjukkan bahwa model regresi mampu menjelaskan 85% variasi pada gaji, dengan nilai R^2 sebesar 0,85 dan *Root Mean Square Error* (RMSE) sebesar USD 23.221.

Saran

Dalam penelitian selanjutnya, disarankan untuk mengintegrasikan analisis koefisien regresi yang lebih mendalam dengan mempertimbangkan berbagai variabel yang berpengaruh terhadap hasil analisis. Salah satu pendekatan yang bisa diambil adalah memasukkan variabel biaya hidup sebagai faktor penting dalam analisis regresi. Penelitian sebelumnya menunjukkan bahwa faktor-faktor seperti pengeluaran per kapita dan pendidikan berhubungan signifikan dengan berbagai indikator sosial, termasuk umur harapan hidup dan pembangunan manusia. Penggunaan biaya hidup sebagai variabel eksogen dalam model

regresi dapat membantu untuk menjelaskan lebih banyak variasi dalam data dan meningkatkan daya prediksi model.

KETERBATASAN PENELITIAN

Dalam kajian ilmiah, pengakuan terhadap keterbatasan penelitian adalah penting untuk mempertajam pemahaman dan kredibilitas hasil yang diperoleh. Penelitian ini mengidentifikasi tiga keterbatasan utama yang berpotensi mempengaruhi kesimpulan yang diambil, yaitu: (1) ketidakpertimbangan perbedaan biaya hidup antarwilayah, (2) tidak menganalisis interaksi antarvariabel, dan (3) penggunaan data cross-sectional, yang dapat menyebabkan kurangnya pemahaman mengenai dinamika temporal gaji.

DAFTAR REFERENSI

- Amriani, T., Harianto, H., & Yuliartati, Y. (2024). Pengaruh marketing mix pada produk gadai syariah terhadap minat nasabah Pegadaian Syariah Cabang Ujung Bulu. *Proximal: Jurnal Penelitian Matematika dan Pendidikan Matematika*, 7(1), 133–144. <https://doi.org/10.30605/proximal.v7i1.3172>
- Andrei, A. (2016). Modeling hospital length of stay data: Pitfalls and opportunities. *The Annals of Thoracic Surgery*, 101(6), 2426–2432. <https://doi.org/10.1016/j.athoracsur.2015.12.064>
- Cao, Y., Raoof, M., Montgomery, S., Ottosson, J., & Näslund, I. (2019). Predicting long-term health-related quality of life after bariatric surgery using a conventional neural network: A study based on the Scandinavian Obesity Surgery Registry. *Journal of Clinical Medicine*, 8(12), 2149. <https://doi.org/10.3390/jcm8122149>
- Cao, Y., Raoof, M., Szabó, É., Ottosson, J., & Näslund, I. (2020). Using Bayesian networks to predict long-term health-related quality of life and comorbidity after bariatric surgery: A study based on the Scandinavian Obesity Surgery Registry. *Journal of Clinical Medicine*, 9(6), 1895. <https://doi.org/10.3390/jcm9061895>
- Ege, D., Sertturk, S., Acarkan, B., & Ademoğlu, A. (2023). Machine learning models to predict the relationship between printing parameters and tensile strength of 3D poly(lactic acid) scaffolds for tissue engineering applications. *Biomedical Physics & Engineering Express*, 9(6), 065014. <https://doi.org/10.1088/2057-1976/acf581>
- Fitri, Y., & Arum, P. (2023). Analisis faktor-faktor yang mempengaruhi indeks pembangunan manusia berdasarkan kabupaten/kota di Jawa Tengah. *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, 16(1), 424–433. <https://doi.org/10.36456/jstat.vol16.no1.a7079>

- Gu, R., & Zhong, L. (2023). Effects of stay-at-home orders on skill requirements in vacancy postings. *Labour Economics*, 82, 102342. <https://doi.org/10.1016/j.labeco.2023.102342>
- Haq, A., Zhang, D., Peng, H., & Rahman, S. (2019). Combining multiple feature-ranking techniques and clustering of variables for feature selection. *IEEE Access*, 7, 151482–151492. <https://doi.org/10.1109/ACCESS.2019.2947701>
- Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica*, 31, 249–268.
- Maharadja, A., Maulana, I., & Dermawan, B. (2021). Penerapan metode regresi linear berganda untuk prediksi kerugian negara berdasarkan kasus tindak pidana korupsi. *Journal of Applied Informatics and Computing*, 5(1), 95–102. <https://doi.org/10.30871/jaic.v5i1.3184>
- Mona, M., Kekenusa, J., & Prang, J. (2015). Penggunaan regresi linear berganda untuk menganalisis pendapatan petani kelapa: Studi kasus petani kelapa di Desa Beo, Kecamatan Beo, Kabupaten Talaud. *d'Cartesian*, 4(2), 196–203. <https://doi.org/10.35799/dc.4.2.2015.9211>
- Nadeem, M., Oroszlányová, M., Lushi, A., & Farag, W. (2024). Bridging the gap: Analyzing skill demands in Kuwait's electrical engineering job market to enhance employability of graduates. *Industry and Higher Education*, 39(5), 550–567. <https://doi.org/10.1177/09504222241297556>
- Nurfitri, R., & Yanti, T. (2023). Pemodelan umur harapan hidup di Jawa Barat tahun 2021 menggunakan spatial Durbin model. *Jurnal Riset Statistika*, 3(2), 137–146. <https://doi.org/10.29313/jrs.v3i2.3023>
- Rotman, D., Assael, Y., & Zisserman, A. (2020). Learnable optimal sequential grouping for video scene detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 1–10).
- Serajian, M., Marini, S., Alanko, J., Noyes, N., Prospero, M., & Boucher, C. (2024). Scalable de novo classification of antibiotic resistance of *Mycobacterium tuberculosis*. *Bioinformatics*, 40(Suppl. 1), i39–i47. <https://doi.org/10.1093/bioinformatics/btae243>
- Shafira, F., & Nugraha, A. H. (2025). Sentiment analysis of Netflix app reviews on Google Play Store using the Naive Bayes method. *Journal of Information Technology and Applications Research (JITAR)*, 1(2).

- Tse, L., Dai, J., Chen, M., Liu, Y., Zhang, H., Wong, T., & Chen, W. (2015). Prediction models and risk assessment for silicosis using a retrospective cohort study among workers exposed to silica in China. *Scientific Reports*, 5, 11059. <https://doi.org/10.1038/srep11059>
- Verdiansyah, M. (2023). Penerapan metode regresi komponen utama kernel untuk prediksi harga rumah. *Bandung Conference Series: Statistics*, 3(2), 653–661. <https://doi.org/10.29313/bcss.v3i2.9084>
- Verma, A., Lamsal, K., & Verma, P. (2021). An investigation of skill requirements in artificial intelligence and machine learning job advertisements. *Industry and Higher Education*, 36(1), 63–73. <https://doi.org/10.1177/0950422221990990>
- Wardani, N. W., Nugraha, P. G. S. C., & Mahendra, G. S. (2024). Implementasi Naive Bayes pada data mining untuk mengklasifikasikan penjualan barang terlaris pada perusahaan ritel. *Jurnal Sains dan Teknologi*, 12(3), 656–668.