

Implementasi *Machine Learning* Dengan Metode Regresi Linear Untuk Prediksi Gaji Karyawan Berdasarkan Masa Kerja

Ismail^{*1}, Ainun Hidayah²

^{1,2}Program Studi Sistem Informasi, Universitas Lamappapoleonro

^{1,2}Jl. Kesatria No.60 Watansoppeng, Soppeng Sulawesi Selatan

Email: ^{*1}ismai@unipol.ac.id, ²ainunhidayah68@gmail.com

(Naskah masuk: 19-04-2025, direvisi: 17-05-2025, diterbitkan: 28-05-2025)

Abstrak

Permasalahan yang muncul di perusahaan PT. Zhafran Putra Utama Kabupaten Soppeng adalah masih banyak karyawan yang sudah lama bekerja namun tidak mendapatkan kenaikan gaji. Hal ini terjadi karena perusahaan tersebut belum memiliki model atau cara untuk menentukan atau memprediksi besaran gaji yang diterima karyawan berdasarkan pengalaman lama bekerja atau masa kerja. Untuk mengetahui besaran gaji yang diterima karyawan berdasarkan lama bekerja dapat diatasi dengan menggunakan penerapan teknologi sistem cerdas yaitu *machine learning*. Teknologi *machine learning* dapat bekerja dengan baik apabila menggunakan metode tertentu sesuai dengan kasus yang akan diselesaikan. Pada kasus prediksi besaran gaji karyawan berdasarkan lama bekerja akan digunakan metode regresi linear sederhana. regresi linier sederhana adalah hubungan secara linear antara satu variabel independen (X) dengan variabel dependen (Y). Hasil implementasi Sistem *machine learning* dengan metode regresi linear untuk prediksi gaji karyawan berdasarkan masa kerja pada PT. Zhafran Putra Utama Kabupaten Soppeng didapatkan hasil prediksi bahwa dengan lama waktu mengabdikan karyawan selama 15 tahun maka diprediksi mendapatkan gaji karyawan sebesar 16700000. dan lama waktu mengabdikan karyawan selama 20 tahun maka diprediksi mendapatkan gaji karyawan sebesar 21373.500.

Kata kunci: *Machine Learning*, Regresi Linier, Prediksi Gaji Karyawan

Abstract

The problem that arises at PT. Zhafran Putra Utama in Soppeng Regency is that many employees who have been working for a long time have not received a salary increase. This occurs because the company does not yet have a model or method for determining or predicting the amount of salary received by employees based on their length of service or work experience. To determine the amount of salary received by employees based on their length of service, this can be addressed by implementing intelligent system technology, namely machine learning. Machine learning technology can work effectively when using specific methods appropriate for the case at hand. In the case of predicting employee salaries based on length of service, the simple linear regression method will be used. Simple linear regression is a linear relationship between one independent variable (X) and one dependent variable (Y). The results of implementing the machine learning system using the linear regression method to predict employee salaries based on length of service at PT. Zhafran Putra Utama in Soppeng Regency indicate that employees with 15 years of service are predicted to receive a salary of 16,700,000, and employees with 20 years of service are predicted to receive a salary of 21,373,500.

Keywords: *Machine Learning, Linear Regression, Employee Salary Prediction*

1. PENDAHULUAN

Trend perkembangan ilmu pengetahuan dan teknologi saat ini adalah revolusi industri 4.0. Integrasi teknologi digital berbasis *cyber physical system* pada hampir semua bidang pekerjaan menjadi ciri hadirnya revolusi industri tersebut (Purba et al., 2021). Teknologi ini memungkinkan terhubungnya manusia dengan manusia, manusia dengan mesin, dan mesin dengan mesin. Perubahan karakteristik pekerjaan menjadi

dampak tersendiri dari hadirnya revolusi industri 4.0. Karakteristik pekerjaan yang berubah akan mendisrupsi pekerjaan yang telah ada dan menggantikannya dengan pekerjaan dengan karakteristik baru (Sulistyanto et al., 2021). Karakteristik baru pada pekerjaan juga membutuhkan kompetensi baru pada para pekerja. Dengan demikian, sangat penting bagi suatu negara untuk mengembangkan kompetensi

berbasis industri 4.0 pada sumber daya manusianya (Ramadhan et al., 2023).

Sekarang ini, dalam dunia bisnis, setiap perusahaan harus bersiap untuk saling bersaing. Manajemen yang efektif dan keunggulan perlu dimiliki oleh setiap perusahaan untuk menyambut persaingan itu. Dalam hal ini, salah satu aspek yang berpengaruh besar terhadap kemajuan dan keberhasilan sebuah perusahaan adalah kinerja dari karyawannya (Saputra, 2022). Faktor yang berpengaruh secara internal dapat terlihat dari penentuan gaji yang tepat oleh sebuah perusahaan kepada karyawan (Saputra, 2022). Di samping itu juga perusahaan juga harus mau mengeluarkan bonus bagi para karyawannya yang berhasil bekerja maksimal sesuai yang dibutuhkan oleh perusahaan. Sementara itu, faktor yang berpengaruh secara eksternal dapat terlihat dari bagaimana sebuah perusahaan mampu memberikan pengembangan dan pelatihan pada para karyawannya (Tuah & Anyan, 2020). Dalam hal ini, perusahaan tidak dapat dipisahkan dengan tenaga kerja. Sekalipun perusahaan mempunyai teknologi canggih dan modal besar. Dengan tidak terdapatnya tenaga manusia sebagai penggerak perusahaan tidak akan mampu meraih tujuan yang direncanakan. Oleh karenanya, diharapkan sumber daya manusia dapat mengalami peningkatan tahun demi tahun, seiring dengan berkembangnya teknologi dan majunya zaman.

Karyawan menjadi bagian yang sangat penting bagi sebuah perusahaan. Tugas manajemen melakukan kegiatan penilaian kinerja karyawannya. Setiap perusahaan mengharapkan karyawan yang memiliki kinerja sesuai dengan standart perusahaan bahkan lebih dari yang diharapkan perusahaan (Luthfiansyah & Wasito, 2023). Penilaian kinerja (performance appraisal) menjadi kegiatan yang penting untuk menentukan apakah seorang karyawan dapat melakukan tugas yang diberikan sesuai dengan tanggung jawab yang diberikan oleh perusahaan dimana kenyataannya sistem penilaian kualitas kinerja karyawan sangat dibutuhkan oleh perusahaan .

Kinerja merupakan hasil kerja secara kualitas dan kuantitas dicapai oleh seorang karyawan dalam melaksanakan tugasnya sesuai dengan tanggung jawab yang diberikan kepadanya

untuk dikerjakan sesuai dengan kebutuhan. Karyawan bagian dari aset berharga untuk menjaga kesejahteraan dan kepuasan sebagai seorang karyawan karena sebuah perusahaan meskipun memiliki sarana dan prasarana yang baik belum tentu bisa berjalan dengan baik tanpa adanya karyawan yang cakap dalam bekerja dan memiliki motivasi kerja yang baik (Abijono et al., 2021). Banyak perusahaan memotivasi karyawannya dengan berbagai cara, seperti perusahaan PT. Zhafran Putra Utama Kabupaten Soppeng memberikan penghargaan dengan cara menaikkan gaji, menaikkan THR ataupun memberikan bonus Tujuan umumnya agar kinerja pegawai semakin baik dan stabil. Namun permasalahan yang muncul di perusahaan PT. Zhafran Putra Utama Kabupaten Soppeng adalah masih banyak karyawan yang sudah lama bekerja namun tidak mendapatkan kenaikan gaji. Hal ini terjadi karena perusahaan tersebut belum memiliki model atau cara untuk menentukan atau memprediksi besaran gaji yang diterima karyawan berdasarkan pengalaman lama bekerja atau masa kerja.

Untuk mengetahui besaran gaji yang diterima karyawan berdasarkan lama bekerja dapat diatasi dengan menggunakan penerapan teknologi sistem cerdas yaitu *machine learning*. Sudah banyak perusahaan atau bisnis yang menerapkan *Machine Learning* untuk membantu menyelesaikan berbagai permasalahan seperti sistem prediksi, klasifikasi, dan pengelompokan data sebagai penunjang pengambilan keputusan (Wirapati et al., 2025). *Machine Learning* adalah ilmu pengembangan algoritme dan model secara statistik yang digunakan sistem komputer untuk menjalankan tugas tanpa instruksi eksplisit, mengandalkan pola serta inferensi sebagai gantinya. Keunggulan *machine learning* dapat mengolah data lebih cepat dan akurat berdasarkan pengalaman pembelajaran komputer dengan pola tertentu (Pradana et al., 2024).

Teknologi *machine learning* dapat bekerja dengan baik apabila menggunakan metode tertentu sesuai dengan kasus yang akan diselesaikan. Pada kasus prediksi besaran gaji karyawan berdasarkan lama bekerja akan

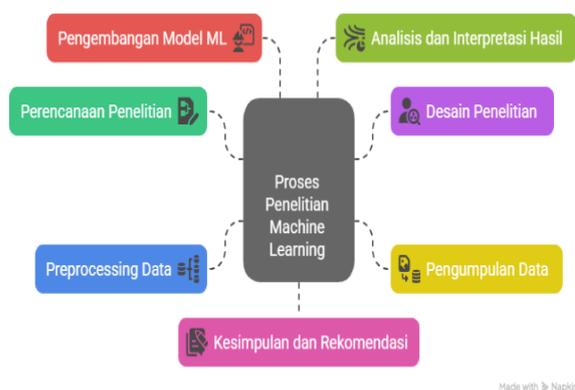
digunakan metode regresi linear sederhana. regresi linier sederhana adalah hubungan secara linear antara satu variabel independen (X) dengan variabel dependen (Y) (Aksa et al., 2025). Analisis regresi sederhana dapat digunakan untuk mengetahui arah dari hubungan antara variabel bebas dengan variabel terikat, apakah memiliki hubungan positif atau negatif serta untuk memprediksi nilai dari variabel terikat apabila nilai variabel bebas mengalami kenaikan ataupun penurunan.

2. METODE PENELITIAN

Untuk mencapai hasil penelitian yang baik, diperlukan metode penelitian yang jelas. Adapun metode tahapan penelitian sebagai berikut:

2.1 Tahapan Penelitian

Sebuah siklus untuk membangun sistem dan memberikannya kepada pengguna melalui tahapan perencanaan, analisa, perancangan dan implementasi dengan cara memahami dan menyeleksi keadaan dan proses yang dilakukan pengguna untuk dapat mendukung kebutuhan pengguna (Desmayani et al., 2021).



Gambar 1. Tahapan Penelitian

Gambar 1 merupakan tahapan-tahapan dalam proses penelitian. Berikut adalah penjelasan untuk setiap bagian yang ada pada gambar tersebut (Anggara et al., 2022):

1) Perencanaan Penelitian

Pada tahap ini, peneliti mengidentifikasi masalah yang ingin dipecahkan, yaitu dalam hal ini adalah prediksi gaji karyawan berdasarkan masa kerja menggunakan regresi linear.

2) Desain Penelitian

Desain ini mencakup pemilihan metode machine learning yang akan diterapkan, dalam hal ini regresi linear.

3) Pengumpulan Data

Pengumpulan data ini bisa dilakukan melalui survei, observasi, atau data sekunder yang sudah ada

4) Preprocessing Data

Sebelum data digunakan untuk membangun model machine learning, data perlu diproses lebih lanjut. Ini bisa mencakup pembersihan data (menghapus data yang hilang atau tidak relevan), transformasi data (misalnya normalisasi), dan membagi data menjadi data pelatihan dan data pengujian

5) Pengembangan Model ML

Proses ini mencakup pemrograman dan pelatihan model menggunakan data pelatihan yang telah diproses. Hasil model kemudian diuji dengan data pengujian untuk mengevaluasi efektivitasnya (Puspita et al., 2025).

6) Analisis dan Interpretasi Hasil

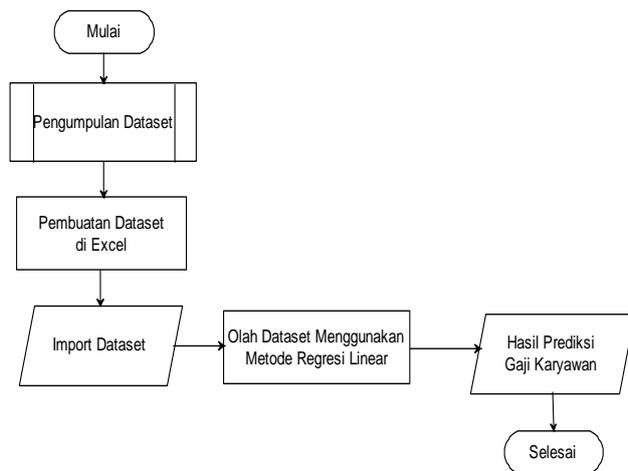
Peneliti menginterpretasikan koefisien regresi, melihat pengaruh variabel independen (masa kerja) terhadap variabel dependen (gaji), dan mengevaluasi akurasi model dengan metrik seperti R-squared atau Mean Squared Error

7) Kesimpulan dan Rekomendasi

Peneliti menarik kesimpulan berdasarkan hasil penelitian dan memberikan rekomendasi untuk organisasi atau untuk penelitian lebih lanjut.

2.2 Perancangan Sistem

Langkah-langkah yang dilakukan pada perancangan sistem ini adalah membuat usulan pemecahan masalah secara logika dan usulan-usulan lainnya. Alat bantu yang digunakan adalah Data flow diagram/Sistem *Flowchart*. Kegunaan data flow diagram ini adalah pemakai dapat dengan mudah mengerti sistem yang akan dikerjakan atau dikembangkan (Ismail Ismail, Rezky Erwin Syah, 2024).



Gambar 2. Dokumen flowchart sistem

Adapun penjelasan gambar diatas yaitu pertama mengumpulkan dataset, kemudian diinput kedalam file excel, setelah itu file dataset diimport edalam bahasa pemrograman python menggunakan jupyter notebook. Selanjutnya dataset akan diproses menggunakan model regresi linear untuk mendapatkan prediksi gaji karyawan berdasarkan masa kerjanya.

3. HASIL DAN PEMBAHASAN

Setelah dilakukan perancangan sistem. Selanjutnya dilakukan implementasi sistem untuk mendapatkan hasil. Penelitian ini bertujuan menghasilkan pola analisis prediksi gaji karyawan dengan penerapan model machine learning regresi linear:

3.1 Pembuatan Dataset

Dataset yang terstruktur dengan baik akan sangat membantu hasil data analysis yang dilakukan, membuat jadi lebih cepat, efisien, dan akurat (Ismail, 2025). Berikut dataset *machine learning* dengan metode regresi linear untuk prediksi gaji karyawan berdasarkan masa kerja pada PT. Zhafran Putra Utama Kabupaten Soppeng:

Tabel 1: Dataset Karyawan

Years, Experience, Salary
1.1,39343
1.3,46205
1.5,37731
2.0,43525
2.2,39891

2.9,56642
3.0,60150
3.2,54445
3.2,64445
3.7,57189
3.9,63218
4.0,55794
4.0,56957
4.1,57081
4.5,61111
4.9,67938
5.1,66029
5.3,83088
5.9,81363
6.0,93940
6.8,91738
7.1,98273
7.9,101302
8.2,113812
8.7,109431
9.0,105582
9.5,116969
9.6,112635
10.3,122391
10.5,121872

Tabel di atas menampilkan data tentang tahun pengalaman kerja dan gaji yang diterima oleh karyawan. Tabel ini menampilkan serangkaian data di mana setiap entri terdiri dari tahun pengalaman kerja seorang karyawan dan gaji yang diterimanya. Misalnya, seorang karyawan dengan 1,1 tahun pengalaman mendapatkan gaji sebesar 39.343, sementara seorang karyawan dengan 8,2 tahun pengalaman mendapatkan gaji sebesar 113.812. Data ini dapat digunakan untuk analisis guna memahami hubungan antara tahun pengalaman dan gaji, misalnya melalui analisis regresi atau teknik machine learning lainnya untuk memprediksi gaji berdasarkan pengalaman.

3.2 Menampilkan Variabel Output (Y)

Pada tahap proses menampilkan variabel output dilakukan untuk melihat format output yang akan digunakan sebagai model proses data

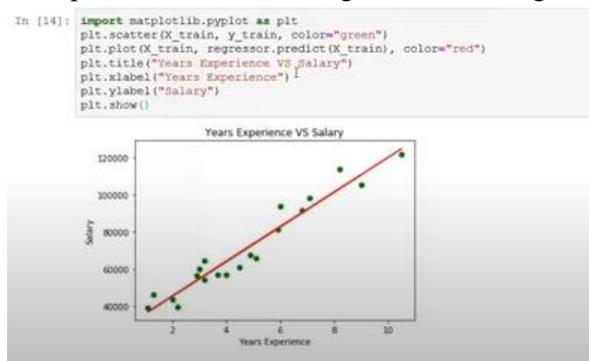
output sistem prediksi. Berikut tampilan algoritma menampilkan variabel output gaji (Y).

```
In [21]: X = df.iloc[:, :-1]
        y = df.iloc[:, 1]
        y
Out[21]: 0    39343
        1    46205
        2    37731
        3    43525
        4    39891
        5    56642
        6    60150
        7    64444
```

Gambar 3. menampilkan variabel output (Y) Gambar diatas merupakan kode Python yang digunakan untuk memisahkan data dalam sebuah DataFrame. Pada kode tersebut, `x = df.iloc[:, :-1]` digunakan untuk memilih semua kolom kecuali kolom terakhir (variabel *YearsExperience*), yang berarti x berisi data fitur atau variabel independen. Sedangkan `y = df.iloc[:, 1]` digunakan untuk memilih kolom kedua, yaitu kolom *Salary*, yang berisi data target atau variabel dependen. Dalam hal ini, x berisi tahun pengalaman kerja dan y berisi gaji yang akan diprediksi. Kode ini merupakan langkah pertama dalam persiapan data untuk pelatihan model machine learning, seperti regresi linear.

3.3 Pembagian data training dan data testing

Pada tahap ini dilakukan pembagian data training dan data testing. Hal ini dilakukan untuk menguji proses prediksi menggunakan metode regresi linear. Berikut tampilan algoritma pembagian data testing dan data training. Berikut hasil pemodelan data training dan data testing



Gambar 4. Hasil pemodelan data

Pada kode tersebut, pertama, fungsi `plt.scatter(x_train, y_train, color="green")` digunakan untuk membuat plot titik (scatter plot)

yang menunjukkan hubungan antara data pelatihan (`x_train: YearsExperience` dan `y_train: Salary`). Titik-titik ini digambarkan dengan warna hijau. Kemudian, `plt.plot(X_train, regressor.predict(X_train), color="red")` menggambarkan garis regresi linear (warna merah) yang dihasilkan oleh model, menunjukkan prediksi gaji berdasarkan tahun pengalaman kerja. Grafik ini memiliki judul "*Years Experience VS Salary*" dan label sumbu yang sesuai, yaitu "*Years Experience*" untuk sumbu x dan "*Salary*" untuk sumbu y. Visualisasi ini membantu untuk mengamati kecocokan antara data aktual dan prediksi model regresi.

3.4 Pengujian Data

Pengujian data baru yang pertama dilakukan dengan memberikan inputan variabel x yang tidak ada pada data set data training maupun data testing. Berikut bentuk algoritma pengujian data:

```
MAKE NEW PREDICTION

In [19]: salary_pred = regressor.predict([[15]])
        print("The salary for that amount of years experience is: ", salary_pred)
        The salary for that amount of years experience is: [167005.32889097]

In [ ]:
```

Gambar 5. Hasil Prediksi

Dari hasil pengujian pertama dengan menggunakan nilai variabel lama kerja ($X=15$) maka hasil prediksi gaji kasyawan adalah 16700000.

pengujian selanjutnya dengan menggunakan nilai variabel lama kerja ($X=20$) maka hasil prediksi gaji kasyawan adalah 21373.500

4. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan tentang *machine learning* dengan metode regresi linear untuk prediksi gaji karyawan berdasarkan masa kerja, maka dapat ditarik kesimpulan yaitu Hasil perancangan sistem dengan menggunakan *machine learning* dapat memudahkan proses prediksi gaji kasyawan. Perancangan sistem dapat membentuk kelompok data yang konsisten

sehingga memudahkan untuk diolah lebih lanjut dalam prediksi menggunakan algoritma regresi linear. Hasil implementasi Sistem *machine learning* dengan metode regresi linear untuk prediksi gaji karyawan berdasarkan masa kerja pada PT. Zhafran Putra Utama Kabupaten Soppeng didapatkan hasil prediksi bahwa dengan lama waktu mengabdikan karyawan selama 15 tahun maka diprediksi mendapatkan gaji karyawan sebesar 16700000. dan lama waktu mengabdikan karyawan selama 20 tahun maka diprediksi mendapatkan gaji karyawan sebesar 21373.500.

5. DAFTAR PUSTAKA

- Abijono, H., Santoso, P., & Anggreini, N. L. (2021). Algoritma Supervised Learning Dan Unsupervised Learning Dalam Pengolahan Data. *Jurnal Teknologi Terapan: G-Tech*, 4(2), 315–318. <https://doi.org/10.33379/gtech.v4i2.635>
- Aksa, A. N., Achmad, A., Arda, A. L., Komputer, S., & Makassar, U. H. (2025). Analisis Metode Decision Tree dan Regresi Logistik Sebagai Sistem Rekomendasi Kenaikan Golongan Berdasarkan Kinerja Pegawai pada Universitas Lamappapoleonro ditingkatkan selain dari kinerja sistem untuk terus bisa mengajukan kenaikan. *Jurnal Saintekom : Sains, Teknologi, Komputer Dan Manajemen*, 15(1), 42–55.
- Anggara, E. D., Widjaja, A., & Suteja, B. R. (2022). Prediksi Kinerja Pegawai sebagai Rekomendasi Kenaikan Golongan dengan Metode Decision Tree dan Regresi Logistik. *Jurnal Teknik Informatika Dan Sistem Informasi*, 8(1), 218–234. <https://doi.org/10.28932/jutisi.v8i1.4479>
- Desmayani, N. M. M. R., Wardani, N. W., Nugraha, P. G. S. C., & Mahendra, G. S. (2021). Sistem Informasi Laporan Keuangan pada Salon Berbasis Website Dengan Metode SDLC. *Jurnal Sistem Informasi Dan Komputer Terapan Indonesia (JSIKTI)*, 4(2), 68–77. <https://doi.org/10.33173/jsikti.118>
- Ismail Ismail, Rezky Erwin Syah, M. A. T. (2024). Klasifikasi Data Mining Pada Tingkat Kepuasan Pengunjung. *Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika (JISTI)*, 7(2), 270–281.
- Ismail, Su. (2025). Data Mining Klasifikasi Penduduk Miskin Menggunakan Metode Support Vektor Machine. *Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika (JISTI)*, 8(April).
- Luthfiansyah, R., & Wasito, B. (2023). Penerapan Teknik Deep Learning (Long Short Term Memory) dan Pendekatan Klasik (Regresi Linier) dalam Prediksi Pergerakan Saham BRI. *Jurnal Informatika Dan Bisnis*, 12(2), 42–54. <https://doi.org/10.46806/jib.v12i2.1059>
- Pradana, R. Y., Nastiti, F. E., & Oktaviani, I. (2024). Machine Learning Pengklasifikasikan Performa Karyawan Direct Sales Force Kartu Prabayar Menggunakan Metode Random Forest Classifier. *JEKIN - Jurnal Teknik Informatika*, 4(3), 590–599. <https://doi.org/10.58794/jekin.v4i3.864>
- Purba, N., Yahya, M., & Nurbaiti. (2021). Revolusi Industri 4.0 : Peran Teknologi Dalam Eksistensi Penguasaan Bisnis Dan Implementasinya. *Jurnal Perilaku Dan Strategi Bisnis*, 9(2), 91–98.
- Puspita, D., Nilam, S., Arifyanto, M. I., Studi, P., & Sains, M. (2025). PREDICTION OF ELECTRICITY BILL PAYMENT DELAYS FOR CUSTOMERS USING A MACHINE LEARNING APPROACH LISTRIK PELANGGAN DENGAN PENDEKATAN MACHINE. *JURNAL INOVTEK POLBENG - SERI INFORMATIKA*, 10(1), 446–457.
- Ramadhan, B., Firdaus, D., & Adiningrum, N. T. R. (2023). Analisis Data Pegawai Untuk Memprediksi Gaji Berdasarkan Faktor-Faktor Spesifik Dengan Pendekatan Machine Learning. *Naratif: Jurnal Nasional Riset, Aplikasi Dan Teknik Informatika*, 5(2), 131–139. <https://doi.org/10.53580/naratif.v5i2.205>
- Saputra, M. A. (2022). Prediksi Gaji Berdasarkan Pengalaman Bekerja Menggunakan Metode Regresi Linear. *Journal of Dinda*, 2(2), 58–63.
- Sulistyanto, S., Mutohhari, F., Kurniawan, A., & Ratnawati, D. (2021). Kebutuhan Kompetensi di Era Revolusi Industri 4.0: Review Perspektif Pendidikan Vokasional. *Jurnal Taman Vokasi*, 9(1), 25–35. <https://doi.org/10.30738/jtv.v9i1.7742>
- Tuah, Y. A. E., & Anyan, A. (2020). Implementasi Model Regresi Linear Sederhana Untuk Prediksi Gaji

Berdasarkan Pengalaman Lama Bekerja.
JUTECH: Journal Education and Technology, 1(2), 56–70.
<https://doi.org/10.31932/jutech.v1i2.1289>

Wirapati, S., Gede, L., Widiartha, I. M., & Santiyasa, I. W. (2025). Prediksi Pengunduran Diri Karyawan Dengan Pendekatan Deep Neural Network Dalam Machine Learning. *Jurnal Elektronik Ilmu Komputer Udayana*, 13(3), 571–576.