



ANALISIS MODEL PREDIKSI KINERJA KARYAWAN PADA SEBUAH PERUSAHAAN GARMEN MENGGUNAKAN *MACHINE LEARNING*

¹⁾Wisnu Naf'ul Huja, ²⁾Muslim Hidayat, ³⁾Muhamad Fuat Asnawi

^{1,2,3)}Manajemen Informatikas, Fakultas Teknik dan Komputer, Universitas sains Al-Qur'an Jawa Tengah

¹⁾ wisnu5149@gmail.com, ²⁾ muslim_h@unsiq.ac.id, ³⁾ fuatasnawi@unsiq.ac.id

INFO ARTIKEL

Riwayat Artikel :

Diterima : 8 Januari 2025

Disetujui : 29 Januari 2025

Kata Kunci :

Model Prediksi, pembelajaran mesin, regresi linier.

ABSTRAK

Penelitian ini bertujuan untuk menganalisis dan memprediksi tingkat efektivitas buruh di pabrik garmen menggunakan metode regresi linier dan teknik machine learning. Efektivitas buruh dipengaruhi oleh faktor-faktor seperti keterampilan, pengalaman, motivasi, kondisi fisik, dan penggunaan teknologi. Model regresi linier, bersama dengan algoritma machine learning, digunakan untuk mengidentifikasi hubungan antara variabel-variabel tersebut dengan tingkat efektivitas buruh. Hasil analisis ini diharapkan dapat membantu manajemen dalam meningkatkan efisiensi operasional, mengurangi biaya produksi, serta melakukan intervensi yang tepat waktu untuk mengatasi penurunan produktivitas. Dengan demikian, perusahaan dapat lebih efektif dalam meningkatkan kinerja buruh dan pengambilan keputusan berbasis data.

ARTICLE INFO

Article History :

Received : January 8, 2025

Accepted : January 29, 2025

Keywords:

prediction, machine learning, linear regression.

ABSTRACT

This study aims to analyze and predict labor effectiveness in garment factories using linear regression methods and machine learning techniques. Labor effectiveness is influenced by factors such as skills, experience, motivation, physical condition, and technology usage. The linear regression model, along with machine learning algorithms, is used to identify the relationship between these variables and labor effectiveness. The analysis results are expected to help management improve operational efficiency, reduce production costs, and implement timely interventions to address productivity declines. Thus, the company can more effectively enhance labor performance and data-driven decision-making.

1. PENDAHULUAN

Pada era kemajuan teknologi ini, industri manufaktur, termasuk industri garmen, mengalami berbagai perubahan yang signifikan dalam hal efisiensi dan efektivitas operasional. Seiring dengan perkembangan teknologi dan otomatisasi, banyak pabrik garmen yang mulai mengintegrasikan mesin dan teknologi canggih dalam proses produksinya. Namun, masih banyak perusahaan yang mengandalkan tenaga manusia dalam sejumlah besar tahap produksi. Salah satu tantangan terbesar yang dihadapi oleh perusahaan garmen adalah bagaimana cara mengelola dan memantau efektivitas buruh yang memiliki kontribusi langsung terhadap proses produksi.

Sebagai seorang Data Scientist, saya diberi tugas untuk membantu manajemen dalam memprediksi dan menganalisa tingkat efektivitas buruh di pabrik garmen. Salah satu teknik yang dapat digunakan untuk memprediksi dan menganalisa efektivitas buruh adalah penerapan machine learning (ML), yang mampu menganalisis data secara lebih mendalam dan memberikan insight yang lebih tajam mengenai faktor-faktor yang mempengaruhi kinerja buruh (Mahesh, 2020).

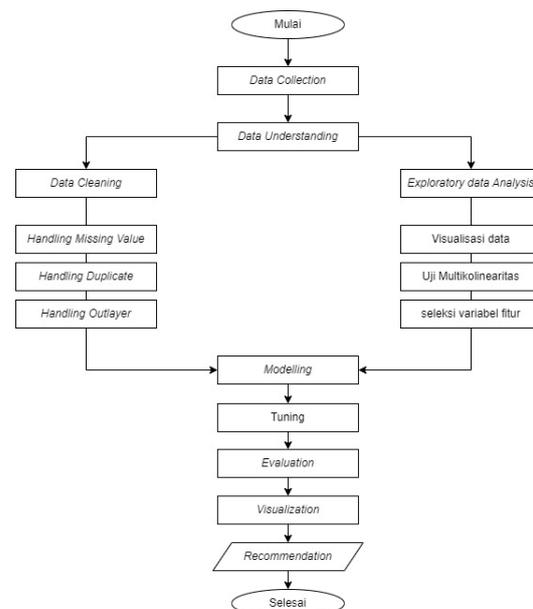
Dengan adanya model prediksi yang akurat, manajemen dapat melakukan intervensi yang lebih tepat waktu apabila terjadi penurunan dalam tingkat efektivitas kerja buruh (Jumlah et al., 2020). Sebagai contoh, jika seorang buruh menunjukkan penurunan dalam produktivitas, analisis berbasis data dapat membantu mengidentifikasi faktor penyebabnya, seperti masalah kesehatan, kelelahan, atau bahkan masalah dalam keterampilan yang dapat diperbaiki melalui pelatihan. Lebih lanjut, dengan memanfaatkan data ini, perusahaan dapat memberikan insentif kepada buruh yang menunjukkan kinerja baik dan konsisten, serta melakukan pembinaan pada buruh yang membutuhkan peningkatan kinerja.

Pentingnya penggunaan data untuk merumuskan kebijakan berbasis bukti yang dapat meningkatkan efektivitas operasional pabrik (Kristiawan & Widjaja, 2021). CEO dan manajemen pabrik perlu memiliki gambaran yang jelas mengenai performa tenaga kerja agar bisa melakukan keputusan yang lebih tepat dalam meningkatkan produktivitas. Dengan

menggunakan teknik analisis data dan machine learning yang sesuai, diharapkan perusahaan dapat meningkatkan efisiensi produksi dan mengurangi biaya operasional secara signifikan. Dengan demikian, melalui analisis data yang komprehensif dan penerapan machine learning, diharapkan perusahaan dapat membangun model prediksi yang efektif untuk memantau dan meningkatkan efektivitas buruh (Nayem & Uddin, 2024). Hal ini akan membantu CEO dan manajemen pabrik untuk mengambil keputusan yang lebih tepat waktu dan strategis, serta berkontribusi pada peningkatan produktivitas dan efisiensi di seluruh lini produksi pabrik garmen.

2. METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan desain deskriptif dan analisis regresi. Pendekatan kuantitatif dipilih karena penelitian ini bertujuan untuk menganalisis hubungan antar variabel secara statistik, khususnya untuk mengukur dan mengevaluasi pengaruh berbagai faktor terhadap kinerja karyawan. Desain penelitian deskriptif digunakan untuk menggambarkan secara sistematis kondisi yang terjadi pada objek penelitian, yaitu kinerja karyawan di perusahaan yang menjadi fokus studi.



Gambar 2. 1. Alur penelitian

Proses pengumpulan data dilakukan melalui kerjasama antara peneliti dan pihak perusahaan, yang telah mengumpulkan dan menyimpan data karyawan dalam sistem internal mereka. Data

beberapa sumber lain menyarankan batas atas yang berbeda, seperti antara 5 hingga 10. Berikut cara untuk mengambil nilai VIF

```
# Menambahkan konstanta (intercept) ke DataFrame untuk menghitung VIF
X = df_no_duplicate.copy()

# Memastikan tidak ada kolom yang berisi string atau non-numerik
X = X.select_dtypes(include=['float64', 'int64'])

# Menambahkan konstanta (intercept) untuk perhitungan VIF
X = add_constant(X)

# Menghitung VIF untuk setiap fitur
vif_data = pd.DataFrame()
vif_data['Variable'] = X.columns
vif_data['VIF'] = [variance_inflation_factor(X.values, i) for i in range(X.shape[1])]
```

Gambar 3. 2. Code pengambilan nilai VIF

Setelah dilakukan pencarian nilai dengan cara di atas menghasilkan nilai VIF pada setiap variabel sebagai berikut.

Pada gambar terlihat beberapa variabel yang memiliki nilai VIF yang seimbang. Dari data tersebut dapat diambil beberapa variabel

	Variable	VIF
0	const	74.429428
1	quarter	1.068719
2	department	9.037680
3	day	1.032927
4	team	1.034672
5	targeted_productivity	1.286623
6	smv	6.343289
7	wip	1.217854
8	over_time	2.485263
9	incentive	1.029618
10	idle_time	1.486236
11	idle_man	1.568151
12	no_of_style_change	1.391819
13	no_of_workers	14.838493
14	actual_productivity	1.351060

Gambar 3. 5. Nilai VIF pada variabel

fitur yang akan digunakan sebagai bahan model prediksi, yang tertera di bawah ini.

1. quarter
2. department
3. day
4. team
5. targeted_productivity
6. smv
7. wip
8. over_time
9. no_of_style_change
10. incentive
11. idle_time
12. idle_man

3.4. Modelling

Pada tahap ini dilakukan proses pembuatan dan pelatihan model menggunakan algoritma machine learning untuk memprediksi atau mengklasifikasikan data berdasarkan data historis.

Berikut tahapan proses modelling.

1. Mempersiapkan data

```
# Memisahkan fitur (X) dan target (y)
X = df_no_duplicate.drop(columns=['actual_productivity'])
y = df_no_duplicate['actual_productivity']
```

Gambar 3. 3. Code mempersiapkan data

Fitur (X) adalah kolom-kolom lain selain actual_productivity, yang digunakan untuk memprediksi actual_productivity (target). Sedangkan target (y) adalah kolom actual_productivity, yang merupakan variabel yang ingin kita prediksi.

2. Pembagian data

```
# Pembagian data menjadi data latih dan data uji (80% latih, 20% uji)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Gambar 3. 4. Code pembagian data

Data dibagi menjadi dua bagian. 80% untuk pelatihan dan 20% untuk pengujian menggunakan train_test_split dari sklearn.model_selection.

3. Melatih Model

Setelah data dibagi, langkah berikutnya adalah melatih model menggunakan data latih. Model machine learning yang digunakan dalam proses ini bisa bervariasi. Beberapa model yang digunakan untuk masalah regresi seperti ini yaitu Linear Regression, Random Forest Regressor, Decision Tree Regressor, Support Vector

```
# Memisahkan fitur (X) dan target (y)
X = df_no_duplicate.drop(columns=['actual_productivity'])
y = df_no_duplicate['actual_productivity']

# Pembagian data menjadi data latih dan data uji (80% latih, 20% uji)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# --- Linear Regression ---
linear_model = LinearRegression()
linear_model.fit(X_train, y_train)

# --- Random Forest Regressor ---
rf_model = RandomForestRegressor(n_estimators=100, random_state=42)
rf_model.fit(X_train, y_train)

# --- Decision Tree Regressor ---
dt_model = DecisionTreeRegressor(random_state=42)
dt_model.fit(X_train, y_train)

# --- Support Vector Regressor (SVR) ---
svr_model = SVR(kernel='rbf')
svr_model.fit(X_train, y_train)

# --- XGBoost Regressor ---
xgb_model = xgb.XGBRegressor(objective='reg:squarederror', random_state=42)
xgb_model.fit(X_train, y_train)
```

Gambar 3. 6. Code melatih model

Regressor (SVR), dan XGBoost Regressor.

3.5. Evaluation

Pada tahap ini dilakukan evaluasi untuk menilai kinerja model setelah dilatih menggunakan data uji yang belum pernah dilihat oleh model sebelumnya. Tahap ini sangat

penting karena membantu kita memahami apakah model tersebut benar-benar dapat menggeneralisasi dengan baik pada data yang baru dan tidak hanya mengingat data pelatihan.

Untuk menilai apakah suatu model regresi baik atau tidak, kita harus mempertimbangkan beberapa hal, yaitu R^2 yang tinggi dan metrics rendah (nilai MAE, MSE dan RMSE).

Berikut hasil evaluasi data yang sebelumnya sudah dilakukan tahap modelling.

Evaluation Results for Models Without Hyperparameter Tuning:				
	MSE	MAE	RMSE	R2
Linear Regression	0.024961	0.122118	0.157990	0.059941
Random Forest	0.017302	0.093447	0.131537	0.348385
K-Nearest Neighbors	0.025548	0.115978	0.159836	0.037846
XGBoost	0.025787	0.124615	0.160585	0.028812
Support Vector Regressor (SVR)	0.026141	0.124378	0.161682	0.015497
Decision Tree Regressor	0.026117	0.110507	0.161606	0.016418

Gambar 3. 7. Hasil evaluasi model prediksi

Berdasar pada gambar di atas, model Random Forest menjadi model yang memiliki kinerja terbaik berdasar pada metrik MSE, MAE dan RMSE cenderung memberikan nilai prediksi yang lebih akurat dan dengan nilai R^2 yang paling tinggi.

3.6. Hyperparameter Tuning

Hyperparameter Tuning merupakan proses penyesuaian hyperparameter dalam sebuah model machine learning untuk meningkatkan kinerjanya.

Tuning dilakukan dengan menggunakan teknik pencarian hyperparameter untuk memilih set nilai terbaik untuk hyperparameter model. Dengan memilih kombinasi hyperparameter yang optimal, model bisa memiliki kinerja yang lebih baik dalam hal akurasi.

Tuning hyperparameter dapat membantu model untuk tidak hanya bekerja dengan baik pada data pelatihan, tetapi juga dapat melakukan generalisasi yang baik pada data yang tidak terlihat (data uji).

Berikut nilai evaluasi setelah dilakukan tuning pada dataset.

Evaluation Results for Models With Hyperparameter Tuning:				
	MSE	MAE	RMSE	R2
Linear Regression	0.024961	0.122118	0.157990	0.059941
Random Forest	0.017204	0.093245	0.131166	0.352057
K-Nearest Neighbors	0.025548	0.115978	0.159836	0.037846
XGBoost	0.025787	0.124615	0.160585	0.028812
Support Vector Regressor (SVR)	0.024218	0.122100	0.155622	0.087913
Decision Tree Regressor	0.026117	0.110507	0.161606	0.016418

Gambar 3. 8. Nilai evaluasi model setelah proses tuning

Terdapat peningkatan yang terlihat pada beberapa model, dan model random forest masih

memiliki nilai terbaik dan semakin baik setelah dilakukan Hyperparameter Tuning pada data. Dengan menghasilkan nilai MSE, MAE dan RMSE semakin kecil, disertai nilai R-squared semakin besar.

4. PENUTUP

4.1. Simpulan

Berdasarkan hasil analisis yang dilakukan, didapatkan kesimpulan dari penerapan machine learning terhadap kinerja karyawan. Ada dua masalah utama yang dibahas yaitu model prediksi yang sesuai. Random Forest menjadi model prediksi yang sesuai untuk memprediksi kinerja karyawan, dengan mendapatkan nilai MSE, MAE, dan RMSE terkecil serta nilai R-squared terbesar.

4.2. Saran

Penambahan variabel dan jumlah data sebagai bahan penelitian dapat meningkatkan nilai prediksi yang lebih baik, serta dapat memberikan rekomendasi untuk meningkatkan produksi dan kinerja karyawan pada perusahaan menjadi lebih akurat.

5. DAFTAR PUSTAKA

- Kashyap, A. A., Raviraj, S., Devarakonda, A., Nayak K, S. R., Santhosh, K. V., & Bhat, S. J. (2022). Traffic flow prediction models—A review of deep learning techniques. In *Cogent Engineering* (Vol. 9, Issue 1). Cogent OA. <https://doi.org/10.1080/23311916.2021.2010510>
- Nayem, Z., & Uddin, M. A. (2024). Unbiased employee performance evaluation using machine learning. *Journal of Open Innovation: Technology, Market, and Complexity*, 10(1). <https://doi.org/10.1016/j.joitmc.2024.100243>
- Sriningsih, M., Hatidja, D., & Prang, J. D. (n.d.). *PENANGANAN MULTIKOLINEARITAS DENGAN MENGGUNAKAN ANALISIS REGRESI KOMPONEN UTAMA PADA KASUS IMPOR BERAS DI PROVINSI SULUT*.
- Winkler, W. E., & Bureau, / U S. (2003). *Data Cleaning Methods*.
- Mahesh, B. (2020). Machine Learning Algorithms - A Review. *International Journal of Science and Research (IJSR)*, 9(1), 381–386. <https://doi.org/10.21275/art20203995>
- Ramadhan, B., & Pane, S. F. (2024). Pengaruh Hyperparameter Tuning untuk Efektivitas pada Pendekatan Hybrid dalam Mendiagnosis Stres dan Depresi: Tinjauan Studi Literatur. *Jurnal Tekno Insentif*, 18(2), 104–118.
- Raparathi, M. (n.d.). Exploratory Data Analysis Techniques-A Comprehensive Review: Reviewing various exploratory data analysis techniques and their applications in uncovering insights from raw data. In *Australian Journal of Machine Learning Research & Applications By Sydney Academics* 215



Australian Journal of Machine Learning Research & Applications (Vol. 4).

Crowley, J. L. (2020). *Pattern Recognition and Machine Learning Generative Networks: EigenSpace Coding, Auto-Encoders, Variational Autoencoders and Generative Adversarial Networks*.

Soebroto, A. A. (2019). *Buku Ajar AI, Machine Learning & Deep Learning*.
<https://www.researchgate.net/publication>

Budi Trisno, I., & Agung Raharja, M. (2023). WEBINAR ARTIFICIAL INTELLIGENCE DAN MACHINE LEARNING. *JPM Jurnal Pengabdian Mandiri*, 2(11).

Rudianto, P., Unit, A., Kartini, T., Wintarti, A., Rakhmawati, L., Hapsari, R., Rochmawati, P. A. T. N., Aribowo, W., & Yustanti, W. (n.d.). *PYTHON PROGRAMMING: MACHINE LEARNING*.

Kristiawan, K., & Widjaja, A. (2021). Perbandingan Algoritma Machine Learning dalam Menilai Sebuah Lokasi Toko Ritel. *Jurnal Teknik Informatika Dan Sistem Informasi*, 7(1).