

Klasifikasi Tingkat Produktivitas Pegawai Garmen Menggunakan Algoritma *Naïve Bayes*

Desak Putu Sri Wulandari^{a1}, I Ketut Gede Suhartana^{a2}

^aProgram Studi Informatika, Universitas Udayana
Bukit Jimbaran, Bali, Indonesia
¹desaksw31@yahoo.com
²ikg.suhartana@unud.ac.id

Abstract

Productivity level of garment employee classification aims to facilitate for companies to give appreciation to employees. In doing classification can use the Naïve Bayes Algorithm, which uses probability and statistical methods to predict opportunities base on experience. The classification of garment employees is categorized into three, such as low, medium, and high. The results of the classification of 1197 data employees, obtained 981 employees have high productivity, 190 employees with moderate productivity, 26 employees with low productivity. Evaluation on the classification obtained an accuracy value of 82,07%.

Keywords: Classification, Garment Employee, Naïve Bayes, Productivity Level

1. Pendahuluan

Industri garmen merupakan jenis industri yang bergerak dalam pembuatan barang tekstil. Dalam operasionalnya, industri ini memiliki banyak proses yang dilakukan sehingga memerlukan banyak tenaga kerja. Berdasarkan hal tersebut, kemajuan industri garmen bergantung pada kinerja tenaga kerja yang dimilikinya. Oleh karena itu, perusahaan pada industri ini akan memberikan banyak apresiasi untuk para tenaga kerja yang memiliki tingkat produktifitas yang baik. Semakin baik tingkat produktifitas dari kinerja tenaga kerja, maka semakin besar pula apresiasi yang diberikan. Akan tetapi, penilaian kinerja tenaga kerja masih banyak dilakukan secara manual, sehingga tenaga kerja membutuhkan waktu yang lama untuk menerima apresiasi yang diberikan.

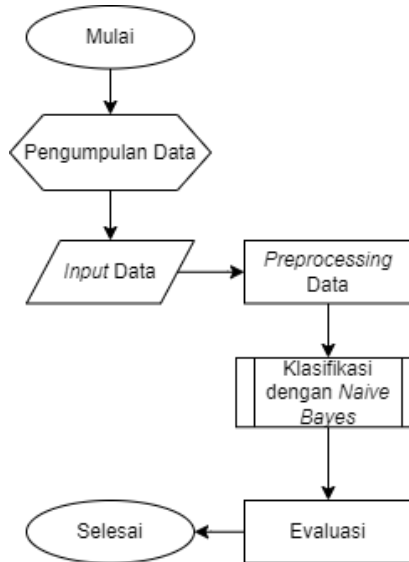
Lambatnya pemberian apresiasi ini akan menurunkan tingkat kepercayaan tenaga kerja pada perusahaan. Hal ini akan berpengaruh pada tingkat kemajuan perusahaan. Oleh karena itu, diperlukan suatu cara dalam mengelompokkan tenaga kerja berdasarkan tingkat produktifitasnya. Salah satunya melakukan klasifikasi menggunakan algoritma *Naïve Bayes*. *Naïve Bayes* merupakan suatu metode klasifikasi yang menggunakan metode probabilitas dan statistik untuk memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. *Naïve Bayes* untuk setiap kelas keputusan menghitung probabilitas dengan syarat bahwa kelas keputusan adalah benar, mengingat vector informasi objek. Algoritma ini mengasumsikan bahwa atribut obyek adalah independen (Olson Delen, 2008).

Algoritma *Naïve Bayes* kerap kali digunakan pada penelitian sebelumnya dengan hasil yang lebih baik dalam menangani masalah klasifikasi. Pada penelitian sebelumnya, algoritma *Naïve Bayes* dibandingkan dengan algoritma *Support Vector Machine* (SVM) dalam klasifikasi pada studi kasus pemberian penerima beasiswa PPA dengan hasil akurasi algoritma *Naïve Bayes* sebesar 90,90% sedangkan algoritma *Support Vector Machine* (SVM) dengan hasil akurasi 89,25% [1]. Selain itu, terdapat penelitian lain yang menggunakan algoritma *Naïve Bayes* dengan algoritma *K-Nearest Neighbor* (KNN) dengan hasil akurasi sebesar 84,12% untuk algoritma *Naïve Bayes* sedangkan 83,18% untuk KNN [2]. Sehingga dapat disimpulkan bahwa algoritma *Naïve Bayes* dapat menjadi solusi yang optimal dalam menyelesaikan masalah klasifikasi.

2. Metode Penelitian

2.1. Desain Penelitian

Tahapan terkait penelitian mengenai klasifikasi tingkat produktivitas pegawai garmen dengan menggunakan algoritma *naïve bayes* dapat dilihat pada gambar 1.



Gambar 1. Desain Penelitian

2.2. Pengumpulan Data

Dataset yang digunakan pada penelitian ini merupakan data prediksi produktivitas pegawai garmen di Bangladesh yang dipublikasikan oleh Abdullah Al Imran pada website repositori *UCI Machine Learning*, diakses pada 25 September 2022 pukul 11:59:50 WITA pada alamat <https://archive.ics.uci.edu/ml/datasets/Productivity+Prediction+of+Garment+Employees>. *Dataset* ini berjumlah 1197 data dengan jumlah fitur data sebanyak 15 fitur [3]. Tipe data dari lima belas fitur tersebut dapat dilihat pada tabel 1.

Tabel 1 Daftar Fitur dan Tipe *Dataset*

Fitur	Tipe Data
date	Object
quarter	Object
department	Object
day	Object
team	Numerik
targeted_productivity	Numerik
smv	Numerik
wip	Numerik
over_time	Numerik
incentive	Numerik
idle_time	Numerik
idle_men	Numerik
no_of_style_change	Numerik

no_of_workers	Numerik
actual_productivity	Numerik

2.3. Pengolahan Data Awal

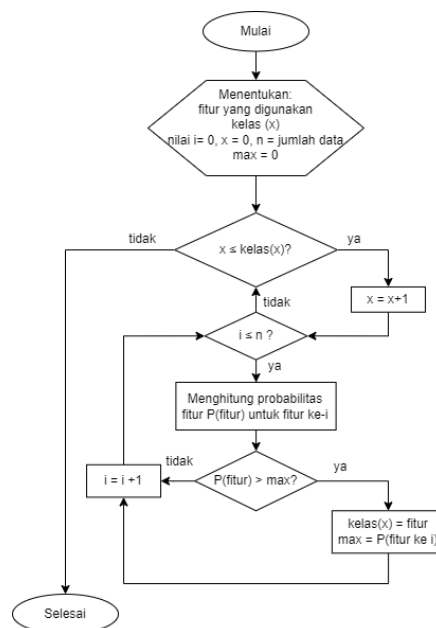
Dataset produktivitas pegawai garmen ini akan dilakukan tahap *preprocessing* yang diantaranya transformasi data, mencari korelasi antar fitur, pengisian nilai kosong (NULL) dan normalisasi data. Transformasi data dilakukan dengan tujuan untuk menyamakan tipe data asli menjadi bentuk lain yang sesuai dengan data asli lainnya sehingga memudahkan dalam memberikan perlakuan pada data. Dikarenakan fitur data yang banyak, maka pengurangan fitur dilakukan dengan mencari korelasi fitur. Berdasarkan hasil korelasi, fitur pada *dataset* yang dihilangkan antara lain *date*, *quarter*, dan *department* sehingga jumlah fitur yang ada menjadi 13 fitur. Oleh karena terdapat banyak data yang memiliki nilai NULL, maka dilakukan pengisian nilai data dengan menggunakan median data. Pengisian ini dilakukan guna menghindari *overfitting*. Pada *dataset* terdapat sebuah fitur yang memiliki nilai yang besar, sehingga rentang nilai data antar fitur serupa. Maka dari itu dilakukan normalisasi dengan menggunakan *Min-Max Normalization* dengan rumus sebagai berikut.

$$N = \frac{x - MinValue}{MaxValue - MinValue} \quad (1)$$

Keterangan:

- N = Hasil normalisasi
- X = Nilai yang dinormalisasi
- MaxValue = Nilai terbesar pada fitur yang dibandingkan
- MinValue = Nilai terkecil pada fitur yang dibandingkan

2.4. Klasifikasi menggunakan *Naïve Bayes*



Gambar 2 Flowchart Klasifikasi dengan Algoritma *Naïve Bayes*

Klasifikasi dengan menggunakan algoritma *naïve bayes* membutuhkan fitur yang memiliki keterhubungan, kelas yang digunakan serta jumlah data yang digunakan. Algoritma ini melakukan klasifikasi menggunakan probabilitas fitur pada setiap data. Adapun tahapan klasifikasi dapat dilihat pada gambar 2.

2.5. Evaluasi

Pada penelitian ini, pengujian dilakukan dengan menghitung akurasi. Akurasi dihitung dengan menggunakan rumus sebagai berikut.

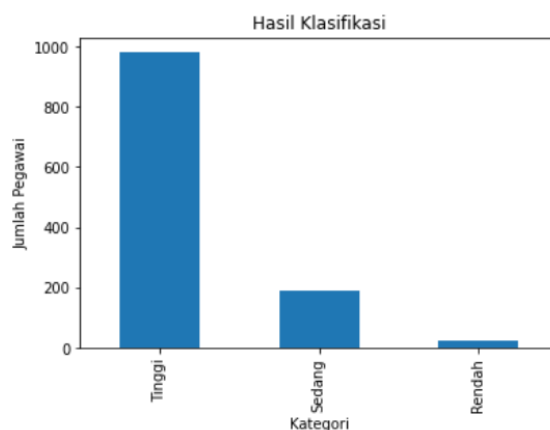
$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2)$$

Keterangan:

- TP = Jumlah positif yang diklasifikasikan sebagai positif
- TN = Jumlah negatif yang diklasifikasikan sebagai negatif
- FP = Jumlah positif yang diklasifikasikan sebagai negatif
- FN = Jumlah negatif yang diklasifikasikan sebagai positif

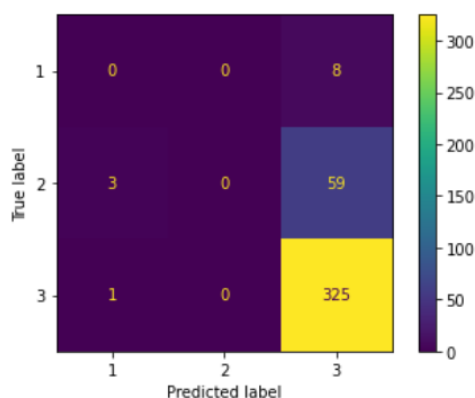
3. Hasil dan Pembahasan

Pada penelitian ini, produktivitas pegawai garmen dibagi menjadi tiga kategori yakni rendah, sedang, dan tinggi. Implementasi algoritma *Naïve Bayes* dalam penelitian ini dilakukan menggunakan Bahasa pemrograman python yang dibantu dengan library *Sklearn* dan *Math* dengan tahapan seperti gambar 2. Berdasarkan hasil klasifikasi, didapatkan data sejumlah 981 pegawai dengan tingkat produktivitas tinggi, 190 pegawai dengan tingkat produktivitas sedang, dan 26 pegawai dengan tingkat produktivitas yang rendah seperti yang tertera pada gambar 3.



Gambar 3. Hasil Klasifikasi Tingkat Produktivitas Pegawai Garmen

Berdasarkan hasil tersebut, dilakukan evaluasi dari model klasifikasi tingkat produktivitas dengan pembagian data latih dan data uji sebesar 70:30. Evaluasi dilakukan dengan menghitung akurasi dengan menggunakan *confusion matrix*. Hasil dari evaluasi terhadap hasil klasifikasi dapat dilihat pada gambar 4.



Gambar 4 Confusion Matrix

Berdasarkan matriks tersebut, didapatkan nilai TP = 325, TN = 0, FP = 59, dan FN = 12 yang kemudian dituangkan pada rumus akurasi berikut.

$$Akurasi = \frac{325}{396} \times 100\% = 82,07\%$$

Sehingga hasil perhitungan akurasi model klasifikasi tingkat produktivitas pegawai garmen sebesar 82,07%.

4. Kesimpulan

Berdasarkan penelitian yang dilakukan, dapat disimpulkan bahwa klasifikasi dengan menggunakan algoritma *Naïve Bayes* pada data produktivitas pegawai garmen menghasilkan 981 pegawai dengan tingkat produktivitas tinggi, 190 pegawai dengan tingkat produktivitas sedang, dan 26 pegawai dengan tingkat produktivitas rendah. Hasil klasifikasi tersebut memiliki menghasilkan akurasi sebesar 82,07%.

References

- [1] S.Linawati, R.A. Safitri, A.R. Alfiah, "Perbandingan Algoritma Klasifikasi Naive Bayes Dan Svm Pada Studi Kasus Pemberian Penerima Beasiswa PPA", *Jurnal Swabumi*, vol.8, no.1, pp 71-75, 2020.
- [2] A. A. Imran, "UCI Machine Learning Repository". [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Productivity+Prediction+of+Garment+Employees>. [22 September 2022]
- [3] Fitinline, "Fitinline", 19 November 2019. [Online]. Available: <https://fitinline.com/article/read/pengertian-industri-garment-dan-istilah-istilah-penting-didalamnya-yang-perlu-anda-ketahui/>. [22 September 2022]
- [4] M. Sarosa, M. Junus, M.U. Hoesny, "Classification Technique of Interviewer-Bot Result using Naïve Bayes and Phrase Reinforcement Algorithms", *International Journal of Emerging Technologies in Learning*, vol.13, no.02, pp 33-47, 2018.
- [5] S. K. Soumik, "Medium", 1 Juli 2020. [Online]. Available: <https://medium.com/analytics-vidhya/how-to-calculate-confusion-matrix-manually-14292c802f52>. [12 Oktober 2022]
- [6] M.H. Widiyanto, "Binus University". [Online]. Available: <https://binus.ac.id/bandung/2019/12/algoritma-naive-bayes/>. [22 September 2022]

Halaman ini sengaja dibiarkan kosong