

Aplikasi Rekrutmen Karyawan Menggunakan Artificial Neural Network Dan Flask

Employee Recruitment Application Using Artificial Neural Network And Flask

Burhanudin Zuhri^{*1}, Nisa Hanum Harani²

^{1,2}Program Studi Teknik Informatika, Universitas Logistik dan Bisnis Internasional, Bandung
e-mail: ^{*1}burhanudinzuhri25@gmail.com , ²nisa@ulbi.ac.id

Abstrak

Pendekatan strategi berbasis Artificial Intelligence telah banyak digunakan dalam berbagai bidang karena membawa dampak positif yaitu meningkatkan mengoptimalkan proses kegiatan pada bidang tersebut. Dengan mengadopsi pendekatan tersebut, perusahaan dapat meningkatkan pengambilan keputusan secara lebih optimal. Saat ini, sebagian besar perusahaan belum mampu memprediksi kandidat karyawan yang akan bergabung atau tidak dalam proses rekrutmen karyawan. Oleh karena itu, dikembangkanlah sebuah sistem klasifikasi rekrutmen untuk mendukung keputusan pada proses perekrutan secara optimal dengan mengklasifikasikan kandidat karyawan yang memiliki peluang akan bergabung dalam suatu perusahaan menggunakan pendekatan Data Science. Sistem ini dirancang menggunakan bahasa pemrograman Python serta menggunakan model prediksi dengan kemampuan interpretasi Deep Learning. Model ini dilatih dengan menggunakan algoritma Artificial Neural Network (ANN) yang memberikan klasifikasi terbaik berdasarkan variabel yang dimiliki oleh kandidat karyawan. Model dibuat menggunakan dataset sebanyak 11018 data yang telah melalui preprocessing data. Hasil evaluasi menggunakan confusion matrix yaitu akurasi sebesar 78%. Dengan dibuatnya sistem ini, diharapkan perusahaan dapat meningkatkan kualitas sumber daya manusia dengan memprediksi kandidat karyawan terbaik yang perpeluang akan bergabung dalam perusahaan tersebut.

Kata kunci: Sistem Klasifikasi, Rekrutmen Karyawan, Artificial Neural Network, Flask.

Abstract

Artificial Intelligence-based strategy approaches have been widely used in various fields because they have a positive impact on optimizing the activity process in these fields. By adopting this approach, companies can improve decision-making in a more optimized manner. Currently, most companies have not been able to predict employee candidates who will join or not in the employee recruitment process. Therefore, a recruitment classification system is developed to support decisions on the recruitment process optimally by classifying employee candidates who have a chance of joining a company using a Data Science approach. This system is designed using Python programming language and uses a prediction model with Deep Learning interpretation capabilities. The model is trained using an Artificial Neural Network (ANN) algorithm that provides the best classification based on variables owned by employee candidates. The model was created using a dataset of 11018 data that had gone through data preprocessing. The evaluation result using confusion matrix is an accuracy of 78%. With the creation of this

system, it is hoped that companies can improve the quality of human resources by predicting the best employee candidates who are likely to join the company.

Keywords: *Classification System, Employee Recruitment, Artificial Neural Network, Flask.*

1. PENDAHULUAN

Salah satu proses dalam perusahaan yang paling penting adalah proses rekrutmen karyawan yang memiliki dampak paling signifikan di bidang sumber daya manusia (SDM) terhadap pertumbuhan pendapatan keuntungan perusahaan dibandingkan dengan fungsi lainnya [1]. Pada kenyataannya memang tidak dapat dipungkiri, bahwa keputusan rekrutmen karyawan yang buruk dapat menyebabkan karyawan memiliki kinerja yang tidak sesuai dengan harapan perusahaan dan juga dapat menyebabkan peningkatan *turnover* atau proses perputaran karyawan dalam mengganti satu karyawan dengan karyawan lain yang didasarkan pada penilaian persentase kinerja karyawan tersebut [2]. Perputaran karyawan ini berkemungkinan akan berdampak langsung terhadap biaya penggantian karyawan seperti biaya wawancara, perekrutan ulang, pelatihan ulang, dan lembur karyawan baru [3]. Kemudian dampak tidak langsung yang akan dialami oleh perusahaan yaitu akan terjadi penurunan tingkat kualitas pelayanan terhadap klien dan akan terjadi juga penurunan antusiasme kerja pada karyawan itu sendiri [4]. Hal ini sangat penting mengingat bahwa karyawan memainkan peran penting ketika bekerja di perusahaan baik bekerja sama dalam tim maupun individu [5]. Ketika setiap anggota tim memiliki tingkat antusiasme individu yang berbeda terhadap tugas yang dikerjakan, maka hal tersebut akan menghalangi kinerja tim yang baik [6]. Berdasarkan permasalahan tersebut, maka perlu dilakukan identifikasi kandidat karyawan yang paling relevan atau paling tepat serta memiliki dampak yang signifikan terhadap kinerja perusahaan dalam rangka meningkatkan tata kelola yang ada pada perusahaan.

Oleh karena itu, dengan melakukan integrasi pemodelan dalam rangka meningkatkan peluang efisiensi pada perusahaan terutama pada pemodelan rekrutmen karyawan yang paling relevan, maka diperlukan pendekatan yang dapat dijadikan panduan dalam membuat keputusan perekrutan karyawan [7]. Kunci dari pendekatan ini adalah menginterpretasikan prediksi dan memberikan penjelasan mengenai bagaimana cara memperoleh prediksi terlepas dari keakuratan model prediksi yang akan didapatkan, kepercayaan perusahaan pada model sering secara langsung dipengaruhi oleh seberapa tinggi tingkat pemahaman mengenai perilaku model itu sendiri [8]. Dengan pemahaman mengenai perilaku model, maka dapat meningkatkan kepercayaan perusahaan dan perusahaan berpotensi untuk bertindak atau mengambil keputusan berdasarkan rekomendasi yang dihasilkan oleh model tersebut. Selain itu, kualitas data yang terdapat pada perusahaan akan sangat mempengaruhi model dan memungkinkan model untuk memberikan prediksi rekomendasi sesuai dengan tingkat pengetahuan model tersebut [9]. Hal tersebut juga berlaku dalam pembuatan keputusan, seperti dalam hal perekrutan karyawan. Dengan kata lain, ketika mengadopsi strategi berbasis pemodelan AI pada perusahaan, maka akan dapat meningkatkan efisiensi dalam pengambilan keputusan, produktivitas, dan kinerja sistem pada perusahaan tersebut [10].

Pada penelitian sebelumnya mengenai algoritma untuk melakukan pengklasifikasian data terdapat pada penelitian prediksi efisiensi energi menggunakan *Artificial Neural Network* (ANN) dengan akurasi sebesar 99.60% [11]. Penelitian mengenai klasifikasi kaca menggunakan *Artificial Neural Network* (ANN) dengan akurasi sebesar 96.70 [12]. Penelitian mengenai prediksi pergantian karyawan menggunakan *Support Vector Machine* (SVM) dengan akurasi sebesar 93.8% [13]. Penelitian mengenai prediksi pembelian mobil menggunakan *Decision Tree* dengan akurasi sebesar 92.36% [14]. Penelitian mengenai teknik penambangan data berbasis *Artificial Neural Network* (ANN) yang optimal untuk prediksi stres pada karyawan yang bekerja dengan akurasi sebesar 90.9% [15]. Penelitian mengenai klasifikasi kualitas air menggunakan *Artificial*

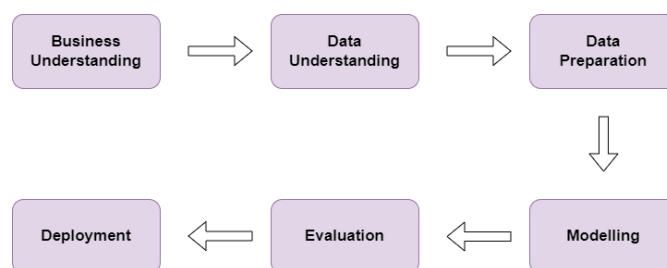
Neural Network (ANN) dengan akurasi sebesar 80.0% [16]. Penelitian mengenai prediksi penyakit diabetes menggunakan *Decision Tree* dengan akurasi sebesar 71% [17].

Tabel 1. Data Preparation Penelitian Sebelumnya dan Penelitian Saat Ini

Penelitian	Model	Data Preparation					
		Data Cleaning	Data Visualization	Encoding	Correlation Analysis	Balancing Data	Standardization
[11]	ANN		✓	✓			✓
[12]	ANN		✓				✓
[13]	SVM			✓	✓	✓	✓
[14]	<i>Decision Tree</i>	✓		✓			✓
[15]	ANN	✓	✓				
[16]	ANN				✓		✓
[17]	<i>Decision Tree</i>	✓	✓		✓		
Model yang diajukan	ANN	✓	✓	✓	✓	✓	✓

Berdasarkan pembahasan yang telah dijabarkan dan referensi dari penelitian sebelumnya yang terdapat pada Tabel 1, menunjukkan bahwa algoritma *Artificial Neural Network* (ANN) merupakan algoritma dengan akurasi tertinggi yang dapat digunakan untuk mengklasifikasikan suatu data tabular. Sehingga, pada penelitian ini akan menggunakan algoritma yang sama yaitu *Artificial Neural Network* (ANN) untuk mengklasifikasikan data rekrutmen kandidat karyawan. Kemudian pada penelitian ini juga terdapat peningkatan (*improvement*) berupa dilakukannya seluruh tahap *data preparation* secara tepat untuk meningkatkan performa model algoritma. Keluaran dari model prediksi perekrutan karyawan ini adalah probabilitas kandidat karyawan akan bergabung dan tidak bergabung pada rekrutmen perusahaan. Klasifikasi rekrutmen kandidat karyawan ini akan dievaluasi menggunakan rumus *confusion matrix* yaitu dengan perhitungan nilai *accuracy*, *precision*, serta *recall*. Manfaat yang diharapkan dari penelitian ini yaitu dapat membantu dalam mengoptimalkan proses rekrutmen karyawan menjadi lebih baik yang dimana hasil dari klasifikasi ini dapat digunakan sebagai alat pendukung dalam pembuatan dan pengambilan keputusan perekrutan kandidat karyawan. Mengingat rekrutmen karyawan memiliki dampak paling signifikan di bidang sumber daya manusia (SDM) terhadap pertumbuhan keuntungan perusahaan dibandingkan dengan fungsi lainnya.

2. METODE PENELITIAN



Gambar 1 Metode CRISP-DM

Penelitian ini didasarkan pada metode CRISP-DM (*Cross-Industry Standard Process for Data Mining*) adalah sebuah metodologi yang digunakan untuk menganalisis data dan membuat model prediktif. CRISP-DM digunakan dalam proyek *data mining* dan analisis data untuk mengidentifikasi tren, pola, dan hubungan dalam data [18]. Metode ini dianggap sebagai prinsip panduan yang paling relevan dan komprehensif untuk melaksanakan proyek analitik [19]. Berdasarkan Gambar 1, proses metode ini terdiri dari beberapa tahap yaitu *business understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation*, dan *deployment*. Adapun penjelasan dari tahapan-tahapan dalam diagram alur metodologi penelitian ini adalah sebagai berikut.

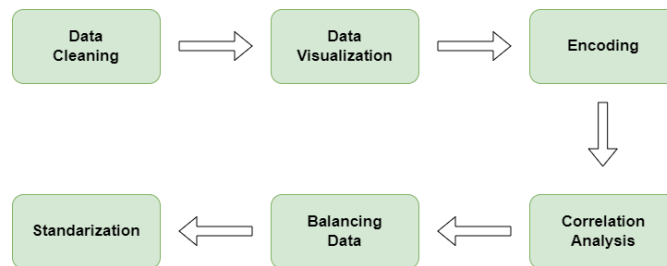
2.1 Business Understanding

Tahap pemahaman bisnis (*business understanding*) akan dilakukan proses memetakan batasan masalah yang ingin diselesaikan agar tujuan penelitian menjadi lebih terarah dan lebih mudah untuk menemukan solusi. Pada tahap ini akan digunakan metode 4W untuk mempermudah proses perumusan masalah [20]. Metode 4W ini berisi siapa saja pihak yang terlibat dalam masalah (*Who*), apa masalah dan faktor pendukung masalah (*What*), kondisi dan tempat masalah diamati (*Where*), serta alasan dan manfaat mengapa masalah tersebut perlu diselesaikan (*Why*). Implementasi metode 4W pada penelitian ini yaitu pihak yang terlibat dalam masalah perekrutan karyawan adalah orang yang ingin mencalonkan diri sebagai kandidat karyawan pada suatu perusahaan dan perusahaan sebagai penyelenggara rekrutmen karyawan. Kemudian masalah yang terjadi yaitu banyak perusahaan yang akan melakukan perekrutan karyawan tetapi perusahaan tersebut belum memiliki sistem untuk memprediksi atau mengklasifikasikan apakah kandidat karyawan akan bergabung atau tidak. Lalu tempat terjadinya masalah tersebut yang diamati yaitu terjadi di kebanyakan perusahaan tempat terjadinya kegiatan perekrutan karyawan. Selanjutnya alasan mengapa masalah tersebut perlu diselesaikan adalah agar perusahaan dapat mengoptimalkan sumber daya manusia mereka dengan menerapkan sistem klasifikasi perekrutan kandidat karyawan berdasarkan beberapa kriteria yang ditetapkan oleh perusahaan agar proses tersebut dapat berjalan dengan tepat dan sesuai dengan tujuan perusahaan.

2.2 Data Understanding

Tahap pemahaman data (*data understanding*) akan dilakukan proses akuisisi data yang dibutuhkan dalam penelitian untuk diolah dan dianalisis sesuai dengan permasalahan agar dapat menghasilkan solusi yang terbaik. Adapun data yang digunakan pada penelitian ini adalah data sekunder yang pengumpulannya tidak dilakukan secara langsung oleh peneliti. Dataset yang digunakan untuk membuat model klasifikasi ini yaitu dataset rekrutmen karyawan yang tersedia di *platform* Kaggle dengan judul HR-Dataset dan berformat csv. Dataset rekrutmen karyawan tersebut memiliki 8995 baris dan 18 kolom. Pada tahap pemahaman data juga akan dilakukan perhitungan statistik deskriptif untuk setiap variabel termasuk menghitung *mean*, *median*, standar deviasi, dan kuartil untuk variabel numerik, serta menghitung frekuensi untuk variabel kategorikal.

2. 3 Data Preparation



Gambar 2 Data Preparation

Berdasarkan Gambar 2, dapat dijelaskan bahwa tahapan persiapan data (*data preparation*) dimulai dengan memahami apa yang terkandung dalam data dan bagaimana data tersebut dikumpulkan termasuk dengan cara mengidentifikasi jumlah variabel hingga melakukan *preprocessing* terhadap data [19]. Pada tahap pemahaman data juga dilakukan proses penyelidikan dataset menggunakan *Exploratory Data Analysis* (EDA) untuk mengidentifikasi pola, memahami karakteristik data, menemukan anomali, dan menguji hipotesis.

2. 3.1 Data Cleaning

Pada tahap *data cleaning* terdapat penanganan *outlier* menggunakan jangkauan interkuartil (IQR) yang merupakan salah satu teknik transformasi statistik terbaik yang dapat membatasi nilai ekstrim dalam data statistik untuk mengurangi efek outlier [21].

2. 3.2 Data Visualization

Pada tahap *data visualization* data akan dilakukan penyajian data dalam bentuk grafis untuk mengeksplorasi distribusi dan hubungan antara variabel [22]. Setelah dilakukan visualisasi data, terdapat jumlah data kandidat karyawan yang berlabel bergabung (*joined*) lebih banyak dibandingkan dengan data kandidat karyawan yang berlabel tidak bergabung (*not joined*).

2. 3.3 Encoding

Pada tahap *data encoding* akan diubah data kategorikal ke dalam bentuk numerik unik yang agar dapat diolah oleh sistem. *Encoding* perlu dilakukan untuk memudahkan dalam melakukan pengolahan data dan membantu algoritma untuk menangani data yang tidak dapat diterima sebagai *input* numerik.

2. 3.4 Correlation Analysis

Pada tahap analisis korelasi akan diukur seberapa kuat hubungan antara dua variabel atau lebih [23]. Pada penelitian ini setiap kolom memiliki korelasi yang cukup lemah terhadap label/target. Sehingga pada pengolahan data ini tidak dilakukan reduksi terhadap kolom pada *dataset*.

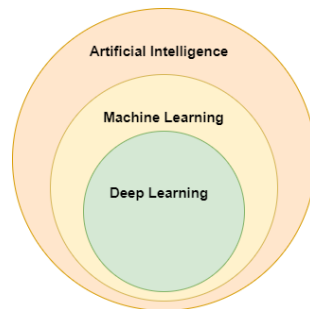
2. 3.5 Balancing Data

Pada tahap *balancing data* akan diubah distribusi data sehingga mendekati proporsi yang seimbang untuk menghindari bias saat pembuatan model. Pada *dataset* rekrutmen karyawan terdapat data dengan label *joined* dan label *not joined* dengan jumlah data yang tidak seimbang (*imbalance data*). Untuk mengatasi hal tersebut, maka akan dilakukan penarikan sampel ulang (*resampling*) menggunakan teknik *over sampling* agar kedua label tersebut menjadi seimbang, sehingga komputasi menjadi lebih efisien dan kinerja data akan meningkat [24].

2. 3.6 Standardization

Pada tahap standarisasi data akan dipastikan bahwa semua data memiliki rentang yang sama. Standarisasi ini harus diimplementasikan ketika terdapat data yang tidak terstruktur dan mengandung rentang nilai yang sangat berbeda [25]. Pada penelitian akan digunakan *MinMaxScaler* untuk menyamakan skala dari data yang ada yaitu data akan berada pada kisaran 0 hingga 1.

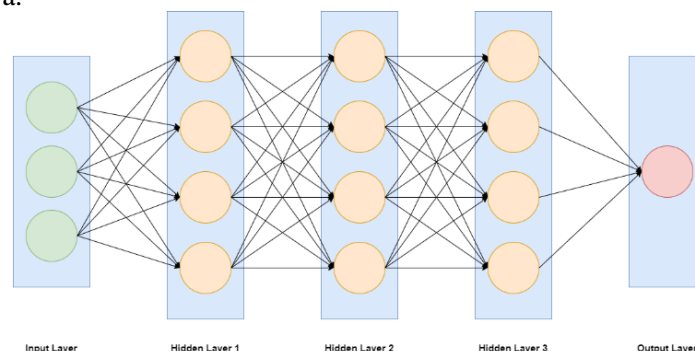
2. 4 Modelling



Gambar 3 Bagan Artificial Intelligence

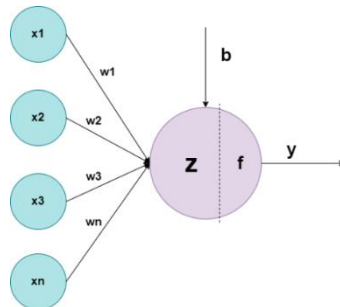
Tahap modelling atau pemodelan akan dilakukan proses penerapan algoritma dalam bahasa pemrograman menggunakan metode *Deep Learning* yang digunakan untuk menemukan pola dalam data sebagai dasar pengetahuan sistem untuk melakukan klasifikasi atau membuat keputusan. Berdasarkan Gambar 3 dapat diketahui bahwa *deep learning* merupakan cabang *machine learning* yang melakukan pembelajaran secara hirarki yaitu melalui banyak *layer* [26]. Proses *training* pada *deep learning* dapat terjadi tanpa pengawasan manusia, yang dimana komputer secara otomatis menemukan pola dan kesamaan dalam data, sehingga *deep learning* sering diaplikasikan pada teknik klasifikasi, clustering, segmentasi ataupun *recognition* [27]. Jadi dengan menggunakan metode ini, maka tidak ada keluaran khusus yang diharapkan dan hal ini sering digunakan sebagai alat eksplorasi untuk mendeteksi variabel dalam data, mengurangi dimensi, atau mengelompokkan kelompok data serupa [28].

Modelling pada rekrutmen karyawan akan menggunakan algoritma *Artificial Neural Networks* (ANN) yang merupakan sebuah model yang dibuat berdasarkan struktur jaringan saraf pada makhluk hidup dan terdiri dari unit-unit yang disebut "*perceptron*" yang terhubung satu sama lain, di mana masing-masing *perceptron* menerima input dari *perceptron* lainnya, mengolahnya dengan menggunakan sebuah fungsi aktivasi, dan mengirimkan *output* ke *perceptron* lainnya.



Gambar 4 Arsitektur Artificial Neural Networks (ANN)

Berdasarkan Gambar 4 dapat disimpulkan bahwa *Artificial Neural Network* (ANN) merupakan salah satu algoritma dalam *deep learning* dengan *multilayer* yang kompleks [29]. *Layer-layer* tersebut dinamakan *multilayer perceptron* (MLP) [30]. Bagian dari *multilayer* yaitu *input layer*, *hidden layer*, dan *output layer* dengan tiap layer memiliki banyak *perceptron* yang memiliki *input* yang sama, namun bobotnya berbeda [31]. Dalam ANN, setiap *perceptron* menerima input dari perceptron lain dan menghasilkan output dengan menggunakan fungsi aktivasi. Fungsi aktivasi ini dapat berupa *sigmoid*, *tangens hiperbolik*, atau ReLU (*Rectified Linear Unit*). Output dari setiap perceptron kemudian dikirim ke perceptron lain melalui jalur yang disebut koneksi dengan bobot yang berbeda-beda. Fungsi aktivasi adalah fungsi yang digunakan untuk mengolah inputan data. Sebagai penggambaran, fungsi aktivasi pada *single perceptron* dapat dilihat pada Gambar 5.



Gambar 5 Single Perceptron

Rumus linear function:

$$z = \sum_{i=1}^n x_i w_i + b$$

Rumus output:

$$y = f(z)$$

Tabel 2. Single Perceptron

No.	Simbol	Keterangan
1.	x	input
2.	w	bobot
3.	b	bias
4.	z	linear function
5.	f	activation function
6.	y	output

Dapat dilihat pada Gambar 5 dan Tabel 2 diatas, terdapat fungsi transfer yang berfungsi untuk mengubah sejumlah *input* (x_1, x_2, \dots, x_n), bobot (w_1, w_2, \dots, w_n), dan bias (b) menjadi nilai *linear function* (z). Kemudian *activation function* (f) akan memproses nilai *linear function* (z) untuk dibandingkan dengan *threshold* yang telah ditentukan. Lalu *activation function* akan mengaktifasi nilai menjadi *output* (y) menggunakan *step function* dengan ketentuan jika $z > 0$, maka *output perceptron* adalah 1, sedangkan jika $z \leq 0$, maka *output perceptron* adalah 0 [32].

2. 5 Evaluation

Tahap *evaluation* atau evaluasi model akan dilakukan proses pengkajian dengan melakukan pemilihan model terbaik yang akan digunakan untuk membuat sistem sesuai dengan tujuan penelitian. Pada tahap ini akan digunakan metode *Confusion Matrix* untuk mendapatkan

beberapa komponen seperti akurasi (*accuracy*), presisi (*precision*), dan *recall* untuk mengetahui kekuatan dan kelemahan model dalam memprediksi kelas yang benar [33].

Tabel 3. Confusion Matrix

	Positive	Negative
True	True Positive (TP)	True Negative (TN)
False	False Positive (FP)	False Negative (FN)

Berdasarkan Tabel 3 dapat dijelaskan bahwa *confusion matrix* memuat informasi tentang *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN). Dari *confusion matrix* tersebut dapat dihitung beberapa metrik yang digunakan untuk mengevaluasi performa model seperti berikut.

Accuracy: mengukur seberapa sering model memprediksi dengan benar. Adapaun rumus *accuracy* adalah sebagai berikut.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

Precision: mengukur seberapa sering model memprediksi positif yang benar. Adapaun rumus *precision* adalah sebagai berikut.

$$Precision = TP / (TP + FP)$$

Recall: mengukur seberapa sering model dapat menemukan semua kelas positif. Adapaun rumus *recall* adalah sebagai berikut.

$$Recall = TP / (TP + FN)$$

2. 5 Deployment

Pada tahap *deployment* atau penyebaran aplikasi akan dilakukan proses penerapan model terhadap sebuah sistem berdasarkan hasil analisa yang telah dilakukan. Tujuan *deployment* ini yaitu dapat menyebarkan ke pengguna dan memudahkan pengguna dalam melakukan klasifikasi data atau menggunakan aplikasi. Pada tahap ini, *deployment* dilakukan menggunakan Ngrok yang dapat mempublikasikan aplikasi *website* atau *server* lokal ke internet dengan mudah.

3. HASIL DAN PEMBAHASAN

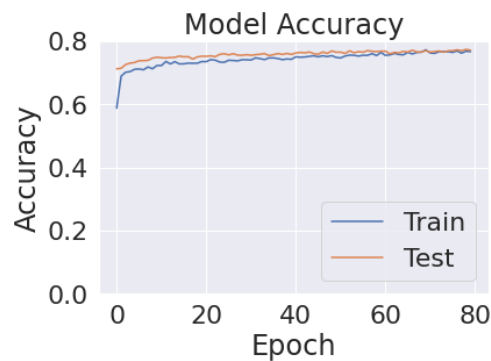
Penelitian ini menggunakan 15 variabel dari 11018 data yang telah diubah dari yang sebelumnya adalah 8995 data setelah melakukan *over sampling* dikarenakan terjadi *imbalance data*. Kemudian dilakukan pembagian data dengan persentase sebesar 80% data *training* dan 20% data *testing*, yaitu 8814 data digunakan sebagai data *training* dan 2204 data digunakan sebagai data *testing*. Uji coba dilakukan dengan melakukan *hyperparameter tuning* pada model dan hasil uji coba pada penelitian ini dapat dilihat pada Tabel 4.

Tabel 4. Tuning Hyperparameter

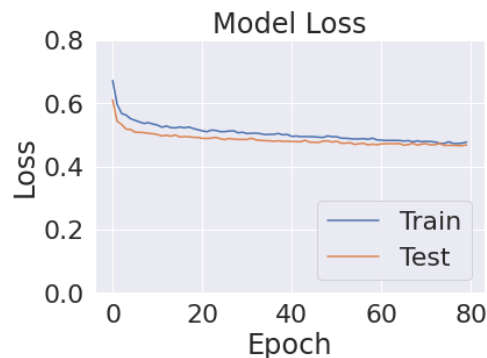
No.	Parameter				Accuracy	Precision	Recall
	Dense	Optimizer	Loss	Epoch			
1.	9-3-3-3	Adam	binary_cross entropy	60	0.74	0.72	0.77

2.	9-3-3-3	Adam	binary_cross entropy	40	0.73	0.71	0.78
3.	9-3-3-3 + Dropout	Adam	binary_cross entropy	60	0.69	0.64	0.86
4.	9-3-3-3	Adam	categorical_c rossentropy	60	0.5	0.0	0.0
5.	128-64- 32-16	SGD	binary_cross entropy	60	0.53	0.54	0.38
6.	128-64- 32-16	Adam	binary_cross entropy	50	0.74	0.72	0.78
7.	128-64- 32-16	Adam	binary_cross entropy	60	0.75	0.73	0.78
8.	128-64- 32-16 + Dropout	Adam	binary_cross entropy	80	0.78	0.75	0.83

Berdasarkan hasil *tuning hyperparameter* pada Tabel 3 didapatkan bahwa tingkat akurasi tertinggi yaitu sebesar 78% dengan jumlah *hidden layer* sebanyak “4” buah, jumlah *perceptron* yaitu “128-64-32-16”, *optimizer* yaitu “Adam”, *loss* yaitu “*binary crossentropy*”, dan *epoch* sebanyak “100”. Selanjutnya dilakukan training berdasarkan struktur model terbaik. Dari proses *training* tersebut didapatkan hasil *epoch* telah berhenti di kisaran angka 0.78 dengan *loss* sebesar 0.45. Hasil dari proses training tersebut dapat dilihat pada Gambar 6 dan Gambar 7.

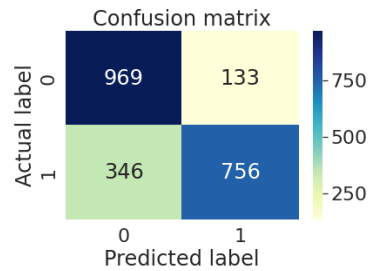


Gambar 6 Grafik Training Model Accuracy



Gambar 7 Grafik Training Model Loss

Selanjutnya adalah tahap evaluasi berfungsi untuk mengukur kinerja klasifikasi yang telah dilakukan. Evaluasi dilakukan menggunakan *confusion matrix*. Berikut merupakan *confusion matrix* dari hasil klasifikasi data menggunakan algoritma ANN dengan persentase 80% data *training* dan 20% data *testing*. Sehingga didapatkan hasil yaitu algoritma ANN memprediksi data dengan benar sebanyak 969 data kelas positif dan 756 data kelas negatif. Hasil tersebut menunjukkan bahwa algoritma ANN cukup baik untuk mengklasifikasikan data kelas positif dan data kelas negatif seperti pada Gambar 8.



Gambar 8 Confusion Matrix

Sehingga dapat dilakukan perhitungan sebagai berikut.

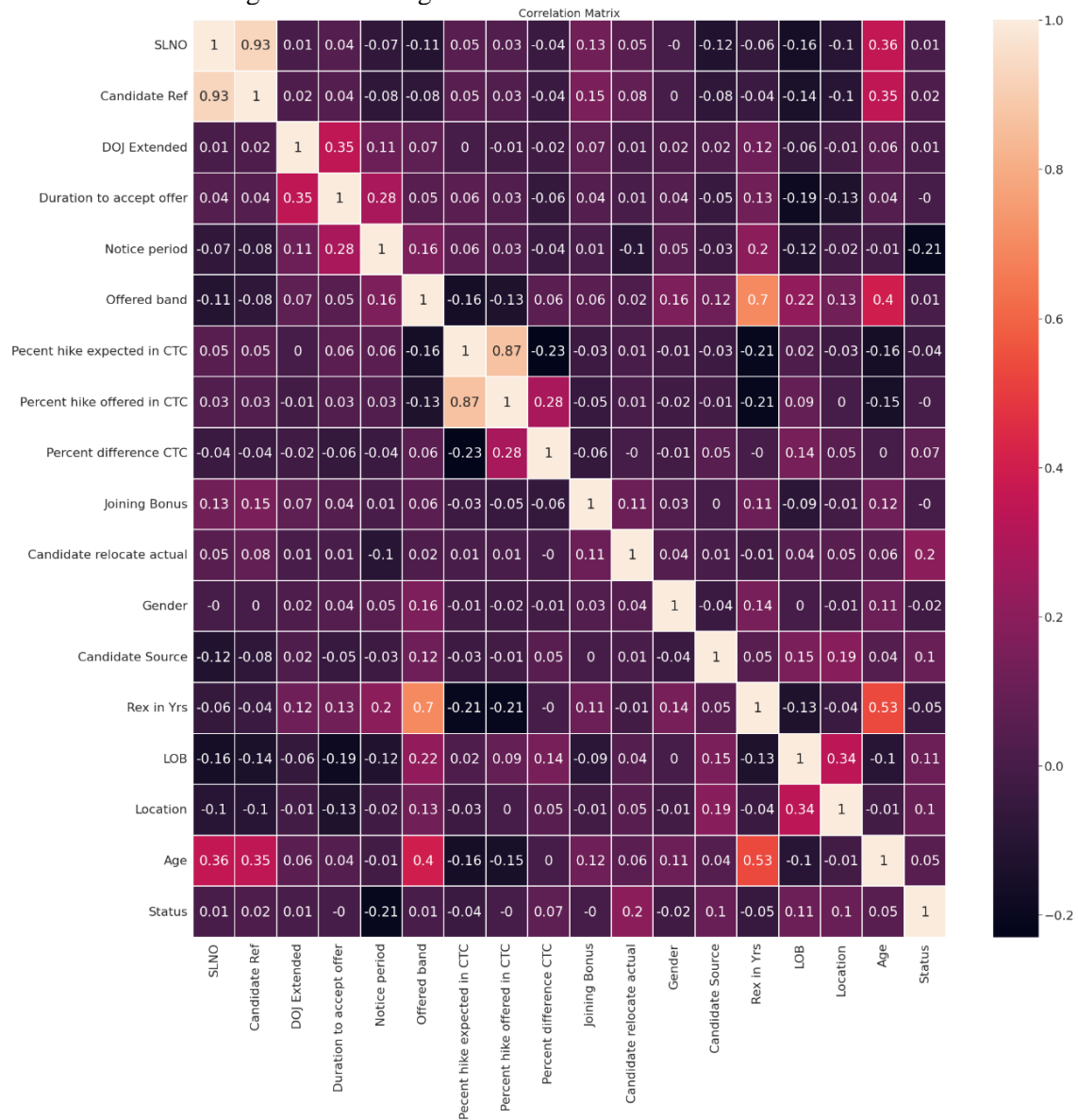
$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) = (969 + 756) / (969 + 756 + 133 + 346) = 0.78$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) = 969 / (969 + 133) = 0.87$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) = 969 / (969 + 346) = 0.73$$

Kemudian dilakukan visualisasi untuk mengetahui variabel/atribut yang dominan dalam menentukan output. Korelasi dari setiap variabel dapat dilihat pada Gambar 9. Variabel yang memiliki korelasi terkuat dengan variabel target/label yaitu dengan memiliki nilai yang mendekati nilai 1 dan sebaliknya. Disini dapat terlihat bahwa variabel *Candidate relocate actual*, *LOB*, *Candidate source*, dan *Location* merupakan variabel dengan korelasi terkuat dengan variabel target/label *Status*. Sedangkan variabel *Notice periode*, *Percent hike expected in CTC*, *Duration*

to accept offer, Percent hike offered in CTC, dan Joining bonus merupakan variabel dengan korelasi terlemah dengan variabel target/label Status.



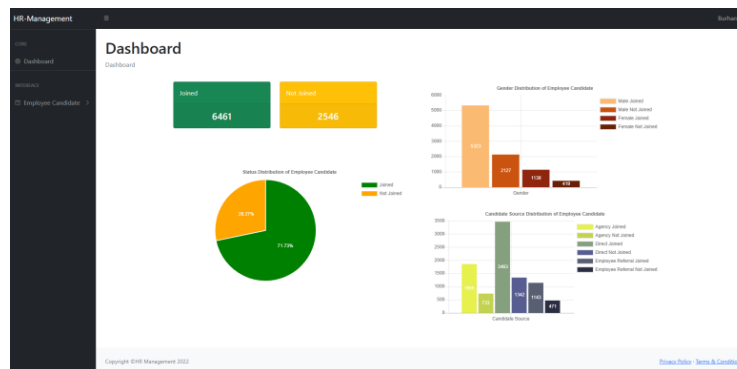
Gambar 9 Visualisasi Korelasi antar Variabel

Setelah mengetahui tingkat akurasi *training* dan *testing* serta korelasi antar variabel selanjutnya yaitu dilakukan *deployment* model. *Deployment* ini berfungsi untuk memudahkan pengguna untuk menggunakan aplikasi rekrutmen karyawan. *Deployment* pada penelitian ini menggunakan Ngrok dapat mempublikasikan aplikasi *website* atau *server* lokal ke internet dengan mudah. *Website* managemen ini dibuat untuk melakukan klasifikasi terhadap data baru yang ditambahkan oleh pengguna. Selain itu, *website* ini juga dapat digunakan untuk mengelola data rekrutment karyawan seperti menampilkan, menambahkan, mengubah, dan menghapus data rekrutment karyawan. Aplikasi ini dikembangkan menggunakan bahasa pemrograman Python dan *framework* Flask. Kemudian untuk tampilannya menggunakan HTML dan CSS. Sebelum pengguna menggunakan aplikasi ini, maka pengguna harus melakukan *register* dan *login* pada aplikasi. Pada Gambar 10 merupakan tampilan halaman kelola data rekrutmen karyawan.

Duration to Accept the Offer	Notice Period	Offered Band	Percent Mile Expected to Cost to Company	Percent Mile Offered to Cost to Company	Percent Difference Cost to Company	Joining Bonus	Candidate Reluctance Actual	Gender of Employee Candidate	Employee Candidate Source	Release Years of Experience	Line of Business	Company Location	Age of Employee Candidate	Status	Probability
10	35	E3	60.0	50.0	20.0	No	No	Male	Employee Referral	7	BP3	Ahmedabad	28	Joined	85.98
20	40	E2	60.0	80.0	20.0	No	No	Male	Employee Referral	8	BP3	Pune	35	Joined	90.03
12	15	E3	60.0	80.0	20.0	No	Yes	Female	Employee Referral	6	BP3	Mumbai	33	Joined	91.02
30	70	E1	60.0	100.0	40.0	No	No	Female	Agency	9	E43	Gurgaon	33	Joined	83.83
44	75	E2	45.0	60.0	10.0	No	No	Male	Direct	8	BP3A	Hyd	34	Joined	85.15
15	40	E1	60.0	100.0	40.0	No	No	Female	Advisor	33	C3AP	Chennai	33	Not	95.53

Gambar 10 Halaman Kelola Data Rekrutmen Karyawan

Pada Gambar 11 merupakan tampilan *dashboard* yang berfungsi untuk menampilkan visualisasi data yang ada pada *database* rekrutmen karyawan dalam bentuk grafik.



Gambar 11 Dashboard Visualisasi Data Rekrutmen Karyawan

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, maka dapat disimpulkan bahwa Aplikasi Klasifikasi Rekrutmen Karyawan Menggunakan *Artificial Neural Network* dan Flask telah berhasil dibangun serta mampu mengklasifikasikan beberapa data rekrutmen karyawan yang akan bergabung dan tidak bergabung pada perusahaan dengan tingkat akurasi sebesar yaitu 78%. Penelitian ini telah dilakukan dengan beberapa peningkatan untuk mengoptimalkan *output* yaitu dengan melakukan semua tahap *data preparation*. Dengan demikian, diharapkan aplikasi dan model ini cukup memungkinkan apabila digunakan sebagai pendukung pengambilan keputusan pada proses perekrutan karyawan pada suatu perusahaan dengan menyesuaikan beberapa variabel sesuai dengan keutuhan perusahaan tersebut. Kemudian diharapkan untuk melakukan penelitian lebih lanjut untuk meningkatkan akurasi pada model.

DAFTAR PUSTAKA

- [1] D. Pessach, G. Singer, D. Avrahami, H. Chalutz Ben-Gal, E. Shmueli, and I. Ben-Gal, "Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming," *Decis Support Syst*, vol. 134, Jul. 2020, doi: 10.1016/j.dss.2020.113290.

-
- [2] Y. Zhao, M. K. Hryniewicki, F. Cheng, B. Fu, and X. Zhu, "Employee turnover prediction with machine learning: A reliable approach," in *Advances in Intelligent Systems and Computing*, Springer Verlag, 2018, pp. 737–758. doi: 10.1007/978-3-030-01057-7_56.
- [3] P. Lyons and R. Bandura, "Employee turnover: features and perspectives," *Development and Learning in Organizations*, vol. 34, no. 1. Emerald Group Holdings Ltd., pp. 1–4, Jan. 20, 2020. doi: 10.1108/DLO-02-2019-0048.
- [4] R. S. Yadav, S. S. Dash, S. Chakraborty, and M. Kumar, "Perceived CSR and Corporate Reputation: The Mediating Role of Employee Trust," *Vikalpa*, vol. 43, no. 3. SAGE Publications Ltd, pp. 139–151, Sep. 01, 2018. doi: 10.1177/0256090918794823.
- [5] Y. Q. Zhu, D. G. Gardner, and H. G. Chen, "Relationships Between Work Team Climate, Individual Motivation, and Creativity," *J Manage*, vol. 44, no. 5, pp. 2094–2115, May 2018, doi: 10.1177/0149206316638161.
- [6] G. Canbaloglu, J. Treur, and P. H. M. P. Roelofsma, "Computational modeling of organisational learning by self-modeling networks," *Cogn Syst Res*, vol. 73, pp. 51–64, Jun. 2022, doi: 10.1016/j.cogsys.2021.12.003.
- [7] W. van Zetten, G. J. Ramackers, and H. H. Hoos, "Increasing trust and fairness in machine learning applications within the mortgage industry," *Machine Learning with Applications*, vol. 10, p. 100406, Dec. 2022, doi: 10.1016/j.mlwa.2022.100406.
- [8] S. Singaravel, J. Suykens, and P. Geyer, "Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction," *Advanced Engineering Informatics*, vol. 38, pp. 81–90, Oct. 2018, doi: 10.1016/j.aei.2018.06.004.
- [9] D. Pavlenko, L. Barykin, S. Nemeshaev, and E. Bezverhny, "Individual approach to knowledge control in learning management system," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 259–263. doi: 10.1016/j.procs.2020.02.162.
- [10] A. Jamwal, R. Agrawal, and M. Sharma, "Deep learning for manufacturing sustainability: Models, applications in Industry 4.0 and implications," *International Journal of Information Management Data Insights*, vol. 2, no. 2, Nov. 2022, doi: 10.1016/j.ijime.2022.100107.
- [11] A. J. Khalil, A. M. Barhoom, B. S. Abu-Nasser, M. M. Musleh, and S. S. Abu-Naser, "Energy Efficiency Prediction using Artificial Neural Network," 2019. [Online]. Available: www.ijeais.org/ijapr
- [12] M. Jamal El-Khatib, B. S. Abu-Nasser, and S. S. Abu-Naser, "Glass Classification Using Artificial Neural Network," 2019. [Online]. Available: www.ijeais.org/ijapr
- [13] Y. Zhao, M. K. Hryniewicki, F. Cheng, B. Fu, and X. Zhu, "Employee turnover prediction with machine learning: A reliable approach," in *Advances in Intelligent Systems and Computing*, Springer Verlag, 2018, pp. 737–758. doi: 10.1007/978-3-030-01057-7_56.
- [14] H. H. Patel and P. Prajapati, "Study and Analysis of Decision Tree Based Classification Algorithms," *International Journal of Computer Sciences and Engineering*, vol. 6, no. 10, pp. 74–78, Oct. 2018, doi: 10.26438/ijcse/v6i10.7478.
- [15] S. Anitha and M. Vanitha, "Optimal artificial neural network-based data mining technique for stress prediction in working employees," *Soft comput*, vol. 25, no. 17, pp. 11523–11534, Sep. 2021, doi: 10.1007/s00500-021-06058-9.
- [16] K. Sulaiman, L. Hakim Ismail, M. Adib Mohammad Razi, M. Shalahuddin Adnan, and R. Ghazali, "Water Quality Classification Using an Artificial Neural Network (ANN)," in *IOP Conference Series: Materials Science and Engineering*, Institute of Physics Publishing, Sep. 2019. doi: 10.1088/1757-899X/601/1/012005.
- [17] T. Dudkina, I. Meniailov, K. Bazilevych, S. Krivtsov, and A. Tkachenko, "Classification and Prediction of Diabetes Disease using Decision Tree Method," *ymposium on Information Technologies & Applied Sciences*, 2021.
-

- [18] C. Schröer, F. Kruse, and J. M. Gómez, “A systematic literature review on applying CRISP-DM process model,” in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 526–534. doi: 10.1016/j.procs.2021.01.199.
- [19] S. Jaggia, A. Kelly, K. Lertwachara, and L. Chen, “Applying the CRISP-DM Framework for Teaching Business Analytics,” 2020.
- [20] F. Azimah, K. Rizky, and N. Wardani, “SISTEM PENDETEKSI GEJALA AWAL COVID-19 DENGAN PENGGUNAAN METODE AI PROJECT CYCLE,” 2022, [Online]. Available: www.kaggle.com.
- [21] J. Yang, S. Rahardja, and P. Fränti, “Outlier detection: How to threshold outlier scores?,” in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Dec. 2019. doi: 10.1145/3371425.3371427.
- [22] M. Waskom, “seaborn: statistical data visualization,” *J Open Source Softw*, vol. 6, no. 60, p. 3021, Apr. 2021, doi: 10.21105/joss.03021.
- [23] D. (Christina) Geng, J. Innes, W. Wu, and G. Wang, “Impacts of COVID-19 pandemic on urban park visitation: a global analysis,” *J For Res (Harbin)*, vol. 32, no. 2, pp. 553–567, Apr. 2021, doi: 10.1007/s11676-020-01249-w.
- [24] S. Maldonado, J. López, and C. Vairetti, “An alternative SMOTE oversampling strategy for high-dimensional datasets,” *Applied Soft Computing Journal*, vol. 76, pp. 380–389, Mar. 2019, doi: 10.1016/j.asoc.2018.12.024.
- [25] B. Deepa and K. Ramesh, “Epileptic seizure detection using deep learning through min max scaler normalization,” *Int J Health Sci (Qassim)*, pp. 10981–10996, May 2022, doi: 10.53730/ijhs.v6ns1.7801.
- [26] H. Liu and B. Lang, “Machine learning and deep learning methods for intrusion detection systems: A survey,” *Applied Sciences (Switzerland)*, vol. 9, no. 20. MDPI AG, Oct. 01, 2019. doi: 10.3390/app9204396.
- [27] S. Christin, É. Hervet, and N. Lecomte, “Applications for deep learning in ecology,” *Methods in Ecology and Evolution*, vol. 10, no. 10. British Ecological Society, pp. 1632–1644, Oct. 01, 2019. doi: 10.1111/2041-210X.13256.
- [28] S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar, “A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning,” *Archives of Computational Methods in Engineering*, vol. 27, no. 4, pp. 1071–1092, Sep. 2020, doi: 10.1007/s11831-019-09344-w.
- [29] O. Isaac Abiodun, A. Jantan, A. Esther Omolara, K. Victoria Dada, N. AbdElatif Mohamed, and H. Arshad, “State-of-the-art in artificial neural network applications: A survey,” *Heliyon*, vol. 4, p. 938, 2018, doi: 10.1016/j.heliyon.2018.
- [30] M. Al-Emran Khaled Shaalan Editors, “Studies in Systems, Decision and Control 335,” 2021. [Online]. Available: <http://www.springer.com/series/13304>
- [31] C. Cao *et al.*, “Deep Learning and Its Applications in Biomedicine,” *Genomics, Proteomics and Bioinformatics*, vol. 16, no. 1. Beijing Genomics Institute, pp. 17–32, Feb. 01, 2018. doi: 10.1016/j.gpb.2017.07.003.
- [32] J. Yang *et al.*, “Quantization Networks,” 2019.
- [33] S. Ruuska, W. Hämäläinen, S. Kajava, M. Mughal, P. Matilainen, and J. Mononen, “Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle,” *Behavioural Processes*, vol. 148, pp. 56–62, Mar. 2018, doi: 10.1016/j.beproc.2018.01.004.