

Penerapan Kombinasi Algoritma SVM-KNN dalam seleksi User SAKTI berdasarkan Hasil Kinerja Pegawai pada Kementerian XYZ

Syaiful Ramadhan^{1*)}, Kusrini²⁾, Kusnawi³⁾

¹⁾²⁾³⁾ Program Studi Magister Teknik Informatika, Universitas Amikom Yogyakarta

^{*)}Correspondence Author: phunk.syaiful@students.amikom.ac.id, Jakarta, Indonesia

DOI: <https://doi.org/10.37012/jtik.v9i2.1672>

Abstrak

Kementerian XYZ merupakan Kementerian dengan jumlah pegawai lebih dari 5.000 pegawai. Pada saat dibentuk tidak dilakukan pemetaan pegawai, hal ini mengakibatkan surplus jumlah pegawai, tidak terkecuali pada Biro Barang Milik Negara (BMN). Bagi sebuah organisasi, SDM yang berlimpah merupakan hal yang baik, namun perlu dilakukan penyeleksian pegawai agar dapat meningkatkan produktivitas sehingga keberhasilan organisasi dapat tercapai. Disamping itu, perbaikan sistem Administrasi Keuangan pemerintahan merupakan suatu keharusan yang diimbangi dengan pengembangan aplikasi terintegrasi Kementerian Keuangan yaitu Sistem Aplikasi Keuangan Tingkat Instansi (SAKTI). Dalam melakukan pengelolaan aset pada Biro BMN, setiap pegawai memiliki role user level kewenangan SAKTI dengan lingkup yang berbeda-beda. Penelitian ini bertujuan melakukan seleksi klasifikasi user berdasarkan hasil penilaian kinerja dengan penerapan metode Kombinasi algoritma SVM dan KNN menggunakan bahasa pemrograman Python. Berdasarkan pengujian dengan sampel data sebesar ± 313 data pegawai dan 18 variabel pegawai dengan atribut target berupa kelayakan yaitu dipertahankan maupun dipertimbangkan, diperoleh hasil akurasi sebesar 94% pada Kernel SVM RBF; nilai $K=5$; metrik Euclidean; Dapat disimpulkan seleksi user aplikasi SAKTI menggunakan kombinasi algoritma SVM dan KNN dapat memberikan prediksi guna meningkatkan efektivitas dan efisiensi organisasi dalam penempatan pegawai yang sesuai dengan kompetensi pada Biro BMN Kementerian XYZ. Penelitian selanjutnya diharapkan dapat membandingkan kombinasi algoritma SVM dan KNN dengan metrik serta parameter yang lebih banyak.

Kata kunci: Seleksi pegawai, algoritma SVM, algoritma KNN, SAKTI

Abstract

XYZ Ministry is a Ministry with more than 5,000 employees. At the time it was formed, no employee mapping was carried out, this resulted in a surplus in the number of employees, including the State Property Bureau (BMN). For an organization, an abundance of human resources is a good thing, but it is necessary to select employees in order to increase productivity so that organizational success can be achieved. In addition, improving the government's Financial Administration system is a must, which is balanced with the development of an integrated application for the Ministry of Finance, namely the Agency Level Financial Application System (SAKTI). In managing assets at the BMN Bureau, each employee has a role user level of SAKTI authority with different scopes. This study aims to select a user classification based on the results of a performance assessment by applying the SVM and KNN combination method using the Python programming language. Based on testing with a data sample of ± 313 employee data and 18 employee variables with target attributes in the form of feasibility, namely being maintained or considered, the results obtained an accuracy of 94% in the SVM RBF Kernel; K value=5; Euclidean metric; It can be concluded that the user selection of the SAKTI application using a combination of the SVM and KNN algorithms can provide predictions to increase organizational effectiveness and efficiency in placing employees according to competence in the XYZ Ministry's BMN Bureau. Future research is expected to be able to compare the combination of SVM and KNN algorithms with more metrics and parameters.

Keywords: Employee selection, SVM algorithm, KNN algorithm, SAKTI

PENDAHULUAN

Sumber Daya Manusia merupakan sebuah aset yang sangat penting didalam sebuah organisasi. Sehingga dibutuhkan sebuah pemilihan pegawai untuk dapat meningkatkan produktivitas agar tujuan organisasi tercapai (Sunandar & Satar, 2020). Pada dasarnya tujuan dari seleksi pegawai adalah mendapatkan sumber daya yang tepat sehingga seseorang dapat bekerja secara optimal dalam hal keterampilan untuk melakukan suatu tugas tertentu (Suwinardi, 2017). Karena dengan skill yang tepat, karyawan akan lebih produktif dalam bekerja. Agar kegiatan organisasi menjadi efektif, perlu dilakukan pemilihan karyawan. Seleksi pegawai dilakukan karena jumlah pegawai Kementerian XYZ lebih dari 5.000 pegawai yang tersebar di unit kerja. Selama integrasi, tidak dilakukan survei kepegawaian, sehingga setiap pegawai dapat dengan leluasa memilih unit kerja yang sesuai, sehingga terjadi kelebihan pegawai di beberapa unit kerja, salah satunya Biro Barang Milik Negara.

Selain itu, prasyarat dalam mencapai prinsip pengelolaan keuangan dan aset Negara yang akuntabel, efisien, efektif dan bertanggung jawab adalah peningkatan administrasi publik melalui teknologi aplikasi terintegrasi untuk seluruh unit kerja di pemerintah pusat Indonesia, yang disebut Sistem Aplikasi Keuangan Tingkat Badan (SAKTI) yang dikembangkan Kementerian Keuangan (Amriani, 2019). Dalam melakukan kegiatan pengelolaan aset Biro BMN, setiap pegawai memiliki otorisasi aplikasi SAKTI dengan user kewenangan mulai dari admin, operator, validator, approver dan komitmen (Nugroho, 2020). Pemanfaatan aplikasi SAKTI pada Biro BMN mencakup sebagian kecil proses pengelolaan keuangan negara yaitu pada tahapan implementasi (modul komitmen, modul aset tetap dan modul persediaan) (Anwar, 2022).

Agar tujuan organisasi yang telah ditetapkan tercapai dalam pengelolaan aset yang efektif, efisien dan bertanggung jawab, diperlukan manajemen SDM yang baik. Kegagalan dalam manajemen sumber daya manusia salah satunya kesalahan penunjukan pegawai sebagai user aplikasi memiliki dampak negatif pada organisasi, karena mengacu pada kompetensi pegawai itu sendiri dalam menentukan efisiensi yang dihasilkan dari pekerjaannya.

Untuk melakukan prediksi berdasarkan data-data masa lalu yang digunakan sebagai data latih dapat menggunakan Teknik data mining dan Machine Learning (Kotu & Deshpande, 2014). Salah satu metode yang dapat digunakan dalam melakukan prediksi

seleksi karyawan adalah klasifikasi, seperti algoritma Naïve Bayes, KNN, Random Forest, Decision Stamp, Decision Tree, Rule Induction, Linear Regresi, Linear Discriminant Analysis, Neural Network, dan SVM dengan menghasilkan akurasi tertinggi pada algoritma SVM yaitu 65% namun masih jauh dari akurasi yang diinginkan (Saifudin, 2018). Penelitian lainnya yang dilakukan Soni Lanka (2021) dalam menilai kinerja klasifikasi perilaku pengemudi dengan teknik SVM dan KNN memberikan hasil masing-masing 90.1% dan 89.4%.

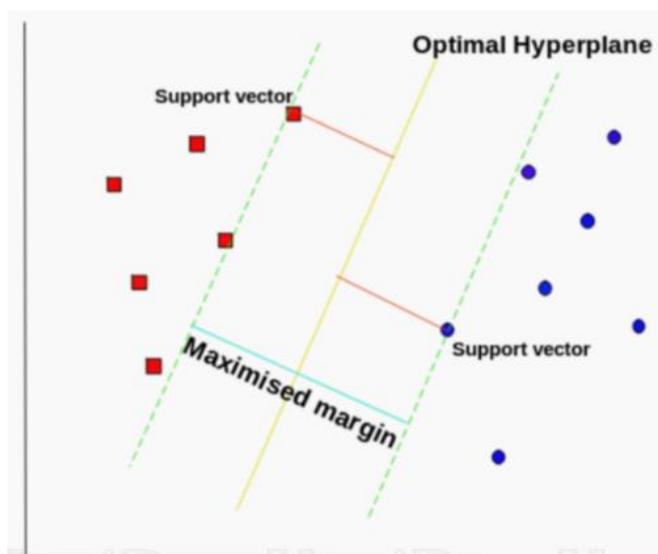
Penelitian-penelitian lain yang dilakukan sebelumnya, salah satunya yang dilakukan Michael, A (2022) dengan teknik kombinasi pre-trained model pada lapisan ekstraksi fitur dan SVM dengan tuning hyperparameter GridsearchCV mampu menghasilkan performa terbaik sebesar 96% pada nilai $C=0.001$ dan kernel Linear. Penelitian serupa yang dilakukan Astrianda (2020) dengan menggunakan ekstraksi fitur pada Support Vector Machine pada klasifikasi dua kelas dengan 54 data uji dan 20 data latih, menghasilkan akurasi sebesar 100%. Pada algoritma Neighbor Weighted KNN dengan tahapan pengujian cross validation pada parameter $K=45$, fitur=9 dan nilai E (eksponen) = 6 menghasilkan rerata nilai akurasi 60%, rerata recall 67%, rata-rata presisi 72%, rata-rata F-measure 66% (Azizah., dkk, 2022). Penelitian lain yang dilakukan Kusriani, dkk (2019) dalam prediksi klasifikasi rekomendasi minat dengan algoritma KNN dengan teknik Case Base Reasoning dengan membagi data latih dan data uji sebesar 70-30 pada metrik Euclidean menghasilkan akurasi 95,89%. Selanjutnya penelitian yang dilakukan Ginting, dkk (2018) berupa kombinasi algoritma Support Vector Machines (SVM) dan K-Nearest Neighbors (KNN) dalam menentukan kelayakan nasabah Bank dengan jumlah data latih 2000 data dan data uji sebanyak 103 data, sistem memberikan hasil performa sebesar 88,35% pada tuning parameter yang optimal, nilai $C=0.1$, $\gamma=2$, sistem mengidentifikasi 1998 support vector, dilanjutkan dengan parameter nilai $K=16$.

METODE

Metode pengumpulan data merupakan bagian tidak terpisahkan dari metode penelitian yang didapatkan dari hasil observasi, arsip data dan bersifat obyektif. Penelitian dilakukan terhadap variabel data pegawai Biro BMN yang dikumpulkan diantaranya variabel input nama, usia, jenis kelamin, tingkat pendidikan, pangkat, lokasi kerja, domisili,

jumlah kehadiran WFO, jumlah kehadiran WFH, role user aplikasi SAKTI, modul SAKTI, hasil penilaian Pegawai, dan kelayakan dengan total 18 variabel menjadi dataset dalam format excel.

Pengolahan data dalam penelitian ini dilakukan menggunakan teknik klasifikasi yaitu kombinasi algoritma SVM dan KNN dalam melakukan seleksi klasifikasi user SAKTI menggunakan python. Klasifikasi adalah kegiatan mengelompokkan data berdasarkan ciri-ciri yang dimiliki terhadap semua dataset menjadi suatu data baru (Nur Widiastui, 2023). Algoritma SVM adalah algoritma klasifikasi yang mengelaskan suatu data baru dalam ruang N-dimensi dengan memisahkan berdasarkan hyperplane (L. Wang, 2005). Tujuan algoritma SVM adalah menemukan optimal hyperplane (fungsi pemisah) yang memisahkan dua buah kelas seperti pada gambar 1, dengan cara memaksimalkan margin antara dua class. Margin adalah jarak antara hyperplane dengan input vectors (data) terdekat (Desiani, 2022). Input vectors yang bersentuhan dengan boundary disebut support vector. SVM dikembangkan berdasarkan Statistical Learning Theory.



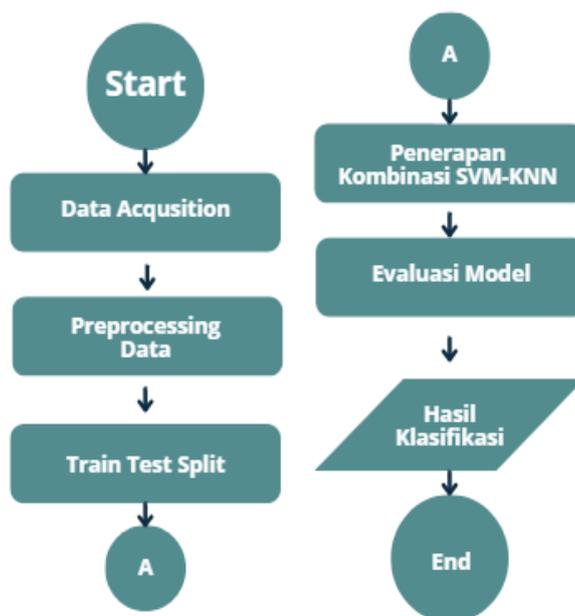
Gambar 1. Hyperplane

Algoritma KNN adalah suatu metode yang bertujuan membuat klasifikasi suatu data yang dilihat dari lokasi terdekat k dengan suatu objek yang didasarkan pada data training (Noviansyah, 2018). Data dibagi menjadi data training dan data testing, setelah itu dilakukan penghitungan jarak dari data testing ke data training. Cara kerja algoritma KNN (Wibowo, 2021) yaitu:

1. Tentukan jumlah tetangga terdekat yaitu parameter K
2. Hitung Query instance atau jarak data baru dari titik objek dataset yang tersedia.
3. Urutkan dan hitung kembali jarak dari terdekat sampai terjauh ke dalam kelas yang memiliki nilai terkecil.
4. Kumpulkan kategori klasifikasi nearest neighbor mayoritas sehingga nilai query instance dapat Diprediksi

Perlu kecermatan dalam menentukan nilai K karena didasarkan pada besarnya jumlah data dan dimensi yang terbentuk dari data tersebut. Nilai K akan semakin kecil jika data semakin banyak. Namun semakin besar dimensi data, angka K juga sebaiknya semakin tinggi.

Agar penelitian berjalan dengan baik maka dibuat langkah-langkah penelitian yang dapat dilihat pada gambar 2, yang menggambarkan alur proses penelitian:



Gambar 2. Alur Penelitian

1. Data Acquisition

Data acquisition adalah proses pertama yang dilakukan peneliti untuk mengumpulkan data-data yang relevan yang akan digunakan untuk melakukan proses analisis. Dalam penelitian ini data yang digunakan terdiri dari berbagai data yaitu SK penempatan pegawai pada Biro BMN, rekap kehadiran selama 1 tahun, data kepegawaian, role user aplikasi SAKTI berdasarkan SK, dan rekap penilaian kinerja.

2. Data Preprocessing

Tahap preprocessing merupakan proses penyesuaian data mentah menjadi format data dengan kriteria data yang dibutuhkan untuk penelitian. Tahapan ini terdiri dari :

- a. Proses Cleaning, yaitu pembersihan data atau variabel yang tidak diperlukan, tidak relevan atau mengandung noise seperti pada gambar 2 penggolongan Tipe Data.
- b. Seleksi Data, yaitu proses mengambil data yang relevan dari data asli dengan cara dilakukan penggabungan, pengolahan data. Proses ini dapat mengurangi representasi dataset dari data asli namun integritas data tetap terjaga agar menghasilkan hasil analisis yang sama.
- c. Transformasi Data, digunakan untuk mengubah atau menggabungkan data dalam bentuk yang sesuai dalam proses data mining sehingga diperoleh pola yang lebih mudah dipahami. Beberapa teknik untuk transformasi data adalah normalisasi, pemilihan atribut, dan discretization. Berdasarkan dari dataset awal, selanjutnya dilakukan transformasi format tipe data menjadi numerik maupun katagorikal agar memudahkan dalam proses pengolahan nantinya.
- d. Train-Test Split
- e. Selanjutnya dataset dibagi menjadi data latih dan data uji dengan pembagian data secara acak menggunakan `train_test_split` yang disediakan pada modul Sklearn dari library Python. Pada penelitian ini pembagian data training dan data latih adalah 70:30. Seperti pada gambar 3.

```
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.3, random_state=0)
```

Gambar 3. Split Data

3. Data Mining

Proses Mining merupakan tahap dalam mengaplikasikan metode-metode tertentu untuk menemukan informasi penting dan tersembunyi dari data. Proses ini menjadi tahap utama dalam data mining. Metode yang digunakan dalam proses ini adalah dengan teknik kombinasi algoritma SVM dan KNN. Tahapan pertama algoritma SVM diimplementasikan pada dataset untuk mencari support vector yang berada dekat dengan hyperplane terbaik yang memisahkan 2 buah class. Hasil Support vector ini selanjutnya digunakan sebagai dataset untuk di proses pada tahapan berikutnya menggunakan

algoritma KNN. Dalam melakukan proses menggunakan algoritma KNN, peneliti akan menentukan nilai K (tetangga terdekat) terlebih dahulu, untuk selanjutnya melakukan proses klasifikasi data. Dalam menentukan nilai K akan didasarkan pada jumlah dataset yang akan diuji. Setelah pengujian menggunakan data training, kemudian dengan menggunakan data testing (uji) untuk menghasilkan prediksi dari seleksi klasifikasi user SAKTI. Parameter-parameter yang digunakan ditunjukkan pada tabel 1.

Tabel 1. Set Parameter SVM-KNN

Algoritma	Parameter Percobaan 1	Parameter Percobaan 2
SVM	RBF	RBF
KNN	K = 5; metrik = euclidean	K= 5; metrik = manhattan

4. Evaluasi Model

Evaluasi model adalah proses untuk mengidentifikasi pola-pola yang telah ditemukan dari dataset yang bersifat relevan dan menarik untuk mengoptimalkan model yang telah dilatih dan diuji agar mendapatkan akurasi. Evaluasi yang dipilih menggunakan metode confusion matrix, karena metode ini melakukan perhitungan dengan menghasilkan 4 keluaran (Huang, 2016), yaitu: recall, precision, accuracy dan error rate. Cara kerja confusion matrix yaitu menilai kinerja model klasifikasi sesuai jumlah objek yang diprediksi dengan benar dan salah (Dewi, 2019). Untuk menghitung nilai akurasi, recall, precision dan error rate dapat dilakukan dengan rumus perhitungan seperti pada rumus 1 (a) – (d).

$$Precision = \frac{TP}{TP+FP} \tag{a}$$

$$Recall = \frac{TP}{TP+FN} \tag{b}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{c}$$

$$F1 - Score = \frac{2 \times recall \times precision}{recall + precision} \tag{d}$$

Rumus 1 (a)-(d). Perhitungan Evaluasi Confusion Matrix

Keterangan:

True positif (TP) merupakan nilai positif, baik nilai prediksi maupun nilai aktualnya.

False positif (FP) merupakan nilai yang prediksinya positif dan nilai aktualnya negatif

False negative (FN) merupakan nilai yang prediksinya negatif dan nilai aktualnya positif

True negative (TN) merupakan nilai negatif, baik nilai prediksinya maupun nilai aktualnya.

HASIL DAN PEMBAHASAN

Sebelum melanjutkan ke tahap pengujian terlebih dahulu dilakukan preprocessing data diantaranya penyiapan dataset dan deklarasi library python yang digunakan. Dalam hal penyiapan dataset hasil pengolahan data asli dengan cara penggabungan dari berbagai data variabel pegawai yang telah dikumpulkan kedalam satu dataset didapatkan 18 kolom dan 313 baris. Dataset berformat .xls ini kemudian dikonversi menjadi .csv agar memudahkan dalam proses pengujian data. Agar dataset yang telah dibuat dapat diolah pada python maka perlu dilakukan deklarasi library python seperti pada gambar 4 baru kemudian hasil dataset acquisition data dapat ditampilkan pada python seperti pada gambar 5.

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.svm import SVC
from sklearn.metrics import confusion_matrix, accuracy_score
from sklearn.neighbors import KNeighborsClassifier
```

Gambar 4. Deklarasi Lybrary Python

NO	NAMA	JENISKELAMIN	PENDIDIKAN	GOLONGAN	DOMISILI	AREA	WFO	WFH	USIA	NILAI	RATINGKERJA	PREDIKATKINERJA	KELAYAKAN	
0	1	ID001	P	Magister	1	Jabodetabek	Jabodetabek	105	111	45	114	DI ATAS EKSPEKTASI	2	2
1	2	ID002	L	Magister	9	Jabodetabek	Jabodetabek	44	154	48	105	SESUAI EKSPEKTASI	1	1
2	3	ID003	L	Diploma	9	Luar Jabodetabek	Jabodetabek	158	88	49	112	SESUAI EKSPEKTASI	1	2
3	4	ID004	L	Sarjana	9	Jabodetabek	Jabodetabek	83	159	57	108	SESUAI EKSPEKTASI	1	2
4	5	ID005	L	SMA	3	Jabodetabek	Jabodetabek	83	159	50	105	SESUAI EKSPEKTASI	1	2
...
308	309	ID309	L	SMA	3	Jabodetabek	Jabodetabek	74	168	44	105	SESUAI EKSPEKTASI	1	1
309	310	ID310	L	SMA	5	Jabodetabek	Jabodetabek	56	188	50	105	SESUAI EKSPEKTASI	1	1
310	311	ID311	L	SMA	7	Jabodetabek	Jabodetabek	35	207	52	104	SESUAI EKSPEKTASI	1	1
311	312	ID312	L	Magister	11	Luar Jabodetabek	Luar Jabodetabek	71	171	58	104	SESUAI EKSPEKTASI	1	2
312	313	ID313	L	SMA	2	Luar Jabodetabek	Luar Jabodetabek	165	77	43	99	BIASA	1	1

313 rows x 18 columns

Gambar 5. Dataset Pegawai

Metode kombinasi klasifikasi SVM dan KNN dalam melakukan pengujian dataset menggunakan tipe data numerik dan katagorikal. Tahapan selanjutnya dari dataset yang telah dibuat dilakukan normalisasi data dengan cara validasi penggolongan tipe data dengan tujuan untuk meminimalisir ketidaktepatan hasil pengujian, baik dengan cara melakukan kroscek missing value (data yang tidak bernilai), pembersihan data yang tidak relevan, penyeleksian data, maupun transformasi data seperti pada gambar 6.

```

#   Column      Non-Null Count  Dtype
---  -
0   NO           313 non-null    int64
1   NAMA          313 non-null    object
2   JENISKELAMIN 313 non-null    object
3   PENDIDIKAN    313 non-null    object
4   GOLONGAN      313 non-null    int64
5   PENUGASAN     313 non-null    object
6   MODULSAKTI   313 non-null    int64
7   USER         313 non-null    int64
8   HOMEBASE      313 non-null    object
9   DOMISILI      313 non-null    object
10  AREA          313 non-null    object
11  WFO           313 non-null    int64
12  WFH           313 non-null    int64
13  USIA          313 non-null    int64
14  NILAI         313 non-null    int64
15  RATINGKERJA  313 non-null    object
16  PREDIKATKINERJA 313 non-null    int64
17  KELAYAKAN    313 non-null    int64
    
```

Gambar 6(a). Data semula

```

#   Column      Non-Null Count  Dtype
---  -
0   NO           313 non-null    int64
1   JENISKELAMIN 313 non-null    int32
2   PENDIDIKAN    313 non-null    int32
3   GOLONGAN      313 non-null    int64
4   PENUGASAN     313 non-null    int32
5   MODULSAKTI   313 non-null    int64
6   USER         313 non-null    int64
7   HOMEBASE      313 non-null    int32
8   DOMISILI      313 non-null    int32
9   AREA          313 non-null    int32
10  WFO           313 non-null    int64
11  WFH           313 non-null    int64
12  USIA          313 non-null    int64
13  NILAI         313 non-null    int64
14  RATINGKERJA  313 non-null    int32
15  PREDIKATKINERJA 313 non-null    int64
16  KELAYAKAN    313 non-null    int64
dtypes: int32(7), int64(10)
    
```

Gambar 6(b). Data setelah normalisasi

NO	JENISKELAMIN	PENDIDIKAN	GOLONGAN	PENUGASAN	MODULSAKTI	USER	HOMEBASE	DOMISILI	AREA	WFO	WFH	USIA	NILAI	RATINGKEF
0	1	1	2	1	0	4	0	3	1	0	135	111	48	114
1	2	0	2	9	29	3	4	2	1	0	44	164	48	106
2	3	0	0	9	18	3	1	2	2	0	158	88	49	112
3	4	0	4	9	31	3	4	2	1	0	83	159	57	108
4	5	0	3	3	18	3	2	2	1	0	83	159	50	108
...
308	309	0	3	3	13	1	1	25	1	0	74	188	44	105
309	310	0	3	5	13	2	1	28	1	0	56	188	50	105
310	311	0	3	7	13	2	1	14	1	0	35	207	52	104
311	312	0	2	11	13	2	1	28	2	1	71	171	58	104
312	313	0	3	2	13	1	1	28	2	1	165	77	43	99

313 rows x 17 columns

Gambar 6(c). Tabel Setelah dilakukan Transformasi Data

Pada gambar 6 (a) dan (b) terlihat perubahan jenis tipe data sehingga mengakibatkan perubahan pada tampilan tabel gambar 6(c). Disamping itu terdapat pengurangan jumlah kolom dari 18 kolom yang terdiri dari ‘no’, ‘nama’, ‘jenis kelamin’, ‘pendidikan’, ‘golongan’, ‘penugasan’, ‘modul sakti’, ‘user’, ‘homebase’, ‘domisili’, ‘area’, ‘wfo’, ‘wfh’, ‘usia’, ‘nilai’, ‘rating kerja’, ‘predikat kinerja’, dan ‘kelayakan’ menjadi 17 kolom yaitu ‘no’, ‘jenis kelamin’, ‘pendidikan’, ‘golongan’, ‘penugasan’, ‘modul sakti’, ‘user’, ‘homebase’,

‘domisili’, ‘area’, ‘wfo’, ‘wfh’, ‘usia’, ‘nilai’, ‘rating kerja’, ‘predikat kinerja’, dan ‘kelayakan’, hal ini terjadi karena kolom ‘nama’ berupa tipe data **object** dan tidak akan mempengaruhi hasil analisis data nantinya.

Pada gambar 6(c) secara konsep proses terjadinya pemberian nilai value katagorikal yang dilakukan secara komputersasi dilakukan seperti halnya pemberian nilai value dengan penggolongan data yang dilakukan secara manual pada tabel 2 berupa penggolongan data untuk kolom ‘Golongan’.

Tabel 2. Katagorikal Data secara manual

No	Golongan	Penggolongan Data
1	Juru Tingkat I – I/d	1
2	Pengatur Muda – II/a	2
3	Pengatur Muda Tingkat I – II/b	3
4	Pengatur – II/c	4
5	Pengatur Tingkat I – II/d	5
6	Penata Muda – III/a	6
7	Penata Muda Tingkat I – III/b	7
8	Penata – III/c	8
9	Penata Tingkat I – III/d	9
10	Pembina – IV/a	10
11	Pembina Tingkat I – IV/b	11

No	MODUL SAKTI	Penggolongan Data
1	Persediaan	1
2	Aset Tetap	2
3	Komitmen	3
4	Administrator	4

No	Role User	Penggolongan Data
1	Admin	0
2	Operator	1
3	Validator	3
4	Approver	4

Tahapan selanjutnya setelah preprocessing selesai adalah membagi data set menjadi data training dan data uji. Dengan pembagian 70% sebagai data training dan 30 % sebagai data uji artinya dengan jumlah data sebesar 313 data, komposisinya menjadi terbagi sebanyak 219 data (baris) sebagai data latih dan 94 data uji. Pada tahap ini juga dilakukan pemisahan variable X dan variabel Y. Variabel X adalah variabel independent meliputi ‘no’, ‘jenis kelamin’, ‘pendidikan’, ‘golongan’, ‘penugasan’, ‘modul sakti’, ‘user’, ‘homebase’, ‘domisili’, ‘area’, ‘wfo’, ‘wfh’, ‘usia’, ‘nilai’, ‘rating kerja’, dan ‘predikat kinerja’. Variabel Y adalah variable dependent meliputi ‘kelayakan’. Pemodelan algoritma yang lebih awal dibuat adalah pemodelan algoritma SVM. Perlu melakukan setting parameter menggunakan kernel ‘RBF’ (Radial Basis Function) seperti pada gambar 7 agar pemodelan dapat dibuat.

```
classifier_svm=SVC(kernel="rbf", random_state=0)
classifier_svm.fit(x_train, y_train)

SVC(random_state=0)

array([[ 2,  0,  2, ..., 106,  2,  0],
       [ 3,  0,  0, ..., 112,  2,  0],
       [ 5,  0,  3, ..., 106,  2,  0],
       ...,
       [213,  1,  4, ..., 115,  1,  1],
       [217,  0,  4, ..., 103,  2,  0],
       [219,  1,  3, ..., 103,  2,  0]], dtype=int64)
```

Gambar 7. Pemodelan SVM kernel RBF

Hasil pembuatan model train dan test pada algoritma klasifikasi SVM memperoleh hasil akurasi yang baik yaitu sebesar 91% yang dapat dilihat pada gambar 8.

```
print(accuracy_score(y_test,y_pred))

0.9148936170212766
```

Gambar 8. Hasil Akurasi Algoritma SVM

Pemodelan algoritma SVM merupakan pemodelan yang digunakan hanya sampai menghasilkan jumlah support vector. Nantinya support vector ini digunakan sebagai dataset baru pada pemodelan berikutnya yaitu KNN. Support Vector yang dihasilkan dengan algoritma SVM seperti pada gambar 9 sejumlah 110 data dengan 16 variabel X dan 1 Variabel sebagai target yaitu Y. Tujuan dari penggunaan support vector adalah untuk mereduksi data yang tidak akan menghasilkan peningkatan akurasi. Karena tujuan dari kombinasi SVM dan KNN adalah untuk melakukan peningkatan akurasi.

(110, 16)

Gambar 9. Jumlah Variabel X yang dihasilkan dari Support Vector

Dalam melakukan pemodelan berikutnya dengan algoritma KNN perlu dilakukan inisiasi variabel x dan y baru dari data yang dihasilkan algoritma SVM dan juga melakukan deklarasi library python yang diperlukan ditunjukkan pada gambar 10. Dalam algoritma KNN, perlu adanya penentuan nilai K dan metrik yang digunakan dalam menghitung jarak dari tetangga terdekat. Pada algoritma KNN, dilakukan dua kali percobaan untuk membandingkan hasil akurasi, pertama setting value yang ditetapkan nilai K=5, dengan metrik Euclidean, dan yang kedua nilai K=5 dengan metrik Manhattan.

KNN

```
x = x[classifier_svm.support_]      m.support_]
y = y[classifier_svm.support_]      m.support_]

from sklearn.neighbors import KNeighborsClassifier
classifier = KNeighborsClassifier(n_neighbors =5, metric = "euclidean") c = "euclidean")

classifier = KNeighborsClassifier(n_neighbors =5, metric = "manhattan")
```

Gambar 10. Inisiasi data baru dan setting value

Pengujian model dilakukan menggunakan parameter terbaik yang dihasilkan dari data latih model. Performa yang dihasilkan dari hasil pengujian dengan menggunakan metode Confusion Matrix dari dua kali pemodelan dalam melakukan klasifikasi terhadap dataset uji ditunjukkan pada gambar 11 dan gambar 12.

	precision	recall	f1-score	support		precision	recall	f1-score	support
1	1.00	0.80	0.89	10	1	0.89	0.80	0.84	10
2	0.92	1.00	0.96	23	2	0.92	0.96	0.94	23
accuracy			0.94	33	accuracy			0.91	33
macro avg	0.96	0.90	0.92	33	macro avg	0.90	0.88	0.89	33
weighted avg	0.94	0.94	0.94	33	weighted avg	0.91	0.91	0.91	33

Gambar 11. Performa Model Percobaan 1 **Gambar 12.** Performa Model Percobaan 2

Dari perhitungan performa confusion matrix dengan memperhatikan nilai akurasi, precision, recall dan F1-Score seperti pada tabel 3 pada pemodelan Algoritma KNN dengan metrik Euclidean dan Manhattan, akurasi terbaik dengan pemodelan kombinasi algoritma SVM dan KNN dengan menggunakan metrik Eculidean sebesar 94%.

Tabel 3. Akurasi pada Pemodelan

Percobaan	Parameter	Akurasi SVM	Akurasi SVM+KNN
1	RBF, K=5, Euclidean	91%	94%
2	RBF, K=5, Manhattan	91%	91%

KESIMPULAN DAN REKOMENDASI

Berdasarkan hasil penelitian yang telah dilakukan dari hasil klasifikasi dengan jumlah dataset 313 pegawai dengan 17 variabel pegawai didapatkan kombinasi algoritma SVM dan KNN dengan metrik Euclidean memiliki kinerja yang lebih baik dan berhasil melakukan peningkatan prediksi seleksi klasifikasi user SAKTI berdasarkan penilaian kinerja dengan nilai akurasi sebesar 94%. Sedangkan akurasi yang dihasilkan algoritma SVM dan SVM+KNN dengan metrik Manhattan menghasilkan akurasi yang sama yaitu 91%. Dalam mendapatkan akurasi yang tinggi, terdapat beberapa proses yang dilakukan diantaranya menghapus variabel yang tidak relevan, normalisasi data dan melakukan transformasi tipe data agar dapat mempermudah proses pemodelan. Pembagian data latih dan uji 70:30 serta penentuan nilai $K=5$ juga bagian dari proses menghasilkan akurasi terbaik.

Penelitian selanjutnya diharapkan dapat membandingkan kombinasi algoritma SVM dan KNN dengan metrik serta parameter yang lebih banyak untuk mengetahui kinerja yang memiliki dampak terbaik untuk menghasilkan akurasi. Selain parameter K dan jenis metrik pada KNN, perbandingan penggunaan parameter kernel pada algoritma SVM seperti nilai C , γ maupun jenis kernel juga direkomendasikan untuk dilakukan pada penelitian-penelitian berikutnya sehingga dapat memberikan kontribusi dalam menyelesaikan masalah-masalah dalam melakukan seleksi klasifikasi pegawai.

REFERENSI

- Amriani, T. N., & Iskandar, A. (2019). Analisis Kesuksesan Implementasi Sistem Aplikasi Keuangan Tingkat Instansi (SAKTI) pada Satuan Kerja di Lingkungan Badan Pendidikan dan Pelatihan Keuangan (BPPK). *Kajian Ekonomi dan Keuangan*, 3(1), 54-74.
- Anwar, A. I., & Hadi, M. (2022). Implementasi aplikasi SAKTI dan SPAN dalam penyusunan laporan keuangan. *Jurnal Informatika*, 1(2), 32-55.
- Astrianda, N. (2020). Klasifikasi Kematangan Buah Tomat Dengan Variasi Model Warna Menggunakan Support Vector Machine. *VOCATECH: Vocational Education and Technology Journal*, 1(2), 110-117.

- Azizah, R. A., Bachtiar, F., & Adinugroho, S. (2022). Klasifikasi Kinerja Akademik Siswa Menggunakan Neighbor Weighted K-Nearest Neighbor dengan Seleksi Fitur Information Gain. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 9(4), 655-664
- Desiani, A., Akbar, M., Irmeilyana, & Amran, A. (2022). Implementasi Algoritma Naïve Bayes Dan Support Vector Machine (SVM) Pada Klasifikasi Penyakit Kardiovaskular. *Jurnal Teknik Elektro Dan Komputasi (ELKOM)*, 4(2), 207–214
- Ginting, S. L. B., & Permana, A. A. (2018). Hybrid Classifier System: Support Vector Machines Dikombinasikan dengan K-Nearest Neighbors untuk Menentukan Kelayakan Nasabah Bank dalam Pengajuan Kredit. *Komputika: Jurnal Sistem Komputer*, 7(1), 9-16.
- Huang, F., Shen, J., Guo, Q., & Shi, Y. (2016). eRFSVM: a hybrid classifier to predict enhancers-integrating random forests with support vector machines. *Hereditas*, 153, 1-11.
- Karri, S. L., De Silva, L. C., Lai, D. T. C., & Yong, S. Y. (2021). Classification and prediction of driving behaviour at a traffic intersection using SVM and KNN. *SN computer science*, 2(3), 1-11.
- Kotu, V., & Deshpande, B. *Predictive analytics and data mining: concepts and practice with rapidminer*. Waltham, USA: Morgan Kaufmann, 2014
- Mardi, Y. (2017) 'Data Mining:Klasifikasi Menggunakan Algoritma c4.5', *Jurnal Edik Informatika*, 2(2), pp 213-219
- Michael, A. (2022). Komparasi Kombinasi Pre-trained Model dengan SVM pada Klasifikasi Kematangan Kopi Berbasis Citra. *Journal Dynamic sainT E-ISSN*, 2722, 5364.
- Noviansyah, M. R., Rismawan, T., & Midyanti, D. M. (2018). Penerapan data mining menggunakan metode k-nearest neighbor untuk klasifikasi indeks cuaca kebakaran berdasarkan data AWS (automatic weather station)(studi kasus: kabupaten Kubu Raya). *Coding Jurnal Komputer dan Aplikasi*, 6(2).
- Nugroho, H. P., & Lestyowati, J. (2020). Analisis Tingkat Kepuasan dan Kepentingan Pengguna Aplikasi SAKTI dengan PIECES Framework. *Indonesian Treasury Review: Jurnal Perbendaharaan, Keuangan Negara dan Kebijakan Publik*, 5(2), 93-104.

-
- Prasetyo, A., Kusriani, K., & Arief, M. R. (2019). Penerapan Algoritma K Nearest Neighbor Untuk Rekomendasi Minat Konsentrasi Di Program Studi Teknik Informatika Universitas PGRI Yogyakarta. *Informasi Interaktif*, 4(1), 1-6
- Sunandar, H. S., & Satar, M. (2020). Tinjauan tentang proses recruitment tenaga kerja di kantor pusat PT. Y Bandung. *Jurnal Industri Elektro dan Penerbangan*, 6(1)
- Suwinardi, S. (2017). Profesionalisme dalam bekerja. *Orbith: Majalah Ilmiah Pengembangan Rekayasa dan Sosial*, 13(2)
- Wibowo, S., & Nababan, A. A. (2021). Penerapan Algoritma K-Nearest Neighbour Dalam Rekrutmen Pegawai Tidak Tetap Di BPJS Kesehatan Cabang Gunungsitoli. *Jurnal Sistem Informasi dan Teknologi Jaringan (SISFOTEKJAR)*, 2(2), 6-11.
- Widiastuti, N., Hermawan, A., & Avianto, D. (2023). Komparasi Algoritma Klasifikasi Datamining Untuk Prediksi Minat Pencari Kerja. *Jurnal Teknoinfo*, 17(1), 219-227.