

Implementasi Algoritma Supervised Learning untuk Pengambilan Keputusan

Jauhan Ahmad

Teknik Informatika / Sains dan Teknologi

Universitas Darussalam Gontor

Email: Jauhanahmad80@student.cs.unida.gontor.acd.id

Abstract—Pengambilan keputusan berbasis data merupakan kebutuhan penting dalam berbagai bidang, seperti perbankan, kesehatan, dan bisnis. Penelitian ini mengimplementasikan tiga algoritma supervised learning, yaitu regresi logistik, Support Vector Machine (SVM), dan Decision Tree untuk mendukung pengambilan keputusan. Studi kasus yang digunakan adalah klasifikasi kelayakan kredit dengan atribut pendapatan, umur, lama bekerja, serta riwayat pinjaman. Hasil pengujian menunjukkan bahwa ketiga algoritma mampu melakukan klasifikasi dengan tingkat akurasi yang berbeda. Decision Tree memberikan keunggulan dalam interpretabilitas aturan, sedangkan SVM dan regresi logistik lebih unggul dalam generalisasi data.

Index Terms—Supervised Learning, Regresi, SVM, Decision Tree, Pengambilan Keputusan

I. PENDAHULUAN

Dalam era digital, pengambilan keputusan tidak lagi hanya mengandalkan intuisi, tetapi membutuhkan pendekatan berbasis data. Supervised learning merupakan salah satu pendekatan dalam machine learning yang menggunakan data berlabel untuk membangun model prediksi. Algoritma yang umum digunakan adalah regresi logistik, Support Vector Machine (SVM), dan Decision Tree. Penelitian ini bertujuan untuk membandingkan kinerja ketiga algoritma tersebut dalam pengambilan keputusan, khususnya pada kasus klasifikasi kelayakan kredit.

II. TINJAUAN PUSTAKA

Supervised learning telah banyak diterapkan dalam sistem pendukung keputusan.

- **Regresi Logistik** digunakan untuk memodelkan probabilitas dari kelas biner.
- **Support Vector Machine (SVM)** bekerja dengan mencari hyperplane optimal untuk memisahkan kelas.
- **Decision Tree** membentuk aturan berbasis pohon keputusan yang mudah diinterpretasikan.

Menurut [1], pemilihan algoritma bergantung pada kebutuhan: interpretasi (Decision Tree), akurasi klasifikasi (SVM), atau probabilitas prediksi (Regresi Logistik).

III. METODOLOGI

A. Dataset

Data simulasi terdiri dari empat fitur: pendapatan, umur, lama bekerja, dan riwayat pinjaman. Label target adalah kelayakan kredit (layak/tidak layak).

B. Pra-Pemrosesan

Data dibagi menjadi training set (70%) dan testing set (30%).

C. Algoritma yang digunakan

- 1) Regresi Logistik → untuk prediksi probabilitas kelayakan.
- 2) SVM (kernel linear) → untuk klasifikasi dengan margin maksimal.
- 3) Decision Tree (max depth = 3) → untuk menghasilkan aturan if-else.

D. Evaluasi

Akurasi dan classification report digunakan sebagai metrik evaluasi.

IV. HASIL DAN PEMBAHASAN

Implementasi menunjukkan hasil sebagai berikut:

TABLE I
PERBANDINGAN HASIL ALGORITMA

Algoritma	Akurasi	Kelebihan	Kekurangan
Regresi Logistik	83%	Mudah diimplementasikan	Kurang fleksibel
SVM (Linear)	83%	Generalisasi baik	Sulit diinterpretasi
Decision Tree	100%	Interpretasi mudah	Rentan overfitting

Decision Tree menghasilkan akurasi tertinggi, namun keunggulannya terutama terletak pada transparansi dalam memberikan aturan keputusan. Sebagai contoh:

Jika pendapatan > 4000 dan lama_bekerja > 5 → Layak.

Jika pendapatan < 3000 dan umur < 25 → Tidak Layak.

Selain itu, hasil *classification report* untuk Decision Tree ditunjukkan pada Tabel II.

TABLE II
CLASSIFICATION REPORT (DECISION TREE)

Kelas	Precision	Recall	F1-score	Support
Tidak Layak (0)	1.00	1.00	1.00	1
Layak (1)	1.00	1.00	1.00	2
Accuracy	1.00			

V. IMPLEMENTASI PROGRAM

Kode Python berikut digunakan untuk implementasi regresi logistik, SVM, dan Decision Tree pada dataset simulasi:

```
1 import pandas as pd
2 from sklearn.model_selection import
  train_test_split
3 from sklearn.linear_model import
  LogisticRegression
4 from sklearn.svm import SVC
5 from sklearn.tree import
  DecisionTreeClassifier
6 from sklearn.metrics import accuracy_score,
  classification_report
7
8 # 1. Dataset
9 data = {
10     "pendapatan": [3000, 4000, 2500, 5000,
11                    10000, 2000, 3500, 8000],
12     "umur": [25, 40, 23, 35, 50, 22, 30, 45],
13     "lama_bekerja": [2, 10, 1, 7, 20, 1, 4,
14                     15],
15     "riwayat_pinjaman": [0, 1, 0, 1, 1, 0, 1,
16                          1],
17     "layak": [0, 1, 0, 1, 1, 0, 1, 1]
18 }
19
20 df = pd.DataFrame(data)
21 X = df.drop("layak", axis=1)
22 y = df["layak"]
23
24 X_train, X_test, y_train, y_test =
25     train_test_split(
26         X, y, test_size=0.3, random_state=42)
27
28 # 2. Logistic Regression
29 logreg = LogisticRegression()
30 logreg.fit(X_train, y_train)
31 y_pred_log = logreg.predict(X_test)
32
33 # 3. Support Vector Machine
34 svm = SVC(kernel='linear')
35 svm.fit(X_train, y_train)
36 y_pred_svm = svm.predict(X_test)
37
38 # 4. Decision Tree
39 tree = DecisionTreeClassifier(max_depth=3,
40                               random_state=42)
41 tree.fit(X_train, y_train)
42 y_pred_tree = tree.predict(X_test)
43
44 # 5. Evaluasi
45 print("Logistic Regression Accuracy:",
46       accuracy_score(y_test, y_pred_log))
47 print("SVM Accuracy:", accuracy_score(y_test,
48                                       y_pred_svm))
49 print("Decision Tree Accuracy:",
50       accuracy_score(y_test, y_pred_tree))
51
52 print("\nClassification Report (Decision Tree)
53 : \n",
54       classification_report(y_test,
55                             y_pred_tree))
```

Listing 1. Implementasi Python untuk Pengambilan Keputusan

VI. KESIMPULAN

Ketiga algoritma supervised learning yang diuji dapat digunakan untuk mendukung pengambilan keputusan. Decision Tree unggul dalam interpretabilitas, SVM unggul dalam generalisasi, dan regresi logistik memberikan informasi probabilitas. Pemilihan algoritma sebaiknya disesuaikan dengan kebutuhan sistem pendukung keputusan yang dikembangkan.

REFERENCES

- [1] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2012.
- [2] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [3] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. CRC Press, 1984.
- [4] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.