

RISIKO MAGANG MAHASISWA DENGAN PENDEKATAN ALGORITMA CONTENT-BASED FILTERING DAN SUPPORT VECTOR MACHINE

Roni Habibi ¹, Darfial Guslan ²

D4 Teknik Informatika, Sekolah Vokasi, Universitas Logistik & Bisnis Internasional ¹
roni.habibi@ulbi.ac.id¹

D4 Logistik Bisnis, Sekolah Vokasi, Universitas Logistik & Bisnis Internasional ²
darfialguslan@ulbi.ac.id²

Abstrak

Penelitian ini menggunakan pendekatan kuantitatif untuk mengkaji penggunaan algoritma *Content-Based Filtering* (CBF) dan *Support Vector Machine* (SVM) dalam manajemen risiko magang mahasiswa. Metode penelitian yang digunakan adalah *Crisp-DM* (Cross-Industry Standard Process for Data Mining) yang membantu peneliti mengorganisasikan dan menganalisis data dengan langkah-langkah yang terstruktur. Dalam penelitian ini, peneliti mengumpulkan data dengan menggunakan kuesioner yang dirancang khusus yang bertujuan untuk memperoleh informasi mengenai risiko-risiko yang mungkin terjadi selama mahasiswa magang di perusahaan, faktor-faktor yang mempengaruhi keberhasilannya, dan bagaimana algoritma CBF dan SVM dapat digunakan untuk menganalisis risiko. Dengan menggunakan pendekatan kuantitatif dan *Crisp-DM*, penelitian ini bertujuan untuk memberikan informasi yang objektif dan terukur tentang penggunaan algoritma CBF dan SVM dalam manajemen risiko magang mahasiswa. Diharapkan penelitian ini dapat memberikan pemahaman yang lebih baik tentang bagaimana algoritma CBF dan SVM dapat membantu mengelola risiko dalam magang mahasiswa. Hasil penelitian ini juga diharapkan dapat memberikan saran praktis kepada organisasi yang lebih baik.

Kata Kunci: Risiko, *Content-Based Filtering*, *Support Vector Machine*, Magang.

Abstract

This study uses a quantitative approach to examine the use of Content-Based Filtering (CBF) and Support Vector Machine (SVM) algorithms in student internship risk management. The research method used is Crisp-DM (Cross-Industry Standard Process for Data Mining) which helps researchers organize and analyze data with structured steps. In this study, researchers collected data using a specially designed questionnaire that aims to obtain information about the risks that may occur during student internships in companies, factors that influence their success, and how CBF and SVM algorithms can be used to analyze risks. By using a quantitative approach and Crisp-DM, this study aims to provide objective and measurable information about the use of CBF and SVM algorithms in student internship risk management. It is hoped that this study can provide a better understanding of how CBF and SVM algorithms can help manage risks in student internships. The results of this study are also expected to provide practical advice to better organizations.

Keywords: Risk, *Content-Based Filtering*, *Support Vector Machine*, Internship.

1. PENDAHULUAN

Saat ini, magang telah menjadi salah satu pengalaman yang sangat diinginkan oleh mahasiswa selama masa studi mereka di perguruan tinggi. Program magang memberikan kesempatan bagi mahasiswa untuk mendapatkan pengalaman kerja praktis sebelum memasuki dunia profesional (“Rethinking Graduate Employability in Context,” 2023) (Galbraith & Mondal, 2020). Program ini memungkinkan mahasiswa untuk menerapkan pengetahuan akademis mereka dalam situasi dunia nyata, meningkatkan keterampilan praktis, dan memperluas jaringan profesional mereka (Bawica, 2021) (Hia et al., 2023). Program ini juga memberikan wawasan yang berharga tentang industri yang ingin mereka masuki, serta membantu mereka dalam memahami dinamika dan tuntutan pekerjaan di bidang mereka (Ismail et al., 2018)(Krechowicz, 2022).

Namun, meskipun program magang memiliki banyak manfaat, terdapat berbagai risiko yang harus dihadapi oleh mahasiswa. Risiko ini mencakup ketidakcocokan antara harapan mahasiswa dan realitas pekerjaan, adaptasi terhadap lingkungan kerja yang baru, dan perbedaan budaya serta nilai-nilai antara mahasiswa dan karyawan perusahaan (Odlin et al., 2022) (Enke et al., 2021). Ketidakcocokan ini dapat menyebabkan ketidakpuasan, penurunan motivasi, dan bahkan kegagalan dalam program magang (Marinakou, 2020; Phuong et al., 2023). Oleh karena itu, penting untuk memiliki strategi yang efektif untuk mengidentifikasi dan mengelola risiko-risiko ini agar program magang dapat berjalan dengan lancar dan sukses.

Untuk mengatasi masalah ini, penelitian ini mengusulkan penggunaan pendekatan berbasis *Content-Based Filtering* (CBF) dan *Support Vector Machine* (SVM) dalam manajemen risiko program magang mahasiswa. CBF adalah teknik yang digunakan untuk membuat rekomendasi berdasarkan karakteristik atau preferensi individu (Hashim & Waden, 2023)(Jasim & Asif, 2023). Dalam konteks magang, CBF dapat digunakan untuk mencocokkan mahasiswa dengan program magang yang sesuai berdasarkan profil dan preferensi mereka (Olasehinde et al., 2022; Poncio, 2024). Dengan memanfaatkan data historis tentang magang sebelumnya dan umpan balik dari mahasiswa, CBF dapat membantu mengurangi risiko ketidakcocokan dan meningkatkan kepuasan.

Di sisi lain, SVM adalah algoritma pembelajaran mesin yang kuat untuk klasifikasi dan regresi (Babatunde et al., 2023)(Rodríguez-Pérez & Bajorath, 2022). SVM dapat digunakan untuk memprediksi dan mengklasifikasikan risiko yang mungkin muncul selama program magang berdasarkan atribut-atribut yang relevan seperti profil mahasiswa, karakteristik perusahaan (A, 2020), dan umpan balik dari pengalaman magang sebelumnya (Saidani et al., 2022). Dengan menggunakan pendekatan CBF dan SVM, penelitian ini bertujuan untuk mengembangkan sebuah *framework* manajemen risiko yang dapat membantu perguruan tinggi dan perusahaan dalam mengidentifikasi, menganalisis, dan mengurangi risiko-risiko yang terkait dengan program magang mahasiswa, sehingga meningkatkan keberhasilan dan kepuasan semua pihak yang terlibat.

2. METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah pendekatan kuantitatif dengan menggunakan kuesioner. Penelitian ini difokuskan pada mahasiswa yang sedang atau telah menjalani program magang. Selain itu, penelitian ini mengintegrasikan kerangka kerja *CRISP-DM* (Cross-Industry Standard Process for Data Mining) untuk memperdalam pemahaman tentang manajemen risiko dalam program magang mahasiswa dengan menggunakan pendekatan CBF dan SVM.



Gambar 1. Metodologi CRISP-DM (Nugroho, 2022)

1. *Business Understanding*: Pada tahap ini, peneliti memperoleh pemahaman yang mendalam tentang konteks manajemen risiko dalam program magang mahasiswa. Tujuan penelitian, kebutuhan informasi, dan tantangan yang dihadapi dalam manajemen risiko magang dipelajari dengan cermat. Hal ini membantu peneliti dalam merumuskan pertanyaan penelitian yang relevan dan menentukan arah penelitian.
2. *Data Understanding*: Pada tahap ini, peneliti memperoleh pemahaman tentang data yang diperlukan dalam penelitian ini. Peneliti menganalisis jenis data yang diperlukan, sumber data yang tersedia, dan potensi data yang relevan untuk memahami manajemen risiko dalam program magang mahasiswa.
3. *Data Preparation*: Tahap ini melibatkan pemrosesan dan pengolahan data yang diperoleh melalui kuesioner. Peneliti membersihkan, mengintegrasikan, dan mempersiapkan data agar siap digunakan dalam tahap analisis lebih lanjut. Langkah-langkah ini termasuk eksplorasi dan persiapan data untuk memastikan kualitas dan konsistensi data.
4. *Modelling*: Pada tahap ini, peneliti menggunakan pendekatan CBF dan SVM untuk menganalisis risiko dalam program magang mahasiswa. Pendekatan ini membantu dalam pemodelan dan identifikasi risiko yang mungkin muncul serta mendukung pengambilan keputusan yang terkait dengan manajemen risiko magang.
5. *Evaluation*: Tahap ini melibatkan penilaian terhadap hasil analisis yang telah dilakukan. Peneliti mengevaluasi hasil analisis risiko, termasuk keakuratan dan kredibilitas model yang digunakan, serta mengukur efektivitas *framework* manajemen risiko yang diusulkan dalam konteks program magang. Evaluasi ini penting untuk memastikan kesesuaian dan keandalan hasil penelitian.
6. *Deployment*: Pada tahap ini, peneliti menyusun laporan penelitian yang mencakup temuan dan hasil dari penelitian ini. Selain itu, peneliti juga menyampaikan temuan kepada pihak-pihak terkait, seperti mahasiswa, universitas, dan perusahaan, untuk memberikan kontribusi dalam pengembangan program magang dan manajemen risikonya. Dengan cara ini, penelitian ini berfungsi sebagai sumber pengetahuan yang dapat digunakan dalam konteks praktis dan akademis.

Penelitian ini berlokasi di Universitas Logistik dan Bisnis Internasional (ULBI) dan beberapa perusahaan yang menjadi tempat magang mahasiswa. Variabel yang diamati dan diukur mencakup persepsi, pengalaman, sikap, dan pendapat mahasiswa terkait manajemen risiko dalam program magang, serta bagaimana pendekatan CBF dan SVM dapat diterapkan untuk mengurangi risiko tersebut.

Dengan menggunakan kerangka kerja *CRISP-DM*, penelitian ini diharapkan dapat menghasilkan temuan yang lebih komprehensif dan mendalam tentang manajemen risiko dalam program magang mahasiswa, serta memberikan kontribusi penting dalam pengembangan bidang ini.

3. HASIL DAN PEMBAHASAN

A. Deskripsi Data

Data yang digunakan dalam penelitian ini merupakan dataset yang menggambarkan pengalaman mahasiswa dan alumni dari ULBI selama mereka menjalani program magang. Data ini merupakan hasil dari survei yang dilakukan terhadap para peserta magang untuk mengevaluasi berbagai aspek pengalaman mereka selama program magang. Dataset ini terdiri dari beberapa kolom yang menyediakan informasi rinci tentang para mahasiswa dan alumni serta bagaimana mereka mengalami program magang di berbagai perusahaan. Berikut adalah rincian kolom yang terdapat dalam dataset :

Tabel 1. Atribut dan artinya dari dataset

No.	Atribut	Deskripsi
1.	Nama Mahasiswa	Kolom ini berisi nama lengkap dari mahasiswa atau alumni ULBI yang telah mengikuti program magang.
2.	Jenis Kelamin	Kolom ini berisi nama lengkap dari mahasiswa atau alumni ULBI yang telah mengikuti program magang.
3.	Angkatan	Tahun masuk mahasiswa ke ULBI. Ini memberikan informasi tentang generasi mahasiswa.
4.	Umur	Usia mahasiswa atau alumni saat mereka menjalani program magang.
5.	Program Studi- Jurusan	Menunjukkan program studi atau jurusan yang diambil oleh mahasiswa atau alumni, seperti "D4-Teknik Informatika," "D3-Manajemen Logistik," dan berbagai program lainnya.
6.	Perusahaan Tempat Magang	Nama perusahaan atau lembaga tempat mahasiswa atau alumni melakukan program magang. Informasi ini membantu mengidentifikasi perusahaan atau organisasi yang terlibat dalam kolaborasi dengan ULBI
7.	Bidang Usaha Perusahaan	Mendeskripsikan jenis bisnis atau bidang usaha dari perusahaan tempat mahasiswa atau alumni melakukan program magang, seperti "Logistik," "Pengembang & Operator Real Estat," atau bidang usaha lainnya.
8.	Jangka Waktu Magang (per)	Kolom ini mencatat berapa lama mahasiswa atau alumni menjalani program magang dalam satuan bulan.
9.	Kesesuaian Minat dan Bakat	Penilaian tingkat kesesuaian program magang dengan minat dan bakat mahasiswa/alumni dalam skala 1-5.
10.	Mudah Memahami Tugas dan Tanggung Jawab	Mengukur kemudahan pemahaman mahasiswa/alumni terhadap tugas dan tanggung jawab yang diberikan selama magang dalam skala 1-5.
11.	Kesesuaian Tugas dan Tanggung Jawab	Menilai tingkat kesesuaian tugas dan tanggung jawab dengan program studi/jurusan mahasiswa/alumni dalam skala 1-5.
12.	Komunikasi dengan Rekan Kerja dan Atasan	Penilaian kemampuan mahasiswa/alumni dalam berkomunikasi dengan rekan kerja dan atasan selama program magang dalam skala 1-5.
13.	Informasi Mengenai Peraturan dan Kebijakan Tempat Kerja	Mengukur pemahaman mahasiswa/alumni terhadap peraturan dan kebijakan tempat magang dalam skala 1-5.
14.	Program Magang Tepat Waktu	Mengindikasikan apakah mahasiswa/alumni menyelesaikan program magang sesuai jadwal yang ditentukan dalam skala 1-5.

No.	Atribut	Deskripsi
15.	Keadilan dalam Evaluasi Kinerja	Penilaian apakah evaluasi kinerja selama program magang dianggap adil oleh mahasiswa/alumni dalam skala 1-5.
16.	Kesesuaian dalam Evaluasi Program Kerja	Menilai sejauh mana program kerja yang diberikan sesuai dengan program studi/jurusan mahasiswa/alumni dalam skala 1-5.
17.	Adanya Diskriminasi atau Pelecehan	Penilaian apakah mahasiswa/alumni mengalami diskriminasi atau pelecehan selama program magang.
18.	Kenyamanan Lingkungan Kerja	Penilaian tentang sejauh mana mahasiswa/alumni merasa nyaman dengan lingkungan kerja selama program magang dalam skala 1-5.
19.	Total	Total akumulasi penilaian dari seluruh aspek di atas.

B. Pengolahan Data

Data yang digunakan dalam penelitian ini telah melalui serangkaian tahap pengolahan untuk memastikan kualitas data yang baik dalam pemodelan machine learning. Berikut adalah tahap-tahap pengolahan data yang telah dilakukan :

- 1) Penghapusan Data Tidak Lengkap: Data yang mengandung nilai kosong (missing values) atau data yang tidak lengkap dihapus dari dataset. Hal ini dilakukan untuk memastikan bahwa data yang digunakan dalam analisis adalah data yang lengkap dan dapat dipercaya.
- 2) Konversi Kolom Teks Menjadi Data Numerik: Beberapa kolom dalam dataset, seperti "Program Studi-Jurusan," "Bidang Usaha Perusahaan," dan "Jenis Kelamin," awalnya berisi data teks. Untuk digunakan dalam model machine learning, kolom-kolom ini diubah menjadi data numerik. Sebagai contoh, "Program Studi-Jurusan" dikonversi menjadi bilangan bulat, di mana setiap program studi diberi nomor yang sesuai.
- 3) Konversi Kolom Target: Kolom "risk_level," yang merupakan target dari model machine learning, berisi tiga kategori tingkat risiko, yaitu "Risiko Rendah," "Risiko Sedang," dan "Risiko Tinggi." Kategori ini diubah menjadi data numerik dengan memberikan label numerik, seperti '1' untuk "Risiko Rendah," '2' untuk "Risiko Sedang," dan '3' untuk "Risiko Tinggi".
- 4) *One-Hot Encoding* dan *Label Encoding*: Ini adalah teknik yang digunakan untuk mengatasi atribut kategorikal (atribut yang memiliki nilai kategori atau label) sebelum menyematkannya ke dalam model pembelajaran mesin. *One-Hot Encoding* mengubah atribut kategorikal menjadi vektor biner yang dapat digunakan dalam model. *Label Encoding*, di sisi lain, mengubah atribut kategorikal menjadi bilangan bulat (integer). Teknik ini diperlukan karena banyak algoritme pembelajaran mesin hanya dapat bekerja dengan data numerik, dan atribut kategorikal perlu dikonversi sebelum digunakan dalam model.
- 5) Pembagian Data: Data dibagi menjadi dua subset, yaitu data latih (train) dan data uji (test). Data latih digunakan untuk melatih model machine learning, sementara data uji digunakan untuk menguji sejauh mana model tersebut mampu membuat prediksi yang akurat.

C. Model Machine Learning

Bagian ini berfokus pada penerapan dua jenis model machine learning yang digunakan dalam penelitian, yaitu *Super Vector Machine* (SVM) dan *Content-Based Filtering* (CBF). Setiap model memiliki keunggulan dan cara kerja yang berbeda. Berikut adalah penjelasan lebih rinci tentang kedua model ini :

1) *Content-Based Filtering* (CBF)

CBF adalah metode rekomendasi yang menggunakan informasi atau atribut dari item yang ada (seperti genre, penulis, atau deskripsi) untuk merekomendasikan item serupa kepada pengguna. Sistem ini menganalisis fitur dari item yang disukai pengguna dan mencari item lain yang memiliki karakteristik serupa. Algoritma ini memiliki beberapa keunggulan:

- a) Personalisasi Rekomendasi: Menyediakan rekomendasi yang disesuaikan dengan preferensi individu berdasarkan riwayat pengguna.
- b) Penanganan Item Baru: Efektif untuk merekomendasikan item baru yang belum pernah diberi peringkat oleh pengguna lain.
- c) Independen dari Pengguna Lain: Rekomendasi tidak bergantung pada preferensi pengguna lain, sehingga lebih fokus pada minat pribadi pengguna.

```
from sklearn.metrics import classification_report, accuracy_score
def predict_risk_cbf(similarity_matrix, y_train):
    predictions = []
    for i in range(similarity_matrix.shape[0]):
        similar_indices = np.argsort(similarity_matrix[i])[-5:]
        similar_scores = y_train.iloc[similar_indices]
        predicted_risk = similar_scores.mode()[0]
        predictions.append(predicted_risk)
    return predictions
y_pred_cbf = predict_risk_cbf(similarity_matrix_test_train, y_train)
cbf_accuracy = accuracy_score(y_test, y_pred_cbf)
cbf_classification_report = classification_report(y_test, y_pred_cbf)
print(f"CBF Accuracy: {cbf_accuracy}")
print("CBF Classification Report:")
print(cbf_classification_report)
```

Output :

```
CBF Accuracy: 0.8717948717948718
CBF Classification Report:
              precision    recall  f1-score   support

     1         0.00        0.00        0.00         4
     2         0.89        0.97        0.93        35

 accuracy          0.87         0.87         0.87         39
 macro avg         0.45         0.49         0.47         39
 weighted avg         0.88         0.87         0.84         39
```

Keterangan :

Dalam evaluasi model Content-Based Filtering, digunakan metode predict untuk set uji (*y_pred_cbf*) dihitung menggunakan matriks kesamaan yang sudah dibuat sebelumnya dengan memanggil fungsi *predict_risk_cbf*. Setelah itu, akurasi model dihitung menggunakan *accuracy_score*, yang membandingkan prediksi (*y_pred_cbf*) dengan nilai risiko aktual (*y_test*). Selain itu, laporan klasifikasi (*classification_report*) juga dihasilkan, memberikan metrik evaluasi seperti precision, recall, dan f1-score untuk setiap kategori risiko.

2) Support Vector Machine (SVM)

SVM adalah salah satu algoritma *machine learning* yang digunakan untuk klasifikasi dan regresi. Algoritma ini bekerja dengan cara menemukan *hyperplane* optimal yang memisahkan data dari dua kelas dengan margin yang terbesar. Hyperplane adalah garis (untuk data dua dimensi) atau bidang (untuk data tiga dimensi atau lebih) yang memisahkan data dari dua kelas. SVM juga dapat digunakan dalam kasus non-linear dengan menerapkan fungsi kernel untuk memetakan data ke dimensi yang lebih tinggi di mana data tersebut dapat dipisahkan dengan lebih mudah. Adapun keunggulannya:

- a) Kinerja yang baik pada data dengan dimensi tinggi: SVM efektif dalam bekerja dengan data yang memiliki banyak fitur (dimensi) karena mencari hyperplane optimal yang memaksimalkan margin antara kelas-kelas.

- b) Efektif untuk kasus non-linear: Dengan menggunakan kernel trick, SVM dapat memetakan data non-linear ke ruang fitur yang lebih tinggi, sehingga memungkinkan pemisahan data yang lebih baik.
- c) Robust terhadap *overfitting*: SVM menggunakan margin maksimal untuk memisahkan kelas-kelas, sehingga lebih robust terhadap *overfitting*, terutama ketika margin antara kelas cukup besar.
- d) Penggunaan kernel yang fleksibel: SVM mendukung berbagai fungsi kernel (linear, polynomial, *radial basis function* (RBF), dll.), yang memungkinkan fleksibilitas dalam menangani berbagai tipe data.

```
from sklearn.metrics import classification_report, accuracy_score

# Initialize the SVM classifier
svm_model_linear = SVC(kernel='linear')
# Fit the model on the training data
svm_model_linear.fit(X_train, y_train)
# Assuming your target variable has two classes named 'Class 0' and 'Class 1'
class_names = ['Class 0', 'Class 1'] # Define the class names
# Evaluate the model
accuracy = accuracy_score(y_test, y_pred)
report = classification_report(y_test, y_pred)
print(f'Accuracy: {accuracy}')
print(f'Classification Report:\n{report}')
```

Output :

```
Accuracy: 0.9565217391304348
Classification Report:
              precision    recall  f1-score   support

     0       0.86       1.00       0.92         6
     1       1.00       0.94       0.97        17

 accuracy          0.96
 macro avg       0.93       0.97       0.95
 weighted avg    0.96       0.96       0.96
```

Keterangan :

model *Decision Tree Classifier* yang telah disesuaikan dengan parameter terbaik dari hasil pencarian *GridSearchCV* digunakan untuk melakukan prediksi pada data uji (x_{test}). Hasil prediksi tersebut kemudian digunakan untuk menghitung beberapa metrik evaluasi, seperti akurasi, presisi, *recall*, dan lainnya.

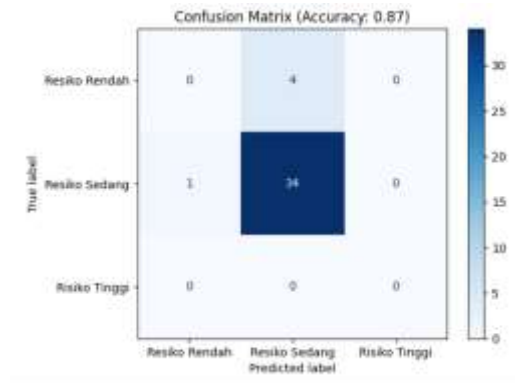
D. Analisis Hasil

Bagian analisis hasil adalah tahap penting dalam penelitian ini karena memberikan wawasan yang berharga tentang bagaimana model *machine learning* telah berhasil dalam mengklasifikasikan mahasiswa dan alumni ULBI ke dalam tiga tingkat risiko. Hasil analisis ini juga memberikan gambaran yang lebih jelas tentang perbandingan kedua algoritma dalam menentukan risiko magang mahasiswa.

1) Hasil Evaluasi Model

Content-Based Filtering:

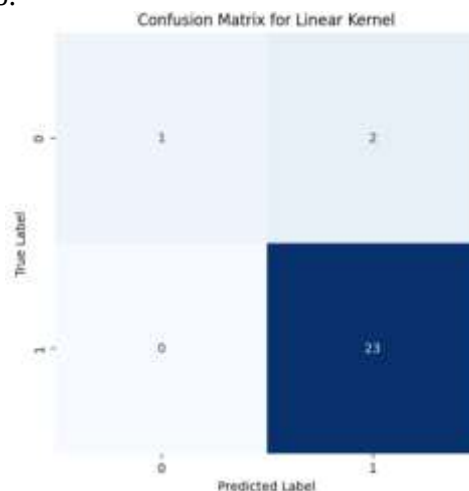
Model ini memiliki akurasi keseluruhan 0.87, yang berarti 87% dari semua prediksi yang dilakukan oleh model adalah benar. Namun, model ini sangat cenderung untuk memprediksi kategori Risiko Sedang, dan kurang mampu dalam membedakan antara Risiko Rendah dan Risiko Tinggi.



Gambar 2. Confusion Matrix Content-Based Filtering

Support Vector Machine:

Model ini memberikan tingkat akurasi sekitar 92%, 88%, 88%, dan 88% pada data uji, yang menggambarkan kemampuannya dalam mengklasifikasikan individu ke dalam tiga kategori risiko dengan tingkat ketepatan yang tinggi. Selain itu, hasil evaluasi menunjukkan bahwa model SVM dengan kernel linear memiliki nilai precision dan recall yang tinggi untuk masing-masing tingkat risiko. Hal ini berarti model ini mampu dengan baik mendeteksi dan mengklasifikasikan individu dengan risiko rendah, sedang, atau tinggi. Hasil ini memberikan indikasi bahwa SVM adalah model yang sangat efektif dalam mengelola risiko.



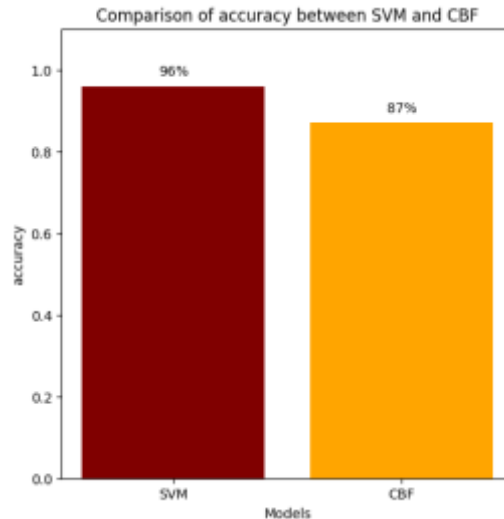
Gambar 3. Confusion Matrix Support Vector Machine

2) Perbandingan Kinerja Model

Dalam analisis perbandingan kinerja model, hasil evaluasi SVM dan CBF dibandingkan.

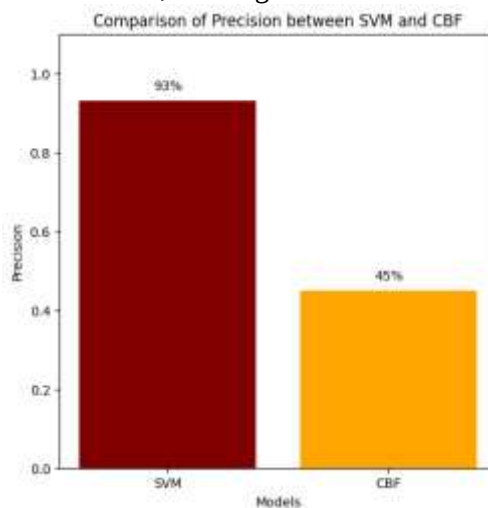
Content-Based Filtering (CBF) vs. Support Vector Machine (SVM):

1. **Akurasi:** memiliki akurasi sebesar 96%, sedangkan model CBF memiliki akurasi sebesar 87%. Perbandingan ini menggambarkan bahwa SVM memberikan hasil yang lebih akurat dibandingkan dengan CBF dalam konteks yang diuji.



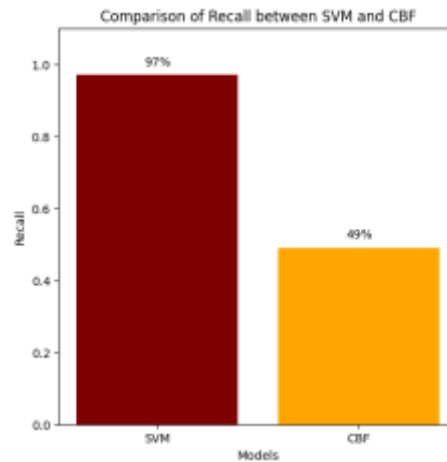
Gambar 4. Perbandingan Nilai Akurasi CBF dan SVM

2. **Precision:** model SVM memiliki tingkat presisi sebesar 93%. Sementara itu, model CBF hanya memiliki tingkat presisi sebesar 45%. Dari perbandingan ini, terlihat bahwa model SVM jauh lebih presisi dibandingkan dengan model CBF. Ini berarti bahwa ketika SVM mengklasifikasikan seseorang ke dalam tingkat risiko tertentu, kemungkinan besar hasil tersebut adalah benar.



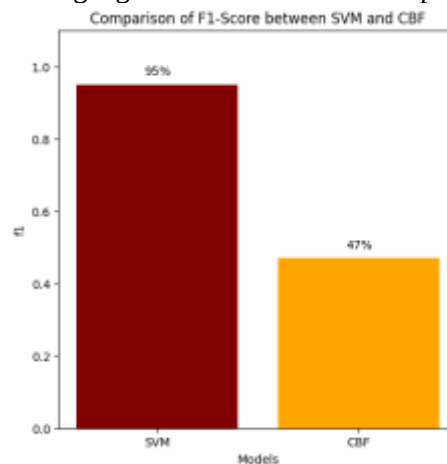
Gambar 5. Perbandingan Nilai *Precision* CBF dan SVM

3. **Recall:** *Recall* mengukur kemampuan model untuk mendeteksi individu dengan risiko tertentu. model SVM memiliki tingkat recall sebesar 97%. Sementara itu, model CBF memiliki tingkat *recall* sebesar 49%. Dari perbandingan ini, terlihat bahwa model SVM memiliki kemampuan yang jauh lebih baik dalam mendeteksi individu berisiko.



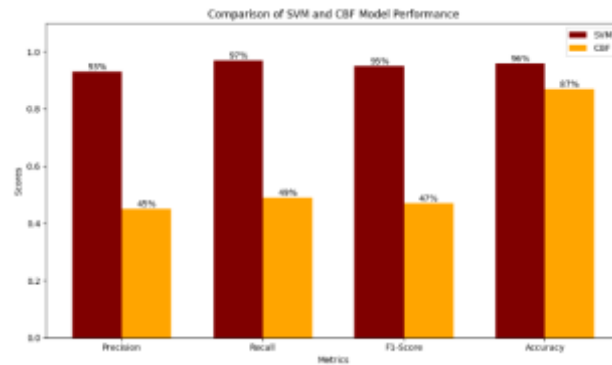
Gambar 6. Perbandingan Nilai Recall CBF dan SVM

4. **F1-Score:** model SVM memiliki hasil *F1-Score* sebesar 95%. Sementara itu, model CBF memiliki hasil *F1-Score* sebesar 47%. Hasil ini mengindikasikan bahwa model CBF tidak seefektif model SVM dalam memprediksi risiko magang mahasiswa dalam kelas positif dan negative.



Gambar 7. Perbandingan Nilai F1-Score CBF dan SVM

Dari hasil keseluruhan perbandingan nilai precision, recall, dan *f1-score* pada kedua model menunjukkan bahwa *Support Vector Machine* (SVM) lebih efektif dalam mengklasifikasikan mahasiswa dan alumni ULBI ke dalam tiga tingkat risiko dibandingkan dengan *Content-Based Filtering* (CBF).



Gambar 8. Perbandingan Model CBF dan SVM

4. KESIMPULAN DAN SARAN

Penelitian ini mengembangkan sebuah *framework* manajemen risiko yang dapat membantu perguruan tinggi dan perusahaan dalam mengidentifikasi, menganalisis, dan mengurangi risiko-risiko yang terkait dengan program magang mahasiswa melalui penggunaan pendekatan *Content-Based Filtering* (CBF) dan *Support Vector Machine* (SVM). Hasil penelitian menunjukkan bahwa algoritma SVM lebih efektif dibandingkan dengan CBF dalam mengklasifikasikan risiko magang, dengan akurasi, presisi, dan kemampuan deteksi yang lebih tinggi. SVM berhasil mengelola risiko secara lebih efisien, membuatnya menjadi pilihan yang lebih unggul dalam konteks analisis risiko magang mahasiswa.

Rekomendasi dari penelitian ini adalah penggunaan algoritma SVM untuk meningkatkan manajemen risiko dalam program magang mahasiswa, karena terbukti lebih efektif dalam klasifikasi risiko dibandingkan dengan CBF. Selain itu, sistem rekomendasi berbasis CBF tetap relevan untuk mencocokkan mahasiswa dengan program magang yang sesuai dengan profil mereka, yang dapat meningkatkan kesesuaian dan kepuasan. Hasil penelitian ini memberikan kontribusi penting dalam pengembangan metode manajemen risiko yang lebih efektif untuk program magang mahasiswa, dengan memanfaatkan algoritma SVM dan CBF. Selain itu, kontribusi ini relevan bagi institusi pendidikan dan perusahaan dalam mengelola risiko secara lebih efisien, yang dapat diterapkan pada berbagai konteks program magang di masa depan.

5. REFERENSI

- A, J. A. (2020). Analyzing and Predicting Career Specialization using Classification Techniques. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(1.3), 342–348. <https://doi.org/10.30534/ijatcse/2020/5391.32020>
- Babatunde, R., Abdulsalam, S. O., Abdulsalam, O. A., & Arowolo, M. O. (2023). Classification of customer churn prediction model for telecommunication industry using analysis of variance. *IAES International Journal of Artificial Intelligence*, 12(3), 1323–1329. <https://doi.org/10.11591/ijai.v12.i3.pp1323-1329>
- Bawica, I. M. (2021). The Effects of Internship Program on the Employability Readiness. *International Journal of Academe and Industry Research*, 2(3), 86–101. <https://doi.org/10.53378/348731>
- Enke, E. J. F. L., Martens, M. L., Martens, C. D. P., Enke, D. B. S., Contador, J. C., Bachim, T., Coutinho, A. dos R., & Satyro, W. C. (2021). Critical success factors in university-industry collaboration: a bibliometric analysis. *Research, Society and Development*, 10(13), e248101321316. <https://doi.org/10.33448/rsd-v10i13.21316>
- Galbraith, D., & Mondal, S. (2020). The potential power of internships and the impact on career preparation. *Research in Higher Education Journal*, 38, 1–9.

- Hashim, S., & Waden, J. (2023). Content-based filtering algorithm in social media. *Wasit Journal of Computer and Mathematics Science*, 2(1), 14–17. <https://doi.org/10.31185/wjcm.112>
- Hia, L. M., Oktavianus, S., Sidebang, S. K., & Kuang, T. M. (2023). Dampak Magang Merdeka Belajar Kampus Merdeka Bagi Mahasiswa Akuntansi. *Owner*, 7(4), 2796–2806. <https://doi.org/10.33395/owner.v7i4.1675>
- Ismail, I., Hasan, H., & Musdalifah, M. (2018). Pengembangan Kompetensi Mahasiswa Melalui Efektivitas Program Magang Kependidikan. *Edumaspul: Jurnal Pendidikan*, 2(1), 124–132. <https://doi.org/10.33487/edumaspul.v2i1.48>
- Jasim, K., & Asif, M. (2023). Movie Recommendation Systems Using Content-Based Filtering. *International Research Journal of Modernization in Engineering Technology and Science*, 06, 3892–3902. <https://doi.org/10.56726/irjmet42626>
- Krechowicz, M. (2022). Towards Sustainable Project Management: Evaluation of Relationship-Specific Risks and Risk Determinants Threatening to Achieve the Intended Benefit of Interorganizational Cooperation in Engineering Projects. *Sustainability (Switzerland)*, 14(5). <https://doi.org/10.3390/su14052961>
- Marinakou, E. (2020). Hotel internships and student satisfaction as key determinant to career intention. *Journal of Tourism Research*, 25, 42.
- Odlin, D., Benson-Rea, M., & Sullivan-Taylor, B. (2022). Student internships and work placements: approaches to risk management in higher education. *Higher Education*, 83(6), 1409–1429. <https://doi.org/10.1007/s10734-021-00749-w>
- Olasehinde, O., Abiona, A. A., Johnson, O. V., & Omilegan, D. (2022). Design and Implementation of a Web-Based Internship Placement Recommendation System: A Case Study of Federal Polytechnic, Ile-Oluji, Nigeria. *International Journal of Scientific & Engineering Research*, 13(March), 398–409.
- Phuong, H. Y., Le, T. T., Ho, P. T., & Do, T. N. (2023). English Major Undergraduates' Unsatisfactory Practicum Experience: A Case Study in Vietnam. *Vietnam Journal of Education*, 7(1), 38–47. <https://doi.org/10.52296/vje.2023.244>
- Poncio, F. P. (2024). Navigating techniques in job recommender systems on internship profile matching: a systematic review. *Journal of Research in Innovative Teaching and Learning*, 17(2), 352–367. <https://doi.org/10.1108/JRIT-01-2024-0016>
- Rethinking Graduate Employability in Context. (2023). In *Rethinking Graduate Employability in Context*. <https://doi.org/10.1007/978-3-031-20653-5>
- Rodríguez-Pérez, R., & Bajorath, J. (2022). Evolution of Support Vector Machine and Regression Modeling in Chemoinformatics and Drug Discovery. *Journal of Computer-Aided Molecular Design*, 36(5), 355–362. <https://doi.org/10.1007/s10822-022-00442-9>
- Saidani, O., Menzli, L. J., Ksibi, A., Alturki, N., & Alluhaidan, A. S. (2022). Predicting Student Employability Through the Internship Context Using Gradient Boosting Models. *IEEE Access*, 10, 46472–46489. <https://doi.org/10.1109/ACCESS.2022.3170421>