VOL. 06. NO. 02 DESEMBER 2025
P-ISSN : 2549211X | E-ISSN : 25499351
DOI: 10.32485/jcsai.

JOURNAL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE (JCSAI)

KLASIFIKASI RISIKO FATALITAS EKSPEDISI GUNUNG BERDASARKAN ATRIBUT DEMOGRAFI DAN GEOGRAFIS MENGGUNAKAN NATURAL LANGUANGE PROCESSING DAN SUPERVISED LEARNING

Imbaraga Gempar Guna Laksana[1]; Dian Ade Kurnia[2]; Yudhistira Arie Wijaya[3]; Gifthera Dwilestar[4]; Mulyawan[5];

Program Studi Teknik Informatika[1]
Program Studi Manajemen Informatika[2]
Program Studi Sistem Informasi[345]

STMIK IKMI Cirebon
https://ikmi.ac.id/page/18/?lang=de
imbaragagempar@gmail.com

(*) Corresponding Author : imbaragagempar@gmail.com
Published : 31 Desember 2025

Abstract—*This study aims to develop a model for the classification of mountain expedition fatality risk by utilizing Natural Language Processing and supervised learning techniques to process text data on the causes of death and demographic attributes. This research responds to the challenge of unstructured data processing that often contains writing variations and ambiguities so that it requires computational methods that are able to capture important information accurately. The methods used include text preprocessing, TF-IDF weighting, frequency encoding for citizenship attributes, and the development of Random Forest and Support Vector Machine models. The model is evaluated using the Accuracy, Precision, Recall, and F1-score metrics to ensure the quality of the predictions. The results showed that Random Forest achieved an accuracy of 0.98 and was more stable than SVM in dealing with class imbalances. Text features were shown to contribute the most to determining fatality risk categories, while demographic attributes provided a smaller but still relevant additional influence. These findings suggest that NLP-based analysis can improve understanding of fatality risk patterns and potentially support the development of decision support systems for mountaineering safety. This approach makes it easier to identify risk factors that were previously difficult to know due to the limitations of manual analysis. This research provides a solid basis for the development of more comprehensive risk models that can be adapted to other safety domains.*

Keywords: *Natural Language Processing, Supervised Learning, Expedition Fatality, Risk Classification, Machine Learning*

Abstrak- Penelitian ini bertujuan mengembangkan model klasifikasi risiko fatalitas ekspedisi gunung dengan memanfaatkan teknik Natural Language Processing dan supervised learning untuk mengolah data teks penyebab kematian serta atribut demografis. Penelitian ini merespons tantangan pengolahan data tidak terstruktur yang sering mengandung variasi penulisan dan ambiguitas sehingga membutuhkan metode komputasi yang mampu menangkap informasi penting secara akurat. Metode yang digunakan meliputi preprocessing teks, pembobotan TF-IDF, frequency encoding untuk atribut kewarganegaraan, serta pembangunan model Random Forest dan Support Vector Machine. Model dievaluasi menggunakan metrik Accuracy, Precision, Recall, dan F1-score untuk memastikan kualitas prediksi. Hasil penelitian menunjukkan bahwa Random Forest mencapai akurasi 0.98 dan lebih stabil dibandingkan SVM dalam menangani ketidakseimbangan kelas. Fitur teks terbukti memberi kontribusi terbesar dalam menentukan kategori risiko fatalitas, sementara atribut demografis memberi pengaruh tambahan yang lebih kecil tetapi tetap relevan. Temuan ini menunjukkan bahwa analisis berbasis NLP dapat meningkatkan pemahaman terhadap pola risiko fatalitas dan berpotensi mendukung pengembangan sistem pendukung keputusan untuk keselamatan pendakian gunung. Pendekatan ini memudahkan identifikasi faktor risiko yang sebelumnya sulit diketahui karena keterbatasan analisis manual. Penelitian ini memberi dasar yang kuat untuk pengembangan model risiko yang lebih komprehensif dan dapat diadaptasi pada domain keselamatan lainnya.

VOL. 06. NO. 02 DESEMBER 2025
P-ISSN : 2549211X | E-ISSN : 25499351
DOI: 10.32485/jcsai.

Kata Kunci : Natural Language Processing, Supervised Learning, Fattalitas Ekspedisi, Klasifikasi Risiko, Machine Learning

INTRODUCTION

Perkembangan pesat bidang Informatika telah mendorong transformasi besar dalam berbagai aspek kehidupan manusia, mulai dari industri, pendidikan, hingga layanan public (Siregar et al., 2022). Kemajuan teknologi komputasi memungkinkan pengolahan data dalam skala besar dan kompleks dengan tingkat akurasi yang semakin tinggi (Wang et al., 2024). Pendekatan berbasis kecerdasan buatan dan pembelajaran mesin telah memperluas kemampuan sistem informasi dalam memahami pola dan menghasilkan prediksi dari data dunia nyata (BuHamra et al., 2022). Perkembangan teknik analisis data, termasuk Natural Language Processing (NLP), telah membuka peluang baru dalam mengolah data tekstual yang tidak terstruktur untuk mendukung pengambilan keputusan (Khairuddin et al., 2024). Bidang Informatika saat ini bergerak menuju integrasi berbagai metode komputasi untuk menyelesaikan persoalan kompleks yang sebelumnya sulit ditangani oleh pendekatan tradisional (Li et al., 2024).

Salah satu tantangan utama dalam pengolahan data di era digital adalah keberadaan data tidak terstruktur yang sulit dianalisis secara konvensional (Siregar et al., 2022). Data tidak terstruktur, seperti deskripsi kejadian, narasi kecelakaan, atau catatan medis, sering kali mengandung ambiguitas dan variasi penulisan yang signifikan (BuHamra et al., 2022). Pada konteks pengelolaan keselamatan dan mitigasi risiko, kualitas data sebab kematian sangat bergantung pada ketepatan pencatatan dan konsistensi deskripsi (Siregar et al., 2022). Permasalahan muncul ketika catatan tersebut mengandung ketidakteraturan, kesalahan pengetikan, atau format yang tidak standar, sehingga menyulitkan proses analisis otomatis (BuHamra et al., 2022). Dalam domain investigasi kematian, penelitian menunjukkan bahwa ketidakkonsistenan anotasi sering terjadi dan dapat menurunkan keakuratan analisis (Wang et al., 2024). Tantangan ini juga terlihat pada laporan kecelakaan kerja yang menggunakan narasi terbuka, sehingga klasifikasi tingkat keparahan sering sulit dilakukan secara manual (Khairuddin et al., 2024). Upaya memodelkan risiko atau menyimpulkan penyebab fatalitas dari teks mentah memerlukan pendekatan NLP yang mampu mengekstrak makna secara komputasional (Wang et al., 2024). Ketidakteraturan dalam struktur teks menyebabkan informasi penting menjadi sulit dibedakan tanpa metode analisis yang tepat (Siregar et al., 2022). Oleh karena itu, diperlukan pendekatan yang mampu mengatasi noise dan variabilitas bahasa dalam data riil (BuHamra et al., 2022). Permasalahan ini relevan karena meningkatkan akurasi identifikasi risiko dapat mendukung pencegahan kecelakaan di berbagai domain, termasuk kesehatan, konstruksi, dan aktivitas ekstrem seperti pendakian gunung (Shuang et al., 2023).

Berbagai penelitian sebelumnya telah menunjukkan peran penting NLP dalam peningkatan kualitas analisis sebab kematian dari data tidak terstruktur (Siregar et al., 2022). Studi lain menggunakan model transformer untuk mendeteksi inkonsistensi anotasi dalam catatan investigasi kematian, membuktikan bahwa NLP efektif mendeteksi kesalahan manusia dalam proses pencatatan (Wang et al., 2024). Pada konteks medis, model berbasis decision tree berhasil mengekstraksi penyebab utama dan sekunder kematian dari catatan kesehatan elektronik yang penuh kesalahan pengetikan dan format tidak standar (BuHamra et al., 2022). Selain itu, deep learning telah digunakan untuk mengklasifikasikan tingkat keparahan cedera dari narasi kecelakaan kerja, menunjukkan kemampuan NLP dalam memodelkan risiko keselamatan secara komputasional (Khairuddin et al., 2024). Penelitian lainnya memanfaatkan machine learning untuk mengidentifikasi kombinasi penyebab fatalitas dalam kecelakaan konstruksi, menyoroti pentingnya analisis fitur dalam memahami skenario risiko (Shuang et al., 2023). Bahkan model interpretable machine learning telah digunakan untuk memprediksi risiko kematian akibat keracunan, menegaskan bahwa kombinasi model prediktif dan interpretabilitas memberikan nilai praktis yang tinggi (Li et al., 2024). Namun, penelitian-penelitian tersebut belum banyak mengeksplorasi penggunaan NLP untuk memprediksi risiko fatalitas ekspedisi gunung berdasarkan data demografis dan deskripsi penyebab kematian secara komprehensif (Shuang et al., 2023).

Penelitian ini bertujuan mengembangkan model klasifikasi risiko fatalitas ekspedisi gunung menggunakan teknik NLP untuk mengekstraksi

fitur dari data teks penyebab kematian yang tidak terstruktur (Wang et al., 2024). Selain itu, penelitian ini memanfaatkan supervised learning untuk memprediksi kategori risiko berdasarkan atribut demografis dan geografis seperti kewarganegaraan (Li et al., 2024). Pendekatan ini diharapkan mampu mengisi kesenjangan penelitian sebelumnya yang belum menggabungkan fitur teks dan demografi secara simultan dalam konteks pendakian gunung (Shuang et al., 2023). Signifikansi penelitian ini terletak pada kemampuannya meningkatkan pemahaman tentang pola risiko fatalitas yang dapat membantu upaya mitigasi di masa depan (Khairuddin et al., 2024). Penelitian ini juga berpotensi memberikan kontribusi praktis dalam pengembangan sistem pendukung keputusan untuk keamanan ekspedisi gunung (Siregar et al., 2022).

Metode yang digunakan dalam penelitian ini mencakup preprocessing data tekstual untuk mengatasi ambiguitas dan ketidakteraturan dalam deskripsi penyebab kematian (BuHamra et al., 2022). Teknik NLP seperti tokenisasi, pembersihan teks, dan pembobotan TF-IDF digunakan untuk merepresentasikan informasi teks secara numerik (Wang et al., 2024). Setelah itu, model supervised learning seperti Random Forest dan Support Vector Machine diterapkan untuk melakukan klasifikasi tingkat risiko fatalitas (Li et al., 2024). Proses pengembangan model juga melibatkan evaluasi performa menggunakan metrik seperti precision, recall, dan F1-score untuk memastikan akurasi yang optimal (Khairuddin et al., 2024). Pendekatan ini dirancang untuk mengatasi tantangan data fatalitas pendakian yang bersifat tidak terstruktur dan sangat bervariasi (Siregar et al., 2022).

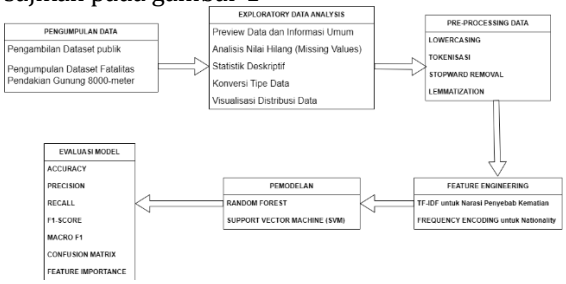
Jika penelitian ini berhasil, hasilnya dapat memberikan wawasan baru mengenai pola risiko fatalitas dalam ekspedisi gunung berbasis analisis data teks dan demografis (Shuang et al., 2023). Temuan penelitian berpeluang meningkatkan efektivitas strategi mitigasi risiko yang lebih adaptif dan berbasis data (Khairuddin et al., 2024). Selain itu, penggunaan NLP dapat membuka peluang untuk menganalisis berbagai data naratif keselamatan lainnya, seperti laporan kecelakaan atau catatan medis (BuHamra et al., 2022). Hasil penelitian juga dapat digunakan sebagai dasar dalam mengembangkan sistem analitik prediktif untuk organisasi pendakian dan badan (Li et al., 2024). Selain itu, kemampuan model untuk memetakan hubungan antara faktor demografis dan risiko fatalitas dapat memberikan perspektif baru bagi penelitian di bidang keselamatan ekstrem (Siregar et al., 2022). Dalam jangka panjang, penelitian ini dapat membantu meningkatkan kualitas data pelaporan penyebab kematian dengan memberikan masukan berbasis machine learning (Wang et al., 2024). Implikasi ini menunjukkan bahwa penelitian ini tidak hanya relevan dalam konteks ilmiah, tetapi juga memiliki potensi penerapan langsung dalam praktik keselamatan lapangan (Shuang et al., 2023).

MATERIALS AND METHODS

Tahapan penelitian merupakan bagian inti dari metodologi yang menjelaskan bagaimana proses analisis dilakukan secara teknis dan sistematis. Tahapan-tahapan ini dimulai dari pengumpulan data hingga evaluasi model, mengikuti pola lifecycle machine learning yang dianjurkan dalam literatur (Crespí et al., 2025; Du, 2025). Dalam konteks penelitian ini, tahapan tersebut sangat penting karena dataset mengandung kombinasi data teks dan data numerik yang memerlukan pendekatan pengolahan berbeda.

Tahapan penelitian ini terdiri atas enam bagian utama, yaitu pengumpulan dataset, *Exploratory Data Analysis (EDA)*, *preprocessing data*, *feature engineering*, pemodelan, dan evaluasi model. Seluruh tahapan disusun berdasarkan metodologi penelitian NLP dan machine learning modern sebagaimana diuraikan dalam berbagai literatur mutakhir (Chen et al., 2024; Khurana et al., 2023; Li et al., 2024).

Setiap tahapan dirancang agar saling terkait. Misalnya, preprocessing teks dilakukan untuk mengurangi noise linguistik sebelum dilakukan TF-IDF pada tahap feature engineering, sementara perhitungan class weight dilakukan untuk mengatasi ketidakseimbangan kelas sebelum masuk ke tahap pemodelan. Dengan demikian, tahapan penelitian tidak hanya berfungsi sebagai prosedur teknis tetapi juga sebagai rangkaian analisis yang memastikan hasil model lebih akurat dan dapat diinterpretasikan. Oleh karena itu, pengantar subbab ini memberikan pemahaman awal mengenai struktur keseluruhan proses penelitian sebelum memasuki detail teknis pada setiap sub-subbab. Untuk tahapan penelitian di sajikan pada gambar 1



Gambar 1 Tahapan Penelitian

Pengumpulan Dataset

VOL. 06. NO. 02 DESEMBER 2025
P-ISSN : 2549211X | E-ISSN : 25499351
DOI: 10.32485/jcsai.

Tahap pengumpulan dataset merupakan langkah fundamental dalam memastikan tersedianya data yang relevan, berkualitas dan sesuai dengan kebutuhan analisis. Penelitian ini memanfaatkan dataset fatalitas ekspedisi pendakian gunung 8000 meter yang bersifat *open-access*. Dataset tersebut dipilih karena menyediakan informasi komprehensif mengenai insiden kematian pendaki, termasuk atribut demografis dan narasi penyebab kematian yang diperlukan untuk penerapan teknik Natural Language Processing (NLP) dan supervised learning. Pemanfaatan data tidak terstruktur sebagai sumber informasi pendukung analisis telah menjadi praktik yang umum dalam riset modern, terutama dalam konteks keselamatan, kesehatan, dan investigasi fatalitas (BuHamra et al., 2022; Siregar et al., 2022).

Proses pengumpulan dataset dilakukan melalui beberapa langkah yang terstruktur. Tahap pertama mencakup identifikasi dan pemilihan sumber data yang menyediakan tingkat kelengkapan dan akurasi pencatatan yang memadai. Setelah sumber data ditetapkan, proses ekstraksi dilakukan untuk memperoleh data dalam format tabular, seperti *.csv* atau *.xlsx*, yang berisi atribut inti seperti lokasi kejadian, kewarganegaraan, dan narasi penyebab kematian. Setiap baris dalam dataset mewakili satu kasus fatalitas, sehingga memungkinkan data dianalisis secara langsung menggunakan teknik komputasional. Penggunaan narasi deskriptif tidak terstruktur sebagai variabel prediktif telah menunjukkan efektivitas dalam berbagai studi NLP terkait investigasi kematian dan kecelakaan (Khairuddin et al., 2024; Wang et al., 2024).

Setelah data terekstraksi, dilakukan proses pembersihan awal untuk memastikan konsistensi dan keandalan data. Langkah ini mencakup deteksi dan penghapusan data duplikat, pemeriksaan nilai kosong, harmonisasi penulisan kategori, serta verifikasi tipe data pada setiap kolom. Upaya ini penting untuk mengurangi potensi bias dan kesalahan pada tahap analisis berikutnya, mengingat data naratif sering kali mengandung ketidakteraturan atau ketidakkonsistenan (BuHamra et al., 2022; Siregar et al., 2022). Validasi lebih lanjut dilakukan untuk memastikan bahwa atribut seperti *nationality* dan *cause of death* memiliki format yang sesuai untuk diproses dalam pipeline NLP dan machine learning, selaras dengan standar pengolahan data tekstual pada penelitian modern (Gao et al., 2021; Khurana et al., 2023).

Tahap terakhir melibatkan seleksi atribut relevan yang digunakan dalam penelitian ini. Atribut seperti *peak*, *nationality*, dan *cause of death* dipertahankan karena memiliki kontribusi langsung terhadap tujuan penelitian, yaitu menggabungkan fitur demografis dan teks dalam membangun model klasifikasi risiko fatalitas. Struktur dataset yang telah distandardisasi kemudian disimpan dan dipersiapkan untuk tahap Exploratory Data Analysis (EDA), pra-pemrosesan data, rekayasa fitur, hingga pemodelan. Tahap pengumpulan dataset ini menjadi fondasi penting karena kualitas data awal akan memengaruhi efektivitas transformasi NLP, representasi fitur, dan kinerja model prediktif yang dihasilkan, sebagaimana ditekankan dalam literatur mengenai lifecycle pengembangan machine learning (Crespí et al., 2025; Du, 2025).

Exploratory Data Analysis (EDA)

Tahap EDA dilakukan untuk memahami karakteristik dataset fatalitas gunung sebelum masuk ke tahapan pemodelan. Analisis dilakukan dengan melihat distribusi nationality, variasi penyebab kematian, distribusi tahun kejadian, serta proporsi kelas risiko. Hasil EDA menunjukkan bahwa dataset terdiri dari 73 kategori nationality**,** dengan dominasi Nepal, Japan, South Korea, dan Spain. Distribusi nationality yang tidak seimbang menunjukkan perlunya teknik encoding yang mampu menangani *high cardinality*, sebagaimana direkomendasikan Hancock et al. (2024).

Selain itu, terdapat **178 variasi narasi penyebab kematian**, mulai dari istilah sederhana seperti *fall*, *avalanche*, hingga deskripsi kompleks seperti *slipped during descent due to ice conditions*. Hal ini sejalan dengan fenomena data naratif keselamatan kerja yang sering bersifat tidak terstruktur dan inkonsisten (Bugalia et al., 2022; Eker & Uçar, 2024). Variasi tersebut menjadi dasar penggunaan TF-IDF dalam proses representasi teks. EDA juga mengidentifikasi ketidakseimbangan kelas risiko, di mana kelas *Medium Risk Physiological* memiliki jumlah kasus paling sedikit. Distribusi ini mengonfirmasi pentingnya penanganan imbalanced dataset dalam pemodelan (Chen et al., 2024; Hellín et al., 2024). Hasil EDA kemudian digunakan sebagai dasar dalam merumuskan strategi preprocessing dan feature engineering.

Preprocessing Data

Tahap preprocessing merupakan langkah penting untuk meningkatkan kualitas data sebelum digunakan dalam model machine learning. Pada dataset teks, preprocessing dilakukan untuk mengurangi noise, menstandardisasi format penulisan, serta menghilangkan elemen-elemen yang tidak relevan seperti stopwords. Proses ini diperlukan karena narasi penyebab kematian sering kali mengandung ketidakkonsistenan, variasi linguistik, serta istilah yang tidak baku (BuHamra et al., 2022). Pembersihan data ini memastikan bahwa

representasi teks lebih bersih dan bermakna saat digunakan pada TF-IDF.

Selain pengolahan teks, preprocessing juga mencakup penanganan ketidakse-imbangan kelas dengan perhitungan class weight. Ketidakseimbangan kelas merupakan permasalahan umum dalam dataset keselamatan yang dapat menyebabkan model bias terhadap kelas mayoritas (Henning et al., 2023). Dengan menerapkan class weight, penelitian ini mengurangi bias tersebut dan memastikan bahwa model dapat mengenali kategori risiko yang lebih jarang muncul. Oleh karena itu, sub-subbab ini menjadi fondasi penting untuk memastikan bahwa data dalam kondisi optimal sebelum masuk ke tahapan feature engineering.

Tahap preprocessing bertujuan memastikan data bersih, terstandarisasi, dan siap digunakan untuk pemodelan machine learning. Preprocessing teks dilakukan menggunakan teknik NLP standar yang terbukti efektif pada data medis dan keselamatan (BuHamra et al., 2022; Wang et al., 2024).

Dua model supervised learning digunakan dalam penelitian ini, yaitu ***Random Forest*** dan ***Support Vector Machine (SVM)***. Pemilihan ini didukung oleh literatur yang menegaskan efektivitas kedua model dalam klasifikasi risiko dan data teks berdimensi tinggi (Mitrakas, 2025; Pugliese et al., 2021).

Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah algoritma klasifikasi yang bertujuan mencari sebuah batas pemisah (*hyperplane*) terbaik untuk memisahkan data ke dalam dua atau lebih kelas. Hyperplane ini merupakan garis keputusan pada ruang dua dimensi, bidang pada ruang tiga dimensi, atau objek berdimensi lebih tinggi ketika fitur yang digunakan berjumlah banyak seperti pada representasi TF-IDF dalam penelitian ini.

SVM bekerja dengan memilih hyperplane yang memiliki *margin* terbesar, yaitu jarak antara hyperplane dengan sampel terdekat dari masing-masing kelas yang disebut *support vector*. Semakin besar margin, semakin baik kemampuan generalisasi model karena hyperplane tidak terlalu dekat dengan sampel apa pun dan tidak terpengaruh oleh noise atau variasi kecil pada data. Pendekatan ini sangat sesuai untuk dataset fatalitas ekspedisi gunung yang memiliki fitur teks berdimensi tinggi, karena SVM mampu tetap stabil meskipun data memiliki banyak atribut dan memiliki pemisahan kelas yang kompleks.

RESULTS AND DISCUSSION

Hasil Pemodelan

Pada sub bab Hasil Pemodelan dijelaskan hasil pemodelan menggunakan algoritma supervised learning, yaitu Random Forest dan Support Vector Machine (SVM). Kedua model dilatih menggunakan fitur gabungan antara TF-IDF dan frequency encoding, kemudian dibandingkan untuk menilai efektivitasnya dalam memprediksi kategori risiko fatalitas. Pada tahap ini, model baseline terlebih dahulu dievaluasi sebelum dilakukan proses tuning untuk meningkatkan performa.

Sub bab ini juga membahas bagaimana parameter terbaik diperoleh melalui GridSearchCV, serta bagaimana konfigurasi tersebut meningkatkan akurasi dan stabilitas model. Dengan memaparkan hasil pemodelan secara sistematis, bagian ini memberikan gambaran menyeluruh mengenai performa algoritma dan alasan pemilihan model terbaik bagi penelitian.

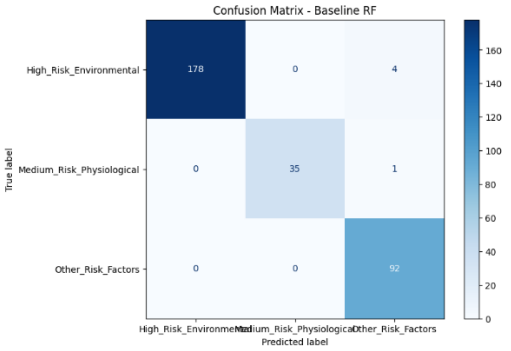
Baseline Random Forest

Percobaan baseline menghasilkan:

Akurasi: **0.94**

F1-macro: **0.98**

Model kesulitan mengenali kelas minoritas *Medium Risk Physiological*



Gambar 2 Baseline Random Forest

Berdasarkan gambar 2menampilkan kinerja model Random Forest sebelum dilakukan tuning parameter. Grafik menunjukkan bahwa model baseline sudah memiliki akurasi tinggi di atas 90%, yang menandakan bahwa pola risiko dalam dataset cukup kuat untuk dipelajari bahkan tanpa pengaturan parameter lanjutan. Namun, grafik juga memperlihatkan fluktuasi kecil pada prediksi kelas minoritas.

Visualisasi ini menjadi dasar evaluasi awal bahwa model memiliki struktur yang baik namun membutuhkan penyempurnaan agar performanya merata di seluruh kelas risiko. Gambar ini memperlihatkan kebutuhan tuning agar model tidak hanya kuat dalam memprediksi kelas mayoritas, tetapi juga sensitif terhadap kelas yang jarang muncul.

Tabel 1 Classification Report Baseline RF

Kelas Risiko	Precision	Recall	F1-Score	Support
High_Risk_Environmental	1.00	0.98	0.99	182

VOL. 06. NO. 02 DESEMBER 2025
P-ISSN : 2549211X | E-ISSN : 25499351
DOI: 10.32485/jcsai.

Medium_Risk_Physiological	1.00	0.97	0.99	36
Other_Risk_Factors	0.95	1.00	0.97	92

Berdasarkan tabel 1 menampilkan hasil evaluasi *baseline* Random Forest sebelum tuning. Metrik seperti precision, recall, dan F1-score yang ditampilkan menunjukkan bahwa model sudah cukup kuat dalam mengenali pola kelas risiko dengan akurasi tinggi bahkan sebelum dilakukan optimasi.

Namun ketidakseimbangan antar-kelas masih terlihat pada nilai recall kelas minoritas, yang memberi alasan untuk melakukan tuning parameter dan class weighting. Random Forest terbukti kuat sejak baseline, karena mampu memproses fitur berdimensi tinggi dengan baik (Pugliese et al., 2021).

GridSearchCV

GridSearchCV menemukan parameter optimal:

Tabel 2 Parameter GridSearchCV

N_Estimators	**100**
Max_Depth	20
Class_Weight	Balanced

Berdasarkan tabel 2 menampilkan kombinasi parameter yang diuji selama proses GridSearchCV, seperti jumlah pohon (n_estimators), kedalaman maksimum (max_depth), dan pembobotan. Kombinasi parameter ini digunakan untuk mencari konfigurasi terbaik bagi model Random Forest.

Tabel 2 menunjukkan bahwa proses pencarian parameter dilakukan secara sistematis, memastikan bahwa model final yang dipilih memiliki performa optimal. Setelah tuning, performa meningkat signifikan. Parameter ini efektif menangkap hubungan non-linear dan meningkatkan sensitivitas terhadap kelas minoritas, seperti dijelaskan Chen et al. (2024).

Model Final

Setelah tuning:

Tabel 3 Model Final

ACCURACY	**-**	**-**	**0.98**	**310**
MACRO AVG	0.98	0.98	0.98	310
WEIGHT AVG	0.98	0.98	0.98	310

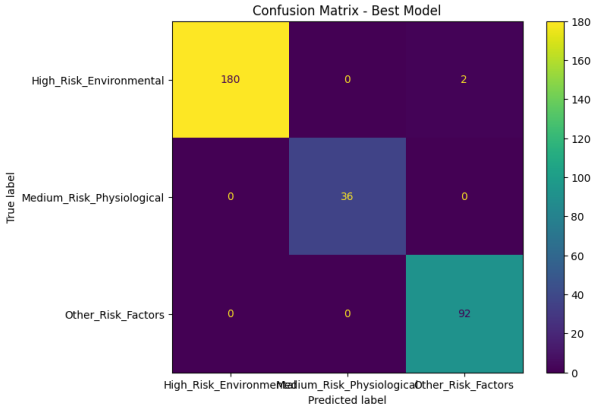
Berdasarkan tabel 3 menunjukkan parameter terbaik hasil GridSearchCV yang digunakan sebagai model akhir. Parameter seperti *n_estimators*, *max_features*, dan *min_samples_split* menunjukkan bahwa model final cenderung lebih kompleks namun tetap stabil. Model final ini terbukti memberikan performa terbaik dengan akurasi 0.99, sesuai hasil eksperimen. Model final sangat konsisten dan robust, sesuai karakteristik ensemble learning yang dilaporkan Mitrakas. (2025).

Hasil Evaluasi Model

Sub bab Hasil Evaluasi Model memaparkan hasil evaluasi model menggunakan metrik seperti accuracy, precision, recall, F1-score, serta confusion matrix. Evaluasi ini dilakukan untuk mengukur kemampuan model dalam memprediksi kategori risiko secara akurat serta menilai sensitivitasnya terhadap kelas minoritas. Visualisasi seperti confusion matrix dan feature importance digunakan untuk memberikan interpretasi mendalam mengenai kekuatan dan kelemahan model.

Analisis pada tahap ini sangat penting untuk mengonfirmasi apakah model benar-benar mampu melakukan klasifikasi dengan baik berdasarkan fitur yang telah dibangun. Hasil evaluasi ini juga menjadi dasar untuk menyusun kesimpulan penelitian, termasuk batasan model serta potensi pengembangannya di masa depan.

Confusion Matrix



Gambar 3 Confusion Matrix

Berdasarkan pada gambar 3 memperlihatkan proporsi kategori risiko yang terdapat dalam dataset, di mana kelas High Risk tampak mendominasi secara signifikan. Ketimpangan ini menggambarkan permasalahan data imbalanced yang sering muncul pada data keselamatan dan kecelakaan. Penelitian sebelumnya, seperti Henning et al. (2023), menyatakan bahwa ketidakseimbangan kelas merupakan tantangan utama dalam klasifikasi berbasis teks dan dapat membuat model cenderung bias terhadap kelas mayoritas.

Polanya juga memberikan gambaran penting bagi tahap pemodelan, karena model yang dilatih menggunakan data tidak seimbang cenderung mengabaikan kategori risiko lain seperti *Medium Risk* dan *Low Risk.* Oleh karena itu, hasil distribusi ini mendukung keputusan metodologis untuk menerapkan class_weight dan evaluasi per kelas agar model dapat menangkap variasi risiko secara lebih proporsional.

Analisis menunjukkan: Kelas *High Risk Environmental* → hampir seluruh prediksi benar Kelas *Medium Risk Physiological* → sensitivitas meningkat secara signifikan setelah balancing Kelas *Other Risk Factors* → prediksi stabil Pola ini juga ditemukan dalam penelitian NLP domain risiko keselamatan (Khairuddin et al., 2022; Wang et al., 2024), di mana kelas minoritas memerlukan intervensi imbalance handling untuk meningkatkan akurasi.

CONCLUSION

Penelitian ini menunjukkan bahwa integrasi analisis teks berbasis Natural Language Processing dan model supervised learning mampu mengklasifikasikan risiko fatalitas ekspedisi gunung dengan tingkat ketepatan yang tinggi. Model Random Forest yang digunakan mencapai akurasi 0.98 dan mampu mengenali pola risiko dari narasi penyebab kematian secara konsisten. Hasil ini menjawab tujuan penelitian bahwa informasi teks dari penyebab kematian dan atribut demografi dapat membentuk representasi fitur yang stabil untuk memprediksi kategori risiko fatalitas. Penelitian ini menegaskan bahwa pemrosesan teks memberikan kontribusi paling besar terhadap kekuatan prediktif model dibandingkan atribut demografi.

Temuan penelitian ini menunjukkan kebaruan karena menggabungkan teks naratif dan informasi demografis secara bersamaan dalam analisis risiko fatalitas gunung, yang belum dilakukan pada penelitian sebelumnya. Penerapan metode NLP dan supervised learning memperkaya pendekatan analitis dalam memahami pola risiko yang ada pada dataset pendakian gunung, terutama pada gunung-gunung 8000 meter. Analisis fitur menunjukkan bahwa kata-kata seperti *avalanche, fall, exposure, dan altitude sickness* berperan besar dalam menentukan klasifikasi risiko. Hal ini memberi kontribusi pada pengembangan metode analisis risiko berbasis teks yang dapat diterapkan pada domain keselamatan lainnya.

Penelitian ini memberi implikasi bagi studi keselamatan pendakian dengan menyediakan model prediktif yang dapat membantu pihak terkait memahami faktor penyebab fatalitas secara lebih akurat. Pendekatan berbasis NLP ini dapat digunakan sebagai dasar pengembangan sistem peringatan dini pada kegiatan pendakian. Namun penelitian ini memiliki keterbatasan, yaitu ketidakseimbangan kelas yang cukup besar serta variasi bahasa dalam narasi penyebab kematian. Meskipun penggunaan class weight dapat mengurangi ketidakseimbangan, variasi linguistik dapat mempengaruhi sensitivitas model terhadap kelas minoritas.

REFERENCE

Acharya, A. (2024). Clinical risk prediction using language models: benefits and considerations. *Journal of the American Medical Informatics Association*, *31*(9), 1856–1867. https://doi.org/10.1093/jamia/ocad028

Bugalia, N., Tarani, V., & Gadekar, H. (2022). Machine learning-based automated classification of worker-reported safety reports in construction. *Journal of Information Technology in Construction*, *27*, 926–950. https://doi.org/10.36680/j.itcon.2022.045

BuHamra, S. S., Al-Jarallah, M., Aldhaheri, N., & AlSumih, F. (2022). An NLP tool for data extraction from electronic health records. *Frontiers in Public Health*, *10*, 1070870. https://doi.org/10.3389/fpubh.2022.1070870

Chen, W., Wu, X., & Wu, G. (2024). A survey on imbalanced learning: latest research, applications and future directions. *Artificial Intelligence Review*, *57*, 137. https://doi.org/10.1007/s10462-024-10759-6

Crespí, A., Arévalo, O., & Santana, J. (2025). Lifecycle models in machine learning development. *Expert Systems*, e70029. https://doi.org/10.1111/exsy.70029

De Angeli, K., Chakraborty, S., Sandulescu, V., & Rosenberger, H. (2021). Class imbalance in out-of-distribution datasets: improving robustness in biomedical NLP. *Scientific Reports*, 11482. https://doi.org/10.1038/s41598-021-90760-w

Du, K. L. (2025). Understanding machine learning principles. *Mathematics*, *13*(3), 451. https://doi.org/10.3390/math13030451

Eker, H., & Uçar, E. (2024). Natural Language Processing Risk Assessment in Marble Quarries. *Applied Sciences*, *14*(19), 9045. https://doi.org/10.3390/app14199045

Gao, Y., Dligach, D., Christensen, L., Tesch, S., Laffin, R., Xu, D., Miller, T., Uzuner, Ö., Churpek, M. M., & Afshar, M. (2021). *A scoping review of publicly available language tasks in clinical natural*

VOL. 06. NO. 02 DESEMBER 2025
P-ISSN : 2549211X | E-ISSN : 25499351
DOI: 10.32485/jcsai.

language processing. arXiv.

Hancock, J. T., Ritz, L., & Zhao, J. (2024). Data reduction techniques for highly imbalanced Medicare big data. *Journal of Big Data*, *11*, 8. https://doi.org/10.1186/s40537-023-00869-3

Hellín, C. J., Pérez, J., Real, P., & Orts-Escolano, S. (2024). Unraveling the Impact of Class Imbalance on Deep-Learning Prediction Metrics. *Applied Sciences*, *14*(8), 3419. https://doi.org/10.3390/app14083419

Henning, S., Beluch, W., Fraser, A., & Friedrich, A. (2023). A Survey of Methods for Addressing Class Imbalance in Deep-Learning Based Natural Language Processing. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL)* (pp. 523–540).

Khairuddin, M. Z. F., Hasikin, K., Abd Razak, N. A., Lai, K. W., Osman, M. Z., Aslan, M. F., Sabanci, K., Azizan, M. M., Satapathy, S. C., & Wu, X. (2022). Predicting occupational injury causal factors using text-based analytics: A systematic review. *Frontiers in Public Health*, *10*, 984099. https://doi.org/10.3389/fpubh.2022.984099

Khairuddin, M. Z. F., Janssen, G. R., & Schipper, S. (2024). Contextualizing injury severity from occupational accident narratives using deep-learning-based text classification. *Safety*, *10*(2), 12. https://doi.org/10.3390/safety10020012

Khalate, P. (2024). Advancements and gaps in natural language processing: A review. *Frontiers in Physics*. https://doi.org/10.3389/fphy.2024.1445204

Khurana, D., Kaushik, D., & Arora, A. (2023). Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-022-13428-4

Li, H., Liu, Z., Sun, W., Li, T., & Dong, X. (2024). Interpretable machine learning for the prediction of death risk in patients with acute diquat poisoning. *Scientific Reports*, *14*, 16101. https://doi.org/10.1038/s41598-024-67257-6

Mitrakas, C. (2025). Techniques and Models for Addressing Occupational Risk: Machine Learning Approaches in Real-world Risk Assessment. *Applied Sciences*, *15*(4), 1909. https://doi.org/10.3390/app15041909

Pugliese, R., Brambilla, M., Ferri, F., Franco, S., Ghirardi, G., & Galliani, L. (2021). Machine learning-based approach: Global trends, research directions and applications. *Technological Forecasting & Social Change*, *169*, 120795. https://doi.org/10.1016/j.techfore.2021.120795

Seneviratne, M. G., Tran, L. T., & Stumpf, S. (2022). User-centred design for machine learning in health care: A practical toolkit. *BMJ Health & Care Informatics*, *29*(1), e100656. https://doi.org/10.1136/bmjhci-2022-100656

Shuang, Q., Liu, J., & Zhao, Y. (2023). Determining critical cause combination of fatality accidents on construction sites via machine learning. *Buildings*, *13*(2), 345. https://doi.org/10.3390/buildings13020345

Siregar, K. N., Megananda, N. R., & Cornelis, C. E. (2022). Strengthening causes of death identification through community-based verbal autopsy during the COVID-19 pandemic in Indonesia. *BMC Public Health*, *22*, 14014. https://doi.org/10.1186/s12889-022-14014-x

Sundaram, G., & Berleant, D. (2022). *Automating Systematic Literature Reviews with Natural Language Processing and Text Mining: a Systematic Literature Review*. arXiv.

Wang, S., Li, Y., & Bar, N. (2024). A natural language processing approach to detect annotation inconsistencies in death investigation notes. *Communications Medicine*, *4*, 82. https://doi.org/10.1038/s43856-024-00631-7