Eduvest – Journal of Universal Studies
Volume 6 Number 5, May, 2026
p- ISSN 2775-3735- e-ISSN 2775-3727

ICD Coding Automation Model of Retinal Detachment Case Using Support Vector Machine and Random Forest

Dyah Kurniawati[1]*, Mieke Nurmalasari[2], Hosizah Markam[3], Dewi Krismawati[4]
Universitas Esa Unggul, Indonesia[1,2,3]
Badan Pusat Statistik, Indonesia[4]
Email: dyah.kurniawati99@gmail.com[1]*, mieke@esaunggul.ac.id[2], hozisah@esaunggul.ac.id[3], dewikrisma@bps.go.id[4]

Keywords:
Clinical Coding Automation; ICD; Machine learning (ML); Natural Language Processing (NLP); Retinal Detachment.

ABSTRACT

Health Information Management (HIM) professionals are responsible for maintaining the consistency of ICD-based clinical codes for the health reimbursement and health analytics through the review of medical documentation. The complexity of coding rules and clinical pathways increases the risk of miscoding, but the implementation of Electronic Medical Record (EMR) opens opportunities for the development of automation of ICD coding. This study aims to build an ICD code automation model for retinal detachment cases from eye referral hospital using artificial intelligence through clinical text classification with Natural Language Processing (NLP) and Machine Learning (ML) algorithms. The dataset includes disease resumes, physical examinations, diagnoses, medical procedures, surgical records, and therapies from 300 inpatients. Text preprocessing uses the NLTK library through sentence splitting, abbreviation expansion, case folding, stop word removal, and tokenization functions. Data preparation involves splitting data (80:20 ratio), feature extraction with TF-IDF Vectorizer, and 5-fold cross validation. Classification modeling uses Support Vector Machine (SVM) and Random Forest (RF). Evaluation of the SVM model showed an accuracy of 0.82 (precision 0.84; recall 0.82; F1-Score 0.82), while the RF model achieved an accuracy of 0.87 (precision 0.88; recall 0.87; F1-Score 0.87). Based on confusion metrics, the correct predictions for classes H33.0, H33.2, and H33.4 on SVM are 79, 87, and 80, while RF reaches 83, 88, and 91. The development of this automation requires HIM professional's role in ensuring the quality of EMR data and accuracy of ICD code as well as intensive model training to handle the complexity of clinical data.

INTRODUCTION

The development of digital technology in society encourages the transformation of the digitalization of health services. Healthcare digital transformation will focus on health data development, healthcare application development, and platform-based sustainable healthcare technology ecosystem enhancement. The goal of digital transformation is to improve the quality of data and its policies and improve the efficiency of health services.

To support the digital transformation of the health sector, one of them is by organizing electronic medical records with the principles of data and information security and confidentiality. Based on the Regulation of the Minister of Health Number 24 of 2022 concerning Medical Records, it is stated that all health service facilities must hold electronic medical records no later than December 31, 2023.

Electronic medical records are medical records made using electronic systems that aim to improve the quality of health services, provide legal certainty in the administration and management of medical records, ensure the security, confidentiality, integrity and availability of medical record data, and realize the implementation and management of digital-based and integrated medical records (Kementerian Kesehatan, 2022).

Along with the widespread use of electronic medical records, large amounts of data can be recorded in the electronic medical record system owned by health facilities. The complexity, heterogeneity, rapid growth of data, and the size of the data stored in electronic medical record systems require big data technology to be able to analyze them (Cyganek et al., 2016).

Big data has characteristics volume, velocity, variety, and Veracity. Magnitude Volume on Big Data requires large-scale and efficient storage, velocity or data speed creates the need for algorithms that are adaptive and can work in a real-time Adjusting to data speed, variety or data diversity i.e. structured data and unstructured data require advanced processing, interoperability, and an integrated platform and Veracity or data correctness emphasizes on data quality assurance (Surur et al., 2025).

The digitization of health service data has an impact on the responsibilities and work of the health information management profession which focuses more on technical roles in terms of data collection, data storage, and use of health service data (Stanfill & Marc, 2019).

Health information management professions that are affected by artificial intelligence technology in practical terms include clinical coding automation and clinical information capture automation, healthcare data management and data regulation, patient privacy and confidentiality, as well as education and training of health information management professionals (Stanfill & Marc, 2019).

A health information management profession is able to establish clinical classifications, codification of diseases, and other health problems, as well as clinical procedures appropriately according to the enforced classifications, which are used for disease statistics and healthcare facility financing systems (Kementerian Kesehatan, 2020).

Clinical coding is the process of encoding a patient's medical data using a specific code standard, such as ICD-10 (International Statistical Classification of Disease and Related Health Problems) for medical diagnosis and procedures. These codes will be used to group patient medical data based on specific types of diseases or medical conditions, which can then be used for reporting and analysis purposes (Anjani, Sylvia ; Tomy Abiyasa, 2023).

Clinical coding is an intense and time-consuming job and requires specific abilities and knowledge. In the coding process, when reading medical records, there can be misunderstandings of the information used in coding. In the future, it can be predicted that clinical coding work will shift towards the process of validating the code generated by the system. This also triggers the need to improve the ability of the health information management profession in terms of critical thinking and coding skills and how to work Clinical Documentation Improvement (Ball, 2014).

Therefore, the role of HIM Professional in the implementation of clinical coding automation is needed to ensure that the code generated by the system is in accordance with the results of the patient's examination listed in the medical record.

Clinical coding automation refers to the concept that clinical coding can be automated by computer systems through the application of artificial intelligence (AI) techniques, such as

Natural Language Processing (NLP) then Machine learning (ML). In recent years, artificial intelligence has been seen as a promising approach in the transformation of healthcare systems by carrying out increasingly complex and large data processing using techniques Machine Learning and Natural Language Processing. Clinical coding automation is a potential application of artificial intelligence to support the administration and management processes of medical records in hospitals and medical research facilities (Dong et al., 2022).

Algorithm Natural Language Processing It is used to process text for the purpose of understanding the meaning of a text. In clinical text analysis, text comprehension depends on the ability to recognize differences between diagnoses associated with comparative diagnoses, disease history, family history, and negation sentences (HIMSS, 2017)

Through the Machine Learning, researchers and practitioners in their industries can conduct big data analysis to improve the user experience. At the practitioner level, strength Machine Learning with Big Data Relying on its ability to extract meaning, identify hidden patterns, and make accurate predictions from large data sets that are too complex to be solved by traditional methods, in this case it supports data-driven decision-making with greater precision and can be applied across a wide range of industries (Surur et al., 2025).

Classification is a process that has the purpose of ascertaining an object into a predetermined group or category. Classification is the process of forming a model that classifies an object according to its attributes. Algorithm Machine Learning allows the system to perform tasks by observing data. Machine learning Able to make hypotheses based on observational data and theories owned by users. Capabilities Machine Learning depending on the amount of data set that has a label (Kaur, Rajvir; Anupama Ginige, Jeewani; Obst, 2023).

Application Machine Learning In this study, it acts as a computational method that is able to study patterns from medical text data to produce automatic classification of ICD codes. In contrast to the rule-based system, Machine Learning It has the ability to tailor models based on patterns found in training data, so it can handle complex and dynamic variations in clinical language. Thus the use of Machine Learning, is expected to provide added value in the form of improved classification accuracy, time efficiency, and adaptive ability to recognize medical contexts that cannot always be addressed by conventional rule-based systems (Kaur, 2018).

AI systems not only treat patient data in a balanced and fair manner, but also do not affect similar groups of people in different ways. In addition, to eliminate bias in research and clinical practice, inclusivity is a concept to be incorporated into the design of AI systems(RI, 2021). Therefore in the development of artificial intelligence models, retinal detachment cases taken from electronic medical records, no distinction was made based on the patient's background.

Retinal ablation or retinal detachment is the separation of the Neu sensory layer of the retina from the epithelial layer of the pigment retina. Based on the process of occurrence, retinal ablation can be classified into rhegmatogenous retinal ablation, tractional retinal ablation, exudative retinal ablation, and tractional-rhegmatogenous retinal ablation. The diagnosis is established based on an anamnesis and ophthalmological examination with a supporting examination of the fundus photo and ultrasound of the eye (Persatuan Dokter Spesialis Mata Indonesia (PERDAMI), 2018).

The operation actions carried out are in the form of pneumatic retinopexy, Installation Scleral buckle, and vitrectomy. The diagnosis of retinal detachment can be clinically enforced, but treatment should be carried out by a retinal consultant ophthalmologist and in a healthcare facility equipped with standard retinal surgical equipment (Persatuan Dokter Spesialis Mata Indonesia (PERDAMI), 2018).

Use of artificial intelligence for detection, prediction and classification Retinal detachment Several previous studies have been conducted. In general, the potential of artificial intelligence in detection and classification Retinal detachment can result in studies Outcome better and lower maintenance costs. Future clinical trial research will require data from different patient groups and public datasets to ensure that the artificial intelligence models developed are generally acceptable (Zaky et al., 2023).

The urgency of this research is underscored by several factors. At the eye referral hospital in this study, retinal detachment cases ranked among the top 10 inpatient diagnoses in 2024, representing a significant documentation burden. Manual clinical coding is time-consuming and error-prone, with miscoding impacting hospital reimbursement, epidemiological statistics, and research data quality. The Indonesian EMR mandate (effective December 31, 2023) makes large-scale clinical text data available for AI applications, creating opportunity for automation development. The novelty of this research lies in: (1) using real-world EMR data (not public datasets or images) from an Indonesian eye referral hospital, (2) focusing specifically on ICD-10 H33 block (retinal detachment: rhegmatogenous H33.0, serous H33.2, tractional H33.4) as the classification target, (3) addressing class imbalance with SMOTE + undersampling, (4) comprehensive preprocessing including abbreviation expansion using hospital-standardized abbreviations, and (5) comparative evaluation of SVM and RF with 5-fold cross-validation and confusion matrix analysis.

Based on a case study at the eye referral hospital, retinal detachment cases are included in the category of the largest diagnosis in the top 10 inpatient diagnosis reports in 2024. Comprehensive electronic medical records have been used at the hospital starting January 2024.

This research focuses on the development and testing of support vector machine and random forest algorithm to automatically classify clinical texts into ICD codes based on medical resumes and surgical reports from electronic medical records. The data set used has gone through the process of classification and codification, diagnosis and action by the hospital coder team consisting of doctors, internal verifiers, and medical recorders. The results of the research are expected to be the basis for further development towards clinical coding automation systems in the future. This study aims to construct an automation model for clinical coding using retinal detachment cases from eye referral hopital through a series of systematic methodological stages.

METHODS

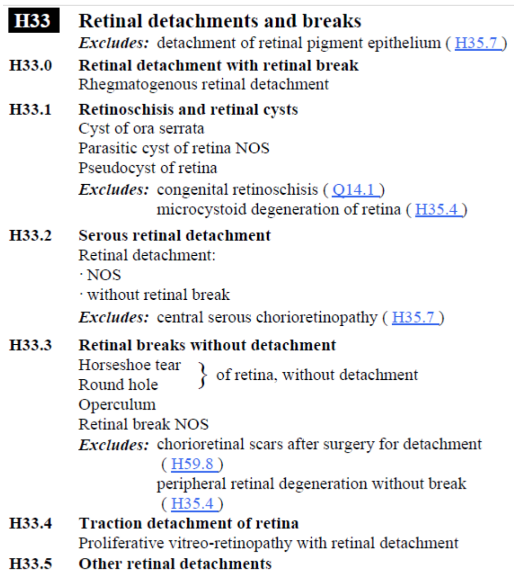
This research and development research aims to construct an automatic clinical text classification model based on machine learning algorithms for the codification of retinal detachment cases into the ICD-10 system using an anonymized secondary dataset from eye referral hospital for the 2024 period. The methodology applied includes data collection through observation of medical resumes and surgical reports, followed by a comprehensive preprocessing phase to ensure data integrity before the modeling process. To overcome the phenomenon of class imbalance in the dataset, this study integrates a cost-sensitive learning approach through the Synthetic Minority Oversampling Technique (SMOTE) and under sampling techniques to achieve class distribution equilibrium. Data preparation including splitting training data and testing data, feature extraction, and cross validation. Computational performance evaluation was carried out on the Support Vector Machine (SVM) and Random Forest with assessment metrics that include accuracy, precision, recall, and f-score to measure the efficacy of the algorithm in relevant zing narrative clinical data into precise ICD code.

Figure 1. Research Method

RESULTS AND DISCUSSION

Data Set Extraction

Data extract through electronic medical record by defining H33 code, the type of inpatient care, primary case, and the period of case from January to December 2024. The result are 157 data of H33.0, 162 data of H33.2, and 49 data of H33.4. Imbalance data set was addressed using cost-sensitive learning through Synthetic Minority Oversampling Technique (SMOTE) dan undersampling to achieve equilibrium of class distribution, thus achieving the target of 100 data set per class.



H33 **Retinal detachments and breaks**
Excludes: detachment of retinal pigment epithelium (H35.7)

H33.0 **Retinal detachment with retinal break**
Rhegmatogenous retinal detachment

H33.1 **Retinoschisis and retinal cysts**
Cyst of ora serrata
Parasitic cyst of retina NOS
Pseudocyst of retina
Excludes: congenital retinoschisis (Q14.1)
microcystoid degeneration of retina (H35.4)

H33.2 **Serous retinal detachment**
Retinal detachment:
· NOS
· without retinal break
Excludes: central serous chorioretinopathy (H35.7)

H33.3 **Retinal breaks without detachment**
Horseshoe tear } of retina, without detachment
Round hole } of retina, without detachment
Operculum
Retinal break NOS
Excludes: chorioretinal scars after surgery for detachment (H59.8)
peripheral retinal degeneration without break (H35.4)

H33.4 **Traction detachment of retina**
Proliferative vitreo-retinopathy with retinal detachment

H33.5 **Other retinal detachments**

Figure 2. H33 Block of ICD-10

Retinal tear is a lesion that is a factor in the occurrence of retinal detachment. Laser retinopexy recommended as a measure to prevent complications from occurring retinal detachment. Patients with Retinal Tear undergoing laser retinopexy in the laser procedure room (Ahsan, 2019). Retinoschisis is a condition of rising layers in the inner layer of the peripheral retina. Retinoschisis has the potential to develop into retinal detachment. In many cases, retinoschisis occurs in the peripheral retina and is asymptomatic, so clinical monitoring is only necessary. (Ness et al., 2022). Vitrectomy with local anesthesia in treating retinal detachment is performed due to its effectiveness with good results, fewer postoperative side effects, and shorter duration of hospitalization so that it is more cost-effective (Dharma et al., 2020).

Data extraction from electronic medical records in this study only obtained data in the categories of H33.0 rhegmatogenous retinal detachment, H33.2 serous retinal detachment, and H33.4 tractional retinal detachment whose treatment requires hospitalization after surgery. Meanwhile, in H33.3 retinal tear, operative procedures were performed in the laser procedure room with an outpatient scheme, and no cases of H33.1 retinoschisis and retinal cysts and H33.5 other retinal detachments were found from data withdrawal on electronic medical records.

Preprocessing Text

Preprocessing text were done using NLP through NLTK library, they are sentence splitting, abbreviation extension, case folding, stop word removal, and tokenization to change the initial data format into a format that is ready to analyzed by machine learning. The process was done using the python programming language in the Visual Code Studio software.

Rachman (2020) in his book states that in doing Preprocessing text There is no need to use all the stages, as they are adapted to the needs of the system (Rachman, 2020). The Natural Language Toolkit Library (NLTK) is one of the most popular libraries in the python programming language for Natural Language Processing (Kedia & Rasu, 2020).

According to the results of Kaur's (2018) research, each clinical document is broken down into individual sentences using Natural Language Toolkit (NLTK) (Kaur, 2018). Sentence splitting is used in this study to break down a paragraph of a clinical text into sentences, i.e. in clinical texts that consist of more than one paragraph that is commonly found in a summary of disease history and physical examination.

According to the results of Kaur's (2018) research, doctors usually prefer to use short names or abbreviations for medical terms that are often difficult to understand without an understanding of medical terms. What's more, in natural language, each word has a variety of meanings that cannot be understood without knowing the structure or phrase in the sentence. Therefore, the dictionary of abbreviated words or abbreviation created based on abbreviations available in medical records (Kaur, 2018).

Abbreviation extension was used in this study to find out the medical abbreviations contained in the dataset. The list of medical abbreviations is made based on the list of official abbreviations available in the eye referral hospital.

Case folding is a technique used in the normalization of data by converting all the words in the corpus text into lowercase letters. This technique helps the system in summoning information (Kedia & Rasu, 2020). Case folding was used in this study to convert all letters in the data set into lowercase letters considering that in practice the writing of medical records by clinical personnel is usually done according to the applicable spelling, namely using uppercase letters at the beginning of sentences.

Stop word are adjectives such as a, in-, two-, from-, and others that often appear in the corpus of texts and do not have much meaningful meaning in context. This word is needed to complete sentences and according to the structure of the word. This word is filtered using NLP to reduce the vocabulary processed. None of stop word that are universally available, but the list of words stop word usually adapts to the language used, and is modified based on the case being worked on (Kedia & Rasu, 2020). The list of stop words in this study was modified to suit the Indonesian language according to the data set.

Tokenization or tokenization is a technique that NLTK libraries can use to separate documents into several units of words, alphabets, and even sentences (Kedia & Rasu, 2020). Tokenization was used in this study to break down long clinical tests on data sets sourced from medical resumes and surgical reports into a word token.

0	[pasien rencana rawat inap persiapan tindakan ...	[[pasien, rencana, rawat, inap, persiapan, tin...
1	[pasien rencana rawat inap persiapan tindakan ...	[[pasien, rencana, rawat, inap, persiapan, tin...
2	[2 minggu masuk rumah sakit mata kiri kilatan ...	[[2, minggu, masuk, rumah, sakit, mata, kiri, ...
3	[pasien rawat inap persiapan tindakan vitrekto...	[[pasien, rawat, inap, persiapan, tindakan, vi...
4	[pandangan mata kanan tertutup tirai 3 masuk r...	[[pandangan, mata, kanan, tertutup, tirai, 3, ...
5	[pasien rawat inap persiapan tindakan vitrekto...	[[pasien, rawat, inap, persiapan, tindakan, vi...
6	[pasien rencana rawat inap persiapan tindakan ...	[[pasien, rencana, rawat, inap, persiapan, tin...
7	[pasien rawat inap pemantauan pasca tindakan v...	[[pasien, rawat, inap, pemantauan, pasca, tind...
8	[keluhan pandangan mata kiri gelap 17 jam masu...	[[keluhan, pandangan, mata, kiri, gelap, 17, j...
9	[pandangan mata kiri buram mendadak 10 masuk r...	[[pandangan, mata, kiri, buram, mendadak, 10, ...
10	[pasien rencana rawat inap persiapan tindakan ...	[[pasien, rencana, rawat, inap, persiapan, tin...
11	[pasien rencana rawat inap persiapan tindakan ...	[[pasien, rencana, rawat, inap, persiapan, tin...
12	[mata kanan buram tertutup tirai 3 masuk rumah...	[[mata, kanan, buram, tertutup, tirai, 3, masu...
13	[pasien rujukan rujukan rumah sakit islam jaka...	[[pasien, rujukan, rujukan, rumah, sakit, isla...
14	[3 jam masuk rumah sakit pandangan mata kanan ...	[[3, jam, masuk, rumah, sakit, pandangan, mata...
15	[1 masuk rumah sakit pandangan mata kanan bura...	[[1, masuk, rumah, sakit, pandangan, mata, kan...

Figure 3. Result of Preprocessing Text

Data Preparation

Data preparation is carried out to convert the text data resulting from preprocessing text into numerical data forms that are ready to be analyzed by machine learning models by using the python programming language in the Visual Code Studio software. The steps are splitting training data and testing data, feature extraction, and cross validation.

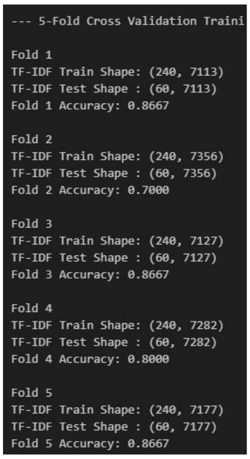
All experiments in Machine Learning Implement an 80:20 ratio with an average of 80% training and 20% testing (Kaur, 2018). Data preparation begins with the process of splitting training data and testing data with an 80:20 ratio thus resulting in 240 training data and 60 testing data.

The majority of study about automation clinical code using *feature extraction TF-IDF* Vectorization to represent feature on high dimensional data. Feature extraction is a process used to assign a weight value to words in a dataset to convert the form of the text data in the data set into numerical data. In this study, TF-IDF Vectorizer is used in feature extraction thus resulting in 7177 feature data.

```
--- TF-IDF Vectorization Status ---
Shape of the training feature matrix: (240, 7177)
Shape of the testing feature matrix: (60, 7177)
```

Figure 4. Result of Data Splitting and Feature Extraction

This study uses 5-Fold Cross Validation, which is a technique in stratified cross validation that is used for data validation by ensuring that each data has the same opportunity as training data. The results of the 5-fold cross validation obtained five feature values in the training data in order 7113, 7356, 7127, 7282, and 7177 and produced sequential variations in accuracy at each fold, namely 0.8667, 0.70, 0.8667, 0.80, and 0.8667.



```
--- 5-Fold Cross Validation Traini

Fold 1
TF-IDF Train Shape: (240, 7113)
TF-IDF Test Shape : (60, 7113)
Fold 1 Accuracy: 0.8667

Fold 2
TF-IDF Train Shape: (240, 7356)
TF-IDF Test Shape : (60, 7356)
Fold 2 Accuracy: 0.7000

Fold 3
TF-IDF Train Shape: (240, 7127)
TF-IDF Test Shape : (60, 7127)
Fold 3 Accuracy: 0.8667

Fold 4
TF-IDF Train Shape: (240, 7282)
TF-IDF Test Shape : (60, 7282)
Fold 4 Accuracy: 0.8000

Fold 5
TF-IDF Train Shape: (240, 7177)
TF-IDF Test Shape : (60, 7177)
Fold 5 Accuracy: 0.8667
```

Figure 5. Result of 5-Fold Cross Validation

Modeling with Support Vector Machine and Random Forest

Machine learning modeling is used to analyze data, recognize patterns, and make predictions on large and complex data sets. In this study, modeling was carried out with two classification models, namely support vector machine and random forest. Modelling process were done using python programming language in the Visual Code Studio software.

Support Vector Machine

Basic concepts Support Vector Machine is looking for a separator field (hyperplane) which is formed by data points (vectors) to maximize the margin or boundary between two classes so that a clear separation is obtained. Models Support Vector Machine commonly used for health data to build robust, accurate, predictive models capable of interpreting the features of clinical variables.

The results of the visualization in this study show the decision limit (hyperplane) of the SVM model in the retinal detachment data. There is an overlap between classes in the two-dimensional feature space as a result of dimension reduction using Principal Component Analysis (PCA). This condition indicates that the feature characteristics of the three classes have a fairly high level of similarity. Nonetheless, the SVM model is still able to form relatively clear decision boundaries, reflecting its ability to handle and separate classes on complex clinical datasets and have a high variety of features.

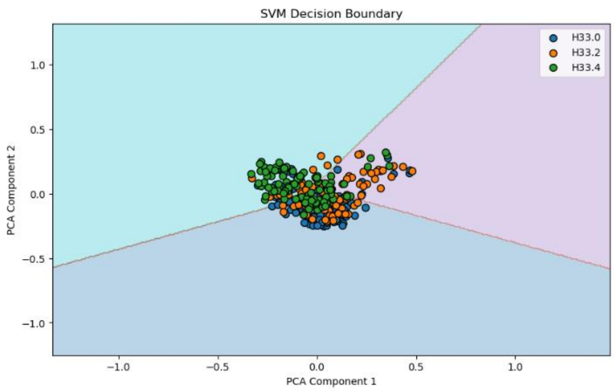


Figure 6. Hyperplane or Decision Boundary of SVM

The figure shows the visualization of the SVM decision boundary model on the dimension-reduced data using Principal Component Analysis (PCA) into two main components. Each dot represents the test data, while the colors indicate classes H33.0, H33.2, and H33.4, respectively.

Random Forest Model

The random forest model is a model that randomly selects features by forming many decision trees and selects classification decisions by a majority vote. Healthcare data is complex in nature so clinicians consider many factors in making decisions. Each clinician takes a different approach in determining the right decisions and treatments even in the same patient. This challenge has similarities with the concept Random Forest that combines many decision trees into a single model and determines the final outcome by majority vote. Individually, the predictions made by a single human or a single decision tree may not be accurate, but by combining many observations, the variance can be reduced and the prediction results reach a value Outcome what is expected (Anthony et al., 2020). In classification using a decision tree, the level of purity node depending on the value This is an index. In doing classification, the decision tree will continue to be separated based on features until it reaches purity in Leaf (Anthony et al., 2020).

In this study, the complexity of the data Retinal detachment that is used has the potential to be analyzed with a model Random Forest. This model is able to handle the variation and complexity of data by combining many decision trees so that the predictions produced are more stable and accurate than predictions from a single decision tree alone.

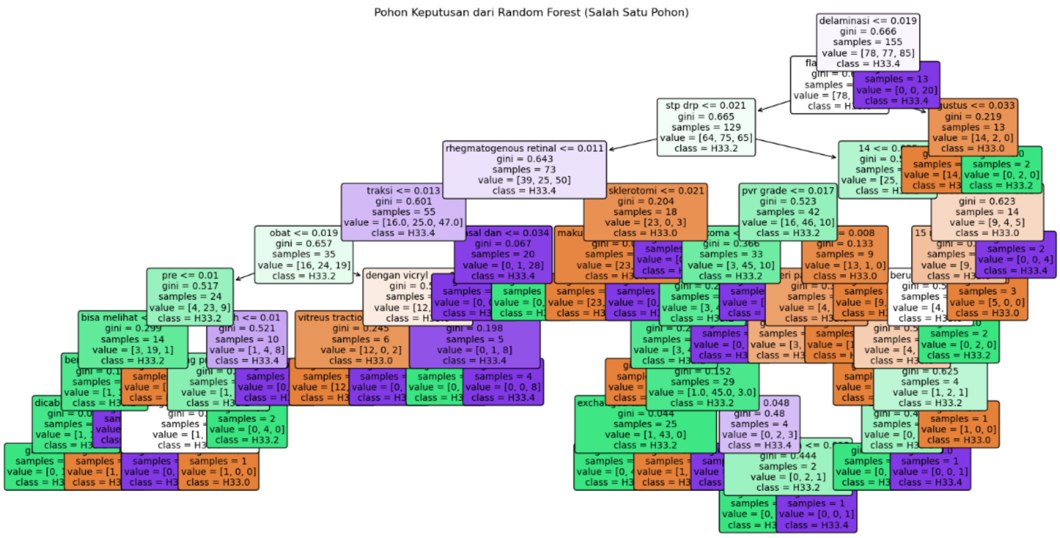


Figure 7. One of Decision Tree from Random Forest

Based on the visualization of the decision tree, the root node value in the early stages indicates a node that is still heterogeneous with a low purity level while the lowest leaf node value that tends to decrease is close to indicating a homogeneous node with a high purity level. Although some leaf nodes have been dominated by a single class, there are still nodes with mixed classes, so Random Forest combines many decision trees to produce a more stable final prediction through majority voting. Based on the visualization of one of the decision trees in the random forest model, the root node showed a Gini value of 0.66, with a sample of 155 data and a class distribution (value) [78, 77, 85]. This condition indicates that in the early stages the data is still heterogeneous with a relatively balanced class distribution and a fairly low level of purity. Meanwhile, the lowest leaf node shows a decreasing Gini value, 1 sample reaches the leaf node, with a class distribution (value) [0, 0, 1]. This condition indicates that the nodes are perfectly homogeneous and have a high level of purity.

Model Evaluation

The evaluation of the clinical coding automation model of retinal detachment cases was carried out to assess the performance of the model in classifying the data set of retinal detachment cases into ICD codes by calculating accuracy, precision, recall, and F1-score values and displaying a visualization of the model's prediction results with a confusion metrics diagram.

Evaluation of the Support Vector Machine Model

Research conducted by Padungweang, et al (2025), the SVM model with a total data set of 72,082 shows the value of Precision 0,89 , recall 0,84 , F1 Score 0.86 , and accuracy 0,90 (Chuabsombat & Padungweang, 2025). The research was carried out by Kaur and Ginige. (2018), SVM model with a total data set of 235 gives a Precision 0,89, Recall 0,55, F1 Score 0.65, and Accuracy 0,54 (Kaur, 2018).

In this study, SVM model showed an accuracy of 0.82. In the H33.0, H33.2, and H33.4 classes, respectively, produced a precision of 0.89, 0.68, and 0.96, followed by a recall of 0.79, 0.87, and 0.80 and an F1-Score of 0.84, 0.76, and 0.87. The performance in 3 different classes was calculated on average on the macro average resulting in a precision of 0.84, recall of 0.82, and F1-Score of 0.82.

--- Classification Report Support Vector Machine (5-Fold CV) ---

	precision	recall	f1-score	support
H33.0	0.89	0.79	0.84	100
H33.2	0.68	0.87	0.76	100
H33.4	0.96	0.80	0.87	100
accuracy			0.82	300
macro avg	0.84	0.82	0.82	300
weighted avg	0.84	0.82	0.82	300

Figure 8. Classification Report of SVM Model

Evaluation of the SVM model with a confusion metrics diagram showed that out of 300 data with 100 data in each class, the model was able to correctly classify H33.0 as many as 79 data, the model was able to correctly classify H33.2 class as many as 87 data, and the model was able to correctly classify H33.4 class as many as 80 data.

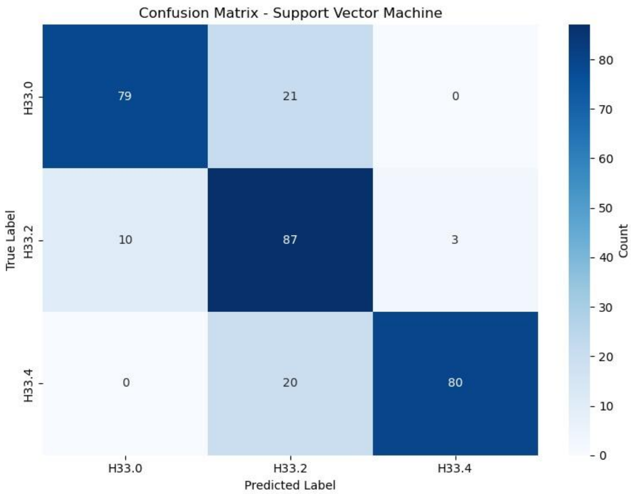


Figure 9. Confusion Metrics of SVM Model

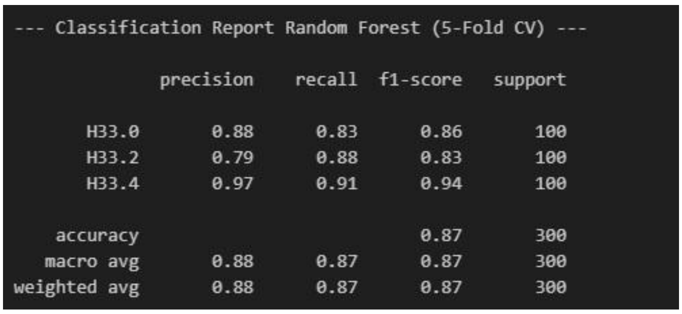
Compare to other similar studies, SVM model in this study showed a precision value of 0.05 lower than Padungweang, et al. (2025) and Kaur and Ginige (2018), a recall 0.02 lower than Padungweang, et al. (2025) but 0.30 higher than Kaur and Ginige (2018), F1-Score 0.04 lower than Padungweang, et al. (2025) but 0.17 higher than Kaur and Ginige (2018), and 0.08 lower accuracy than Padungweang, et al. (2025) but 0.28 higher than Kaur and Ginige (2018).

Based on the results of these tests, it can be concluded that the SVM model in this study produces good performance, although it is still lower than Padungweang et al. (2025) who use a much larger and higher dataset than Kaur and Ginige (2018) who use a smaller dataset. This suggests that the size of the dataset can affect the performance of SVM models in ICD coding automation.

Random Forest Model Evaluation

Research conducted by Padungweang, et al (2025), model Random Forest with a total data set of 72,082 showing the value of Precision 0,91 , recall 0,84 , F1 Score 0.87 , and accuracy 0,91 (Chuabsombat & Padungweang, 2025). The research was carried out by Kaur and Ginige. (2018), Models Random Forest with a total of 235 data sets providing a value Precision 0,66 , Recall 0,30, F1 Score 0.39, and Accuracy 0,29 (Kaur, 2018).

In this study, random forest model showed an accuracy of 0.87. In the H33.0, H33.2, and H33.4 classes, respectively, produced a precision of 0.88, 0.79, and 0.97, followed by a recall of 0.83, 0.88, and 0.91 and an F1-Score of 0.86, 0.83, and 0.94. The performance in 3 different classes was calculated on average on the macro average resulting in a precision of 0.88, recall of 0.87, and an F1-Score of 0.87.


--- Classification Report Random Forest (5-Fold CV) ---

	precision	recall	f1-score	support
H33.0	0.88	0.83	0.86	100
H33.2	0.79	0.88	0.83	100
H33.4	0.97	0.91	0.94	100
accuracy			0.87	300
macro avg	0.88	0.87	0.87	300
weighted avg	0.88	0.87	0.87	300

Figure 10. Classification Report of Random Forest Model

Evaluation of the random forest model with a confusion metrics diagram showed that from 300 data with 100 data in each class, the model was able to correctly classify H33.0 class as many as 83 data, the model was able to correctly classify H33.2 class as many as 88 data, and the model was able to classify H33.4 class as many as 91 data.

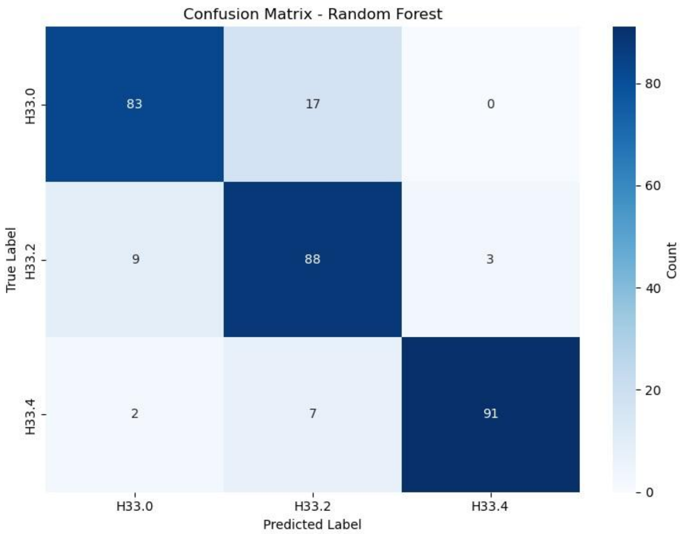


Figure 11. Confusion Metrics of Random Forest Model

Compare to other similar studies, random forest model in this study showed a precision value of 0.03 lower than Padungweang et al. (2025) and 0.22 higher than Kaur and Ginige (2018), a recall 0.03 lower than Padungweang, et al. (2025) but 0.57 higher than Kaur and Ginige (2018), The F1-Score is the same as Padungweang, et al. (2025) but 0.48 higher than Kaur and Ginige (2018), and 0.04 accuracy lower than Padungweang, et al. (2025) but 0.58 higher than Kaur and Ginige (2018).

Based on the results of these tests, it can be concluded that the random forest modeling in this study produced good performance, although slightly lower than Padungweang et al. (2025) who used a much larger and higher dataset than Kaur and Ginige (2018) who used a smaller dataset. This suggests that the size of the dataset can affect the performance of SVM models in clinical coding automation.

CONCLUSION

This study successfully implemented the ICD coding automation flow for retinal detachment cases (H33.0, H33.2, H33.4) through the integration of Natural Language Processing (NLP) techniques and machine learning algorithms on electronic medical record data from eye referral hospital. Through systematic stages that include granular data extraction, textual processing using NLTK libraries, and numerical transformation via TF-IDF Vectorizer, the developed model shows significant classification performance. The comparison results showed that the Random Forest algorithm outperformed the Support Vector Machine (SVM) with an accuracy of 0.87 and an F1-Score value of 0.87, where the confusion matrix visualization confirmed the model's efficacy in predicting the H33.4 class most optimally with 91 correct predictions. Thus, the use of ensemble learning has been shown to have superior capabilities in interpreting complex clinical narratives into accurate ICD codes compared to the single hyperplane approach in SVM. Future development of ICD coding automation requires HIM professional's role in ensuring the quality of EMR data and accuracy of ICD code as well as intensive model training to handle the complexity of clinical data.

BIBLIOGRAPHY

Ahsan, H. (2019). Karakteristik Laser Retinopexy pada Pasiendengan Tear Retina di Divisi Vitreoretina RS Cipto Mangunkusomo Periode Januari – Desember 2018. *Health Anf Medical Journal*, *1*(2), 47–52. https://doi.org/https://doi.org/10.33854/heme.v1i2.237

Anjani, Sylvia ; Tomy Abiyasa, M. (2023). *Disrupsi Digital dan Masa Depan Rekam Medis*. Selat Media Partners.

Anthony, L., Maimuna, C., Ordóñez, P., & Sebastian, J. (2020). *Leveraging Data Science for Global Health*. Springer Nature.

Ball, T. G. B. B. C. et al. (2014). *Computer-Assisted Coding Toolkit*. AHIMA Press.

Chuabsombat, K., & Padungweang, P. (2025). Automated ICD-9 and ICD-10 Coding with Machine Learning : A Real-World Study Using Electronic Medical Record Text from Udon Thani Cancer. *2025 11th International Conference on Computing and Artificial Intelligence (ICCAI)*, 794–801. https://doi.org/10.1109/ICCAI66501.2025.00123

Cyganek, B., Graña, M., & Krawczyk, B. (2016). A Survey of Big Data Issues in Electronic Health Record Analysis. *Applied Artificial Inteliigence International Journal*. https://doi.org/10.1080/08839514.2016.1193714

Dharma, A. G., Djatikusumo, A., Adriono, G. A., Yudantha, A. R., Hutapea, M. M., & Victor, A. A. (2020). Vitrektomi dengan Anestesi Lokal pada Ablasio Retina Rhegmatogen di Rumah Sakit Cipto Mangunkusumo. *Opthalmologica Indonesiana*, *46*(2), 131–136.

Dong, H., Falis, M., Whiteley, W., Alex, B., Matterson, J., & ... (2022). Automated clinical coding: what, why, and where we are? In *NPJ digital …*. nature.com.

HIMSS. (2017). *Demystifying Big Data and Machine Learning for Healthcare*. Taylor & Francis Group.

Kaur, Rajvir; Anupama Ginige, Jeewani; Obst, O. (2023). AI-Based ICD Coding and Classification Approaches Using Discharge Summaries: A Systematic Literature Review. *Expert Systems with Applications*, *213*. https://doi.org/https://doi.org/10.1016/j.eswa.2022.118997

Kaur, R. (2018). *A Comparative Analysis of Selected Set of Natural Language Processing (NLP) And Machine Learning (ML) Algorithms For Clinical Coding Using Clinical Classification Standars*. Pubmed.

Kedia, A., & Rasu, M. (2020). Hands-On - Python Natural Language Processing. In *Packt Publishing*.

Kementerian Kesehatan. (2020). *Keputusan Menteri Kesehatan Republik Indonesia Nomor Hk 01.07/Menkes/312/2020 tentang Standar Profesi Perekam Medis dan Informasi Kesehatan*.

Kementerian Kesehatan. (2022). *Peraturan Menteri Kesehatan Republik Indonesia Nomor 24 Tahun 2022 tentang Rekam Medis*.

Ness, S., Subramanian, M. L., & Chen, X. (2022). Diagnosis and Management Degenerative of Retinoschisis and Related Complications.pdf. *Survey of Opthalmology*, *67*(4), 892–927. https://doi.org/https://doi.org/10.1016/j.survophthal.2021.12.004

Persatuan Dokter Spesialis Mata Indonesia (PERDAMI). (2018). *Pedoman Nasional Pelayanan Kedoteran Ablasio Retina Regmatogen*.

Rachman, F. H. (2020). *Buku Ajar Komputasi Bahasa Alami* (1st ed.). Media Nusa Creative (MNC Publisihing).

RI, K. E. P. dan P. K. N. K. K. (2021). *Pedoman dan Standar Etik Penelitian dan*

Pengembangan Kesehatan Nasional. Lembaga Penerbit Badan Penelitian dan Pengembangan Kesehatan.
Stanfill, M. H., & Marc, D. T. (2019). Health information management: implications of artificial intelligence on healthcare data and information management. In *Yearbook of medical informatics*. thieme-connect.com. https://doi.org/10.1055/s-0039-1677913
Surur, F. M., Mamo, A. A., Gebresilassie, B. G., Mekonen, K. A., Golda, A., Behera, R. K., & Kumar, K. (2025). Unlocking The Power of Machine Learning in Big Data: a Scoping Survey. *Data Science and Management*. https://doi.org/10.1016/j.dsm.2025.02.004
Zaky, H., Salem, A., Alzubaidi, M., Shah, H. A., Alam, T., Shah, Z., & Househ, M. (2023). Using AI for Detection, Prediction and Classification of Retinal Detachment. *Studies in Health Technology and Informatics, 305*, 636–639. https://doi.org/10.3233/SHTI230578