**Journal of E-Komtek**
**Vol. 9, No. 1. (2025) pp. 163-168**
**https://jurnal.politeknik-kebumen.ac.id/index.php/E-KOMTEK**
**p-ISSN : 2580-3719 e-ISSN : 2622-3066**

# Application of Random Forest Classification Method in Determining the Best Quality Service in the Implementation of International Certification at ITCC ITPLN

Hendra Jatnika, Luqman Luqman, Mochamad Farid Rifai, Nasya Miranda Umar
Department of Technical Information, Institut Teknologi PLN, Jakarta, Indonesia, 11750

h.jatnika@itpln.ac.id

https://doi.org/10.37339/e-komtek.v9i1.2349

**Published by Piksi Ganesha Indonesia Polytechnic**

***Article Info***
*Submitted:*
03-03-2025
*Revised:*
26-05-2025
*Accepted:*
18-06-2025
*Online first :*
30-06-2025

**Abstract**

*Information Technology Certification Center or also known as ITCC is one of the work units owned by the Institute Technology of PLN which is a unit that organizes training and international certification. In order to improve the quality of service from the activities that have been organized by ITCC, the ITCC committee always prepares a link for participants to write feedback which will later become material for evaluation by ITCC. In this study, 2,720 data were used which were divided into 2 categories, namely 1,884 data with positive sentiment categories and 836 data with negative sentiment categories. The data is processed using the Random Forest method in order to find out the optimality of knowing the method. The final result obtained from the application of the Random Forest classification method is an accuracy percentage of 88.97% with a precision value of 0.92, recall of 0.91, and f1 score of 0.92*

**Keywords**: *ITCC, Random Forest Classification, Quality Service, Sentiment Analysis, Machine Learning, Data Mining.*

***Abstract***

*Pusat Sertifikasi Teknologi Informasi atau yang juga dikenal dengan ITCC merupakan salah satu unit kerja yang dimiliki oleh Institut Teknologi PLN yang merupakan unit penyelenggara pelatihan dan sertifikasi internasional. Dalam rangka meningkatkan kualitas pelayanan dari kegiatan yang telah diselenggarakan oleh ITCC, panitia ITCC selalu menyiapkan link bagi peserta untuk menulis feedback yang nantinya akan digunakan sebagai bahan evaluasi oleh ITCC. Dalam penelitian ini, digunakan sebanyak 2.720 data yang dibagi menjadi 2 kategori, yaitu 1.884 data pada kategori sentimen positif dan 836 data pada kategori sentimen negatif. Data diolah dengan menggunakan metode Random Forest untuk mengetahui optimalitas mengetahui metode tersebut. Hasil akhir yang diperoleh dari penerapan metode klasifikasi Random Forest adalah persentase akurasi sebesar 88,97% dengan nilai presisi 0,92, recall 0,91, dan skor f1 0,92.*

**Keywords**:*ITCC, Klasifikasi Random Forest, Layanan Berkualitas, Sentiment Analysis, Machine Learning, Data Mining.*

## 1. Introduction

*Information Technology Certification Center* or also known as ITCC is one of the work units owned by the PLN Institute of Technology and has been established since 2016 [1]. ITCC has been authorized by Certiport Inc as Certiport Authorized Testing Center (CATC) and also MikroTik – SIA Mikrotikls to organize international training and certification. In order to improve the quality of service from the activities organized by ITCC, the committee always prepares a link for participants to write their feedback using one of Microsoft's features, namely Microsoft Form, where after the certification activity is completed, participants will be asked to fill in the link to provide feedback, one of the purposes of which is to get criticism and suggestions for certification activities [2].

The main purpose of creating this feedback filling facility is so that ITCC can evaluate the activities that have taken place, but in fact, up until this research was conducted, the evaluation stage had not been carried out due to the large amount of data received from each activity carried out which had an impact on the quality of ITCC service quality not being able to develop [3]. Therefore, an initial step is needed, namely a proper analysis in order to process text data so that it can provide accurate information. In this study, sentiment analysis will provide high-quality information from text for classification or categorization of the results of participant feedback who have taken certification at ITCC. The Random Forest method will later be used to build a model so that it will classify feedback text based on predetermined input and categories. This aims to determine the level of accuracy of the application of the Random Forest method in classifying feedback [4].

Based on the background above, the researcher took the title of the thesis "Application of Random Forest Classification Method in Determining the Best Quality Service in the Implementation of International Certification at ITCC ITPLN".

## 2. Method

### 2.1 Data Mining

Data mining is a series of automatic search actions for information. Data mining itself is used to examine large databases in order to find new and useful patterns [5]. In

this study, a process called CRISP-DM (Cross-Industry Standard Process for Data Mining) was used **[6]**.

## 2.2 Classification

Classification itself is the grouping of objects into one of several categories.haspredetermined. In Classification,donethe evaluation process of the classification model performance. For this classification problem, a confusion matrix is usually used, as illustrated in **Figure 1. [7]**.

| | | *Predicted Class* | |
|---|---|---|---|
| | | Negatif | Positive |
| *Actual Class* | Negatif | *True Negative (TN)* | *False Positive (FP)* |
| | Positive | *False Negative (FN)* | *True Positive (TP)* |

**Figure 1.** Confusion matrix

From the figure, we can calculate:

1. *Accuracy*(Total number of times the model is correct)

$$\frac{TP + TN}{Total}$$

2. *Precision*(when the model predicts positive, how often is the prediction correct)

$$\frac{TP}{FP + TP}$$

3. *Recall*(when the actual class is positive, how often the model predicts positive)

$$\frac{TP}{FN + TP}$$

4. *F1-Score*(is the average of Precision and Recall)

$$2 * \frac{Precision * Recall}{Precision + Recall}$$

## 2.3 Random Forest

The Random Forest algorithm begins with a basic data mining technique, namely the decision tree. In this decision tree, we enter input at the top (root) then go down to the bottom (leaf) to determine whether the data is included in class A or B or so on. Random Forest is a classification consisting of a collection of structured tree classifications where

each tree throws a voting unit for the most popular class on the input x [8].

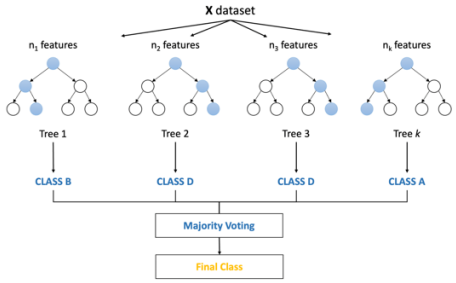


**Figure 2.** Random Forest Illustration

Random Forest is also a supervised classification method [9]. As the name suggests, this method creates a forest with a number of trees. In general, the more trees in a forest, the greater the accuracy obtained [10].

## 3. Results and Discussion

The results obtained in the stages that have been passed are from the data that the author has analyzed that the total data with positive labels is at 1,884 and for negative labels it is at 836. In the process, data preparation is carried out starting with case folding, tokenizing, stopword remover, and stemming with the results of the process as in the **Table 1.**

**Table 1.** Case Folding

| Feedback Data | Results *Case Folding* |
|---|---|
| The material presented by the speaker is very useful | The material presented by the speaker is very useful |

**Table 2.** Tokenization

| Results *Case Folding* | Results *Tokenization* |
|---|---|
| The material presented by the speaker is very useful | 'material', 'which', 'brought', 'by', 'presenter', 'very', 'useful' |

**Table 3.** Stop Wordsremover

| Results *Tokenization* | Results*stopword remover* |
|---|---|
| The material presented by the speaker is very useful | The material presented by the speaker was very useful |

**Table 4.** Stemming

| Results *stopword remover* | Results *stemming* |
|---|---|
| The material presented by the speaker is very useful | The material brought by the speaker is very useful |

After preprocessing the data, the following is the sum of positive sentiment and negative sentiment that has been split into two parts.

```
[ ] kalkulasi_test = y_train.value_counts()
    kalkulasi_train = y_test.value_counts()
    print('Total label Test: \n', kalkulasi_test)
    print('Total label Latih: \n', kalkulasi_train)

    Total label Test:
      1    1504
    -1     672
    Name: score, dtype: int64
    Total label Latih:
      1    380
    -1    164
    Name: score, dtype: int64
```

**Figure 1.** Label Results and Split Data

After the Random Forest Classification method was applied, the results were obtained.evaluation results with The indicators used are precision, recall, f1-score, and accuracy values [11].

```
[141] print("RFF Accuracy Score: ", accuracy_score(y_predict_rfc, y_test)*100)

      RFF Accuracy Score:  88.97058823529412

[144] from sklearn.ensemble import RandomForestClassifier
      from sklearn.metrics import classification_report

      model = RandomForestClassifier().fit(x_train, y_train)

      pred_rf = model.predict(x_test)
      print(classification_report(pred_rf, y_test))

                    precision    recall  f1-score   support

                -1       0.79      0.82      0.80       159
                 1       0.92      0.91      0.92       385

          accuracy                           0.88       544
         macro avg       0.86      0.86      0.86       544
      weighted avg       0.89      0.88      0.88       544
```

**Figure 2.** Accuracy results

The precision value obtained a value of 0.92, the recall value was at a value of 0.91, the f1 score obtained a value of 0.92, and for accuracy the percentage obtained was 88.97%.

## 4. Conclusion

FromBased on the research that the author has conducted, the following conclusions were obtained:

1. The quality service that has been running at ITCC ITPLN is considered quite good when seen from the feedback response of certification participants. Only a few things need to be improved.
2. Based on the results of this study, which used the Random Forest classification method by calculating the weight of each word, from 2,720 feedback data from certification participants, the results were quite optimal.
3. The Random Forest classification obtained a precision value of 0.92, a recall of 0.91, an f1 score of 0.92, and an accuracy percentage of 88.97%, which means that the Random Forest classification method is very good.

**Reference**

[1] "HOME - ITCC," INFORMATION TECHNOLOGI CERTIFICATION CENTER.

[2] H. Jatnika, Rakhmadi Irfansyah Putra, Muhammad Zaid Al-Khair, Arfani Lovina Br. Stendel, and Mohamad Tanwirul Akbar, "Application of Genetic Algorithm with Roulette Wheel Selection Method for International Certification Scheduling Optimization at ITCC ITPLN," *Jurnal E-Komtek (Elektro-Komputer-Teknik)*, vol. 6, no. 2, pp. 360–369, Dec. 2022, doi: 10.37339/e-komtek.v6i2.2325.

[3] A. Nasihi and T. A. R. Hapsari, "Monitoring dan evaluasi kebijakan pendidikan," *Indonesian Journal of Teaching and Learning (INTEL)*, vol. 1, no. 1, pp. 77–88, 2022.

[4] A. Ramadhan and B. Susetyo, "Penerapan Metode Klasifikasi Random Forest Dalam Mengidentifikasi Faktor Penting Penilaian Mutu Pendidikan," *Jurnal Pendidikan dan Kebudayaan*, vol. 4, no. 2, pp. 169–182, 2019.

[5] K. Agustini, G. A. Pradnyana, I. G. M. Darmawiguna, and I. N. S. W. Wijaya, *DATA MINING*, vol. 1. Universitas Terbuka, 2022.

[6] U. Shafique and H. Qaiser, "A comparative study of data mining process models (KDD, CRISP-DM and SEMMA)," *International Journal of Innovation and Scientific Research*, vol. 12, no. 1, pp. 217–222, 2014.

[7] "International Conference on Instrumentation, Communication, Information Technology, and Biomedical Engineering 2009 : 23-25 November 2009.," New York: IEEE, 2010.

[8] L. Breiman, "Random Forests," 2001.

[9] A. Primajaya and B. N. Sari, "Random forest algorithm for prediction of precipitation," *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 1, no. 1, pp. 27–31, 2018.

[10] D. Polamuri, K. Srinivas, and A. Mohan, "Stock Market Prices Prediction using Random Forest and Extra Tree Regression," *International Journal of Recent Technology and Engineering*, vol. 8, pp. 1224–1228, Sep. 2019, doi: 10.35940/ijrte.C4314.098319.

[11] A. Nasihi, T. Asihati Ratna Hapsari, and K. Kota Jakarta Selatan, "Indonesian Journal of Teaching and Learning," vol. 1, no. 1, pp. 77–88, 2022, doi: 10.56855/intel.v1i1.112.