

Enhancing ITIL Incident Management: Innovative Machine Learning Approaches for Efficient Incident Prioritization and Resolution

Alifia Ayu Zahrothul Ain¹, Cutifa Safitri²

^{1,2} Master of Informatics, Faculty of Computing, President University

^{1,2}Ki Hajar Dewantara Street Jababeka Education Park, Cikarang, West Java, Indonesia.

ABSTRACT

Article:

Accepted: August 17, 2023
Revised: May 23, 2023
Issued: October 28, 2023

© 2023 The Author(s).



This is an open-access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

*Correspondence Address:

alifiazahrothulain@gmail.com

Incident Management in ITIL requires an effective process so the incidents do not disrupt business processes for too long. This research aims to automate decision-making in Incident Management process. To perform the automation in decision-making process requires machine learning algorithms. The development of machine learning method in this research will bring significance result such as a new technique of decision-making process in Incident Management, accelerate decision-making process in Incident Management by implementing machine learning to determine the category, group, and priority. By training historical full description, short description, and title, machine learning can classify the new incident. In this research different classification algorithms are used to automate decision making process. Performances of automated decision-making are evaluated with accuracy, precision, recall, and f1-score. Based on the result of various performance metrics, classifier based on K-Nearest Neighbor performed well on predicting group, and both category and priority get the best performance with Support Vector Machine.

Keywords: *IT service management, IT Infrastructure library, incident management, machine learning, support vector machine (SVM), k-nearest neighbor (KNN), naïve bayes*

1. INTRODUCTION

Information Systems (IS) must be developed in a system with good performance, has an advanced level of availability, has a good security, and is capable to compete in the market. So, IT service providers need to focus more on service quality and user satisfaction compared to technology and internal affairs. [1] Information Technology Service Management (ITSM) with reference to the principle of “good practices” has been widely used to overcome this. [2]

ITSM in providing guidelines for managing IT services effectively, several frameworks have been developed. The most used framework is Information Technology Infrastructure Library - ITIL. [1] ITIL successfully become de facto standard for ITSM in industry. ITIL guides and describe how IT operations can handle the main tasks, such as incident management, financial management, IT service continuity, risk management, service level management, capacity management, release management, and change management processes. [3]

One of main tasks in IT Operation is to handle unplanned interruption in services. This process called Incident Management. Incident Management (IM) focuses on the incident process from the incident is received until the incident is resolved. An incident should have a solution as soon as possible so the business system restored properly. [4]

Incident Management Process is a very complicated and time-consuming process. As time goes by and the existing technology, of course, the number of incidents that come in is also increasing becoming a very large data. This will make the IT team difficult to gain useful insights from the incident with using the conventional technique. The wrong result from the process will lead to a decrease in service quality and give bad affect in productivity. Taking these two points in mind, with the current manual processes cannot ensure the response time and accuracy of incident management. So, to analyze the large data in this era, it is needed a new smart technology to perform an efficient decision-making support method. [5] [6] [7]

Several studies discussed the use of machine learning to automate the incident management process, including research conducted by [5]

which discusses about determining the priority of incidents using Fuzzy Mamdani. Research [8] [9] [10] also discusses the use of machine learning to determine incident categories. Each study evaluates several supervised learning algorithms with different approaches.

In this research, with the purpose of preventing the negative impact on the quality of services, the authors proposed Machine Learning method to optimize the decision- making process in Incident Management. It uses the incident history in training phase to finally can predict the new one. The authors will find which machine algorithm can optimize the decision-making the most and how the algorithm can optimize the decision-making process.

2. METHODS

2.1. Model Preparation and Development

The very first step to do after preparing the dataset is preparing the model : cleansing data with identify of the missing values, filter out the attributes to contribute for the decision- making process, filter out the data that not necessary for the research, vector representation using TFIDF Vectorizer.

Another thing that is also important is preparing different Optimization parameters in each machine learning algorithm to improve the performance of the models. Table 1 show optimizer for each algorithm

Table 1. Optimization for each algorithm

Naïve Bayes	Model Gaussian
	Model Multinomial
	Model Bernoulli
Support Vector Machine	Kernel RBF
	Kernel Sigmoid
K-Nearest Neighbor	N Neighbors
	p Parameters

The next process is build classifier model, training data, and testing data. In this research use three algorithms with each algorithm has 3 different optimization parameters to be compared. The models must be able to determine the category, group, and priority of the incident. Figure 1 show the Classifier Model process.

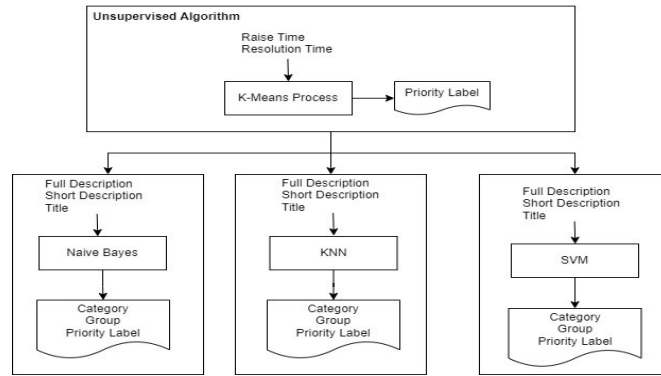


Figure 1. Classifier model process

a. *K-Means*

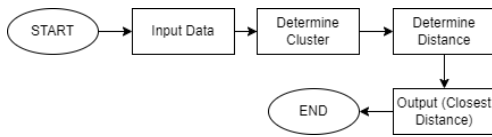


Figure 2. K-means process

K-Means will be used to find priority labels. K-Means categorize a group of objects according to the same attributes or characteristics as other data. In this research K- Means will categorize the incident data into 3 labels output (High, Medium, Low). Figure 2 show how K-Means determine priority labels of the incident.

b. *Naïve Bayes*

Naive Bayes classifies by probability and statistical methods. To make classifications, we need to use Full Description, Short Description, and Title to predict Category, Group, and Priority. For figure out the detail process of Naïve Bayes take a look in Figure 3

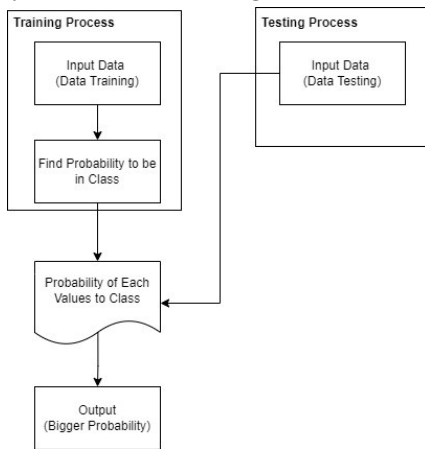


Figure 3. Naïve bayes model

c. *Support Vector Machine*

Details process of SVM in this research showed in Figure 4.

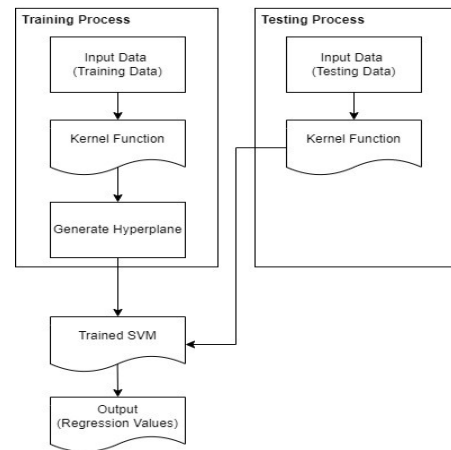


Figure 4. SVM Models

d. *K-Nearest Neighbor*

The k-Nearest Neighbor algorithm uses the Neighborhood classification as the predicted value of the new instance value. [12] k Neighbor with the highest score will be chosen which class with the most number. Details of K- Nearest Neighbor Process showed in Figure 5

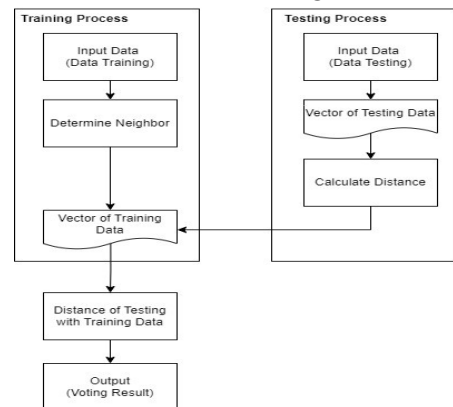


Figure 5. K-nearest neighbor

2.2. Dataset Preparation

In this research, the data were obtained from IT Operation Division of a company. The dataset consisted of several specific records. This event log was extracted from data gathered from the Service Desk platform used by the company. The event log consist of two types : Incident and Request. For the objective of the research, we worked only with the incident dataset. [13]

A real-time enterprise IT infrastructure service desk ticket collected over a period of year provide 7,629 tickets for this research purposes. The incident dataset is presented in a*.csv. Each incident contains 21 fields, but only tickets full description, short description, title, category, group, and priority were used for development process and rest of the fields are not being used. This research selection of attributes is based on the attribute related to determining the category, group and priority.

2.3. Model Evaluation

To evaluate the predicted results of a model, the indicators used in this study are accuracy, recall, precision, and f-1 score. These indicators are calculated based on the confusion matrix of the model. The confusion matrix will be determined based on each output parameter in the research: Category, Group, Priority. [14] [15]

3. RESULTS AND DISCUSSION

3.1. Experimental Result

The model development process is carried out to show how the machine learning algorithm can classify the incident based on Full Description, Short Description, and Title. The training data were correctly labeled manually for Category and Group. The dataset would be split using 80:20 percentage split ratio with 80% of tickets used for training the classifier and the remaining instances used for testing the classification accuracy.

K-Means clustering categorize the incident data into 3 priority labels (High, Medium, Low). Figure 6 show the result of K- Means. There are 6163 data with labels “High”, 1266 data “Medium”, and 198 data “Low”.

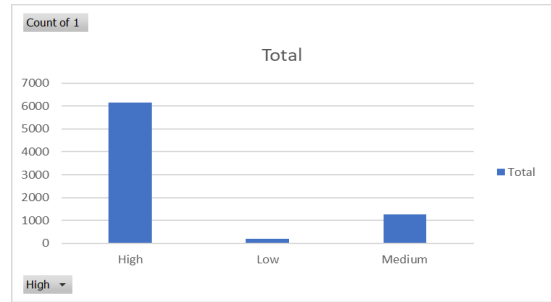


Figure 6.K-means result

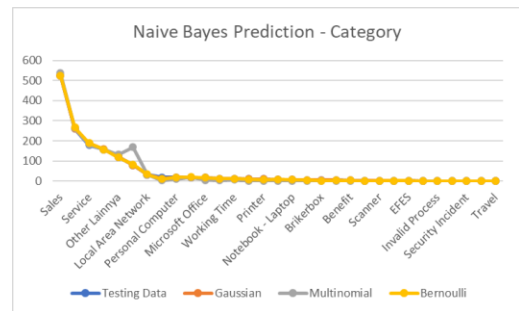


Figure 7.Naive bayes prediction result of category

Figure 7 show the comparison between actual category data in Testing Data and predicted data in Naïve Bayes. Almost all category data has the same number of rows between actual data and predictive data in Gaussian Naïve Bayes. The gap between the prediction results and the testing data can be seen in Table 2.

Table 2.Gaussian naïve bayes confusion matrix for category

Actual/- Prediction	AR	Sales	Insurance	Service
AR	260	1	0	0
Sales	1	524	0	0
Insurance	0	0	12	8

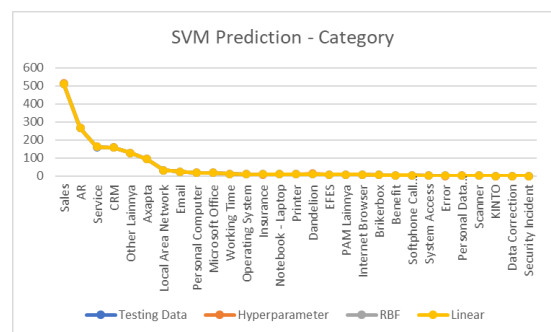


Figure 8.SVM result of category

Figure 8 show the comparison between actual category data in Testing Data and predicted data in Support Vector Machine. Almost all category data has the same number of rows between actual data and predictive data. The details distribution of prediction data show in 3, Table 4, and Table 5.

Table 3. Hyperparameter SVM confusion matrix for category

Actual/ Prediction	Sales	Insuranc e	Service
Sales	511	0	0
Insurance	2	0	1
Service	0	0	161

Table 4.RBF SVM confusion matrix for category

Actual/ Prediction	Sales	Insuranc e	Service
Sales	511	0	0
Insurance	0	10	1
Service	0	0	161

Table 5.Linear SVM confusion matrix for category

Actual/ Prediction	Sales	Insuranc e	Service
Sales	511	0	0
Insurance	0	10	1
Service	0	0	161

Figure 9 show the comparison between actual category data in Testing Data and predicted data in K-Nearest Neighbor. Table 6 and Table 7 show the details distribution of prediction data.

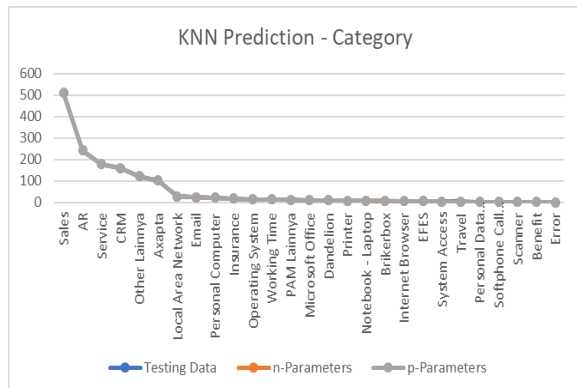


Figure 9.KNN Result of Category

Table 6.n-neighbor knn confusion matrix for category

Actual/ Prediction	System Access	Browse dan Portal	Trav el	Benefit
System Access	3	1	0	0
Browser dan Portal	0	0	0	0
Travel	0	0	3	1
Benefit	0	0	0	2

Table 7.p-Parameter KNN Confusion Matrix for Category

Actual/ Prediction	Brik ebox	Softphone Call Center	Trav el	Benefit
Brikerbox	7	0	0	0
Softphone Call Center	1	2	0	0
Travel	0	0	3	1
Benefit	0	0	0	2

Figure 10 and Table 8 show the comparison between actual group data in Testing Data and predicted data in Naïve Bayes. Table 9, Table 10, and Table 11 show the details distribution of prediction data.

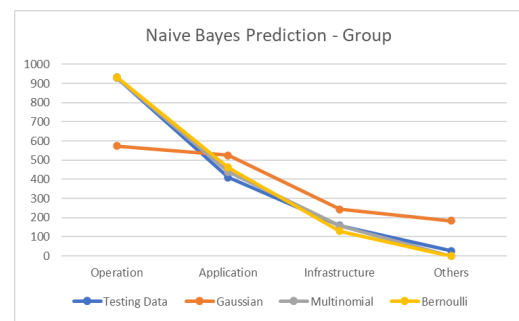


Figure 10.Naïve bayes result of group

Table 8.Naive bayes prediction result of group

Group	Testing Data	Gau- ssian	Multi- nomial	Bernoulli
Operation	931	574	930	934
Applicatio n	410	525	439	462
Infrastruc ture	159	243	157	130
Others	26	184	0	0

Table 9. Gaussian naïve bayes confusion matrix for group

	Application	Infrastructure	Operation	Others
Application	357	27	20	6
Infrastructure	6	149	2	2
Operation	159	66	542	164
Others	3	1	10	12

Table 10. Multinomial naïve bayes confusion matrix for group

	Application	Infrastructure	Operation	Others
Application	387	3	20	0
Infrastructure	12	144	3	0
Operation	38	0	893	0
Others	2	10	14	0

Table 11. Bernoulli naïve bayes confusion matrix for group

	Application	Infrastructure	Operation	Others
Application	386	0	24	0
Infrastructure	27	129	3	0
Operation	43	0	888	0
Others	6	1	19	0

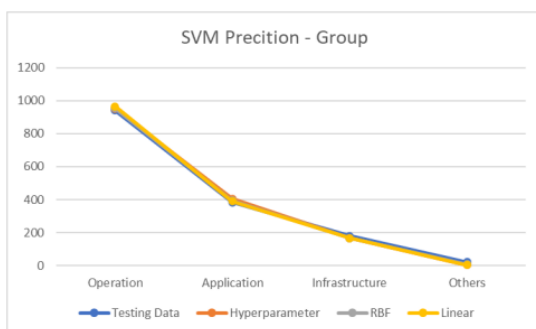


Figure 11. Support vector machine result of group

Figure 11 show the comparison between actual category data in Testing Data and predicted data in Support Vector Machine. The details distribution of prediction data show in Table 12, Table 13, Table 14.

Table 12. Hyperparameter SVM confusion matrix for group

Actual/Prediction	Application	Infrastructure	Operation	Others
Application	356	1	26	0
Infrastructure	9	165	7	0
Operation	38	0	903	0
Others	2	0	15	4

Table 13. RBF SVM confusion matrix for group

Actual/Prediction	Application	Infrastructure	Operation	Others
Application	357	2	24	0
Infrastructure	8	166	7	0
Operation	28	0	908	5
Others	1	0	16	4

Table 14. Linear SVM confusion matrix for group

Actual/Prediction	Application	Infrastructure	Operation	Others
Application	355	1	27	0
Infrastructure	8	166	7	0
Operation	26	0	915	0
Others	1	0	16	4

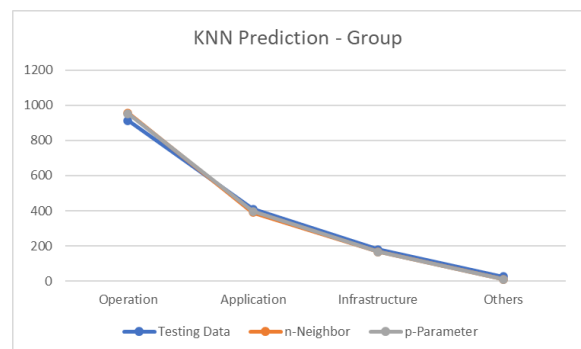


Figure 12. K-nearest neighbor of group

Figure 12 show the comparison between actual group data in Testing Data and predicted data in K-Nearest Neighbor. There is no single

group with the same number of prediction results as the testing data. Table 15 and Table 16 show the details distribution of prediction data.

Table 15.K-nearest neighbor prediction result of group

Actual / Prediction	Appl ication	Infras tructure	Oper ation	Other s
Applicatio n	365	1	43	0
Infrastruct ure	7	166	6	0
Operation	16	0	897	0
Others	2	1	11	11

Table 16.p-Parameter KNN confusion matrix for category

Actual / Prediction	Appl ication	Infras tructure	Oper ation	Other s
Applicatio n	365	1	43	0
Infrastruct ure	7	166	6	0
Operation	16	0	897	0
Others	2	1	11	11

Figure 13 show the comparison between actual priority data in Testing Data and predicted data in Naïve Bayes. Table 17, Table 18, and Table 19 show the details distribution of prediction data.

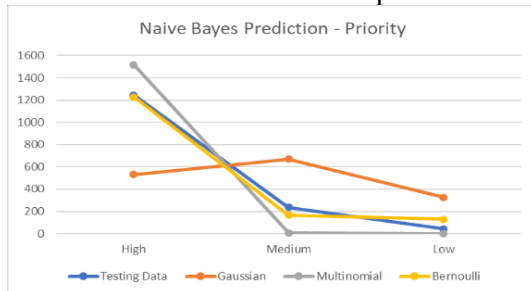


Figure 13.Naïve bayes prediction result of priority

Table 17.Gaussian naïve bayes confusion matrix for priority

	Medium	High	Low
Medium	93	70	74
High	555	447	244
Low	21	13	9

Table 18.Multinomial naïve bayes confusion matrix for priority

	Medium	High	Low
Medium	2	235	0
High	4	1242	0
Low	2	41	0

Table 19.Bernoulli naïve bayes confusion matrix for priority

	Medium	High	Low
Medium	48	158	31
High	114	1036	96
Low	6	35	2

Figure 14 show the comparison between actual group data in Testing Data and predicted data in Support Vector Machine. Table 20, Table 21, and Table 22 show the details distribution of prediction data.

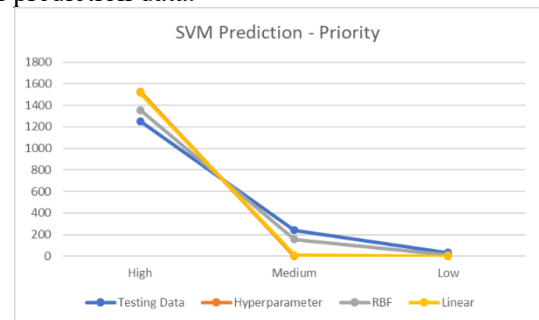


Figure 14.Support vector machine result of priority

Table 20.Hyperparameter SVM confusion matrix for priority

Actual/ Prediction	Medium	High	Low
Medium	0	241	0
High	0	1251	0
Low	0	34	0

Table 21.RBF SVM confusion matrix for priority

Actual/ Prediction	Medium	High	Low
Medium	39	197	5
High	110	1136	5
Low	8	23	3

Table 22. Linear kernel SVM confusion matrix for priority

Actual/ Prediction	Medium	High	Low
Medium	5	236	0
High	5	1246	0
Low	0	33	1

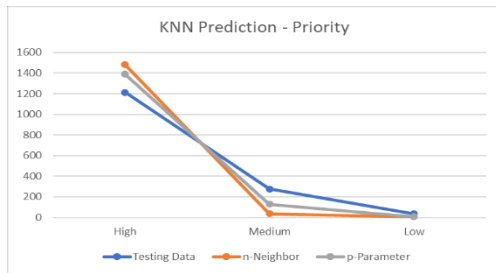


Figure 15. K-nearest neighbour result of priority

Figure 15 show the comparison between actual priority data in Testing Data and predicted data in K-Nearest Neighbor. There is no single priority with the same number of prediction results as the testing data. Table 23 and Table 24 show the details distribution of prediction data.

Table 23. n-neighbor KNN confusion matrix for priority

Actual / Prediction	Medium	High	Low
Medium	14	259	4
High	20	1191	2
Low	3	32	1

Table 24. -Parameter KNN confusion matrix for priority

Actual / Prediction	Medium	High	Low
Medium	49	224	4
High	76	1135	2
Low	4	31	1

3.2. Evaluation Result

All three naive Bayes algorithms has good performance. Bernoulli Naive bayes is good at dealing with boolean/binary attributes, while Multinomial Naive bayes is good at dealing with discrete values and Gaussian Naive Bayes is good at dealing with continuous values.

One of the features that can be used in SVM is that it can even work with non-linear data sets. To achieve this, we use the “Kernel Trick” which makes it easier to classify points. Which kernel to use is purely determined by hyperparameter tuning. The RBF kernel combines multiple polynomial kernels multiple times with different degrees to project non-linearly separable data into a higher-dimensional space so that it can be separated using a hyperplane. The linear kernel is simply different in case of making the hyperplane decision boundary between the classes. Usually linear is less time consuming and provides less accuracy than the rbf.

a. Category

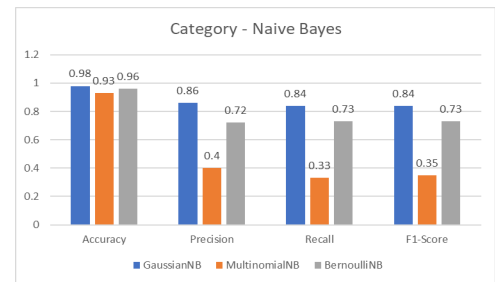


Figure 16. Evaluation of naïve bayes for category

Figure 16 show the result of an evaluation of decision-making process for predicting category using the Naive Bayes method. The accuracy results show that Gaussian Naive Bayes (98%) has a higher level of prediction compared to Multinomial Naive Bayes and Bernoulli Naive Bayes (93% and 96%). This result is in line with the results of precision, recall, and also the f1 score, Gaussian Naive Bayes also has the highest result.

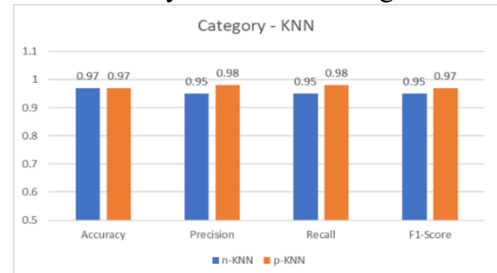


Figure 17. Evaluation of k-nearest neighbor for category

Figure 17 show the result of an evaluation of decision-making process for predicting category using K-Nearest Neighbour algorithm. The accuracy results show that both algorithms have the same level for prediction. The result of precision, recall, and f1-score give the different inside, n-Neighbour KNN has better result for predicting the category.

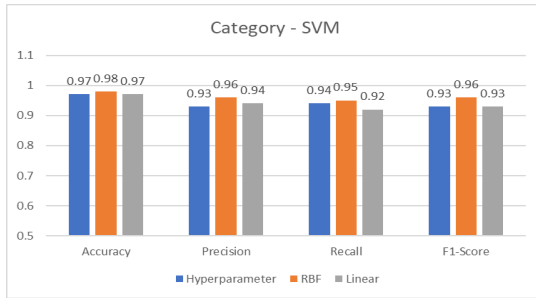


Figure 18. Evaluation of support vector machine for category

Figure 18 show the result of an evaluation of decision-making process for predicting category using the Support Vector Machine algorithm. The accuracy results show that RBF Kernel SVM (98%) has a higher level of prediction compared to Hyperparameter SVM and Linear SVM (97%). This result is in line with the results of precision, recall, and also the f1 score, RBF Kernel SVM also has the highest result.

b. Group

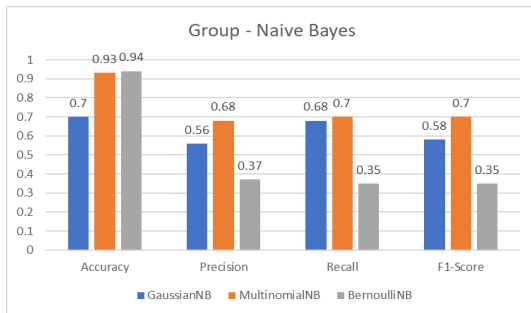


Figure 19. Evaluation of naïve bayes for group

Figure 19 show the result of an evaluation of decision-making process for predicting group using the Naive Bayes algorithm. The accuracy results show that Bernoulli Naive Bayes (94%) has a higher level of prediction compared to Multinomial Naive Bayes and Gaussian Naive Bayes (93% and 70%). Unfortunately, the precision, recall, and f1-score results show the opposite. Bernoulli Naive Bayes has the lowest result for recall, precision, and f1-score. (35%, 37%, and 35%)

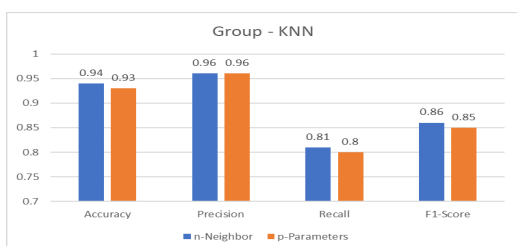


Figure 20. Evaluation of k-nearest neighbor for group

Figure 20 show the result of an evaluation of decision-making process for predicting group using the K-nearest Neighbor Algorithm. The accuracy results show that n- Neighbor KNN (94%) has a higher level of prediction compared to p-Parameter KNN (93%). This result is in line with the results of precision, recall, and also the f1 score, n- Neighbor KNN also has the highest result.

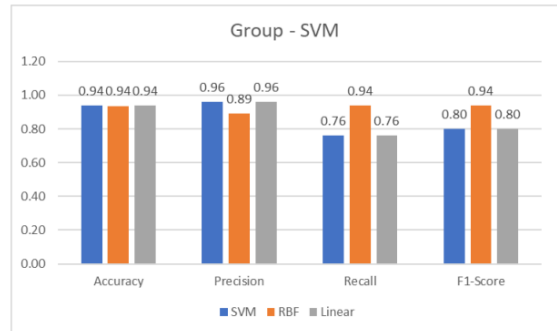


Figure 21. Evaluation of support vector machine for group

Figure 21 show the result of an evaluation of decision-making process for predicting group using the Support Vector Machine Algorithm. The accuracy results show that all three algorithms can reach the same level of prediction. The result of precision, recall, and f1-score show different. Hyperparameter SVM and Linear SVM (96%) has higher precision than RBF Kernel SVM (89%). RBF Kernel SVM (94%) has higher recall than Hyperparameter SVM and Linear SVM (76%).

c. Priority

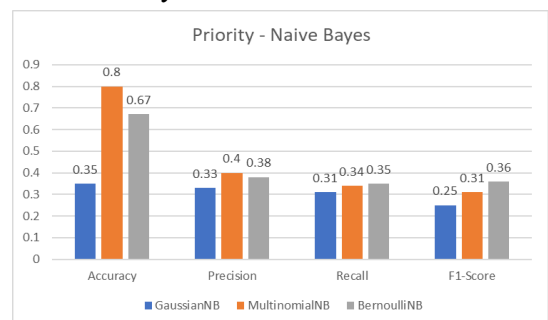


Figure 22. Evaluation of naïve bayes for priority

Figure 22 show the result of an evaluation of decision-making process for predicting priority using the Naive Bayes algorithm. The accuracy results show that Multinomial Naive Bayes (80%) has a higher level of prediction compared to Bernoulli Naive Bayes and Gaussian Naive Bayes (67% and 35%).

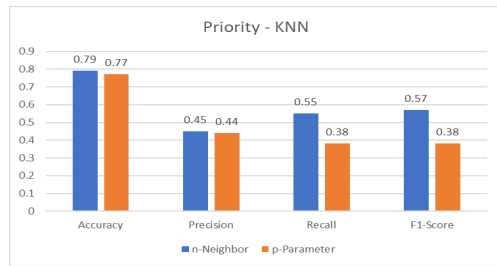


Figure 23. Evaluation of k-nearest neighbor for priority

Figure 23 show the result of an evaluation of decision-making process for predicting priority using the Support Vector Machine algorithm. The accuracy results show that Hyperparameter SVM (81%) has a higher level of prediction compared to RBF Kernel SVM and Linear Kernel SVM (76% and 79%).

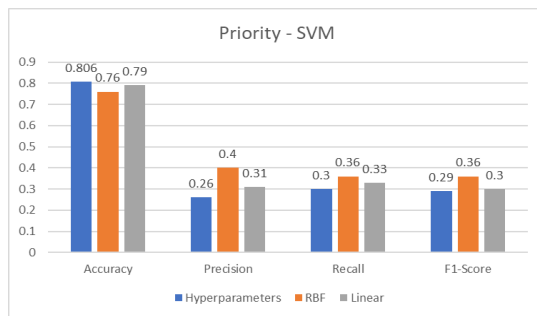


Figure 24. Evaluation of support vector machine for priority

Figure 24 show the result of an evaluation of decision-making process for predicting priority using the Support Vector Machine algorithm. The accuracy results show that Hyperparameter SVM (81%) has a higher level of prediction compared to RBF Kernel SVM and Linear Kernel SVM (76% and 79%).

CONCLUSION

The decision-making process in incident management is a very complex and time-consuming process that will have an impact on the quality of the services provided. The implementation of the proposed method is to automate the decision-making process, which includes: categories, groups, and outputs. A collection of historical data for a certain period from land desk system of a Financial Company is used as data collection or knowledge in this research. Full Description, Short Description, and Title are used as input parameters to predict Category, Group, and Priority.

This study uses three different well-known text classification algorithms (Naïve Bayes, KNN, and SVM) and evaluates the classifier performance on ticket data with several different metrics (accuracy, precision, recall, and f1-score). Automation in decision-making process in this research consists of 2 (two) important parts: training and testing process. Training process aims to provide clues through the algorithm so that the machine can find the own correlation. After the models or the machine can find the correlation between parameter input and the output, the machine can easily predict the result after we input the parameter. Testing process aims to see its accuracy, or in other words see its performance. The higher the accuracy, then in the future the model will potentially provide accurate prediction results.

The implementation of proposed method in this research show that each output (Category, Group, and Priority) requires a different algorithm to achieve the highest predictive level. Category and Group get best result when predicted using SVM RBF Kernel, but Priority get best result when predicted using KNN n-Neighbors. Evaluation matrix of Category and Group shows good result with balanced accuracy, precision, recall, and f-1 score. This means that the input parameters are suitable for predicting Category and Group. Evaluation matrix of Priority show unbalance accuracy, precision, recall, and f-1 score. The accuracy shows a good result, but the precision, recall, and f-1 score has a big gap with the accuracy. This unbalanced result show that we need more suitable parameter input to predict priority.

REFERENCES

- [1] E. Orta and M. Ruiz, "Met4ITIL: A process management and simulation-based method for implementing ITIL," *Comput Stand Interfaces*, vol. 61, pp. 1–19, Jan. 2019, doi: 10.1016/j.csi.2018.01.006.
- [2] T. J. Winkler and J. Wulf, "Effectiveness of IT Service Management Capability: Value Co-Creation and Value Facilitation Mechanisms," *Journal of Management Information Systems*, vol. 36, no. 2, pp. 639–675, Apr. 2019, doi:

- 10.1080/07421222.2019.1599513.
- [3] A. Mahalle, J. Yong, and X. Tao, "ITIL Processes to Control Operational Risk in Cloud Architecture Infrastructure For Banking and Financial Services Industry" doi: 10.1109/BESC.2018.00048.
- [4] S. Silva, R. Pereira, and R. Ribeiro, "Machine Learning in Incident Categorization Automation."
- [5] I. Lee and K. Lee, "The Internet of Things (IoT): Applications, investments, and challenges for enterprises," *Bus Horiz*, vol. 58, no. 4, pp. 431–440, Jul. 2015, doi: 10.1016/j.bushor.2015.03.008.
- [6] M. Yun, Y. Lan, and T. Han, "Automate Incident Management by Decision-making Model," 2017.
- [7] D. Hoorpah, S. Kishnah, and S. Pudaruth, "Development of an incident prioritization model using fuzzy logic to improve quality and productivity in IT support services," in *Advances in Intelligent Systems and Computing*, 2019, vol. 863, pp. 67–77. doi: 10.1007/978-981-13-3338-5_7.
- [8] S. Silva, R. Ribeiro, and R. Pereira, "Less is more in incident categorization," in *OpenAccess Series in Informatics*, Jul. 2018, vol. 62. doi: 10.4230/OASIS.SLATE.2018.17.
- [9] F. Al-Hawari and H. Barham, "A machine learning based help desk system for IT service management," *Journal of King Saud University - Computer and Information Sciences*, 2019, doi: 10.1016/j.jksuci.2019.04.001.
- [10] S. P. Paramesh and K. S. Shreedhara, "Automated IT service desk systems using machine learning techniques," in *Lecture Notes in Networks and Systems*, vol. 43, Springer, 2019, pp. 331–346. doi: 10.1007/978-981-13-2514-4_28.
- [11] L. Kyung Oh, M. Hugo Athallah Hardy, and Y. Wijaya, "Comparative Analysis of KNN, Naïve Bayes and SVM Algorithms for Movie Genres Classification Based on Synopsis," *Jurnal Teknik Informatika*, vol. 15, no. 2, pp. 169–177, 2022, doi: 10.15408/jti.v15i2.29302.
- [12] S. Sudianto, J. Arton Masheli, N. Nugroho, R. Wika, A. Rumpoko, and Z. Akhmad, "Comparison of Support Vector Machines and K-Nearest Neighbor Algorithm Analysis of Spam Comments on Youtube Covid Omicron," *Jurnal Teknik Informatika*, vol. 15, no. 2, pp. 110–118, 2022, doi: 10.15408/jti.v15i2.24996.
- [13] L. D. Fitriani and R. V. H. Ginardi, "Analysis Improvement of Helpdesk System Services Based on Framework COBIT 5 and ITIL 3rd Version (Case Study: DSIK Airlangga University)," 2018.
- [14] M. Sarnovsky and J. Surma, "PREDICTIVE MODELS FOR SUPPORT OF INCIDENT MANAGEMENT PROCESS IN IT SERVICE MANAGEMENT," *Acta Electrotechnica et Informatica*, vol. 18, no. 1, pp. 57–62, Mar. 2018, doi: 10.15546/aei-2018-0009
- [15] Ruby and D. Balkishan, "Fuzzy Logic based Requirement Prioritization (FLRP)-An Approach," vol. 6