International Conference on Global Optimization and Its Applications 2021 (ICoGOIA 2021), Bogor, December 23, 2021

Prediction System For QoS Level of Network Infrastructure Using Naive Bayes Classifier

Nizirwan Anwar [1], Dewanto Rosian Adhy [2*], Siti Maesroh[3*], Muhammad Abdullah Hadi[4], Rudi Hermawan[5*]

[1,4]*Esa Unggul University, Jakarta, Indonesia*
[2,3,5]*College of Technology YBS International, Tasikmalaya, Indonesia*

**Corresponding author email: dewanto_ra@sttybsi.ac.id, sitimaesaroh40@gmail.com, rudihermawan567@sttybsi.ac.id*

Abstract

The implementation of the Internet of Vehicle (IoV) demands the reliability of the network infrastructure because it involves the safety and security of motorists. The rapid development of IoV is sometimes not matched by the development of the quality and availability of network infrastructure. Coverage area constraints, geographical conditions and business are some of the things that hinder the scale-up of network infrastructure. A concept is needed to ensure that the path to be traversed by IoV-based vehicles has a minimum QoS level. In the IT field there is the concept of prediction by utilizing Data Mining technology. The prediction concept is built based on existing historical data and uses data mining algorithms. This predictive capability will provide information to users whether to use IoV services during the trip or not, referring to the QoS prediction results. The algorithm used is tested with data sets obtained from the results of mapping and tracking in certain areas as samples. The system is able to provide predictions with fairly good accuracy. Testing the prediction results is carried out by several methods.

Keywords: Internet of Vehicle, Machine Learning, Prediction , Quality of Service (QoS)

1. Introduction

The Internet of Things (IoT) is one of the technologies that are becoming a trend nowadays. Needs and implementation develop very fast. These conditions encourage increased infrastructure requirements and good data management. The number of points connected to IoT increases rapidly, as does the quantity and quality of data sent. IoT is a technology that connects objects (Things) or devices in digital communication. It is a development from the previous internet connection that connects only computer devices.

Referring to the needs above, before building the need for transportation facilities in providing technology-based plus services, it is necessary to first make sure of the quality of the communication system. The problems that arise from the data communication system in transportation are:

a) The limited coverage of BTS and the mobility of vehicles make the potential for loss of connection increase.
b) The movement of vehicles will pass through high-rise buildings and large buildings, which are very likely to block communication signals.
c) The accumulation of vehicles at a point due to traffic jams, crossroads, accidents, or other activities will cause an overload of the communication network so that connection loss will occur. (Anwar et al., n.d.)

Internet of Vehicle is a subset of the Internet of Things that is developing in-depth and more precisely. The Internet of Vehicles (IoV) is a convergence of the mobile Internet and the Internet of Things (IoT), where vehicles function as smart moving intelligent nodes or objects within the sensing network (Zantalis et al., 2019) .The Internet of Vehicles (IoV) is an application of the Internet of Things that provides a solution for traffic and safety management

Like the Internet of things, implementation in the transportation sector (IoV) also has the same problems and is even more complicated. The development of services that arise by society's dynamics, the growth of equipment and technology, and transportation technology itself make the challenges and problems even more significant. The IoT (or IoV) layer is growing and getting more detailed. From 3 layers (things, communication, computer) to 5 layers (Kaiwartya et al., 2016) and even up to 7 layers (Contreras-Castillo et al., 2018).

The unique characteristic of the Internet of Vehicle is the vehicle component's mobility (or sensing/perception layer). Mobility at high speeds creates connection problems. Transportation dynamics that give rise to congestion and unpredictable conditions complicate IoV services. (Xu et al., 2018). The challenges are getting more significant due to problems with non-standard equipment, uneven infrastructure and large gaps between users, and unbalanced development funds with service improvements (Kumar & Singh, 2020)

The need to obtain a useful (reliable and fast) Internet of vehicle configuration or service architecture is necessary. A concept that can provide IoV design quickly or an IoV system capable of compiling its structure according to field conditions needs. (Abosaif & Hamza, 2020a). This concept develops with the name Service Composer or Service Recommendation. In this concept, there is a core unit, namely Service Selection. This concept is how an IoV arranges architecture options from the lowest to the top layers. Using an algorithm, the selection is carried out to get optimal service by referring to the specified Quality of Service

The IoV architecture has been done a lot of research. The research results are increasingly detailed about the components and layers of IoV. This condition requires the ability to choose the right architecture according to infrastructure and QoS. Programming algorithms are overgrowing, one of which is in Machine learning with increasingly extraordinary skills. A challenge to implement this algorithm into the IoV architecture selection. The implementation of machine learning in IoV is more about developing the functions of IoV components. There is still little use of Machine Learning as one of the IoV prediction algorithms. (Sirohi et al., 2020)

2. Literature Review

2.1. Quality of Service

As time goes on, more information is collected and stored. The addition is due to data entry from the running system and additional vehicle units and system area coverage. Adding data is not only in percentages but increases rapidly or multiply. Problems arise regarding the data collected, namely:

a. Validation or correctness of data does not guarantee due to untrusted sources of information or the existence of connection disruptions that cause a decrease in data quality.
b. The importance of information on the results of data processing that can lead to the loss.
c. Data that is time-dependent and there is no guarantee that data will update or no time delay
d. Increasing coverage area and the number of connected vehicles connected requires setting up the communication network so that there is no overload of traffic.

These problems will lead to a decline in the quality of data produced by the IoV. The decrease is the cause:

a. Trust in the correctness of the data
b. Centralized data reduces network quality and process speed and produces latency
c. Data lost due to connection loss

2.2. Quality of Service in IoV

Quality of Service from IoV is not much different from IoT in general. The difference is only at a critical level. Time delay and data loss are things that are avoided by IoV. Data delays will result in an incorrect action from the Vehicle that can trigger an accident.

There has been a lot of research on the measurement and development of QoS from IoV. Using a Likert Scale to improve performance in terms of time (Abosaif & Hamza, 2020b)(Abu-safe & Elrofai, 2020), QoS development with Preventive Model for IoV devices can reduce time delay and data loss (T. Manivannan et al., 2020). Mesh Network development to improve the IoV network's quality (Ligo & Peha, 2017) Implementation of Operational Systematic Model to reduce time delay and bandwidth problems (Brogi & Forti, 2017) Algorithm development to improve QoS in several parameters (Butt et al., 2018)

An IoV system before implementation requires a certain level of QoS assurance. The high risk of failure requires a high QoS level. With the characteristics of the need for scalability, it is a difficulty in itself to guarantee QoS. Referring to these conditions, we need a Model of IoV. The model includes the network system's behaviour, data transmission, the anticipation of device mobility, congestion conditions where there are many active nodes at one point.

Scale-up in IoV will occur when there are additional vehicles and other needs are followed (such as seamless, secure, robust, scalable information exchange among vehicles, humans, and roadside infrastructures) (Vaidya & Mouftah, 2020)(Hussain et al., 2019) (Guan, 2018). Several models have been built but are limited in the movement of the vehicle. The mathematical formula for tracking vehicle movement models. (W. Zhang & Xi, 2016). A model to predict the vehicle's state, such as speed, position, and motion acceleration (Liu et al., 2018). Another model is describing the dynamics of the communication spectrum of IoV devices.(Hussain et al., 2019). More detailed modelling uses time-series behaviour to model collecting data from Vehicle (T. Zhang et al., 2019).

Table 1. A summary of the Quality of Service parameters of the Internet of Vehicle

No	Parameter	Description
1	Response Time	The lapse time between a user request the service and the response received by the user
2	Reliability	The probability that a user request is handled correctly within the minimum response time
3	Availability	The probability that a service is available for handling the request of a user.
4	Cost	The work need to do by resource for the execution of service
5	Latency	Time is taken by a service request to travel across the network from the requester start point to the provider endpoint

2.3. Machine Learning

Machine learning is a technology that is part of artificial intelligence. Learning is how to get knowledge and apply it. How is human learning adopted by machines or computer devices. Machine learning is growing very rapidly from the side of the algorithm to the implementation of various things. Machine learning is able to replace the functions performed by statistics with better results. The modeling, classification, prediction and functions that statistics normally do can be replaced by Machine learning. Taxonomy algorithm and the use of machine learning can be described as follows

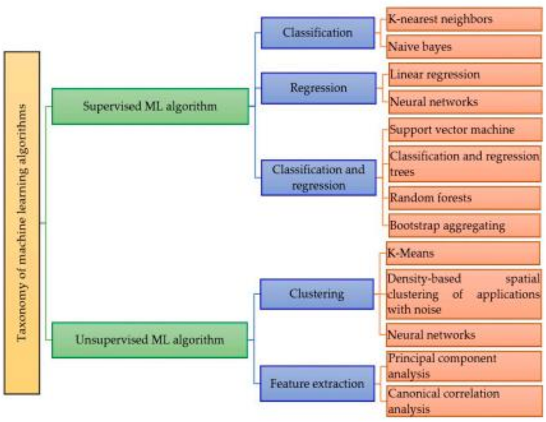

Figure 1. Machine Learning Taxonomy (Alsharif et al., 2020)

The use of machine learning on the internet of vehicles has been widely used. The implementation of machine learning in IoV, among others, is a system to determine the characteristics of traffic (congestion prediction, finding the shortest path). Machine learning is used in managing traffic light, determining the identity of the vehicle and other functions.(Haghighat et al., 2020)

Prediction and classification capabilities of machine learning algorithms can provide a solution to the Quality of Service problem from the Internet of Vehicle. The condition of network infrastructure is not evenly distributed, the quality still needs to be improved, making IoV implementation quite worrying. when a guarantee of IoV reliability is needed, of course, a high level of QoS is needed from the network infrastructure.

3. Materials and Methods

3.1. Hypothesis

IoV in its implementation aims to provide security and comfort for vehicles, drivers and passengers. Security means avoiding accidents caused by other drivers or natural disasters (rain, floods, landslides and others). Convenience means that the vehicle, driver and passenger can reach their destination in a fast time, travel is comfortable and the supply of entertainment and information is well-available.

Diverse needs certainly demand support from infrastructure. GPS devices require communication that is always available to obtain accurate position information. Entertainment supplies require connections capable of streaming large capacities of data. Controlling the ECU (Electronic Control Unit) requires a high transfer rate, because a slight delay will interfere with engine action.(Hussain et al., 2019)

To get security and convenience, reliability and availability of network infrastructure are needed. With the real condition that the infrastructure is not evenly distributed, while transportation does not know the position and location. Need a guarantee of the QoS of the network infrastructure or a warning if the vehicle unit will pass through a path where the QoS of the network infrastructure is low. This is necessary to avoid the risk of IoV failure and also as a reference for scaling up networking.

The method that can be used is a prediction system that is able to provide QoS information on the path to be passed by IoV. These predictions can be built using Machine Learning

3.2. Methods

Machine learning is like a human mindset. The first is to study what happened, then do an analysis to make a decision so that a solution is obtained. In the case of QoS prediction of a network infrastructure, there are several references, namely whether the location has no signal or blank spot or whether the location is overloaded.

The human mindset will see whether the location is far from settlements, whether it is busy time, whether the weather is good, whether it is blocked by buildings or hills and other information.

Referring to this, the first step of the research is to build the parameters that affect the quality of the network infrastructure. A number of parameters are determined that are commonly used to judge that the quality of the network infrastructure is good or not.

The second stage is in compiling a data set or a collection of information that captures real data that links the parameters with QoS. The more data collected, the more accurate the prediction will be. The second step is to observe and collect data directly from the field. The data collected is the capital that will be used to build machine learning or prediction systems

After the data set is collected, the next step is to test whether the data set is valid as a reference for building a machine learning. the test method can use K Fold Validation. This test is to determine the method that will be used to build a prediction system.

Based on the taxonomy in Figure 1, the QoS prediction case falls into the supervised learning category because there is a data set that helps the pattern formation process. The function of machine learning is to classify whether the location points with the parameters attached to them are classified as good points for the network infrastructure or not.

The next step is to determine the algorithm to be used. According to the taxonomy in Figure 1, the algorithm that can be used is Naive Bayes or K Neigborhood. In this research, the Naive Bayes algorithm will be used. This selection is based on the format of the existing data set. All variables use levels or labels that represent values/numbers.(Portugal et al., 2018)

The last stage is the implementation of the prediction system that has been built with Naive Bayes. Implementation is done using the rapid miner application. The resulting naive Bayes rule or algorithm is then tested using existing data sets to determine the accuracy of predictions. Naive Bayes is compute the probability of each category for the given data based on the assumption of the independence between features.(Binkhonain & Zhao, 2019)

The naive Bayesian classifier implies that the classes of the learning dataset are known and provided, hence the supervised nature of the tool. This classifier is one of the simplest methods of supervised learning based on Bayes' theorem as explained in. (Messaoud et al., 2020)

4. Results and Discussion

The data used for training in making rules QoS level prediction in this study consisted of 5 variables referring to on the availability of data and journal references. Variable These are location, vehicle density, topography, signal and weather. (Muhammad Mudjiono, n.d.)

Table 2 . Variable Definition

Location	the position of the vehicle in an open or closed area
Vehicle density	density of vehicles due to traffic jam
Topografy	vehicle location in mountains or valleys or normal
Signal	the signal strength received by the device at the location point
weather	weather conditions like rain or cloudy or sunny

The next step is to classify each variable. This classification is needed to be able to calculate the probability of the condition of each variable. In the Naive Bayes algorithm, this probability is used to reduce the target variable, which is connected or not connected.

Table 3. Classification

Variable	Class
Location	- Open Area - Closed Area
Vehicle density	- Heavy traffic - Traffic Jam - Lite traffic
Topografy	- Mountain - Normal - Valley
Signal	- Strong - Medium - Weak
weather	- Rain - Cloudy - Sunny

The next stage is compiling data sets from surveys in the field. One data pair (record) contains 5 (five) entity variables and one target variable. The target variable is the target you want to predict. The target variable in question is connected or unconnected. The amount of data collected is 526 records resulting from testing through the installation of GPS Tracking and entering weather conditions, topography and vehicle locations. The sample data are as follows:

Table 4. Sample Data Set

Location	Density	Topografy	Signal	Weather	Connection
Open	Lite	Normal	Weak	Rain	Not
Closed	Lite	Valley	Weak	Rain	Not
Open	Heavy	Valley	Strong	Cloud	Connect
Open	lite	Normal	Medium	Sunny	Connect

The data set is a sample of 5 records from 526 collected data. Furthermore, to test the validity of the data whether it can be used as training data, testing is carried out first. The method used is *K Fold Cross Validation*. Using K Fold Cross Validation, the test results are shown in table 4.

Table 5. K Fold Cross Validation Result

No	K	Test Data	Train Data	Accuracy
1	10	50	450	75%
2	20	25	475	72%

The test results using K Fold Cross Validation obtained a fairly good accuracy so that the existing data set can be used to train naive bayes.

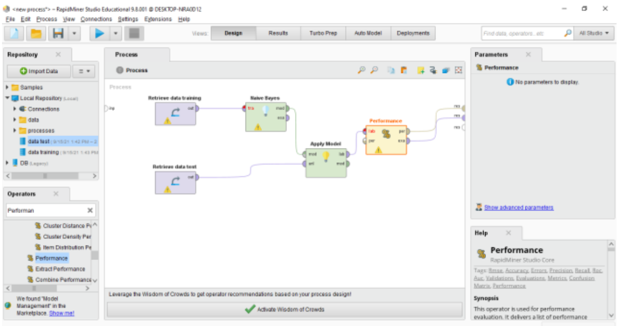

Figure 2. Rapid Miner

In this study, the Rapidminer application was used to test Data Training. Through this application, you can easily measure the accuracy (performance) of naive bayes.
After the data set is tested and proven to be valid and has sufficient accuracy, then an application is made using the Naive Bayes probability formula and the data set that has been tested. This application is used to predict potential connection loss at a certain location and time.

5. Conclussion

The research conducted resulted in several conclusions and suggestions for development. The naive bayes algorithm is able to calculate the prediction of 5 variables problems owned by the Internet of Vehicle (IoV). The collected data sets are able to provide a fairly good prediction accuracy.
Naive Bayes algorithm and the collected data set can be relied upon in predicting the IoV implementation at a certain point in time.
The development of this system is very possible to do. Integration with the user interface, big data and sensors will provide a real time prediction system for the Internet of Vehicles. By looking at the display, the driver can decide whether or not to go through the lane, related to the predicted QoS in that lane.

References

Abosaif, A. N., & Hamza, H. S. (2020a). Quality of service-aware service selection algorithms for the internet of things environment: A review paper. *Array*, *8*(September), 100041. https://doi.org/10.1016/j.array.2020.100041

Abosaif, A. N., & Hamza, H. S. (2020b). Quality of service-aware service selection algorithms for the internet of things environment: A review paper. *Array*, *8*(September), 100041. https://doi.org/10.1016/j.array.2020.100041

Abu-safe, A. N., & Elrofai, S. E. (2020). QOS – Aware Meta-Heuristic Services Selection Algorithm and Likert Scale Measurement for IOT Environment. *International Journal of Computer Science Trends and Technology (IJCST)*, *8*(1), 1–8.

Alsharif, M. H., Kelechi, A. H., Yahya, K., & Chaudhry, S. A. (2020). Machine learning algorithms for smart data analysis in internet of things environment: Taxonomies and research trends. *Symmetry*, *12*(1). https://doi.org/10.3390/SYM12010088

Anwar, N., Adhy, D. R., Tjahjono, B., & Hermawan, R. (n.d.). *Reliability Analysis of Communication Network Service Quality For Internet of Vehicles (IoV)*.

Binkhonain, M., & Zhao, L. (2019). A review of machine learning algorithms for identification and classification of non-functional requirements. *Expert Systems with Applications: X*, *1*. https://doi.org/10.1016/j.eswax.2019.100001

Brogi, A., & Forti, S. (2017). QoS-aware deployment of IoT applications through the fog. *IEEE Internet of Things Journal*, *4*(5), 1185–1192. https://doi.org/10.1109/JIOT.2017.2701408

Butt, T. A., Iqbal, R., Shah, S. C., & Umar, T. (2018). Social Internet of Vehicles: Architecture and enabling technologies. *Computers and Electrical Engineering*, *69*, 68–84. https://doi.org/10.1016/j.compeleceng.2018.05.023

Contreras-Castillo, J., Zeadally, S., & Guerrero-Ibanez, J. A. (2018). Internet of Vehicles: Architecture, Protocols, and Security. *IEEE Internet of Things Journal*, *5*(5), 3701–3709. https://doi.org/10.1109/JIOT.2017.2690902

Guan, W. (2018). Reliability Analysis of the Internet of Things Based on Ordered Binary Decision Diagram. *International*

Journal of Online Engineering, *14*(8), 20–35.

Haghighat, A. K., Ravichandra-Mouli, V., Chakraborty, P., Esfandiari, Y., Arabi, S., & Sharma, A. (2020). Applications of Deep Learning in Intelligent Transportation Systems. In *Journal of Big Data Analytics in Transportation* (Vol. 2, Issue 2). Springer Singapore. https://doi.org/10.1007/s42421-020-00020-1

Hussain, S. A., Yusof, K. M., Hussain, S. M., & Singh, A. V. (2019). A Review of Quality of Service Issues in Internet of Vehicles (IoV). *Proceedings - 2019 Amity International Conference on Artificial Intelligence, AICAI 2019*, 380–383. https://doi.org/10.1109/AICAI.2019.8701299

Kaiwartya, O., Abdullah, A. H., Cao, Y., Altameem, A., Prasad, M., Lin, C. T., & Liu, X. (2016). Internet of Vehicles: Motivation, Layered Architecture, Network Model, Challenges, and Future Aspects. *IEEE Access*, *4*, 5356–5373. https://doi.org/10.1109/ACCESS.2016.2603219

Kumar, S., & Singh, J. (2020). Internet of Vehicle Over Vanets: Smart and Secure Communication Using IoT. *Scalable Computing: Practice and Experience*, *21*, 425–440.

Ligo, A. K., & Peha, J. M. (2017). Spectrum Policies for Intelligent Transportation Systems. *TPRC - Research Conference on Communications, Information and Internet Policyet Policy*, *1547237*, 1–23.

Liu, Y., Cheng, D., Wang, Y., Cheng, J., & Gao, S. (2018). A Novel Method for Predicting Vehicle State in Internet of Vehicles. *Mobile Information Systems*, *2018*. https://doi.org/10.1155/2018/9728328

Messaoud, S., Bradai, A., Hashim, S., Bukhari, R., Tran, P., Qung, A., Ahmed, O. Ben, & Atri, M. (2020). A Survey on Machine Learning in Internet of Things : Algorithms , Strategies , and Applications Affiliations : *Internet of Things*, 100314. https://doi.org/10.1016/j.iot.2020.100314

Muhammad Mudjiono. (n.d.). *ANALIS PREDIKSI PENERIMAAN LEVEL SINYAL PADA DAERAH SUB URBAN TERHADAP UNJUK KERJA JARINGAN GSM Program Studi Teknik Elektro Jurusan Teknik Elektro Fakultas Teknik Universitas Tanjungpura Pontianak Email : mudji.ingback@gmail.com*.

Portugal, I., Alencar, P., & Cowan, D. (2018). The use of machine learning algorithms in recommender systems: A systematic review. *Expert Systems with Applications*, *97*, 205–227. https://doi.org/10.1016/j.eswa.2017.12.020

Sirohi, D., Kumar, N., & Rana, P. S. (2020). Convolutional neural networks for 5G-enabled Intelligent Transportation System: A systematic review. *Computer Communications*, *153*(February), 459–498. https://doi.org/10.1016/j.comcom.2020.01.058

T. Manivannan et al., T. M. et al. (2020). Preventive Model on Quality of Service in IOT Applications. *International Journal of Mechanical and Production Engineering Research and Development*, *10*(3), 1247–1264. https://doi.org/10.24247/ijmperdjun2020109

Vaidya, B., & Mouftah, H. T. (2020). IoT Applications and Services for Connected and Autonomous Electric Vehicles. *Arabian Journal for Science and Engineering*, *45*(4), 2559–2569. https://doi.org/10.1007/s13369-019-04216-8

Xu, W., Zhou, H., Cheng, N., Lyu, F., Shi, W., Chen, J., & Shen, X. (2018). Internet of vehicles in big data era. *IEEE/CAA Journal of Automatica Sinica*, *5*(1), 19–35. https://doi.org/10.1109/JAS.2017.7510736

Zantalis, F., Koulouras, G., Karabetsos, S., & Kandris, D. (2019). A review of machine learning and IoT in smart transportation. *Future Internet*, *11*(4), 1–23. https://doi.org/10.3390/FI11040094

Zhang, T., Liu, X., Luo, Z., Dong, F., & Jiang, Y. (2019). Time series behavior modeling with digital twin for Internet of Vehicles. *Eurasip Journal on Wireless Communications and Networking*, *2019*(1), 1–11. https://doi.org/10.1186/s13638-019-1589-8

Zhang, W., & Xi, X. (2016). The innovation and development of Internet of Vehicles. *China Communications*, *13*(5), 122–127. https://doi.org/10.1109/CC.2016.7489980