Study on information system infrastructure in Probolinggo by using Self Organizing Map

E Rohadi¹,*, A Amalia², R F D Qoyyimah¹, R Ardiansyah¹ and I Siradjuddin²

¹Information Technology Department, State Polytechnic of Malang, Jl. Soekarno Hatta No. 9 Malang 65141, Indonesia
²Electrical Engineering Department, State Polytechnic of Malang, Jl. Soekarno Hatta No. 9 Malang 65141, Indonesia

*erfanr@polinema.ac.id

Abstract. The infrastructure of information systems becomes one of the main components that support government managements system. This study is addressed to achieve the infrastructure information mapping in District of Probolinggo. The proposed of the research is getting an infrastructure information map of 29 districts administered that become its reliability recommendation. The system information as website is developed for operational purpose. The Self Organizing Map (SOM) as one of neural networks is used. The clustering results by in field observation data show that the infrastructure information system of 29 districts administered are defined by 3 clusters, specifically 3 districts administered have good infrastructure, 20 have fair infrastructure and 6 districts administered have less infrastructure. Furthermore, the clustering by using SOM shows that have 1 good, 22 fair and 5 less, respectively. As a result, the accuracy of both clustering is less than 70 %. That means, SOM still need in field observation data to evaluate the recommendation.

1. Introduction
The role of information systems is very important in an organization. Practically, all organizations use computers in their activities, even in the smallest businesses, technology usually has a role in the administration, marketing, or in maintaining customer relationships. Currently, organizations rely on high-level technology in its implementation, even small technical damage can have an impact on operations and have a large impact on the entire organization. In many cases, operations stop altogether when technology does not function as expected [1-3].

The need for information systems infrastructure with robust, reliable and smoothly operating information systems in all business operations and activities becomes important. Infrastructure is the general term for the basic physical systems of a business, region, or nation, for instance, transportation systems, communication networks, sewage, water, and electric systems are all examples of infrastructure. These systems tend to be capital intensive and high-cost investments. The development of infrastructure aims to improve the welfare of the community [2].

In this study, the information systems management specifically on information systems infrastructure in Probolinggo discussed. Probolinggo as one of the cities experiencing growth in the infrastructure sector wants to improve the quality of its services by developing telecommunications infrastructure due to the high capability IT network development. In this city, the process of reviewing IT infrastructures such as the internet network and information systems of each district in government
institutes still done by checking manually. Therefore, the development of both the telecommunications network sector and the information system are very necessary. This information system utilizes data clustering or grouping the data obtained from data mining that can develop and monitor infrastructure development not only in urban areas but also in rural areas, especially in border areas.

Neural networks considered one of the most popular techniques in this field of research. The Kohonen card or Self Organizing Map (SOM) forms a very well-structured neural network model [4-6]. In previous study stated that this method has a pretty good accurate level for classifying a place, region, region, object, etc. [7]. This study is addressed to achieve the infrastructure information mapping in District of Probolinggo using the Self Organizing Map (SOM). The proposed study is getting an infrastructure information map of 29 districts administered that become its reliability recommendation. The system information as website is developed for operational purpose. By creating an infrastructure mapping information system in Probolinggo, it assist local governments in grouping an area in real-time, also it can assist local governments in utilizing the development of an area based on the fiscal year budgeting.

2. Methods

The Self Organizing Map (SOM) is one of the most popular neural network models. It belongs to the category of competitive learning networks. The SOM based on unsupervised learning in order to replacing human intervention during the learning and that little need to be known about the characteristics of the input data. In general applications, the SOM for clustering data is used without knowing the class memberships of the input data. The SOM can be used to detect features inherent to the problem and thus has also been called SOFM, the Self Organizing Feature Map. The Self-Organizing Map was developed by professor Kohonen [4-6]. The SOM has been proven useful in many applications [8].

The system uses the SOM method to classify an area based on the spread of internet access data in Probolinggo. Internet access distribution data used for an input vector from this clustering process, then SOM will form output neurons according to the expected number of clusters into a grid. After the training process completed, each input vector will be mapped to the cluster according to the closest weight. The advantage of using SOM is that the method can automatically cluster nodes [9,10].

![System design](image_url)

Figure 1. System design.

The purpose of this study is to create an information system that can provide information in the form of grouping and mapping infrastructure in real-time manner. By using this information system, it can facilitate the local government in the distribution of IT infrastructure in Probolinggo. Figure 1 shows the design of the information system infrastructure in Probolinggo by using SOM. The real data from Probolinggo Office of Communication and Information used are internet network stability data in each urban district administered, information system data and infrastructure data. Furthermore, the system
will classify areas based on urban district administered that have the results of equitable distribution of infrastructure, information systems, and internet stability networks that are good, fair and less. The flowchart of the information system infrastructure is shown in Figure 2.

![Flowchart](image_url)

**Figure 2.** The information system infrastructure flowchart.

3. Results and discussion

3.1. The implementation of Self Organizing Map (SOM)

For the purpose of this study, data were collected from 29 districts administered including the information on internet network stability data, information system data and infrastructure. The data must be normalized for clustering calculations. Then the system initiates the initial cluster centre point value. The value of the learning rate and the determination of the starting point of the cluster are done randomly. The learning rate itself is a function of decreasing the level of learning as time changes. Euclidean Distance is used to find the closest distance to each cluster by adding up the value of each attribute in the data with the specified cluster weight value, then looking for which weight has the smallest value of the sum [11]. The smallest distance is chosen as the winning neuron. Each winning neuron and its neighbours will carry out an adaptation process that is renewing the weight value.

The system will determine the results of the clustering when the results of the clustered index are the same as the previous iteration. If not, it will carry out a calculation process like the initial steps until the results of clustering from the previous iteration and the current iteration is the same. The iteration process will stop if the clustered index from the previous iteration has the same result. Furthermore, how accurately the system can work is by making comparisons between in field observation and by using the SOM method as shown in Table 1. The clustering results by in field observation show that the infrastructure information system of 29 districts administered are defined by 3 clusters, specifically 3 urban districts administered have good infrastructure, 20 have fair infrastructure and 6 have less infrastructure. At the same time, the clustering by using SOM calculations show that 1 urban district administered has good infrastructure, 22 have fair infrastructure and 5 have less infrastructure. As a result, the accuracy of both clustering is less than 70%.
| No | District Administered | Network Stability | Number of Application | Infrastructure | Field Observation | SOM | Comparable |
|----|----------------------|-------------------|-----------------------|----------------|------------------|-----|------------|
| 1  | Kademangan           | 82.83             | 4                     | 10.20          | 2                | 2   | Match      |
| 2  | Kedung Lor           | 90.91             | 4                     | 95.92          | 2                | 3   | Not Match  |
| 3  | Poh Sangit Kidul     | 71.72             | 4                     | 10.20          | 3                | 2   | Not Match  |
| 4  | Triwung Kidul        | 98.99             | 4                     | 10.20          | 2                | 2   | Match      |
| 5  | Triwung Lor          | 72.73             | 4                     | 10.20          | 3                | 2   | Not Match  |
| 6  | Curah Grinting       | 60.61             | 8                     | 10.20          | 3                | 2   | Not Match  |
| 7  | Kanigaran            | 100               | 24                    | 10.20          | 2                | 2   | Match      |
| 8  | Tisnonegaran         | 96.97             | 60                    | 70.41          | 1                | 3   | Not Match  |
| 9  | Kebonsari Kulon      | 56.57             | 100                   | 100            | 1                | 1   | Match      |
| 10 | Kebonsari Wetan      | 64.65             | 97.96                 | 2              | 3                |     | Not Match  |
| 11 | Sukoharjo            | 98.99             | 4                     | 20.41          | 2                | 2   | Match      |
| 12 | Jrebeng Kulon        | 83.84             | 4                     | 10.20          | 2                | 2   | Match      |
| 13 | Jrebeng Lor          | 71.72             | 4                     | 10.20          | 3                | 2   | Not Match  |
| 14 | Jrebeng Wetan        | 60.61             | 12                    | 10.20          | 3                | 2   | Not Match  |
| 15 | Karedeng Lor         | 100               | 4                     | 10.20          | 2                | 2   | Match      |
| 16 | Kedung Wetan         | 100               | 97.96                 | 1              | 3                |     | Not Match  |
| 17 | Sumber Wetan         | 82.83             | 4                     | 10.20          | 3                | 2   | Not Match  |
| 18 | Jati                 | 100               | 8                     | 77.55          | 2                | 3   | Not Match  |
| 19 | Mangunharjo          | 100               | 8                     | 30.61          | 2                | 2   | Not Match  |
| 20 | Mayangan             | 90.91             | 68                    | 30.61          | 2                | 2   | Match      |
| 21 | Sukabumi             | 98.99             | 28                    | 20.41          | 2                | 2   | Match      |
| 22 | Wiroborang           | 85.86             | 36                    | 30.61          | 2                | 2   | Match      |
| 23 | Jrebeng Kidul        | 87.88             | 4                     | 10.20          | 2                | 2   | Match      |
| 24 | Kedung Asem          | 95.96             | 4                     | 10.20          | 2                | 2   | Match      |
| 25 | Kedung Galeng        | 82.83             | 4                     | 10.20          | 2                | 2   | Match      |
| 26 | Pakis Taji           | 92.93             | 4                     | 10.20          | 2                | 2   | Match      |
| 27 | Sumber Taman         | 77.78             | 4                     | 10.20          | 2                | 2   | Match      |
| 28 | Wonoasih             | 100               | 4                     | 10.20          | 2                | 2   | Match      |

In this section, the results of system implementation and system testing discussed. Figure 3 shows the clustering calculation where the user can see the results of the clustering calculation in detail, from the normalization results, the calculation on each iteration that contains distance data, cluster, and weight updates. While the results of clustering using SOM on the Probolinggo map can be seen by marking the area based on cluster results as shown in Figure 4. The infrastructure information system defined by 3 clusters, where cluster 1 is green which means it has good infrastructure, cluster 2 is blue which means it has fair infrastructure and cluster 3 is red which means it has a less infrastructure.
4. Conclusion
The study on information system infrastructure in Probolinggo by using Self Organizing Map (SOM) has been designed and evaluated. This system has successfully implemented SOM method to determine the information system infrastructure by marking the area based on cluster results. The clustering results show that the infrastructure information system of 29 districts administered are defined by 3 clusters, specifically 1 urban district administered has good infrastructure, 22 have fair infrastructure and 6 have less infrastructure. While the clustering results by in field observation shows that have 3 good, 20 fair and 6 less, respectively. As a result, the accuracy of both clustering is less than 70 % and can be used as an application in determining information system infrastructure in Probolinggo.

Acknowledgments
The authors would like to thank State Polytechnic of Malang for supporting this research.
References

[1] Wraikat H, Bellamy A and Tang H 2017 Exploring organizational readiness factors for new technology implementation within non-profit organizations Open J. of Social Sciences 5 pp 1-13

[2] Kanyuga L W 2019 The strategic role of information technology on organization financial performance: a case of xyz paper industry limited, Kenya J. Inf. and Tech. 3 pp 1-14

[3] Sirkemaa S 2017 Information systems infrastructure–importance of robustness Int. Conf. on Information Technology Science Springer, Cham pp 241-244

[4] Lasri R 2016 Clustering and classification using a self-organizing MAP: The main flaw and the improvement perspectives SAI Computing Conf. (SAI) pp 1315-1318

[5] AghaeiRad A, Chen N and Ribeiro B 2017 Improve credit scoring using transfer of learned knowledge from self-organizing map Neural Computing and Applications 28 pp 1329-1342

[6] Ben Khalifa K., Blaiech A G and Bedoui M H 2019 A novel hardware systolic architecture of a self-organizing map neural network Computational intelligence and neuroscience 2019

[7] Wulandari K and Fauzy A 2016 Implementasi self organizing maps untuk clustering ketahanan dan kerentanan pangan desa di kabupaten magetan 2014 Jurnal Matematika 1

[8] An W 2018 Approach to golf swing model using hadoop and self-organizing map in big data environment Int.Inf. Institute (Tokyo) 21 pp 1897-1905

[9] Sharma A, Sharma M K and Dwivedi R K 2018 Analyzing facebook data set using self-organizing map Int. Conf. on System Modeling & Advancement in Research Trends (SMART) pp 109-112

[10] Delgado S, Higuera C, Calle-Espinosa J, Morán F and Montero F 2017 A SOM prototype-based cluster analysis methodology Expert Systems with Applications 88 pp 14-28

[11] Lamb D S, Downs J and Reader S 2020 Space-time hierarchical clustering for identifying clusters in spatiotemporal point data ISPRS Int. J. of Geo-Information 9 p 85