Jurnal Nasional Teknik Elektro dan Teknologi Informasi
Volume 13 Number 3 August 2024

DOI: 10.22146/jnteti.v13i3.9872

The Development of Intelligent Web for Rural Social E-Learning

Seno Adi Putra[1], Timmie Siswandi[1], Dessy Yussela[1], Rinez Asprinola[1], Erin Karina[1], Mega Candra Dewi[1], Santi Al-arif[1]

[1] Laboratory for Enterprise Intelligent System, School of Industrial and System Engineering, Telkom University, Bandung, Jawa Barat 40257, Indonesia

[Received: 26 September 2023, Revised: 8 December 2023, Accepted: 5 July 2024]
Corresponding Author: Seno Adi Putra (email: adiputra@telkomuniversity.ac.id)

ABSTRACT — Social media technology affects the learning paradigm change towards social media-based learning, known as social e-learning. Social e-learning regards a person as a center of learning, dubbed people-centered learning. Here, people are encouraged to interact or communicate with others and produce their learning content. This work attempted to provide a solution model for rural e-learning social learning empowered with intelligent web technology. The proposed social e-learning includes several modules for development, such as personal space, collaboration space, and communication space modules. It also leverages intelligent web technologies currently implemented in today's social media applications, such as article search, article recommendations, friendship recommendations, and document classification. In the searching module, the PageRank method was used to calculate the relevance score to determine the rating of the documents or articles. The similarity-based element calculation method was utilized to create articles' suggestions and recommendations. The naïve Bayes algorithm, decision tree, and neural network were compared to find the best solution for article classification in agriculture, fisheries, animal husbandry, and plantations. When comparing these three algorithms, the result showed that the neural network was the most accurate classification, reaching 95.2% accuracy. A clustering algorithm, namely robust clustering using links (ROCK), was utilized for rural friendship recommendation. Thus, these algorithms (the PageRank, the similarity-based element, neural network, and the ROCK) were suitable and recommended for supporting intelligent web paradigms in social e-learning applications.

KEYWORDS — Social E-Learning, Intelligent Web, PageRank, Similarity Score, ROCK Algorithm, Naïve Bayes, Decision Tree, Neural Network.

I. INTRODUCTION

Internet has significant impacts on social community learning. The evolution of e-learning enables people to gather information on the Internet anytime and anywhere. Today, e-learning platforms, such as ELM-ART and AHA, have been widely used by people worldwide to support their learning [1]. Here, the learning approach has changed from teacher-centered learning to learner-centered learning using various methods, e.g. e-learning, mobile learning, or blended learning. This approach allows learners to access broad learning resources and peer interaction, providing flexible learning methods. At the same time, it increases learners' motivation by encouraging them to be more responsible for their learning and challenging them to share their knowledge and experience during the learning process [2].

Today, e-learning is considered as a prominent technology for enhancing the quality of learning. The term e-learning has become very popular and has been accepted as a tool for the teaching and learning process [3]. E-learning is becoming an essential and widespread technology that provides easy and cost-effective resource sharing. It is implemented in many educational institutions and universities worldwide. It offers many benefits, including accessible content material, team collaboration, and appropriate joint discussions [4].

A large number of social media users gives potential opportunities to develop informal online learning applications, especially for rural communities where all users can communicate with others and generate learning content related to their expertise and experience. Communication and collaboration among users to generate learning content in specific knowledge or experience are the terminology of the future e-learning, called social e-learning, which allows learners to develop social interaction in the learning process [5].

The e-learning framework refers to the knowledge-sharing process among people regardless of Internet technologies' geographic boundaries and limitations [6]. Several popular e-learning platforms are used today. Pluralsight is one of the top digital platforms for learning technical skills. It provides complete courses developed by experts in several fields. It also has a wide range of information technology and computer science subjects. Learners can pursue careers in information security, telecommunications, coding, software engineering, web design, etc., in this e-learning platform. They will receive a certificate after completing the course. Pluralsight has improved its video content and audio quality requirements [7]. Another e-learning platform is CodeAcadeny, which is one of the largest online learning sites for programmers. It is a website where people can learn all programming languages. At CodeAcademy, learners can communicate with CodeAcademy members and teachers.

Thinkific is a digital learning platform that allows users to build and sell courses quickly. This platform aims to facilitate users who like teaching rather than learning. It offers services to simplify course views and add content like videos, quizzes, digital tutorials, assessments, documents, audio, and complete reviews [7]. Udemy has some specialized courses that make it different from others. It provides easy services for learners. The user can find their preferred course skills in video format.

However, most advanced programs require payment. Like other platforms, users receive a Udemy certificate upon completion of course although it will not be admitted by the company.

Coursera is an online learning that offers free courses. Here, leading universities and companies offer courses that can be enrolled by everyone. It is an ideal platform for academics, practitioners, and business people. Registered users can access a large number of courses and offer quizzes as well as homework assistance from other classmates. Coursera is very well organized and works closely with the world's leading academic institutions and companies [7].

Skillshare is an ideal e-learning platform for designers and visual artists. Here, large groups and new courses are introduced all the time. Nevertheless, in order to access all of the available courses, users must pay a monthly fee. Skillshare classes are classified into four main types: business, technology, creative, and lifestyle. This e-learning does not provide qualifications and certifications because it focuses on providing practical information that can be applied later in the career rather than certification. Since Skillshare is a group, most courses are developed by the participants. Users can communicate with other members and obtain support and advice from them [7].

Tiger's eLearning systems can enhance distance education by providing online lecture materials such as presentations, documentation, and videos [6]. The system is efficient and flexible and provides both external and internal libraries. In 2009, this e-learning system was developed into an application called CalStateLA. It offers data or note storage service, uploading a video, image, audio, file service, bulletin board, calendar, email, account management, Wikipedia search, and course management. Today, machine learning is used in e-learning systems. The summaries of artificial intelligence that can be implemented in e-learning are described in Table I [8].

TABLE I
LIST OF ARTIFICIAL INTELLIGENT IMPLEMENTED IN E-LEARNING

Method	Implementation
Hidden Markov model	User opinion
ADTree algorithm, Apriori association rule algorithm, and simple k-means algorithm	Course recommendations
Genetic algorithm	System-to-learners response optimization
Support vector machine (SVM) algorithm	Student performance knowledge
K-nearest neighbor (K-NN) algorithm	To determine student's emotions
Ensemble classifier bagging algorithm embedded with machine learning (ML)	Online session assessment
Error-correcting output codes (ECOC) combined classifier	Student ranking credits
Bayesian estimation algorithm	Learning styles and learning objects
Fuzzy cluster technique algorithm	Learner behavior Sequence
K-means clustering	Learner sequence and pattern Student engagement
Perceptron artificial neural network (ANN) algorithm	Student graduation results
Felder Silverman model algorithm	AUI Features Course
Levenberg–Marquardt algorithm (LMA)	Course information classification
Conv-GRU-AvgP in P-xNN	Learning processing data
Deep Learning TensorFlow engine	Students' assessment
Random forest algorithm	Student test result

Today, e-learning applications still implement conventional paradigms that do not reflect natural learning, like social, personal, open, and dynamic. In addition, today's e-learning applications have not accommodated the trend toward social networking principles and mobility solutions. Social networking collaboration changes online communities and mobility into an interesting delivery channel. This trend affects mobile social network enablement, embedding on-demand social networking and collaboration capabilities. Finally, current e-learning applications have yet to be supported by the intelligent web and collective intelligence paradigm to develop new ways to gain insights, opinions, and perceptions and open interesting community discussions. Therefore, social networks and collaborative solutions can be applied in e-learning. Adopting social networking and collaborative solutions can help increase interaction between people in communities, generating collaboration, expanding networks, and developing a culture of information sharing.

The first contribution of this work is providing a solution model for social e-learning for rural communities, which leverages the social media paradigm. It provides an informal facility for communities to collaborate in learning and share their knowledge and experiences. The personal space module inspired by Facebook, Edmodo, Twitter, LinkedIn, and Instagram was developed. The personal space module is a facility that can be used to establish users' personal knowledge and interactions. The collaboration and communication space module that facilitates user communication and collaboration when they generate learning content according to their communities was also developed.

The second contribution of this work is to empower social e-learning platforms with intelligent web technologies. Users in rural communities are encouraged to generate information and learning content in blogs or wikis. Here, article search is an important feature that should be provided by applying the PageRank method with a relevance score algorithm. Article recommendation is the next feature in e-social learning, which calculates the frequency of reading articles using the similarity-based element method.

The clustering method using the ROCK algorithm is considered to support friendship recommendations based on the similarity of articles' interests in writing and reading. Here, the system automatically groups users who have a similar interest in writing and reading articles so that they can connect with each other as friends. Finally, the article classification using neural networks automatically categorizes articles into the fields of rural expertise, such as agriculture, fisheries, animal husbandry, and plantations.

This paper is divided into five sections. Section II describes social e-learning. Section III describes the proposed solution and model utilized in social e-learning. Section IV is the system evaluation. Section V is the conclusion.

II. SOCIAL E-LEARNING

The effect of today's system and information technology in education overcomes the limitations of space and time for learning. The e-learning system has gained popularity. It is distance learning which involves the use of Internet and computer. E-learning automates parts of the face-to-face learning process; through e-learning, teachers and students do not need to be in the same place simultaneously. However, even though a lot of money has been spent on developing interactive multimedia systems, an essential element, called the presence of other people, has been missing. In Addition, most of the existing e-learning only facilitates users in accessing material without providing facilities to perform social interactions with others and generate users' learning content [9]. This problem is solved by blended learning, although face-to-face interactions are often conducted traditionally [10]. The informal learning process has gained popularity; these types of learning include observing other people, asking a friend next door, calling the help desk, performing trial and error, or working with people who understand the subjects. The academic learning process also uses various informal channels, such as games, simulations, experiments, stories, and discovery.

Over the last few years, the web has shifted from a medium to transmit information to a platform where contents are created, shared, combined, and delivered [10]. This latest generation, which is referred to as Web 2.0, has the characteristics of a user-centric, open, dynamic web, with peer production, sharing, collaboration, collective intelligence, distributed content, and decentralized authority. Social software has emerged as a major component of Web 2.0 development. It can be defined as a tool to enhance human social and collaborative capabilities, as a medium to facilitate social relations and information exchange.

The social software approach provides more advantages to online learning than traditional approaches [11]. This approach reflects the nature of learning, which is called social, personal, distributed, flexible, dynamic, and complex. In addition, this approach will represent a shift in the learning management system model toward more social, personal, open, and dynamic. An interesting trend is the rapid development of social networking and mobility solutions. Social networking collaboration creates online communities, and mobility becomes the preferred delivery channel. It is not only improving a return on investment but also expanding global coverage and improving the operational efficiency of workers in companies or institutions [12]. It enables mobile social networks by embedding on-demand social networking and collaboration capabilities into mobile applications.

Mobile services and social collaboration applications should provide social and collaborative experiences to users, create collective intelligence, and develop new ways to gain insights, opinions, perceptions and open interesting discussions with all communities. Therefore, social networking applications should not be stand-alone applications but also be integrated with applications in the organization [12]. These social networks and collaborative solutions can be applied to mobile learning methods. Adopting social networking and collaborative solutions can help users perform better interactions, expand networks, and share information.

The implementation of social software in e-learning then creates a new generation of e-learning called e-learning 2.0 with the following features [12]: social or collaborative learning environment; users develop content, rather than relying solely on teachers; learning process involving the activity of developing content (generating content) and communicating with others; the presence of aggregating using really simple syndication (RSS) and tagging features; the existence of knowledge sharing (knowledge sharing); the personalization (personal learning environments); the

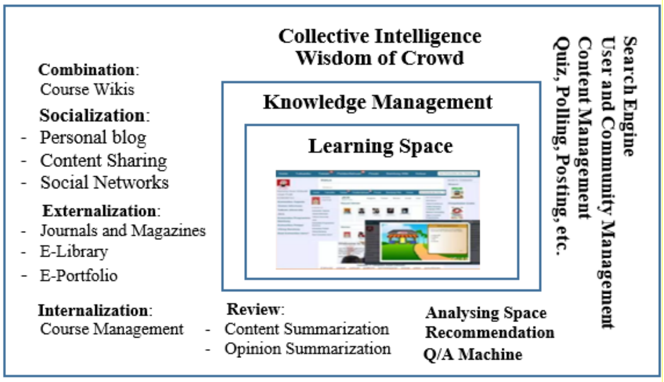

Figure 1. E-Learning 2.0 ecosystem.

collective intelligence (wisdom of the crowd); the use of networks with various technologies; and creativity and innovation development.

E-learning ecosystem 2.0 defines three aspects that social e-learning should provide: learning space, knowledge management, and collective intelligence. The Learning space aspect has five facilities: personal space, collaboration and communication space, media library, search engine, and analyzing space. The e-learning ecosystem 2.0 is shown in Figure 1. The e-Learning 2.0 ecosystem includes the following aspects: (a) learning space consists of personal space (RSS feed, e-portfolio, and personal blog), collaboration and communication space (course blog, course wiki, bookmarking, messaging, content sharing, podcasting, social network, and mashup), media library, search engine, and analyzing space; (b) knowledge management; and (c) collective intelligence.

The intelligent web is a feature that supports social e-learning. It is designed to be trained automatically to understand user input, user behavior, or both, as well as the response. There are three basic elements of the intelligent web [13].

1. Aggregated contents are mostly related to data from a specific application. The collected contents are more dynamic than static, and the content includes the origin and storage locations. These storage locations can be geographically dispersed, presenting potential challenges in data management. Every piece of information in the content is usually associated with other information.
2. A reference structure that provides one or more structural and semantic interpretations of the content. Knowledge is divided into three categories: dictionaries, knowledge bases, and ontologies.
3. Algorithms that allow applications to identify hidden information are employed for abstraction (generalization), prediction, and enhancing interaction with its users. The algorithm is applied to the collected contents, and occasionally necessitates the existence of a reference structure.

An intelligent web in a social e-learning application is classified into three categories: explicit, implicit, and derived intelligence [14]. Explicit intelligence is intelligence obtained from users' explicit information, such as reviews, recommendations, tagging, and voting. Implicit intelligence is intelligence obtained from users' indirect information, such as blogs. Derived intelligence is intelligence obtained from collected user data, such as data and text mining, predictive clustering and analysis, and search or recommendation engines.

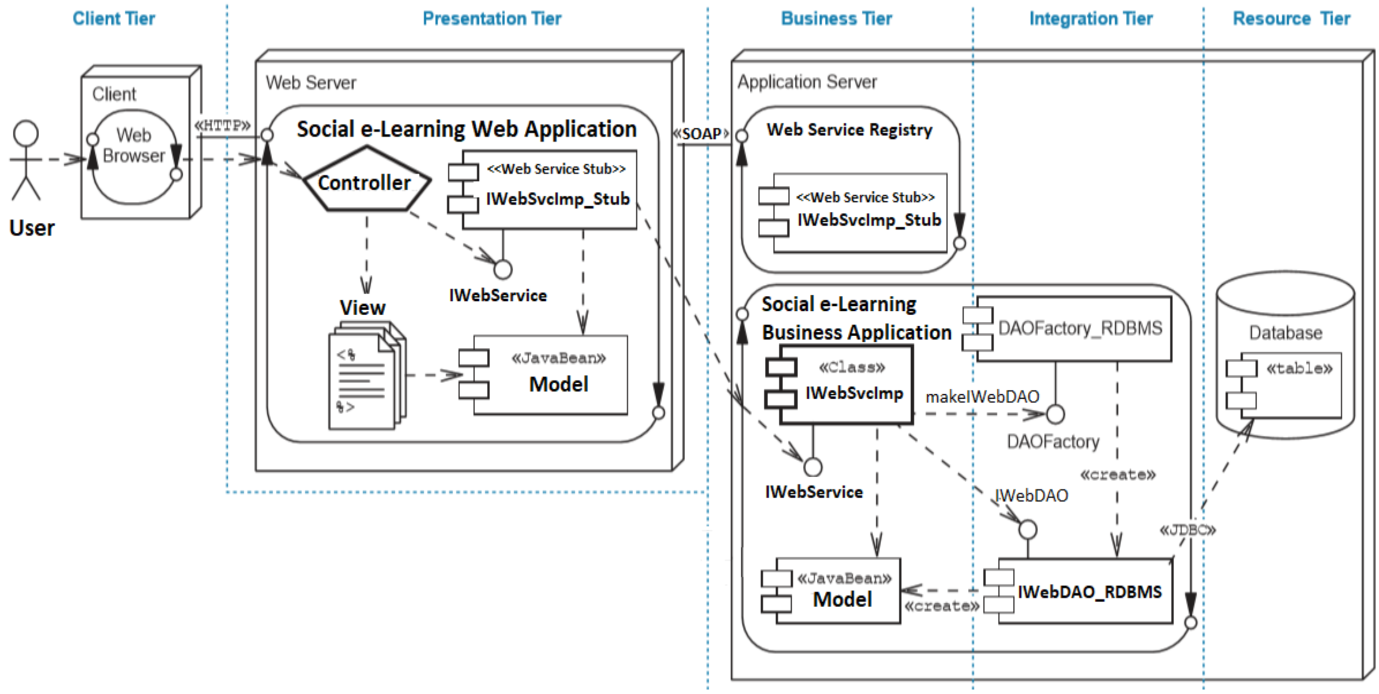

Figure 2. Proposed architecture model in social e-learning system.

III. PROPOSED MODEL AND SOCIAL E-LEARNING

A. PROPOSED FEATURES OF SOCIAL E-LEARNING

Social e-learning was built based on the e-learning 2.0 ecosystem, as shown in Figure 2. The basic features applied to social e-learning applications are collaboration and communication space modules and personal space. The personal space is the following.

1. A personal blog facilitates users sharing their knowledge or experience in the form of a user blog so that everyone can read or learn what they write. Users can write, view, change, add comments, and delete blogs or posts here.
2. Like blogs, users can contribute to writing a public article in Wiki communities. The admin in its community is responsible for checking all content in its Wiki by creating, editing, or deleting irrelevant Wikis.
3. Profile management changes profile photos, passwords, and other data. Each user can view and change their personal data. The data that can be changed are personal data.
4. Timeline and notifications that view and delete notifications. This notice contains activities related to that person. In addition, users can view the notification in the form of a timeline. Each user can communicate using messaging, in which they can send, receive, and delete messages. News management facilitates authored users in a group or application admin to publish news.

The following are the main functions in collaboration and communication space.

1. User management considers the management of users along with their friendships. There are several functions under user management, for example, adding friends, rejecting friend requests, and approving friend requests.
2. Group management manages the groups or communities that users follow. It has several functions, such as creating a group, making an admin, inviting members, leaving the group, kicking members from the group, and asking them to join the group.
3. Course management facilitates knowledge sharing for users in the group they follow. There are several functions under course management, for example, writing blogs, editing blogs, viewing blogs, adding comments, and deleting blogs. Along with course management, there is task management, which manages trivial quizzes in a group. It includes creating quizzes, adding comments, editing quizzes, viewing results, deleting quizzes, and accomplishing quizzes.
4. In content management, users can share images, videos, and presentation slides in the group or community they follow. There are several functions under content management, for example, uploading content, deleting content, viewing content, adding comments, and editing content.
5. Event management is used by users when they want to publish an event in a group or community. Event management includes creating events, deleting events, viewing events, adding comments, editing events, and confirming events.

The system architecture proposed in social e-learning implemented a multitier architecture (see Figure 2). Here, a multitier architecture for separating the concerns was developed to build a high-performance system. Five tiers consisting of the client tier, the presentation tier, the business tier, the integration tier, and the resource tier were then defined. The client tier is the front end, where users directly access the application. The presentation tier handles user requests and delegates tasks to specific business logic or services. Servlet technology and Java Server Page (JSP) were utilized in this work. The architecture was developed based on the model-view-controller (MVC). The MVC pattern is widely used by web-based applications [15]. The MVC pattern can be used to build mobile applications that help developers build cheaper and cleaner code. The framework is based on Java technology.

The business tier primarily contains entity components and service components. Entity components represent the data stored in local storage or sent to a remote system. Service components validate business rules and update entity components. The business tier called the IWebSvcImpl class, implements an intelligence web algorithm, e.g., Page Rank, article similarity method, naïve Bayes algorithm, decision tree, neural networks, and ROCK algorithm. This class is accessed remotely using web service protocol. The integration tier separates the entity components from the resources. Here, a resource can be a database. Data access object (DAO) was used in this work to implement create read use delete (CRUD) operations.

Cloud technology was considered when deploying this system. A cloud system provides a collection of services through virtualized shared resources, which makes it cheaper to operate [16]. The cloud architecture defines an abstract distributed system model with appropriate elements that represent application components and their interrelationship.

B. PROPOSED INTELLIGENT WEB IN E-SOCIAL LEARNING

This section describes intelligent web features that are embedded in the system. The intelligent module includes document search, article recommendation, article classification, and user clustering module.

1) DOCUMENT SEARCH

The PageRank method was utilized for document search. It is a ranking search algorithm on the web that uses hyperlinks. It is known as one of the main features of the Google search engine. In this method, the ranking process was conducted recursively, and the ranking of the web page was determined by ranking other web pages with links to that page.

In this work, a calculation called the relevance score [17] should be performed. This calculation calculated the score for each article after the user searched based on the criteria the application had determined. The relevant article consisted of three parameters: keyword, indexed content, and document quality. Equation (1) was used to calculate the iterative vector value of the relevance score.

$$score = \sqrt{\frac{((loglog\,(o) * log(df) * ssw) + (qw * qs) + (ucw * ucs)}{(maxdq + maxuc + (loglog\,(df) * loglog\,(max)\,))}} \qquad (1)$$

where o is keyword occurrences, df is document frequency or total document containing a keyword, ssw is score sum weight or keyword weight, qw is quality weight or the latest article weight, qs is quality score or score for the latest article, ucw is user click weight, ucs is user click score (value from 1 to 10), $maxdq$ is the maximum document quality or maximum score of the latest article (ranging from 1 to 10), $maxuc$ is maximum user click (ranging from 1 to 10), and max is the maximum value for keyword occurrence (value from 1 to 10).

The PageRank is a page ranking algorithm that can handle a large amount of data. It can show that a page is relevant and important. When more users visit a page, the page will receive a higher rank. In the proposed system, not only articles stored in the database were retrieved, but also articles outside the database by using the crawler.

2) DOCUMENT SUGGESTION AND RECOMMENDATION

The similarity method for article suggestion and recommendation was used. Machine learning algorithm is used to identify similar data [18]. To increase the probability, n is the total element that will be compared, x is the first compared element, and y is the second compared element. This method compares each element, called average first-passage time (one-way), meaning that it can perform the average time on the first pass and calculate the similarity score between the elements in the database. Similarity calculation was conducted by remerging all data features into a single numeric value. In this step, root mean square error (RMSE) was utilized as it calculates the error between the values of observed value and the predicted value. Equation (2) was applied to determine RMSE.

$$RMSE = \sqrt{\frac{1}{n}\left(\sum_{1}^{n}(x^2 + y^2)\right)}. \qquad (2)$$

The recommendation system used the content attribute to define several article recommendations [19]. The similarity algorithm method was utilized to compare the similarities between groups of recommendations. The similarity between two articles was determined by identifying similar data and comparing the number of elements against recommended articles.

3) DOCUMENT CLASSIFICATION

The proposed social e-learning will be implemented in rural communities. The contents published in the system should be rural activities, such as agriculture, animal husbandry, fishery, and plantation.

In this work, the naïve Bayes, decision tree, and neural networks were compared to identify the best classification approach. Naïve Bayes is a method that analyzes simple probabilities by calculating a set of probabilities by combining frequencies and values from certain datasets [20] and [21]. In (3), $p(X)$ denotes prior probability, $p(Y)$ denotes evidence, and $p(Y|X)$ denotes likelihood.

$$p(Y) = \frac{p(Y)p(X)}{p(Y)}. \qquad (3)$$

Decision tree is a machine learning technique for solving both classification and regression problems. It helps to identify relationships between data points in a dataset by constructing a tree structure. This tree-like shape is used to make accurate predictions about invisible data. The dataset is divided into several subsets and therefore generates decision node branches or decision nodes. The first node is called the root node, and the final decision node with no branch is called the leaf node. To determine the best attributes from the selection criteria, the information gain is defined to select attributes for a node. Entropy is also defined to determine the variance of the sample data [22]. The entropy is calculated using (4).

$$Entropy\,(D) = -\sum_{i=1}^{n} p(x_i) log p(x_i) \qquad (4)$$

where D is a discrete random variable, x_i is possible outcomes, and $p(x_i)$ is the probability of x_i occurrence.

Neural networks can also be used to classify things. It is composed of a neuron node connected to other nodes. Each node is responsible for receiving data from the input node or other neuron nodes and sending data to another node [23]. In this case, six neural network inputs were proposed to represent a number of agriculture keywords found in the article title, number of animal husbandry keywords found in the article title, number of fisheries keywords found in the article title, number of agriculture keywords found in the article content, number of animal husbandry keywords found in the article content, and number of fisheries keywords found in the article content (5).

$$y_{1,j} = \sum_{i=1}^{6} w_{1,(3(i-1)+1)} x_i + b_{1,j}$$
$$y_{2,j} = \sum_{i=1}^{3} w_{2,(3(i-1)+1)} y_{1,i} + b_{2,j}. \quad (5)$$

In addition, artificial neural networks (ANN) consist of *n* input layer with 6 inputs, a hidden layer with 3 neurons, and an output layer with 3 neurons connected to 3 outputs. The variable of w_1 is the weight of each line connecting from input to hidden layer, in this case, there were a total of 18 lines. There were three biases of b1 for each neuron in the buried layer. From the hidden layer, there were nine lines of y_1, called the outputs of the hidden layer. In the same lines, there are nine weights called w_2. Finally, the output layer consists of 3 biases of b_2 and 3 outputs of y_2.

4) USERS CLUSTERING

The ROCK clustering algorithm is one of the most famous algorithms for clustering categorical datasets. It is very suitable for categorical data, such as keywords, Boolean attributes, and enumerations. This algorithm works well on large datasets.

The ROCK algorithm was implemented using a dendrogram to determine the cluster structure. The main point of this algorithm was to utilize links as a measure of similarity, rather than relying on distance as a measure. Distance metric (Jaccard coefficient) was used to determine the point connected to a certain point. It was used to compare the number of common terms between descriptions [13].

Dendrogram is the basic structure used during encapsulated clustering. Its structure is a tree data structure that helps capture hierarchical cluster formation. Jaccard similarity or Jaccard coefficient is an algorithm for comparing two documents based on the strings they have (6).

$$JC(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (6)$$

where $JC(A,B)$ is Jaccard coefficient similarity of A and B, *A* and *B* are documents or sets of words.

Goodness measurement was utilized to measure goodness, evaluating whether it was necessary to merge two clusters or not. The best ROCK cluster is the cluster that maximizes the goodness measure value (7)–(9).

$$f = \frac{1.0 - th}{1.0 + th} \quad (7)$$

$$p = 1.0 + 2.0x\, f(th) \quad (8)$$

$$g = \frac{nLinks(X,Y)}{((nX)+n(Y)))^p - (nX)^p - (nY)^p} \quad (9)$$

where *th* is the link threshold, *f* ensures that the link threshold is the threshold for the similarity measure, *p* calculates whether the merge goodness measure is necessary to carry out, *g* is the goodness measure, *X* is the cluster X, and *Y* is the cluster Y.

In this work, clustering was carried out for the friendship recommendation starting from writing an article first on the blog. The articles in the database would be clustered. The results were articles clustered considering the similarities in article keywords written by users. These similarity articles were considered as friendship recommendations.

5) ACCURACY MEASUREMENT

Confusion matrix can be used for measuring the performance of a classification method. Basically, the confusion matrix contains information that compares the classification results carried out by the system with the real classification results.

There are four measurements, namely true positive (TP), true negative (TN), false positive (FP), and false negative (FN). For multiclass, confusion matrix with class K has a confusion matrix K × K. Confusion matrix was applied to evaluate the performance of the classifier on the dataset. It was used to distinguish predicted values from real values of model elements in software engineering, referred to as accuracy. The accuracy value for the K × K matrix was obtained using (10).

$$Accuracy = \frac{TF}{Total\ Prediction} \times 100. \quad (10)$$

IV. SYSTEM EVALUATION

The system will be implemented in all rural areas of Indonesia. The performance evaluation is the primary concern due to the large number of users who may potentially utilize the system.

A. SYSTEM PERFORMANCE EVALUATION

The stress test was conducted to determine the ability of social e-learning applications in dealing with abnormal conditions in terms of the number of users. The system was tested using the JMeter application with several trials of requests or threads accessing the system. The stress test was carried out on two-tier and three-tier architectures to determine the best performance of these two architectures. The social e-learning application was deployed on a server with Intel Xeon CPU E3-1230 V2 3.30 GHz (CPUs) ~3.3GHz, 16 GB memory, and ran on a Windows server.

In this work, the stress test compared the average value or average response time, throughput or total requests that could be handled by the system in seconds, and the percentage of error when handling user requests or sending responses to users. When the number of requests reached 2,000 threads, the system could not handle the requests. It is indicated by the error shown in each architecture. The error rate was 8.22% and 0.73% for the two-tier architecture and the three-tier architecture, respectively.

Figure 3 shows the test result where two types of architectures are examined. Figure 3(a) shows the average value or response time in the two different architectures. The results showed that the average response time of two-tier architecture was smaller than that of the three-tier architecture. The web server and database server were deployed on the same machine in the two-tier architecture. Therefore, the time required to handle the user request was faster. Meanwhile, the user requests that a three-tier architecture could handle were higher than that of the two-tier architecture (Figure 3(b)). In the three-tier architecture, three different machines focused on different concerns. In the two-tier architecture, however, only the web server and database were on the same machine.

B. DOCUMENT SEARCH MODULE EVALUATION

This work implemented the document search module to search documents and articles. The PageRank method using the

p-ISSN 2301–4156 | e-ISSN 2460–5719

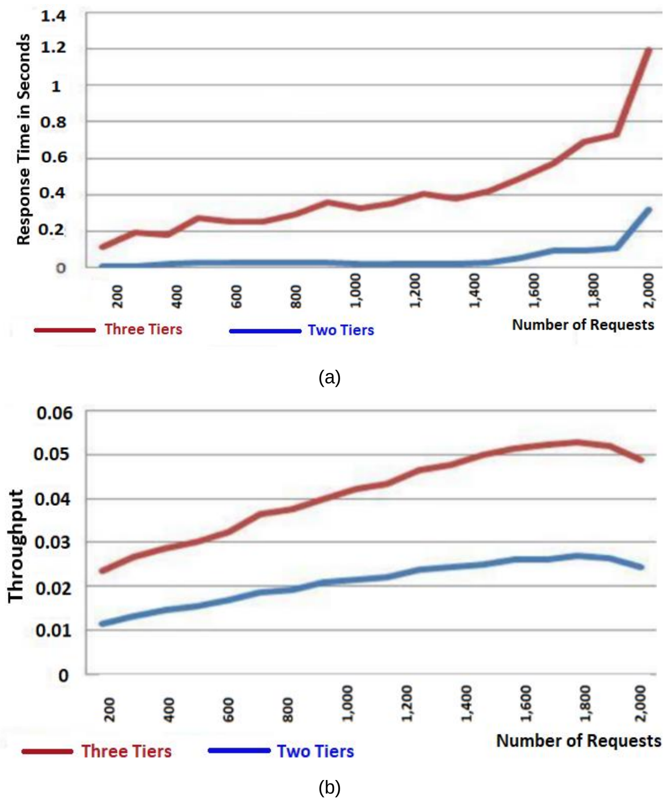

Figure 3. Performance test, (a) response time and (b) throughput.

relevance score algorithm was examined. Users can search articles by inputting keywords in the search field. The search results were articles about users' articles, other people's articles, and news. It facilitates villagers' finding their information, especially about social activities, in the village social e-learning application.

Figure 4 shows search results in the village social e-learning application. This figure shows Desa Wiki, whose English translation is Wiki Village. In the search results, several articles appeared based on the keywords inputted and the calculation results of the relevance score for each article. The relevance score was based on three parameters: keywords, indexed content, and article publication date. The relevance scores were ordered from the highest to the lowest, showing the rank of the articles based on relevant keywords that the users had inputted in the "Masukkan Kata Kunci" (input the keyword) field.

C. DOCUMENT SUGGESTION AND RECOMMENDATION EVALUATION

The proposed document suggestion and recommendation page displayed articles that were grouped into seven categories: *terkini* (latest news), *populer* (popular), *pencarian* (search), *dibaca* (read), *tulisan anda* (user's articles), *tematik* (thematic), and *tempat lahir* (user's birthplace). Similarity calculations were carried out by comparing all keywords written on the article recommendation menu. Then, the page showed two parts: the recommended articles and the similarity score. The similarity score was calculated based on the element's differences among the compared articles. The score ranged from 0 to 1. Figure 5 shows the suggestion and recommendation page.

The "Populer" menu contained popular articles read by most users. Users could select one of the articles that were frequently visited. The "Dibaca" menu showed the articles that users had read. It could display articles and recommend articles

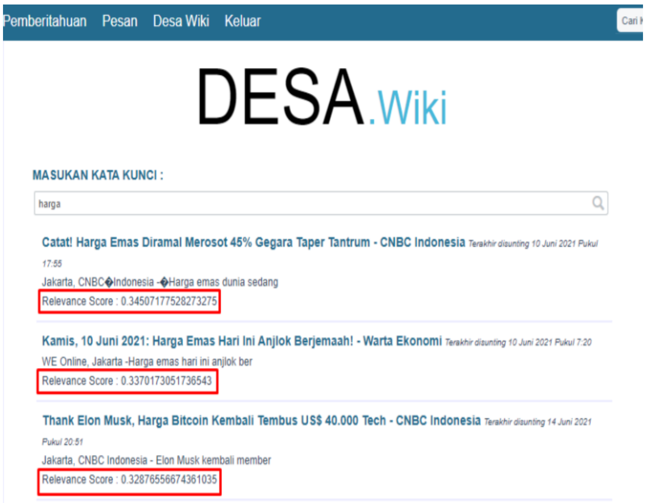

Figure 4. Article search page.

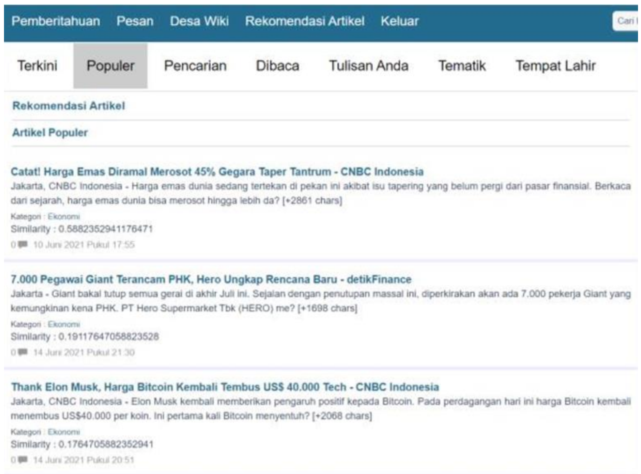

Figure 5. Article suggestion and recommendation page.

that had been read; then, the users could select one of these articles and display the contents of the entire article. The "Tulisan Anda" menu contained articles that users had written. It also showed recommended articles with topics similar to or related to user login articles. The "Tematik" menu contained interesting articles that users log in to like. It showed recommended thematic articles based on topics that users always read. The similarity score was displayed along with each article. The similarity was calculated based on the elements in the article. Subsequently, these elements were compared based on the differences related to existing articles. The following methods were used to determine the differences between the two articles.

1. Creating an element scale from 0 to 1.
2. Giving a score to the elements in both articles with a scale of 3.
3. Calculating the differences in element values in the two articles.
4. Displaying the similarity value on the recommendation page.

A calculation evaluation was carried out to determine the comparison between other articles according to the selected category. The similarity calculation based on categories required five articles. Each article was then compared with the others. The following are the results of similarity calculations based on five popular articles recommended by the system. When the article was read, the value of the number of articles

```
17
1.0
8
0.47058823529411764
7
0.4117647058823529
7
0.4117647058823529
6
0.35294117647058826
0.5882352941176471
0.19117647058823528
0.1764705882352941
0.1764705882352941
0.22058823529411764
CALL viewPopularArticle ();
```

Figure 6. Example of similarity calculation on popular articles.

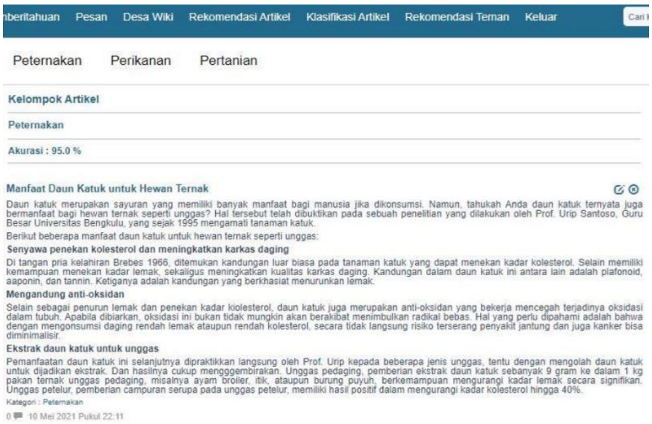

Figure 7. Articles classification page.

TABLE II
COMPARISON OF ARTICLES CLASSIFICATION ALGORITHM

Algorithm	Articles Correctly Classified	Articles Incorrectly Classified	Accuracy
Neural Network	243	12	95.2%
Decision Tree	217	38	87.0%
Naive Bayes	163	92	63.9%

was scaled from 0 to 1, and the relative similarity results can be seen in Figure 6.

According to Figure 6, the similarity calculation results from evaluating the implementation of the similarity-based element algorithm by producing an article similarity score for the number of elements being compared. The similarity score results can be seen in the first row, showing a score of 17, indicating the number of clicks or times to read. The second row shows a score of 1.0 as the number of click values that have been scaled. The similarity value in the first article is 0.5882 (see row 11). This calculation was applied to calculations in the second article, whose similarity score is 0.1911 (see row 12), the third article is 0.1764 (see row 13), the fourth article is 0.1764 (see row 14), and the fifth article is 0.2205 (see row 15). A conclusion that can be drawn is that the more articles are read by users, the more popular articles appear at the top. In this case, popularity was calculated based on the number of articles read.

D. DOCUMENT CLASSIFICATION EVALUATION

A confusion matrix was used to test the naïve Bayes algorithm, decision tree algorithm, and neural network. After comparing the naïve Bayes algorithm, decision tree algorithm, and neural network, the result found that the neural network yielded the best result. Therefore, the article classification module, shown in Figure 7, uses the neural network.

Table II shows the test results comparing the performance of the naïve Bayes, decision tree, and neural network algorithms. Naïve Bayes showed lower accuracy due to its drawbacks in probability-based decision-making when compared with neural networks and decision trees. Naïve Bayes had zero probability problems when the test data for a particular class were not in the training data. Hence, it would result in zero frequency probability.


```
{user7:Pupuk yang Cocok untuk Kolam Ikan Nila(pupuk,kolam,dedak,ikan),
user9:Herbal untuk Mengatasi Ikan Lele Mendadak Mati(kipahit,ikan,tanaman,bunga),
user9:Jenis Penyakit Ikan Cupang yang Perlu Diwaspadai(sudah,cupang,air,ikan),
user9:Kriteria Kolam Ikan Koi yang Baik(kolam,koi,air,ikan),
user5:Teori Dasar Akuaponik yang Perlu Dipahami(budidaya,air,tanaman,ikan),
user7:Teori Dasar Akuaponik yang Perlu Dipahami(budidaya,air,tanaman,ikan),
user2:Ternak Lele dengan Air Payau(payau,budidaya,air,ikan),
user8:Ternak Lele dengan Air Payau Lanjutan(lele,budidaya,air,ikan)}
```

Figure 8. User clustering page.

TABLE III
FRIEND RECOMMENDATION FOR EACH USER

User ID	Recommended Friends
2	User 5, User 7, User 8, and User 9
5	User 2, User 7, User 8, and User 9
7	User 2, User 5, User 8, and User 9
8	User 2, User 5, User 7, and User 9
9	User 2, User 5, User 7, and User 8

TABLE IV
GOODNESS MEASURE IN EACH CLUSTER LEVEL

Level	Goodness Measure	Level	Goodness Measure
1	Infinity	10	0.86
2	2.30	11	0.65
3	1.97	12	0.65
4	1.97	13	0.65
5	1.56	14	0.51
6	1.37	15	0.43
7	0.98	16	0.35
8	0.98	17	0.21

E. USER CLUSTERING EVALUATION

The clustering method was applied to the friend recommendation feature. Here, users received friend recommendations based on interesting writing and reading articles. When logging in to the social e-learning application, the user could click the friend recommendation menu, and then the application would run the ROCK clustering algorithm.

Clustering was carried out based on the title and content of the articles written by each user. The algorithm calculated how often the keywords appeared in an article; in this stage, the

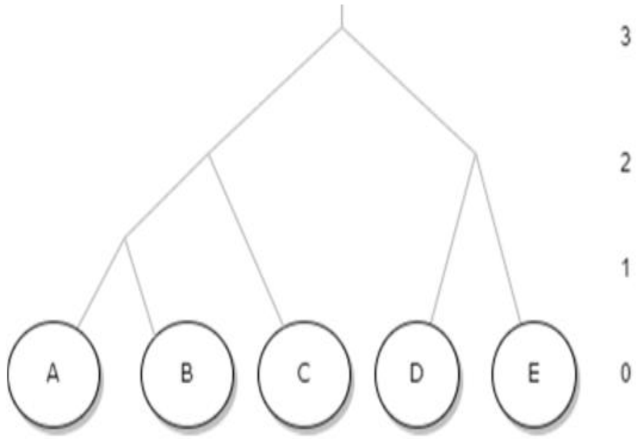

Figure 9. Visualization of dendrogram cluster hierarchy.

topNterms parameter was defined. This calculation was carried out on all existing articles. After obtaining topNterms, articles with similar keywords were clustered, forming a number of similar article groups. The generated clusters were used to consider friend recommendations. Figure 8 shows the friend recommendation page.

Figure 8 also shows the log behind the screen printing the clustering result. In this case, there were ten users writing five articles: users 1, 2, 3, and 10 wrote animal husbandry articles; user 2 wrote fisheries articles; users 4, 5, and 6 wrote agriculture articles; users 7, 8, and 9 wrote fisheries articles; and user 7 wrote an agriculture article. The result of the friend recommendations can be seen in Table III. As can be seen in Table IV, the advantage measure of clustering for this case was 0.10, reaching 18 levels of the cluster.

A simple explanation of leveling is described below. The dendrogram structure is a tree data structure that helps capture the formation of hierarchical clusters. Figure 9 depicts the visualization of a dendrogram.

dendrogram is interpreted as a triple set [d, k, {…}]. The first element is the closeness threshold (d), the second element is the number of clusters (k), and the third element is the set of clusters. Figure 9 depicts the visualization of a dendrogram with four levels where the clusters are divided as {[0,5,{{A},{B},{C},{D},{E}}], [1, 3, {{A ,B},{C},{D,E}}], [2, 2, {{A,B,C}, {D,E}}], [3,1,{A,B,C, D,E}]}. Her e, there were four cluster levels starting from level 0 to level 3.

V. CONCLUSION

The intelligent web algorithms were implemented in finding the ranking of an article, the article suggestion and recommendation, document classification, and friend recommendation. The PageRank algorithm was used for document search. The similarity method was utilized for document suggestion and recommendation. The classification algorithm, i.e., the naïve Bayes, decision tree, and neural network were compared to find the most accurate result. The result showed that the neural network is the best choice for document classification. Finally, the ROCK algorithm was utilized for friend recommendation.

In future work, the development of social learning will be conducted toward the project-based learning module in which villagers share their ongoing or history of agriculture or aquaculture project along with environment conservation with its artificial intelligence implementation.

CONFLICTS OF INTEREST

The authors declare that this work was conducted and written without conflict of interest.

AUTHORS' CONTRIBUTIONS

Methodology, Seno Adi Putra; clustering program, Timmie Siswandi; article searching program, Dessy Yussela; article suggestion and recommendation program, Rinez Asprinola; Naïve Bayed-based classification program, Erin Karina; neural network-based classification program, Mega Candra Dewi; decision tree-based classification program, Santi Al-arif; writing, Seno Adi Putra.

REFERENCES

[1] S. Wan and Z. Niu, "A hybrid e-learning recommendation approach based on learners' influence propagation," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 5, pp. 827–840, May 2020, doi:10.1109/TKDE.2019.2895033.

[2] J.J. Walcutt and S. Schatz, *Modernizing Learning: Building the Future Learning Ecosystem.* Washington, DC, USA: Government Publishing Office, 2019.

[3] S.A. Salloum *et al.*, "Exploring students' acceptance of e-learning through the development of a comprehensive technology acceptance model," *IEEE Access*, vol. 7, pp. 128445–128462, Sep. 2019, doi: 10.1109/ACCESS.2019.2939467.

[4] M. Hajli, H. Bugshan, X. Lin, and M. Featherman, "From e-learning to social learning – A health care study," *Eur. J. Train. Dev.*, vol. 37, no. 9, pp. 851–863, Nov. 2013, doi: 10.1108/EJTD-10-2012-0062.

[5] R. Moraghebi, J. Guo, and G.F. Laksmono, "Tiger e-learning system," *J. Comput. Sci. Coll.*, vol. 24, no. 4, pp. 221–228, Apr. 2009.

[6] B. Alojaiman, "Toward selection of trustworthy and efficient e-learning platform," *IEEE Access*, vol. 9, pp. 133889–133901, Sep. 2021, doi: 10.1109/ACCESS.2021.3114150.

[7] S.M. Aslam, A.K. Jilani, J. Sultana, and L. Almutairi, "Feature evaluation of emerging e-learning systems using machine learning: An extensive survey," *IEEE Access*, vol. 9, pp. 69573–69587, May 2021, doi: 10.1109/ACCESS.2021.3077663.

[8] A.J. Henderson, *The E-learning Question and Answer Book: A Survival Guide for Trainers and Business Managers*. New York, NY, USA: Amacom, 2003.

[9] M.A. Chatti, M. Jarke, and D. Frosch-Wilke, "The future of e-learning: A shift to knowledge networking and social software," *Int. J. Knowl. Learn. (IJLK)*, vol. 3, no. 4–5, pp. 404–420, 2007, doi: 10.1504/IJKL.2007.016702.

[10] J. Maan, "A connected enterprise - Transformation through mobility and social network," *Int. J. Manag. Inf. Technol. (IJMIT)*, vol. 4, no. 3, pp. 89–96, Aug. 2012, doi: 10.5121/ijmit.2012.4308.

[11] G. Dafoulas and A. Shokri, "Investigating the educational value of social learning networks: A quantitative analysis," *Interac. Technol. Smart Educ.*, vol. 13, no. 4, pp. 305–322, Nov. 2016, doi: 10.1108/ITSE-09-2016-0034.

[12] H.H. Yang and S.C.-Y. Yuen, *Collective Intelligence and E-learning 2.0: Implications of Web-Based Communities and Networking.* Hershey, PA, USA: IGI Global, 2010.

[13] H. Marmanis and D. Babenko, *Algorithms of the Intelligent Web*. Shelter Island, NY, USA: Manning, 2009.

[14] S. Alag, *Collective Intelligence in Action*. Shelter Island, NY, USA: Manning, 2008.

[15] W. Ngaogate, "Integrating flyweight design pattern and MVC in development of web application," in *ITCC '20, Proc. 2020 2nd Int. Conf. Inf. Technol. Comput. Commun.*, 2020, pp. 27–31, doi: 10.1145/3417473.3417478.

[16] C. Pahl, P. Jamshidi, and O. Zimmermann, "Architectural principles for cloud software," *ACM Trans. Internet Technol. (TOIT)*, vol. 18, no. 2, pp. 1–23, May 2018, doi: 10.1145/3104028.

[17] Z. Zhu *et al.*, "Fast PageRank computation based on network decomposition and DAG structure," *IEEE Access*, vol. 6, pp. 41760–41770, Jun. 2018, doi: 10.1109/ACCESS.2018.2851604.

[18] Y. Mo *et al.*, "Cloud-based mobile multimedia recommendation system with user behavior information," *IEEE Syst. J.*, vol. 8, no. 1, pp. 184–193, Mar. 2014, doi: 10.1109/JSYST.2013.2279732.

[19] I.H. Mwinyi, H.S. Narman, K.-C. Fang, and W.-S. Yoo, "Predictive self-learning content recommendation system for multimedia contents," in *2018 Wireless Telecommun. Symp. (WTS)*, 2018, pp. 1–6, doi: 10.1109/WTS.2018.8363949.

[20] C.K. Aridas *et al.*, "Uncertainty based under-sampling for learning naive Bayes classifiers under imbalanced data sets," *IEEE Access*, vol. 8, pp. 2122–2133, Jan. 2020, doi: 10.1109/ACCESS.2019.2961784.

[21] S. Wang, J. Ren, and R. Bai, "A regularized attribute weighting framework for naive Bayes," *IEEE Access*, vol. 8, pp. 225639–225649, Dec. 2020, doi: 10.1109/ACCESS.2020.3044946.

[22] W. Xiaohu, W. Lele, and L. Nianfeng, "An application of decision tree based on ID3," *Phys. Procedia*, vol. 25, pp. 1017–1021, 2012, doi: 10.1016/j.phpro.2012.03.193.

[23] R. Xin, J. Zhang, and Y. Shao, "Complex network classification with convolutional neural network," *Tsinghua Sci. Technol.*, vol. 25, no. 4, pp. 447–457, Aug. 2020, doi: 10.26599/TST.2019.9010055.