Improving the Platform in the Lecture Scope with the Implementation of the TF-IDF Algorithm

Alfred Widjaja
Bina Nusantara University

Fahmi Efendy
Bina Nusantara University

Sugiono Kurniawan
Bina Nusantara University

Abdul Haris Rangkuti
Bina Nusantara University

Abstract: Information plays a crucial role in the lives of everyone, including students. One of the most important types of information that students need is related to lecture events. However, students often have difficulty finding suitable lecture event information, and they sometimes forget to attend events they have registered for. As a result, this research focuses on designing and implementing a website-based recommendation system for lectures. The recommendation system utilizes the TF-IDF (Term Frequency-Inverse Document Frequency) algorithm in its development. The primary goal of developing this recommendation system is to help students easily find event information and receive event activity notifications. In alignment with the established objectives, a significant number of respondents agree that the recommendation system simplifies event registration. Consequently, with the aid of this system, students can become more proactive and participate more readily in events. Following a survey conducted with 47 respondents, it was found that 84.4% of them had never used a recommendation system for college activities. Furthermore, 53.1% expressed satisfaction with this recommendation system. It is hoped that the recommendation system, employing the TF-IDF algorithm, can be further optimized to yield even better results.

1. INTRODUCTION

Information is important and inseparable from everyday human life. Information itself means notification, news, or news about something [1]. Information has a very important role for a student, one of which is information about lecture events such as seminars from within and outside the university, academic and non-academic competitions, social activities, and workshops. However, students sometimes experience obstacles in the form of limited information on existing lecture events. One example of a problem is that lecture events are only shared via social media platforms in group chats, and some students sometimes do not open the group chat so that information about the lecture event is not received by students. In fact, with the use of the right information technology, students can access a variety of events easily [2]. To overcome the problems that have been described previously, this research wants to create a website that contains events about lectures equipped with a recommendation system. A recommendation system is a step approach to existing problems in providing solutions in handling the selection of things that are suitable according to user preferences from the many existing items [3]. The website will be built using the React and Express frameworks with TF-IDF as the main method for the recommendation system, and all data will be stored in MongoDB. This application is expected to be a solution for every student to become more active and easier to participate in an event because students can get event information more efficiently and easily. In order to enhance the platform within the lecture scope, the implementation of the TF-IDF (Term Frequency-Inverse Document Frequency) algorithm is proposed. This algorithm will be utilized to improve the platform's functionality and effectiveness in providing relevant information to users. By implementing the TF-IDF algorithm, the platform will be able to analyze and

evaluate the importance of different terms within the lecture materials and documents. This will allow the platform to provide more accurate and relevant recommendations to users based on their specific needs and preferences.

2. LITERATURE REVIEW

A recommendation system is a step approach to existing problems in providing solutions to handle the selection of things that are suitable according to user preferences from many existing items [3]. One example of recommendation system methods that can be used is TF-IDF. TF-IDF (Term Frequency-Inverse Document Frequency) is an algorithm used to calculate the weight of the statistical value of a word to determine how important a word is in an existing text collection. TF represents the number of similar words that appear in a group of texts. While IDF represents the weight of the importance of a word in a collection of texts as a whole [4]. One example of implementing the TF-IDF algorithm can be done in the information retrieval process. Information retrieval is a system that is used to search for related or relevant information from a set of information automatically based on user-defined search keywords [5]. In implementing a recommendation system, it takes a display of the computer system to be used. The science that studies the interaction between humans and computers from theory, design, implementation, and evaluation is called HCI (Human Computer Interaction) [6]. That way, HCI is very important in making an application in order to create a good user experience. Application developers must pay attention to the design contained in the application; otherwise, the appearance of the application becomes laborious and reduces the user experience [7]. There are four important components in HCI, including user, purpose, appearance, and context.

To create an application that can satisfy in terms of user experience and also appearance, several components are needed in making the system, such as the React framework and also a database like MongoDB for data storage. Framework itself is a function that assists developers in developing applications by creating predefined methods such as URL routing, state management, bundling, and others [8]. Framework has the advantage that it can be used repeatedly and also has better security because it has been identified by many expert developers and continues to be developed [9]. React is a JavaScript library developed by Facebook that can be used to create website applications. React can divide a complex user interface into smaller parts called components [8]. Some of the advantages of React compared to other website application development methods are:

1) Composable

   React divides complex web applications into smaller components that can be built to produce many combinations of components

2) Declarative

   Application development only needs to mention the desired display, it doesn't need to include the steps in it

3) Multi-Platform

   The code in React can be extended and used on various platform

4) Easy to use and learn

   Application developers only need to master the HTML, CSS, JavaScript programming languages where HTML and JavaScript programming in React can be combined into JSX (JavaScript XML) to facilitate development.

HTML (Hypertext Markup Language) is a markup language used by developers in making websites [10]. CSS (Cascading Style Sheets) is also needed to determine how the HTML elements on the website will be displayed. This is because without CSS, elements in HTML will only give an ordinary and unattractive appearance. CSS can also determine the layout and position of the HTML elements used on the website [11]. While JavaScript is used to adjust the flow of logic on websites that previously have been designed using HTML and CSS. JavaScript itself can be used for various ways of programming, such as imperative programming, object-oriented programming, and also functional programming [12].

To handle data displayed from the frontend side, a backend side such as Node.js is needed. Node.js is a fast and reliable platform for handling heavy files and networks in making applications because the approach is event-driven, uses non-blocking types, and uses an asynchronous method [13]. One example of a node.js-based framework is Express.js. Express.js is a flexible node.js backend framework useful for making it easier to create node.js-based applications in managing the functionality of a website [14]. Some of the advantages of Express.js compared to other backend platforms are [15].

1) Supports Google V8 Engine

   With Google V8 Engine, it increases website performance and ensures that any errors can be handled quickly during development process

2) Provides server-side caching

   Doing server only need to respond once for each client request for the same page to improve user experiences

3) Dynamic and more powerful

   Has lot of modules available on NPM (Node Package Manager)

4) Reducing development costs

   Express is an open source based node.js framework and also doesn't require a lot of resources because API used in it is very light.

Data that has been processed from the backend side is then stored in the database. One example of an existing database is MongoDB. MongoDB is a NoSQL based database that is used to store data in a different format than Relational Database Management System (RDBMS) tables that focus on scalability, query speed, and compatibility with application changes [16]. MongoDB provides faster performance, does not require a complicated table structure, is able to accommodate a lot of varied data, and manage queries better [17].
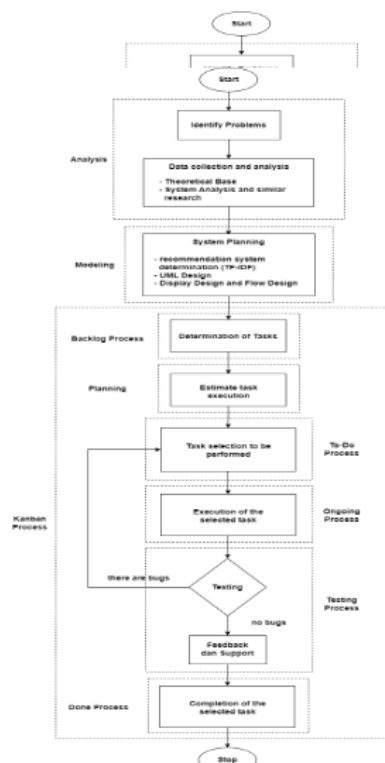
3. METHOD

Figure 1. Application Development Process

281

Process in Figure 1 starts with the identification of the problems that were the cause of this research being carried out first. After the problem is found, the process of data collection and analysis will be carried out to determine application requirements such as the algorithm, programming language, and also database to be used. Methods of data collection will be carried out by means of reference theory, literature study, and also analysis of similar research. The study of literature and reference theory is intended for the purpose of collecting topical reference data that supports system development. Then, a similar research analysis is used to determine the recommendation system algorithm used. After the process of collecting data is carried out, the process continues with the system design stage. At this stage, a system recommendation algorithm has been established, namely the TF-IDF algorithm. Modeling of the flow design and appearance design is done at this stage. The flow design is described using UML diagrams, which consist of a use case diagram, an activity diagram, a sequence diagram, and a class diagram. For display design, the author used Figma as a place to make it. Display design starts from low-fidelity design to finished display. After all modeling has been completed, the process will then implement the Kanban Method. Kanban method starts with the backlog process. The backlog process is used to make a determination of how many tasks are needed to finish the project. A task will be created in the form of a card. Tasks can be broken down into smaller parts to ease the application development process. After every task has been made, a planning process will be carried out to determine the estimated length of time every task can be completed. After the estimated processing time for each task has been completed, Process will proceed to the next stage. The next stage consists of a series of stages starting from the to-do stage, which will be iterated continuously until the cards previously set in the backlog run out. The to-do is a stage for determining which tasks will be carried out in one iteration. At this stage, the cards in the backlog column will be shifted to the To-Do column to list the tasks to be carried out. After the listing is completed, the tasks will be carried out at the ongoing process stage. At this stage, the tasks in the To-Do column will be done one by one. The way to make changes to the Kanban board is still the same, namely by shifting the card task that you want to do to the destination column. After a task in the ongoing process is finished, it will be shifted to the testing column. At this stage, the task will be tested by means of black box testing. If the task passes the requirements, then the task card will be shifted to the done column on the Kanban board to indicate that the task has been completed. If there are still bugs, the card will be returned to the To-Do column. After all tasks have been completed, the development process is finally complete.

4. RESULT AND DISCUSSION

System recommendation event platform development starts with every step the author previously mentioned, with each step as a follow.

A. Requirement Definition

System recommendation event platform application starts with specifying and defining all feature requirements needed for the application. Following is a list of requirements needed to develop this application:

1) Creating an event recommendation system so that students can easily find an event based on their preference
2) Create a system that can help students remember the event they are participating in by giving notifications to students
3) Create a system that can help organizers manage an event.

B. Software and Design System

After all of the requirements listed above have been defined, the author starts with designing the system using a use case diagram.

1) System Design

At the beginning of the system design, a use case diagram was made with as many as 13 use cases to inform about all activities, which can be seen in figure 2. There are 3 actors in the system environment, consisting of students, organizers, and admins.

- Use Case Diagram

Figure 2. Use Case Diagram System Recommendation Event Platform
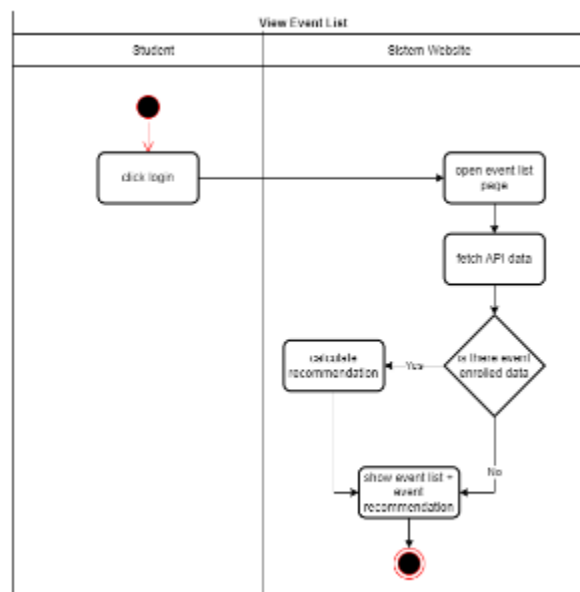
- Activity Diagram

Figure 3. Activity Diagram for View Event List in Student Role
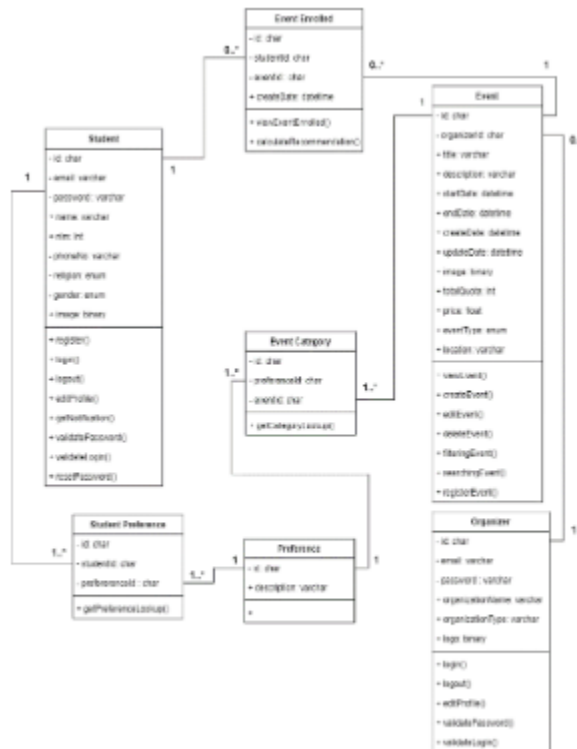
- Class Diagram

Figure 4. Activity Diagram for View Event List in Student Role

Figure 4 describes the class diagrams used to create a database for the System Recommendation Event Platform application. There are a total of 7 tables, with 3 of them being the main table in the diagram, which is the student table, organizer table, and event table. Admin can directly access the class in the database mentioned above.
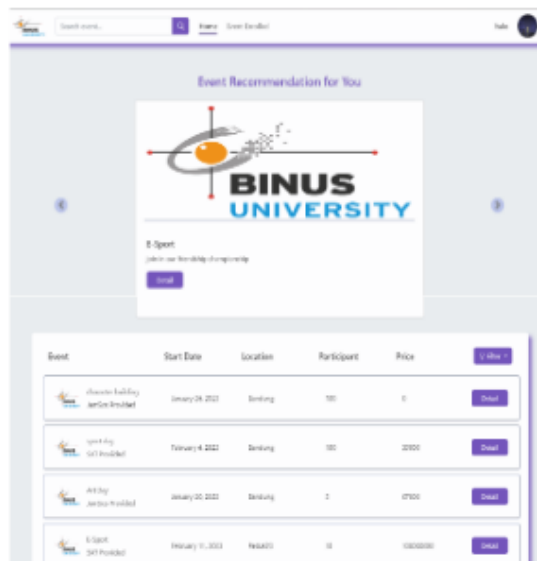
2) User Interface

Figure 5. User Interface of Student Role Main Page

Figure 5 is a student role view main page of the System Recommendation Event application. In this page, there are 2 components: the first component is a carousel to show student events recommended, and the event table list to show students every event listed in the application.

C. Application Development

After system design for the application is completed, the author begins with creating a base server for the application. Below are details of the server used to deploy both React for the frontend application and the Express for the backend:
- Host Name: 89.116.229.148
- Cores: 2 CPU
- RAM: 2GB
- Storage: 39.06 GB

After the server is deployed, the Express and React application can be deployed to the server. In the Backend, there are 2 services that are deployed, which are binus event authentication and binus event service. The service list server can be seen in figure 6.
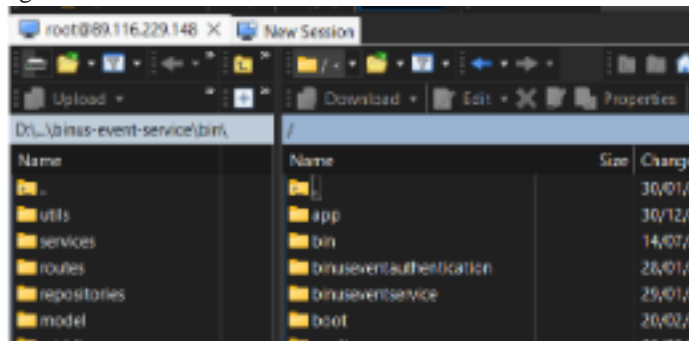

Figure 6. Express.js backend service list used for the application

```
const termFrequency = {};
tokens.forEach((token) => {
  if (!termFrequency[token]) {
    termFrequency[token] = 1;
  } else {
    termFrequency[token]++;
  }
});
```

```
    termFrequency[token] = 1;
  } else {
    termFrequency[token]++;
  }
```

```
const inverseDocumentFrequency = {};
tokens.forEach((token) => {
    if (!inverseDocumentFrequency[token]) {
        inverseDocumentFrequency[token] =
            Math.log(data.length / tokens.filter(
            (t) => t === token
        ).length);
    }
});
```

```
Object.keys(termFrequency).forEach((token) => {
    if (!scores[token]) {
        scores[token] = termFrequency[token] *
inverseDocumentFrequency[token];
    } else {
        scores[token] += termFrequency[token] *
inverseDocumentFrequency[token];
    }
});
```

D. Testing

After developing the System Recommendation Event Platform application, testing is carried out to ensure the application created runs smoothly and safely. The type of testing that author used was black box testing. Black Box testing is used to ensure every function of the application is running as planned. The System Recommendation Event Platform application was tested with a total of 47 test cases, with the distribution of 25 test cases for the student role, 21 test cases for the organizer role, and 1 test case for the admin role. The overall test result from 47 test cases went well and was declared a pass.

E. Questionnaires

Questionnaires are one of the most important things in evaluating system recommendation event platform applications. Questionnaires can be used to get valuable input so that the author can improve the applications that have been made. The results of this question are also useful for determining whether the design of the application is suitable for its purpose or not. Here are some questions asked to students about the application:
1) Have you ever used a recommendation system application for lecture activities before?

Have you ever used a recommendation system application for lecture activities before?
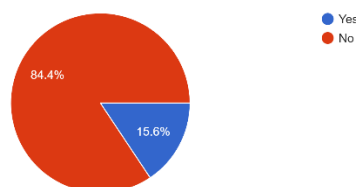32 responses

● Yes
● No

84.4%

15.6%

Figure 7. Student once use about recommendation system for lecture activities

286

Figure 7 shows that 27 respondents (84.4%) had never used the recommendation system application for lecture activities before, while only 5 respondents (15.6%) had used the recommendation system application for lecture activities. So, most respondents never had the opportunity to use an application recommendation system before in the scope of lectures.

2) How consistent do you think the look and theme of these recommendation system apps are?

How consistent do you think the look and theme of these recommendation system apps are?
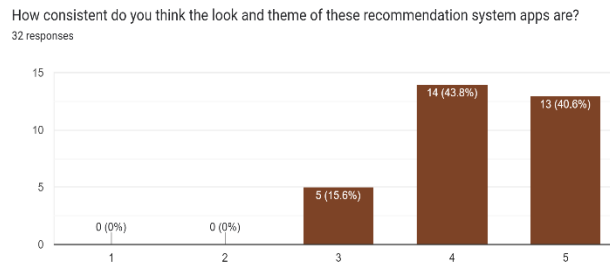32 responses

Figure 8. Student opinion about application design and theme consistency

Figure 8 shows that as many as 27 respondents (84.4%) agreed that the system recommendation event platform application design was consistent, and 5 respondents (15.6%) were neutral about it. It shows that the system recommendation event platform application has achieved good design consistency in its appearance.

3) How satisfied are you in future use of this app?

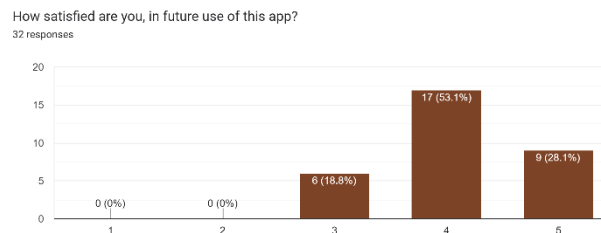How satisfied are you, in future use of this app?
32 responses

Figure 9 shows that in general, respondents felt quite satisfied with the results of the development and design of the system recommendation event platform application that had been made with the criteria of 9 respondents (28.1%) feeling very satisfied, 17 respondents (53.1%) feeling satisfied, and 6 respondents (18.8%) feeling neutral.

5. CONCLUSION

1) By implementing the TF-IDF algorithm, the platform will be able to analyze and evaluate the importance of different terms within the lecture materials and documents. This will allow the platform to provide more accurate and relevant recommendations to users based on their specific needs and preferences.
2) Overall, the implementation of the TF-IDF algorithm represents a significant step towards improving the platform's performance and user experience within the lecture scope. By leveraging this algorithm, the platform will be able to provide more personalized and accurate recommendations, ultimately benefiting users and enhancing their learning experience.

3) With this event recommendation system, it will simplify students' ability to determine which events they are interested in. Then, students can also more quickly check the types of activity events based on the preferences they want to follow.

4) The research and development of this activity event recommendation system has produced an application that can be used to facilitate students who want to register for an event and to simplify the sharing of information from organizers to students.

5) The UI of the system recommendation event platform application is suitable for use, but there can still be many improvements, such as improving the color theme and also making improvements like code refactoring to reduce unnecessary loading time, which would enhance the UX in the application.

SUGGESTION

1) In further development, it can be done by adding multiple organization accounts to make it easier for various organizations and university administrators to add, edit, and delete events without waiting for the organizer role provided by the administrator to proceed.

2) In further development, the application can add features like a like button, a dislike button, and chat features to make it easier for students to find out whether an event is in based on student preference.

3) Future system development will not only focus on computer resolution but on phone resolution too. This is done because many people access websites via smartphones.

REFERENCES

Ariffudin, M. (2021, December 11). Mengenal Express.js: Pengertian, Cara Kerja, Keunggulan, Tutorial. NIAGAHOSTER Blog. Retrieved December 18, 2022, from https://www.niagahoster.co.id/blog/express-js-adalah/

Attardi, J. (2020). *Modern CSS*. Springer.

Babich, N. (2020, July 28). Man and Machine: A Guide to Human-Computer Interaction. Adobe. Retrieved December 16, 2022, from https://xd.adobe.com/ideas/principles/human-computer-interaction/man-and-machine-guide-to-human-computer-interaction/

Chiny, M., Chihab, M., Bencharef, O., & Chihab, Y. (2022). Netflix recommendation system based on TF-IDF and cosine similarity algorithms. *no. Bml*, 15-20.

Das, D., Sahoo, L., & Datta, S. (2017). A survey on recommendation system. *International Journal of Computer Applications*, *160*(7), 6-10.

Fast, unopinionated, minimalist web framework for Node.js. (2017). Strongloop, IBM, and o. e. contributors. Retrieved December 17, 2022, from https://expressjs.com

Humaini, I., Wulandari, L., Ikasari, D., & Yusnitasari, T. (2020, March). Penerapan Algoritma TF-IDF Vector Space Model (VSM) Pada Information Retrieval Terjemahan Al Quran Surat 1 Samai Dengan Surat 16 Berdasarkan Kesamaan Makna. In *Prosiding Seminar Nasional Teknik Elektro UIN Sunan Gunung Djati Bandung* (pp. 525-534).

Jamsa, K. (2013). *Introduction to web development using HTML 5*. Jones & Bartlett Publishers.

Kamus versi online/daring. (2016). KBBI. Retrieved December 13, 2022, from https://kbbi.web.id/informasi

Kim, G. J. (2015). *Human-computer interaction: fundamentals and practice*. CRC press.

Kiswanto, H. (2022, September 12). Pemanfaatan Teknologi Informasi dalam Pendidikan Masa Kini. Sevima. Retrieved December 13, 2022, from https://sevima.com/pemanfaatan-teknologi-informasi-dalam-pendidikan-masa-kini/

Kurniawan, D. (2020, May 13). Apa Itu Framework? Yuk Kenali Pengertian dan Fungsinya!. Niagahoster. Retrieved December 17, 2022, from https://www.niagahoster.co.id/blog/apa-itu-framework/

Schaefer, L. (2023). NoSQL vs. SQL Databases. MongoDB. Retrieved December 18, 2022, from
https://www.mongodb.com/nosql-explained/nosql-vs-sql

Schaefer, L. (2023). What is NoSQL?. MongoDB. Retrieved December 18, 2022, from
https://www.mongodb.com/nosql-explained

Shah, H., & Soomro, T. R. (2017). Node. js challenges in implementation. *Global Journal of Computer Science and Technology*, *17*(2), 73-83.

Wohlgethan, E. (2018). *Supportingweb development decisions by comparing three major javascript frameworks: Angular, react and vue. js* (Doctoral dissertation, Hochschule für Angewandte Wissenschaften Hamburg).

Zaidi, R. (2017). *JavaScript Essentials for SAP ABAP Developers*. Apres.