Development of e-learning and statistical simulation for explorative data analysis based on android

Nia Gracelita, Lutfi Rahmatuti Maghfiroh
Politeknik Statistika STIS
1niagracelita@gmail.com, 2lutfirm@stis.ac.id

Abstract. Exploratory Data Analysis (EDA) is one of the basic statistics subjects that needs to be studied as a basis for higher processing and analyzing data. This subject is taught in class, but these activities are not necessarily enough to understand the subject. This problem can be solved by learning independently. However, in the process of learning independently there are obstacles such as the unavailability of suitable material so that a search is needed which may not produce anything or if the material is available in book or textbook format, the material is less flexible in its use. An e-learning system needs to be developed to solve these problems. In addition, simulation can also help students' understanding in building knowledge base. Furthermore, the use of smartphones with the Android operating system is the most used by Politeknik Statistika STIS students. Based on these problems, an e-learning system and statistical simulation will be built for this Android-based EDA subject. In its development, we use Java programming language and integrate with R Shiny for statistical simulations in the form of data processing. Evaluation shows that the system can be accepted by users as indicated by the Black Box test results and the SUS value of 84.13. The system still consist of basic features and still need to be improved with advanced features so can support the learning activities of EDA.

1. Introduction
Politeknik Statistika STIS is one of Indonesia's official colleges managed by the Statistics of Indonesia (BPS) which educates students to become statisticians. Graduates from Politeknik Statistika STIS who will work in BPS and in other work environments are expected to master the field of statistical science that is useful for data processing and analysis. For this reason, learning at the Politeknik Statistika STIS is inseparable from statistics courses that teach students to collect, process and present data, as well as other subjects that can support student knowledge in special fields such as economics, population, and computing. One of the basic statistics courses that are important to master for data processing is Explorative Data Analysis (EDA). Exploratory data analysis isolates data patterns and features and expresses them firmly to analysts [1]. In other words, EDA is an analysis to recognize data patterns through diagrams or graphs, detect extreme values so that analyzes are made that are not affected by extreme effects, determine patterns of relationships between variables using scatter diagrams, and create equation lines and smoothing data. EDA is needed to support the study and to find out the characteristics of the data which will be useful for selecting appropriate statistical models in the advanced data processing.

Learning activities at the Politeknik Statistika STIS so far, including those for EDA courses, are still conventional, where the lecturer explains in class using books, presentation materials, and modules. For some students, this is enough to understand the subject. However, for some students, this
way of learning is lacking to really understand the course. According to Riding [2], it is natural to happen because everyone has different characteristics in learning, such as existing knowledge, intelligence, memory efficiency, gender, style in analyzing, and style in imagining a thing. This difference in comprehension encourages students to ask questions or discuss with lecturers or friends for things that are still not understood or self-study outside of class hours.

Based on a preliminary survey conducted on 271 Politeknik Statistika STIS students in the academic year 2018/2019, 88.6% of students had experienced learning difficulties while in class. There are several causes of this learning difficulty, such as lack of focus when learning activities, lecturers are too fast in explaining, use of terms that are less understood, and have not mastered the previous material so they cannot follow the material being taught. When experiencing learning difficulties in class, as many as 84% of students will try to understand the material by carrying out learning activities in private. Personal learning activities will be very helpful when learning materials are available. When conducting private learning activities, students usually learn to use books or learning modules provided by lecturers and it is not uncommon to use the internet to find material related to EDA courses because they feel they will immediately answer the difficulties that occur. However, these searches sometimes produce a lot of files or data that are not useful and also can not get results at all. Therefore, we need a learning media that can support learning activities personally for students.

Prawiradilaga [3] in his book entitled Mozaik Technology Education: E-Learning said that e-learning is a set of innovative applications that are used to conduct learning using technology. This application consists of material presented as attractive as possible consisting of text, video, and audio. This shows that e-learning can be used by everyone with different learning styles and can also motivate them to do personal learning to increase knowledge. In addition, the availability of learning materials that are already included in this application will assist in the process of gathering learning materials that are suitable for use.

In addition to learning media consisting of text, video, and audio, a learning system based on simulation will be very helpful for learning activities, especially when studying on your own. Computer simulation-based learning systems can help students to build a knowledge base and reduce errors in the introduction of concepts [4]. According to [5], learning systems based on computer simulations can not only develop spatial learning but also abilities in the perception of students. For this reason, it is necessary to build an e-learning system that can also simulate data processing for EDA courses. This aims to facilitate the preliminary analysis of the data, be a means to build knowledge about EDA courses, and to share information about EDA courses.

Moreover, from the preliminary survey of the use of multimedia learning in students of Politeknik Statistika STIS obtained data shown in figure 1 and figure 2. In Figure 1, 60.5% of students use smartphones for an average of 3-8 hours per day and 36.9% of students use smartphones for an
average of more than 8 hours per day. This means that the use of smartphones is very familiar with the daily life of Politeknik Statistika STIS students. In Figure 2, 93% of students use smartphones with the Android operating system. This situation makes the opportunity for the use of smartphone devices with the Android operating system as a learning medium even greater. This can be useful to help complete learning in class because teaching materials can be easily accessed anywhere as long as the device is attached to the application and connected to the internet, so learning is more efficient and saves money. Thus, in this research, we build an e-learning system for EDA courses in the android-based environment.

2. Methodology
Explorative data analysis (EDA) is a technique that uses mathematical rationale that is based on logic and is useful in verifying that certain procedures actually optimize the model to be used, regardless of whether the procedure can actually be used in actual circumstances [1]. Presentation of data based on EDA can be done in several ways, such as tables, five number summaries (largest, smallest, median, upper quartile, and lower quartile), steam and leaf diagrams, scatter plots, residual plots, and outlier searches.

Other According to [6], in his journal entitled Mobile Learning Application Based on Mobile Application Technology Running on Android Smartphones and Blackberry, mobile learning can be interpreted as the use of IT-based mobile devices, such as PDAs, telephones, and smartphones, which are used to support learning activities and teach. The use of m-learning can help users in increasing motivation for self-learning and also cognitive activities [7]. In terms of increasing motivation for self-learning, users will be more charged with responsibility for learning activities and doing the exercises that have been designed and provided. In [8], there are several advantages of using m-learning, namely flexibility in accessing content, wherever and whenever due to the use of mobile devices and ease in internet access. In addition, there are videos and audio that are very useful in helping learning activities that can facilitate users in imagining and understanding things or learning objects.

The research in [9] is about developing learning systems and web-based simulations to improve understanding of the concepts of the Linked-list structure in the Data Structure course. Other research related in [10] which built an android application that helps EFL (English as a Foreign Language or teaching English where the main language is not English) in learning reduced form (words that are not written in English but are often used by native speakers, like gonna, wanna, dunno, etc.). The system is expected to be able to improve cognitive abilities in the English listening section where there is a problem in the use of reduced forms for people whose primary language is not English. The journal in [11] built an android application that helps users in learning about ASEAN countries. The system built is expected to increase knowledge about ASEAN so that it can be more proud to be part of ASEAN.

In this research, an e-learning system will be built for Android-based EDA courses. In addition to containing learning material in the form of text, some videos can better explain the material, processing applications for data summaries, and discussion forums that can help to understand the material through information sharing questions. It is expected to be able to answer the problems that occur in EDA learning activities in Politeknik Statistika STIS.

2.1. A Method of Collecting Data
Data collection to support the development of this e-learning system uses the following methods:

2.1.1. Literature Study
A literature study is carried out by reading and gathering information from various written sources such as journals, books, articles, etc. related to the development of this system. In the development of the system, three main keywords are used in library research, namely mobile-learning, exploratory-data-analysis, and android.
2.1.2. **Interview**

In this study, interviews were conducted with Explorative Data Analysis (EDA) lecturers and 7 students from the Politeknik Statistika STIS students in the academic year 2018/2019. Interviews were conducted by asking questions related to the needs of the system.

2.1.3. **Questionnaire**

The questionnaire consisted of a list of questions compiled to ask respondents. In this study, the questionnaire was used in the needs analysis and evaluation stages of the product produced from the development process. Specifically, the questionnaire used at the requirements analysis stage was the Preliminary Survey Questionnaire on Multimedia Learning questionnaire for Politeknik Statistika STIS students. The questionnaire used in the evaluation stage was the System Usability Scale (SUS) questionnaire and the Black Box Testing. SUS questionnaire is used to determine user satisfaction with the system built. According to [12], System Usability Scale (SUS) is a method for evaluating the usefulness of a system that has been designed. SUS has the advantage of being able to be done at a low cost but reliable in evaluation. The black box testing questionnaire is used to determine the success of the functions contained in the system when it is run. According to [13], Black Box Testing is a testing of functional specifications of software which is conducted to see inconsistencies or ambiguities in the requirements specifications.

2.2. **System Development Methods**

The system development method used is the System Development Life Cycle (SDLC), namely the waterfall method. The stages in the waterfall method are requirements analysis, system design, implementation (writing program code), program testing, program implementation and maintenance [14].

2.2.1. **Running System Analysis.**

Based on the results of interviews with lecturers and students who have received EDA courses, an analysis of the running system is performed as shown in Figure 3. The activity carried out by lecturers is to provide material related to EDA courses. After giving the material, the lecturer gives a task in the form of a re-presentation of material that has been delivered along with exposure to the results of BPS data processing or other data in accordance with the material presented. Then the lecturer assesses the assignment and uploads the assignment value to SIP ADU. For the activities carried out by the lecturer, there are no obstacles. Every activity can be done effectively, but less efficiently. This can be seen from the assignments given by the lecturer in the form of a re-presentation by students in the form of material that has been delivered by the lecturer.

Activities undertaken by students are to receive EDA material that has been taught by lecturers. However, not all students are able to accept the material taught by lecturers. This happens because the abilities and learning styles of each student are different. When students already understand the material that has been given by the lecturer, students will be able to directly do the assignments given.

However, there are obstacles when students have not been able to understand the material provided. For this problem, students will usually ask fellow friends, ask the lecturer, or look for other sources. When asking fellow friends or lecturers, the question may be asked more than once, for example, asked by lecturers by students from different classes. When searching for this other source, students often experience problems because too much online material is circulating in search engines like Google.

The material found is often not related at all and this causes efficiency in learning to decrease. For assignments in the form of material presentations, students can find references that match the material from textbooks, digital books, and other sources. When working on the data processing section related to the application of the material being taught, students must use software for appropriate processing so that the data that is processed can produce a good and appropriate interpretation. For data processing software that is often used by students in processing data related to EDA courses are
Microsoft Office Excel. However, a comparison is needed in processing to find out whether data processing is good and right. There is other software as a comparison, such as E-views or R Studio, but the software is not simple and not all students have it.

![Business Process of Running Systems](image)

**Figure 3. Business Process of Running Systems**

2.2.2. Problem Analysis.

Based on the analysis of the running system that has been done and the facts obtained from the results of interviews, questionnaires, and literature studies, then obtained problems that can be described in the Ishikawa diagram in Figure 4. In Figure 4 it is found that the main problems faced in the system are independent learning activities and the implementation of EDA learning which is still less effective. The problem is caused by three main causes, namely learning materials and applications, procedures, and human resources.

In terms of teaching materials and applications found that personal learning materials are less flexible when used because it is limited by the text in the book. For desktop processing applications are also less flexible related to the use of learning that can be anywhere and anytime. From the procedural side, it was found that learning that was limited to campus meetings still lacked the ability to learn and the questions were asked repeatedly at different times and places. In terms of human resources, it was found that student learning styles differed. That causes limitations in participating in learning in class.

![Ishikawa Diagram](image)

**Figure 4. Ishikawa Diagram**
2.3. System Requirements Analysis

| Requirements | Problem | Solution |
|--------------|---------|----------|
| **Performance** | Face-to-face learning is only done in class and is limited to 14 meetings. | Creating a system that facilitates learning so that it is not limited by time and place (class). |
| **Information** | Searching for learning materials from various sources with different results and may not get any information at all. | Provide a system with materials that are already integrated with learning resources such as text, photos, and videos. |
| **Economic – Cost** | Inefficient use of time for gathering learning material. | Provide a system with appropriate learning materials. |
| **Control** | Information access security in the form of asking questions and answers has not yet been authenticated. | Provides a system that is able to maintain the security of information access in the form of asking questions and answers through account authentication. |
| **Efficiency** | Questions asked by students to lecturers or fellow students are similar or the same but must still be answered even though they have been answered at different times or places. | Provide a system that can accommodate questions and answers that can be accessed so that there is no repetition in asking questions. |
| **Service** | The teaching materials available are not flexible and the data processing software as a comparison is still not flexible. | Provides applications containing teaching materials, processing applications and discussion forums on an Android-based system. |

3. Implementation and Result

3.1. Proposed System Design

![Business Process of Proposal Systems](image)

**Figure 5. Business Process of Proposal Systems**

To meet the needs that are the solution to problems in the running system, a proposed business process system is designed that can be seen in Figure 5. The proposed system in the development of
this learning system is an Android-based system. The use of an android-based system aims to eliminate problems in the form of a lack of flexibility of learning materials as well as data processing applications for EDA courses because most still use the desktop. In addition, the high use of Android-based smartphones among Politeknik Statistika STIS students makes this system very easy to use and does not bother students to carry and access the EDA learning system anywhere. In Figure 5, the gray box shows the process to be carried out using an Android-based learning system that will be built.

In the Android-based learning system for EDA courses, there are several main functions, namely material, data processing applications, and discussion forums. The material function is a function that handles the process of searching material related to EDA courses and other references on the material (P1). The function of the discussion forum is a function that handles the discussion process or the submission of questions and answers to questions (P2). The data processing application function is a function that handles data processing and displays statistical methods found in EDA courses (P3).

3.2. Architectural Design
The system architecture used is the development of android-based applications that are connected to the internet network. Display interface is divided into two types, namely interfaces that can be accessed through the android application which was developed using java programming language and web interfaces that are accessed through a web browser from the android application.

The architectural design of the proposed system is built using two types of servers, namely Apache Server and Shiny Server. Apache Server is used to run android-based applications which are the main framework, while Shiny Server is used to run data processing pages that are built using R in the form of R Shiny.

![Diagram of Architectural Design](image)

**Figure 6. Architectural Design**

3.3. Interface Implementation
Figure 7-17 are the implementation of system design. Figure 7-15 are user interface for student and lecturer in android-based applications. Figure 16 and Figure 17 are user interfaces for admin in the web-based application. The admin side is developed using a responsive framework that is bootstrap which can be accessed through the various device while connected to the internet. The user side developed using native android.
Figure 7. Implementation of Material Pages

Figure 8. Implementation of Material Content Pages

Figure 9. Implementation of Processing Pages

Figure 10. Implementation of the Processed Data Results Pages

Figure 11. Implementation of Discussion Forum Pages

Figure 12. Implementation of the Questions List Pages
3.4. System Test Result

System testing is done by using the Usability Scale Survey (SUS) and black-box testing. Testing using SUS was conducted on Politeknik Statistika STIS Statistics Polytechnic students who had been taught Explorative Data Analysis courses. SUS testing has been carried out on 20 respondents with a value of 84.13. This means the system is in a good category and can be accepted by users. This is done based on the user's perspective, so no basic skills regarding programming languages are needed. The highest score is obtained in statements number 3 and 8 from the SUS questionnaire. For statement 3 which has a positive value, this system is considered to be very easy to use by users, whereas for statement 8 which has a negative value, the system is not felt to be confusing when using it.

Black box testing is done by checking the functions of the application being built. Black box testing is done by testing the scenarios of functions that are built. Table 2 shows the results. The results of the black box test state that the functions contained in the system are running well and the desired output is appropriate. This system can be accessed as long as an internet connection is available and specifically for material that can be accessed offline. The admin side is developed using a responsive framework that is bootstrap which can be accessed through various devices while...
connected to the internet. The user side is developed using native android. Especially for processing features is developed using R Shiny.

Table 2. Black Box Test Result

| Function         | Status  | Function         | Status  |
|------------------|---------|------------------|---------|
| Material access  | Success | Update material  | Success |
| Data processing  | Success | Manage forums    | Success |
| Discussion forum | Success | Login            | Success |
| Manage users     | Success | Logout           | Success |

4. Conclusion

In this research, we build an e-learning system and statistical simulation for EDA courses in the android-based environment. In its development, we use the Java programming language and integrate with R Shiny for statistical simulations in the form of data processing. Black Box Test and SUS score of 84.13 show that the system developed has been accepted by the user. However, the system is still in the early stages of development so that the features contained in this system still consist of basic features, such as material access, data processing, and discussion forums. This system can continue to be developed so that it can become a better system with features that can further support the learning activities of EDA.

References
[1] Tukey J W 1993 Exploratory Data Analysis, Past, Present and Future, Technical Report (Princeton University)
[2] Riding R 2005 Individual Differences and Educational Performance (Educational Psychology, vol 25 no 6) pp 659 – 672
[3] Prawiradilaga D S 2013. Mozaik Teknologi Pendidikan: E-Learning (Jakarta: Kencana)
[4] Chen D J and Lai A F 2005 Network and Education (New Taipei City: Open University)
[5] Lindgren R and Schwartz D L 2009 Spatial learning and computer simulations in science (International Journal of Science Education vol 31 no 3) pp 419-438.
[6] Setiabudi D 2013 Mobile Learning Application Based on Hybrid Mobile Application Technology Running on Android Smartphone and Blackberry (International Conference on ICT for Smart Society)
[7] Sapargaliyev D 2012 Development of Mobile Learning in Higher Education of Russia (Proceedings of 2012 International Conference on Interactive Mobile and Computer Aided Learning)
[8] The Excellence Gateway. What is Mobile Learning? accessed on 15 Januari 2019. https://www.excellencegateway.org.uk/content/eg4692
[9] Ah-Fur Lai and Gon-Yi Lee 2015 Developing a Web-based Simulation-based Learning System for Enhancing Concepts of Linked-list Structures in Data Structures Curriculum (International Conference on Artificial Intelligence, Modelling and Simulation) pp 419-438.
[10] Yeh H and Chang C 2017 Android App Development for Teaching Reduced Forms of EFL Listening Comprehension to Decrease Cognitive Load (International Conference of Educational Innovation through Technology) pp 316-321
[11] Siboopimpa Y et al 2014 “I’m in ASEAN”: A Learning Media Application on Android Operating System (Third ICT International Student Project Conference)
[12] Brooke J 1996 SUS: a “quick and dirty” usability scale (Usability Evaluation in Industry) (London: Taylor & Francis) pp 189-194
[13] Nidhra S and Dondeti J 2012 Blackbox and Whitebox Testing Techniques – A Literature Review. (International Journal of Embedded Systems and Applications vol 2 no 2)
[14] Dwivedi S 2016 Software Development Life Cycle Models – A Comparative analysis (International Journal of Advanced Research in Computer and Communication Engineering vol 5 issue 2) pp 232-233