Development of an Android-based Machine Learning Student Problem Identification Tool Application at YPT Banjarmasin VHS

Muhammad Arisandy Rizky¹, Handaru Jati¹

¹Department of Electronics Technology and Informatics Education, Yogyakarta State University, Yogyakarta, Indonesia.

Abstract: The era of the 5.0 Industrial Revolution demands that we develop automation and digitalization technologies in various aspects of life, including education. Even Guidance and Counseling teachers who manually analyze counseling instrument items need assistance in swiftly and accurately analyzing instruments for hundreds of students. This research aims to support counselors in analyzing the Student Problem Identification Tool Instrument, which consists of 225 items, through student's Android devices, thereby enabling the prompt resolution of student issues. Through the stages of Research and Development (R&D), the Student Problem Identification Tool Application is developed using the Multinomial Logistic Regression method within Machine Learning. Research outcomes reveal that the application achieves an accuracy rate of 100% when compared to manual analysis by counselors and application-based analysis for 30 students. The average performance test result is 85.00%, and the feasibility test result is 96.30%, categorizing it as "Highly Feasible." In conclusion, Machine Learning facilitates the effective and efficient analysis of extensive data when supported by quality training data and the appropriate method selection for problem-solving.

Keywords: Android-based; Machine Learning; Multinomial Logistic Regression; Student Problem Identification Tool

Introduction

Vocational High Schools (VHS) are secondary education institutions that aim to prepare students for employment in specific fields. Irwanto (2021) suggests that VHS should focus on integrating classroom learning with real-world work situations in companies or industries. Yayasan Pendidikan Teknologi (YPT) VHS is a private vocational high school located in Banjarmasin. Similar to other schools, YPT Banjarmasin VHS enrolls students from diverse backgrounds who may face various challenges. It is crucial for Guidance and Counseling Teachers to address these student problems promptly, ensuring that they do not hinder the teaching and learning process for students.

Guidance and Counseling Teachers play a vital role in monitoring and supporting students' personal development through a range of service programs and accountable counseling procedures. According to Azahari et al. (2022), counseling involves providing students with guidance, information, and advice to help them reach their goals. For Guidance and Counseling Teachers to provide effective counseling services, they need to have a thorough understanding of their students' problems. Rosiani et al. (2022) explained that counseling guidance for new students can be provided through orientation services to help students adapt to and understand their new environment, thereby preventing problems that may hinder their learning and harm them.

How to Cite:
The availability of well-designed counseling instruments is crucial to accurately measure the intended aspects. Trisnawati et al. (2023) emphasizes that instilling and nurturing personality and character from an early age is crucial. Anshari (2019) explains that counseling management involves planning various counseling programs with clear objectives. Planning counseling services is crucial to establish the desired outcomes. Counseling instruments encompass both test and non-test tools. As stated by Winarno et al. (2019), test instruments are utilized to measure students' cognitive abilities, while non-test instruments are employed to assess affective and psychomotor aspects. Non-test standard inventory instruments commonly used in counseling services include the Problem Identification Tool, interest scales, and self-assessment scales.

The Problem Identification Tool is the counseling service instrument utilized by Guidance and Counseling Teachers at YPT Banjarmasin VHS. According to Putri et al. (2022), this questionnaire-based tool is employed to uncover common student problems. Fitril et al. (2022) explain that the Problem Identification Tool, commonly used by teenagers, helps them feel acknowledged regarding their current circumstances. Setyana et al. (2018) found that Guidance and Counseling Teachers require significant time to process the results of the Problem Identification Tool questionnaire, given the extensive range of problem areas and statement items included.

Ifdil et al. (2017) revealed that school counselors have been conducting suboptimal problem identification due to the manual processing and administration of the Problem Identification Tool, leading to significant effort and inefficiency in time consumption. However, Guidance and Counseling Teachers at YPT Banjarmasin VHS encounter challenges in manually analyzing the extensive Problem Identification Tool questionnaire, which comprises 225 instrument items, for numerous students. These difficulties arise due to limited numbers of Guidance and Counseling Teachers and the lack of supporting applications provided by the school, primarily due to budget constraints. The existing Problem Identification Tool applications available in the market exhibit shortcomings in accuracy and necessitate the use of computers, which is not feasible given the student population at YPT Banjarmasin VHS.

Based on the problems found and the considerations above, the development of an Android-based Machine Learning Student Problem Identification Tool Application is deemed necessary for YPT Banjarmasin VHS. Utilizing the Android devices owned by students offers a viable platform for this application. The employed Machine Learning method can emulate the capabilities of the Guidance and Counseling teachers, thereby providing accurate and automated analysis. Consequently, this application can assist counselors in making effective and efficient guidance decisions based on the urgency levels of the numerous students facing issues.

**Method**

This research adopts the Research and Development (R&D) method, which enables researchers to identify field problems, propose solutions, and determine suitable supporting tools for implementation. According to Oktarina et al. (2023), R&D aims to produce a beneficial product. In this case, the research aims to address the analysis instrument Problem-Revealing Tool at YPT Banjarmasin VHS by proposing the development of an application using the Multinomial Logistic Regression method in Machine Learning. Fahrizal et al. (2020) define Machine Learning as a branch of Artificial Intelligence that focuses on developing systems with the ability to learn automatically from datasets, enabling them to adapt to changes. Machine Learning can assist in analyzing and streamlining human work in various fields. It utilizes algorithms that learn from data to predict outcomes. As the algorithm absorbs more data, it can generate more accurate models or predictions based on that data (Slimani et al., 2022). Machine Learning as a field of computer science that utilizes statistical techniques to enable computer systems to learn from data without explicit programming, as conveyed by Patel et al. (2020), machine learning is a subcategory of Artificial Intelligence that involves constructing a framework capable of learning from experience without being explicitly programmed.

Kriswantara et al. (2022) highlight that Machine Learning is a statistical computational model primarily focused on predicting datasets. Tiyar et al. (2021) emphasize the importance of data collection for training the system, as the dataset serves as input for the learning process, enabling the machine to generate accurate analyses. Furthermore, Roihan et al. (2020) note that Machine Learning finds applications in various domains such as traffic, industry, medicine, technology, and reasoning, as it offers effective solutions to diverse problems. To demonstrate the effectiveness of Machine Learning in the Problem Identification Tool application, the research needs to undergo a series of development, testing, and usability assessments. By utilizing Machine Learning, the application can automatically handle analysis problems, assisting Guidance and Counseling Teachers in making informed counseling decisions for students facing challenges.
The application's user interface is developed based on the Android platform to ensure accessibility for students at any time and from anywhere. Sasmitha et al. (2023) explains that Android-based smartphones are easily portable and can be used independently, both at home and at school. Susanto et al. (2022) also elaborate on how smartphone users, including students, tend to spend a significant amount of time with their Android smartphones, which contributes to the increasing popularity of Android. Pioke et al. (2023) convey that students' interest in smartphones represents a significant impact of technological development on human lifestyle patterns. It is observed that most students and teachers at YPT Banjarmasin VHS already utilize Android-based smartphones for school exams and learning activities. Therefore, students have become accustomed to using smartphones as a learning tool.

![Figure 1. Waterfall Method (Pressman et al., 2015)](image-url)

Elhaloui et al. (2023) explain that there are three types of Machine Learning. Supervised learning aims to establish behavioral rules from a dataset containing previously labeled examples. Unsupervised learning aims to discover hidden patterns in unlabeled datasets. Reinforcement learning aims to enable an agent to develop in an environment and learn from its own experiences. The environment in which it operates must be measurable and have a function that evaluates the agent's performance. According to Kaewchada et al. (2023), there are two types of Supervised Learning, Classification (where the dependent variable is discrete) and Regression (where the dependent variable is continuous). In the Problem-Revealing Tool Application, the suitable method for handling more than two categories (multinomial or polychotomous) is Multinomial Logistic Regression (Supervised Learning). Multinomial logistic regression analyzes the relationship between a categorical response variable and multiple predictor variables. Kotha et al. (2023) describe that Multinomial Logistic Regression allows for multiple categories, and it utilizes maximum likelihood estimation to determine the probability of membership in these categories. Hosmer et al. (2000) describe Equation (1) used in multinomial logistic regression as follows:

\[
\text{Logit}P(Y = 1) = \alpha + \beta_1X_1 + \beta_2X_2 + \ldots + \beta_nX_n
\]  

(1)

By using logit transformation, we obtain the equations for the logit functions (2) and (3).

\[
P_1(x) = \ln \left[ \frac{P(Y = 1)|x}{P(Y = 0)|x} \right] \\
= \beta_{10} + \beta_{11}X_1 + \beta_{12}X_2 + \ldots + \beta_{1n}X_n
\]  

(2)

\[
P_2(x) = \ln \left[ \frac{P(Y = 2)|x}{P(Y = 0)|x} \right] \\
= \beta_{20} + \beta_{21}X_1 + \beta_{22}X_2 + \ldots + \beta_{2n}X_n
\]  

(3)

Based on both logit functions, the multinomial logistic regression model (4) and (5) can be obtained as follows:

\[
\pi_0(x) = \frac{1}{1 + \exp P_1(x) + \cdots + \exp P_n(x)}
\]  

(4)

\[
\pi_n(x) = \frac{\exp P_n(x)}{1 + \exp P_1(x) + \cdots + \exp P_n(x)}
\]  

(5)

The procedure used in the research and development of the Problem-Revealing Tool...
Application, which utilizes Android-based Machine Learning, employs the Waterfall method. This method is implemented with a systematic approach, where each step must be completed one by one and sequentially. This is why it is referred to as the waterfall method. Pressman et al. (2015) explained in figure 1 that there are five stages in the Waterfall method: Communication, Planning, Modeling, Construction, and Deployment. Almazaydeh et al. (2022) also elaborate that the waterfall model requires researchers to comprehensively define and document a stable set of requirements at the beginning of the project. The study involved 30 students from the 10th and 11th grades of the Computer and Network Engineering program at YPT Banjarmasin Vocational High School during the academic year 2022/2023 as the research subjects. Data were obtained through interviews with the Guidance and Counseling Coordinator and the collection of archival data from the previous year's Problem Identification Tools. Observation and questionnaires were used to observe student issues and identify supporting tool requirements. The research instruments were validated by relevant experts.

According to Djollong (2014), the percentage of feasibility can be obtained by summing up the quantitative data of the measurement results and then comparing it with the expected total. The application is considered valid if the minimum result reaches 75% with the predicate "feasible." Data analysis in this research utilizes descriptive statistical analysis techniques. Descriptive statistics provide a description of the research object through the obtained data without conducting analysis and making general conclusions.

**Result and Discussion**

The student Problem Identification Tool Application is done in several parts using Android Studio and WEKA. Ntobuo et al. (2023) explain that the choice of Android-based learning models is suitable for use in the learning process because the use of Android as a learning medium is engaging for students. Android Studios serves as an Integrated Development Environment (IDE), essentially being the official integrated development environment crafted specifically for creating applications on the Google Android operating system. Testing and verification are performed on the functionality of the developed program, which is then integrated in the final stage. Rihyanti et al. (2020) note that Android runs not only on smartphones but also on tablets, laptops, and other devices. Therefore, it can be concluded that the Android system exhibits a high level of compatibility with various devices.

The creation of the Machine Learning model is done using the WEKA application. Zahidi et al. (2021) convey that the Weka workbench can group algorithms and serve as a visualization tool for predictive modeling and data analysis. According to Özdemir et al. (2019), Weka has a collection of machine learning algorithms and standards for data mining processes. Evangelista et al. (2022) also explain that WEKA is capable of parallel experimentation with various algorithms, datasets, and different parameters, allowing for the construction of models in significantly less time and the collection of performance statistics comparisons. Abuhaija et al. (2023) explained WEKA aims to assist users in handling data, starting from the initial processing stage all the way through to data modeling. The number of data used to create this model is 432 from the Problem Identification Tool. Multinomial Logistic Regression is used with 30 iterations, utilizing the formula "weka.classifiers.functions.Logistic -R 1.0E-8 -M 30 -num-decimal-places 4" due to the categorical nature of the class labels of the Problem Identification Tool.

![Figure 2. The Home page](image)

The first splash logo is located in activity_splash.xml, displaying the UNY logo. After 3 seconds, it will redirect the user to the second splash logo. The second splash logo is located in activity_splash2.xml, displaying the YPT Banjarmasin VHS logo. After 3 seconds, it will redirect the user to the
Home page. The Home page is located in activity_cover.xml. In figure 2, there is a "Start" button that will redirect the user to the Content page. On this page, there are also "Help," "About," and "Profile" buttons.

The content page (Figure 3) located in activity_main.xml, contains a student biodata form and 225 items of the Problem Identification Tool with selectable responses for the students. At the bottom of the content page, there are input fields for "Other Issues Faced" and "Whom the problem wants to consult with." The "Classify Problem" button will redirect the user to the Results page. All activities on this page will be processed by MainActivity.java.

The Results page is located in activity_hasil.xml. In figure 4, the analysis results from the previous page are displayed. The analysis results can be saved by pressing the "Save As PDF" button, which will generate a PDF file located in the main File Manager page. All activities on the Results page will be processed by HasilActivity.java.

The completed Problem Identification Tool application has been evaluated by two expert judgments. In the performance test, expert judgment 1 achieved a performance percentage of 82.50%, while expert judgment 2 achieved a performance percentage of 87.50%. The average of these two performance results is 85%, which falls into the "Very Good" category. The performance results can be visualized through the graph in figure 5, where the x-axis represents expert judgment 1 and 2, and the y-axis represents the range of performance test percentage ratings.

The accuracy and quality of the Problem Identification Tool application, using the Multinomial Logistic Regression method, can be observed from the case findings presented in figure 6. Manual analysis results by the Guidance and Counseling Teacher on student issues indicate a similarity in the average scores for Personal Identity (DPI) and Social Relationships (HSO), both valued at 0.80. In this case, based on the number of rounded items selected by the student, the Guidance and Counseling Teacher determines that the main issue being experienced by the student is Personal Identity (DPI). The analysis results from the Problem Identification Tool application also indicate a similar outcome, which is Personal Identity (DPI). Therefore, it can be concluded that the Problem Identification Tool application consistently provides measurement and analysis results for students experiencing problems. Based on the comparison between the accuracy of the manual analysis conducted by the Guidance and...
Counseling Teacher and the analysis performed through the Problem Identification Tool application on 30 students, it can be determined that the accuracy of the Problem Identification Tool application reaches 100%. The validation results from the Guidance and Counseling Expert yielded 99.43%, while the Application Development Expert provided a validation result of 93.18%.

Figure 6. Comparison of Manual Analysis by Guidance Counselor and Analysis by Problem Identification Tool Application

The average of these two validation results is 96.30%, which falls into the "Very Good" category. The feasibility test results can be explained through the graph in figure 7, where the x-axis represents the validators, namely the Guidance and Counseling Expert and the Application Development Expert, while the y-axis represents the range of feasibility percentage ratings.

Figure 7. Graph of Feasibility Testing Results

The Problem Identification Tool application developed using the Multinomial Logistic Regression method has met the criteria of a good measurement tool. Reliability, the Problem Identification Tool application provides consistent measurement and analysis results for students experiencing problems. Standardization, the instrument used in the Problem Identification Tool application is aligned with the standard Problem Identification Tool for high school or vocational school level. Validity, the Problem Identification Tool application yields the same measurement and analysis results as the manual analysis conducted by the Guidance and Counseling Teacher. Bakhri et al. (2023) explained that Android-based learning media aims to introduce teachers to the use of technology and contribute to enhancing their digital literacy. Practicality, the Problem Identification Tool application file can be easily shared through the school’s WhatsApp group by the Guidance and Counseling Teacher. This aligns with what was stated by Adelia et al. (2023), that the utilization of Android in education greatly facilitates students in accessing it from anywhere and at any time. The analysis results from the Problem Identification Tool application in PDF format also facilitate administrative tasks for the Guidance and Counseling Teacher, reducing paper usage and file clutter. As conveyed by Indahsari et al. (2023), mobile learning simplifies the learning process, is practical, accessible anywhere and anytime, and is also paperless.

Conclusion

Based on the research and development of the Problem Identification Tool application, The analysis of the Problem Identification Tool instrument at YPT Banjarmasin VHS can be performed more easily through the implementation of the Multinomial Logistic Regression method in the Problem Identification Tool application, achieving an accuracy of 100%. The accuracy comparison is obtained from the manual analysis by the Guidance and Counseling Teacher and the analysis results from the Problem Identification Tool application on 30 students. The performance test of the Problem Identification Tool application at YPT Banjarmasin VHS is evidenced by an average percentage result of 85.00% for the functionality aspect. The feasibility test of the Problem Identification Tool application at YPT Banjarmasin VHS is evidenced by an average validation feasibility result of 96.30%, categorized as "Very Good".

Acknowledgments

All authors are grateful thanks to the Electronics Technology and Informatics Education, Yogyakarta State University.
Author Contribution
All authors participated in the creation of this entire article.

Funding
This research was self-funded and did not receive assistance from any third party.

Conflicts of interest
All authors declare that there are no conflicts of interest related to the publication of this paper.

References


