Journal of Theoretical and Applied Information Technology
15th January 2025. Vol.103. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

UTILIZING EXPLORATORY DATA ANALYSIS AND MACHINE LEARNING TO ENHANCE LEARNING QUALITY IN THERMODYNAMICS

ARWIZET KARUDIN[1], DESMARITA LENI[2*], REMON LAPISA[3], YUDA PERDANA KUSUMA[4]

[1,3]Department of Mechanical Engineering, Faculty of Engineering, Universitas Negeri Padang
[2] Departement of Mechanical Engineering, Faculty of Engineering, Universitas Muhammadiyah Sumatera Barat
[4]IJIMCE, Faculty of Engineering, Universitas Negeri Padang
E-mail: [1]arwizet@ft.unp.ac.id, [2]desmaritaleni@gmail.com

ABSTRACT

Students often face difficulties in understanding abstract concepts in thermodynamics, such as thermal system efficiency, temperature distribution, and inter-variable interactions in thermal phenomena. These challenges are compounded by traditional teaching methods, which typically rely only on text and mathematical calculations. This research aims to develop a lab data-based visualization model to enhance students' understanding of thermodynamics concepts. The model utilizes experimental data collected by Mechanical Engineering students at Universitas Negeri Padang on various instructional apparatus, including steam power plants, internal combustion engines, wind turbines, and crank mechanisms, gathered between 2022 and 2024. This visualization model includes several key features: a CRUD-based data storage system using Pymongo, 2D data visualization with Pygwalker, 3D visualizations in contour and surface diagrams, correlation analysis using heatmaps, and machine learning-based predictions with PyCaret. Evaluation results indicate that this model significantly improves students' comprehension of abstract thermodynamics concepts. Based on assessments from four thermodynamics expert lecturers, the information accuracy aspect received an average score of 4 (Good), visualization quality received an average score of 4.25 (Very Good), and ease of use received the highest average score of 4.5 (Very Good). The relevance of the model to learning objectives received an average score of 3.75, slightly below the Good category, indicating room for improvement in terms of educational relevance. Overall, this research demonstrates that integrating EDA and machine learning through the developed visualization model is effective in supporting more interactive and data-driven thermodynamics learning, aligned with the needs of education in the digital era.
Keywords: *Exploratory Data Analysis, thermodynamics, machine learning, correlation, analysis*

1. INTRODUCTION

Thermodynamics is a branch of physics that studies the relationships between heat, energy, and work within a system. It focuses on how energy is transferred and transformed, as well as how these processes affect the properties of matter, such as pressure, temperature, and volume. In mechanical engineering, thermodynamics is crucial for analyzing and designing machines and systems that utilize energy, such as combustion engines, turbines, and cooling systems. The fundamental principles of thermodynamics include the laws of thermodynamics, thermodynamic cycles, and thermodynamic processes. A solid understanding of the laws of thermodynamics enables students to enhance their ability to design and optimize more efficient and environmentally friendly systems, as well as to predict and address technical challenges encountered in real-world applications [1]. Despite its importance, thermodynamics is known as a challenging subject for many engineering students. Previous studies have found that thermodynamics is often perceived as one of the most complex and difficult subjects to grasp [2], [3].

Traditional teaching methods that rely on lectures and theoretical explanations tend to be less effective in helping students comprehend abstract and complex concepts. These methods often fail to provide students with opportunities to observe the practical applications of the theories being taught, resulting in difficulties in connecting theoretical concepts to real-world scenarios [4]. Consequently, many students struggle to understand and effectively apply thermodynamic principles. This is consistent with the findings of a study [5] involving 200 engineering students from various programs at the University of Mataram, Indonesia. The results

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

showed that most students faced difficulties in understanding abstract concepts in thermodynamics, such as the laws of thermodynamics, energy transformations, and thermodynamic cycles. The data revealed that only 40% of respondents could accurately explain the concepts of the first and second laws of thermodynamics. Furthermore, more than 60% of students experienced challenges in applying these concepts to real-world situations, such as analyzing thermal systems or heat engines. Therefore, there is a need for more interactive teaching methods and effective learning media to visualize thermodynamics concepts.

Exploratory Data Analysis (EDA) is a data analysis technique used to explore and visualize data to make it easier to understand. EDA helps identify patterns, trends, and relationships in data that might not be apparent through traditional analysis [6]. In the context of thermodynamics education, EDA can transform students' experimental data into more comprehensible graphs and diagrams, helping them connect theory with practice. Machine Learning, on the other hand, is a technology that enables computers to learn from data and make predictions or recommendations without explicit programming. In thermodynamics education, machine learning can be employed to simulate various scenarios in thermodynamic systems based on experimental data. This allows students to test how changes in parameters like temperature or pressure affect system performance, enhancing their understanding of the real-world impact of the theories they study. Currently, experimental data generated by students is often underutilized in the learning process. By employing EDA, this data can be visualized to provide students with direct insights into how thermodynamic theories are applied in real-world scenarios. Machine learning can then be used to create predictive models, enabling students to test various scenarios virtually, thereby deepening their understanding of abstract concepts such as thermodynamic cycles or the laws of thermodynamics. By integrating EDA and machine learning, thermodynamics education becomes more interactive. This creates a more effective and dynamic learning environment where experimental data is not merely an end product but a tool for interactive learning.

The use of data-driven technologies such as Exploratory Data Analysis (EDA) and Machine Learning in education has shown positive results in recent years. Research [7] demonstrates that the application of EDA in online statistics courses facilitates students' understanding of hidden patterns in data through visualization, leading to improved comprehension of previously abstract materials. Simplified data visualization through EDA allows students to see relationships between variables, making it easier to apply theoretical concepts. Research [8] on science education further explores how the implementation of EDA in exact sciences, including thermodynamics, enables students to more easily interpret laboratory experiment results. EDA transforms raw data into more accessible graphs and visualizations, helping students map the taught theories to the practical results they observe, effectively bridging the gap between theory and practice.

In addition to EDA, Machine Learning has also been introduced as a tool in thermodynamics education to simulate and estimate thermodynamic phenomena [9]. In thermodynamics learning, Machine Learning can be used to predict how changes in variables such as temperature or pressure affect the efficiency of thermal systems. This helps students understand the impact of these variable changes through simulations without the need for physical experiments. Consistent with research [10], Machine Learning can be used to develop predictive models that assist students in simulating various scenarios in thermodynamic systems. Through these simulations, students can explore the effects of parameter changes in real-time without conducting laboratory experiments. Another study [11] highlights that the combination of EDA and Machine Learning provides significant benefits in data-driven learning, particularly in engineering disciplines. The integration of these two methods not only accelerates students' understanding of complex concepts but also promotes interactivity and active participation in the learning process. With the help of EDA, students can visually analyze data, while Machine Learning enables them to predict and test theories under various conditions practically.

Based on previous studies, the application of EDA and Machine Learning has been proven to not only enhance students' understanding but also transform monotonous learning into a more interactive and dynamic process, aligned with current technological advancements. This research aims to develop a visualization-based learning model utilizing students' experimental data in thermodynamics education. Through this approach, it is expected that students' understanding of abstract thermodynamics concepts will improve through more concrete and easily comprehensible visual illustrations.

2. RESEARCH METHODS

The approach used in this research is a quantitative and experimental approach [8], [12], aimed at developing and testing a data visualization-based learning model to enhance students' understanding of thermodynamics concepts. This approach involves the collection of students' experimental data, the application of Exploratory Data Analysis (EDA) techniques for data visualization, and the use of Machine Learning to simulate various variable-change scenarios. This enables students to comprehend how these parameters influence thermodynamic systems. The final stage involves evaluating the model through internal testing and expert assessment to ensure the accuracy, visualization quality, learning relevance, and ease of use of the model.

1. Collection of Students' Experimental Data

The data used in this research comes from laboratory thermodynamics experiments conducted by students. Experimental data was collected from laboratory practices and independent tests carried out by students as part of thermodynamics learning from 2022 to 2024. This data covers various topics in thermodynamics, such as steam power plants, internal combustion engines, wind turbines, and crank mechanisms. This stage also includes data preprocessing aimed at cleaning anomalies and incomplete data to facilitate further analysis using EDA and machine learning [13].

2. EDA Visualization Design

The design of the visualization system based on Exploratory Data Analysis (EDA) aims to make it easier for students to understand experimental data more intuitively through interactive graphics [14]. This visualization includes several main features, such as Automated Visualization (Sweetviz), which automatically generates a comprehensive visual report from the data, enabling students to quickly see summaries and important insights. The 2D Visualization feature is used to display relationships between variables in a simpler manner, while 3D Visualization allows for data exploration with richer and more interactive visual depth, especially for datasets with more than two dimensions [15]. In addition, this visualization system is also equipped with correlation analysis using a correlation heatmap, which helps students understand relationships between variables through the visual representation of correlations. This feature enables the identification of significantly related variables in thermodynamic processes. The correlation heatmap uses the Pearson correlation equation, which can be calculated using the following formula:

$$r_{xy} = \frac{\Sigma xy}{(n-1)s_x.s_y} \quad (2)$$

where r_xy represents the Pearson correlation coefficient, Σxy denotes the sum of the products of x and y, n indicates the sample size, x stands for the independent variable, y represents the depen-dent variable, and S signifies the standard deviation [16]. The correlation coefficient ranges from -1 to 1. A value of -1 indicates a strong negative correlation between the two variables, a value of 0 indicates no correlation, and a value of 1 indicates a strong positive correlation.

By combining these various types of visualizations, students can explore data independently and test hypotheses more effectively, ultimately improving their understanding of abstract concepts in thermodynamics. This integration of dynamic visualizations provides an interactive learning experience, where students can observe the impact of each variable change in real-time, making previously abstract concepts more concrete and easier to understand.

3. Machine Learning Implementation

Machine Learning is utilized to build predictive models based on experimental data. These models are used to simulate various scenarios of variable changes, allowing students to understand how these parameters affect thermodynamic systems. The machine learning models in this study were developed using the PyCaret library. This library functions as an automation tool that enables users to quickly and easily build, train, and evaluate Machine Learning models without requiring deep knowledge of algorithms or complex programming [17], [18]. Using PyCaret, users can perform data preprocessing, select the best model from various available algorithms, and optimize the model through parameter tuning [19]. PyCaret is highly useful for predictive modeling in the context of data-driven education and research. The data used for training and testing the models is divided into two parts: 80% training data and 20% testing data. Each machine learning model is validated using cross-validation. This technique allows the training data to be divided into multiple subsets or folds, and iterations are performed on each subset, where one subset is used as testing data while the others serve as training data. The evaluation metrics used in this study consist of [20], [21]:

1. Mean absolute error (MAE)

   Measures the average absolute error between predicted values and actual values. The smaller the MAE, the better the model's performance.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(yi - \widehat{yi}) \quad (2)$$

2. Mean Squared Error (MSE)
   Measures the average squared difference between predicted values and actual target values in a dataset.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(yi - \widehat{yi})^2 \quad (3)$$

3. Root Mean Squared Error (RMSE)
   RMSE is the square root of MSE, used to interpret errors in the same units as the data.

$$MSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - \widehat{yi})^2} \quad (4)$$

4. R^2
   The coefficient of determination measures how well the predictor variables explain the variance in the actual data. Values close to 1 indicate a better model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(yi-\widehat{yi})^2}{\sum_{i=1}^{n}(yi-\overline{yi})^2} \quad (5)$$

5. Root Mean Squared Logarithmic Error (RMSLE)
   This metric is used to measure the error between predicted values and actual values in regression problems.

$$MSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\log(1+yi) - \log(1+\widehat{yi}))^2} \quad (6)$$

6. Mean Absolute Percentage Error (MAPE)
   This metric is useful for measuring absolute error as a percentage, making it suitable for evaluating relative error.

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{yi-\hat{y}}{yi}\right| \quad (7)$$

Where n is the total number of data points, yi is the actual value for the i-th data point, and $\widehat{yi}$ is the predicted value for the i-th data point. $\hat{y}$ represents the mean of all actual values. $\log(1 + yi)$ denotes the logarithm of the actual value incremented by 1, and $\log(1 + \widehat{yi})$ denotes the logarithm of the predicted value incremented by 1.

4. System Testing and Validation
   The visualization-based learning model for thermodynamics education will be evaluated through a series of tests to ensure its effectiveness. This evaluation involves two main aspects [12], [22], [23]:
   1. Internal Testing
      This testing aims to ensure that all functions and components within the model work as expected. Each feature, including data input, visualization, and prediction, will be thoroughly tested to verify that the system operates smoothly and without errors. This process also helps identify areas that may require further refinement to make the model more effective and user-friendly.
   2. Expert Assessment
      After completing internal testing, the model will be evaluated by experts in the field of thermodynamics. This assessment will cover several key aspects, including the accuracy of the information displayed, the quality of the visualizations, the alignment of the model with educational objectives, and its ease of use for students. The evaluation will be conducted using a Likert scale, enabling experts to objectively rate each aspect and provide a comprehensive overview of the model's performance and quality.

3. RESULTS AND DISCUSSION

In project-based thermodynamics learning, the testing of thermodynamic simulation tools such as steam power plants, internal combustion engines, wind turbines, and crank mechanisms is often conducted. However, so far, this test data has only been used to fulfill assignments given by lecturers. In fact, this data could be leveraged for more interactive learning. A study [24] shows that learning platforms combining real-time data from laboratories with visualization tools can create a more dynamic and relevant learning experience. In the digital era, data has become a critical element in decision-making and analysis across various fields. According to study [25], data serves as an interactive medium that helps students visualize experimental results. Modern visualization technologies enable students to better understand complex patterns and parameter relationships in thermodynamics. Research [26] found that the use of interactive visualization software in thermodynamics courses enhances students' analytical abilities, particularly in understanding the Rankine cycle and steam turbine efficiency. Additionally, study [27] demonstrates that data-driven simulations help students grasp isentropic and isobaric processes more effectively. Therefore, integrating data and visualization not only strengthens thermodynamics learning but also prepares students for challenges in the workforce, which increasingly emphasizes technology and data.

1. Data Collection
The data used in this study consists of various tests conducted on thermodynamic teaching

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

aids by Mechanical Engineering Education students at Universitas Negeri Padang. The tests include experiments on steam power plants, internal combustion engines, wind turbines, and crank mechanisms. During thermodynamics learning, students were divided into small groups of 4 to 5 members. Each group was assigned to observe and collect data from various teaching aids, such as:

a. Internal Combustion Engine Performance Dataset

In the internal combustion engine tests, students were trained to understand the concept of thermal efficiency, the impact of fuel consumption on performance, and strategies to improve energy efficiency. This also helped them grasp real-world applications of thermodynamic principles in the design and operation of combustion engines. Table 1 presents the statistics from the internal combustion engine performance tests collected from the practical experiments conducted by Mechanical Engineering Education students at Universitas Negeri Padang between 2022 and 2024. After data cleaning, the dataset comprises 157 samples with seven input variables and one output variable.

Table 1. Internal Combustion Engine Performance Dataset

Parameter	Variable	Min	Max	Mean
Engine Speed (rpm)	Input	2000	3500	2785.701
Fuel Volume (cc)	Input	7	15	11.701
Fuel Mass (gram)	Input	5.25	11.25	9.025
Manometer Reading (H) (m)	Input	2.5	10	5.285
Fuel Mass Flow Rate (Kg/s)	Input	0.00052	0.00067	0.00061
Specific Fuel Consumption (SFC)	Input	0.0024	0.0076	0.0047
Combustion Heat Energy (KJ/kg)	Input	177409.5	369603	286801.896
Thermal Efficiency (%)	Output	3.25	4.16	3.627

b. Savonius Wind Turbine Testing Dataset

This testing aims to provide students with a deep understanding of applying thermodynamic principles in renewable energy systems, particularly through the utilization of wind energy. Through this experiment, students learn about converting wind kinetic energy into mechanical energy and evaluating the performance efficiency of Savonius wind turbines. Parameters such as wind speed, turbine rotation, and braking force are used to measure the system's performance. In addition to helping students understand energy efficiency, this test also prepares them to face challenges in developing renewable energy technologies, such as designing more efficient and environmentally friendly wind turbines. During data collection, 112 samples were initially gathered, but after data cleaning, 75 valid samples remained. This dataset includes three main input variables and one output variable, as shown in Table 2. Students learn the importance of data quality through cleaning and analysis processes, as well as how to use valid data to model and visualize turbine performance accurately.

Table 2. Savonius Wind Turbine Dataset

Parameter	Variable	Min	Max	Mean
Wind Speed (V) (m/s)	Input	1.8	6.5	4.03333
Turbine Rotation (n) (rpm)	Input	675.7	1090.5	884.525
Braking Force (F) (N)	Input	0.2	3	1.525
Velocity Ratio (λ)	Output	1.45	3.24	2.05833

Journal of Theoretical and Applied Information Technology
15th January 2025. Vol.103. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

c. Crank Mechanism Dataset

The crank mechanism testing aims to measure the efficiency and accuracy of converting rotational motion into linear motion. This test also helps students understand the factors that influence the performance of the mechanism, such as friction and the discrepancies between ideal and actual conditions. The crank mechanism plays a crucial role in internal combustion engines and pumps, directly related to fundamental concepts in thermodynamics and fluid mechanics. Students gain insights into how crank motion drives pistons, forming the basis for understanding thermodynamic cycles, such as the Otto and Diesel cycles, used in internal combustion engines. Furthermore, this testing allows students to observe how friction and other factors reduce engine efficiency and how to analyze and mitigate these effects in real-world systems. The dataset used for this testing initially consisted of 98 samples, but after data cleaning, 77 valid samples remained. It includes seven input variables, such as crank angle, theoretical and actual piston stroke, piston mass, actual torque, and ideal torque. By utilizing this dataset, students can perform in-depth analyses of friction losses and understand the differences between theory and practice, which are invaluable for designing more efficient machines.

Table 3. Crank Mechanism Dataset

Parameter	Variable	Min	Max	Mean
Crank Angle (θ)	Input	0	180	90
Practical Piston Stroke (mm)	Input	0.0005	0.07	0.03732
Theoretical Piston Stroke (mm)	Input	0.105	0.245	0.17453
Error (%)	Input	99	99.9	99.6526
Mass (m) Kg (10^-3)	Input	5	370	212.316
Actual Torque (Ta)	Input	0.01	0.58	0.33368
Ideal Torque (Ti)	Input	0	112.01	58.1437
Friction Loss (%)	Output	93	100	98.2105
Friction Loss (%)	Output	93	100	98.2105

d. Steam Power Plant Dataset

The data was collected from previous student experiments and simulation assignments related to steam turbine trainers. The initial dataset consisted of 312 samples with 9 variables, including 8 input variables and 1 output variable, as listed in Table 4. After data cleaning, the number of samples was reduced to 300 due to missing values, duplicates, and significant scale discrepancies. Additionally, data on energy flow in each component of the steam power plant, such as the boiler, turbine, condenser, and generator, was also collected. This data cleaning process is essential to ensure reliable analysis and modeling.

Tabel 4. Steam Power Plant Dataset

Parameter	Variabel	Min	Max	Mean
Steam pressure inside the boiler (bar)	Input	3.44	4.36	3.9
Boiler steam temperature (°C)	Input	141.37	151.31	147.62
Fuel consumption (L/h)	Input	50	51	50
Turbine RPM	Input	1247.77	1285.164	1260.44
Inlet turbine temperature (°C)	Input	109	123	113
Outlet turbine temperature (°C)	Input	96	106	100
Inlet turbine pressure (bar)	Input	2.79	3.58	2.72
Outlet turbine pressure (bar)	Input	0.05	1	0.4
Generator Output (Watt)	Output	2.22	3.41	2.73

2. EDA Visualization Design

Exploratory Data Analysis (EDA) is a data analysis technique used to explore and visualize data to make it easier to understand. EDA helps identify patterns, trends, and relationships in data that may not be apparent through traditional analysis. To maximize students' learning experience in understanding thermodynamic concepts, this EDA-based visualization system is equipped with key features systematically designed to enhance interactivity and student engagement in the learning process. Each feature plays a strategic role, both in experimental data storage and in visually presenting the data.

1. Student Experimental Data Storage Feature

   In the field of science and engineering education, the ability to effectively store, access, and analyze experimental data has become a key aspect of enhancing students' understanding and engagement. Numerous previous studies have shown that effective data management plays a crucial role in enriching the learning process and facilitating the exploration of theoretical concepts through empirical data [28], [29]. In thermodynamics education, structured data storage from various experiments, such as tests on steam power plants, internal combustion engines, wind turbines, and crank mechanisms, provides students with continuous access to information that can be used for further analysis, including model development and comparison of results across practical sessions. Consistent with prior research [30], effective data storage systems encourage students to engage in data-driven learning, enabling them to revisit historical data and identify relevant patterns or trends in a scientific context. Easy access to historical data helps students understand the variability of experimental results and allows them to evaluate and compare outcomes under different testing conditions. An example of the student experimental data storage feature is shown in Figure 1.

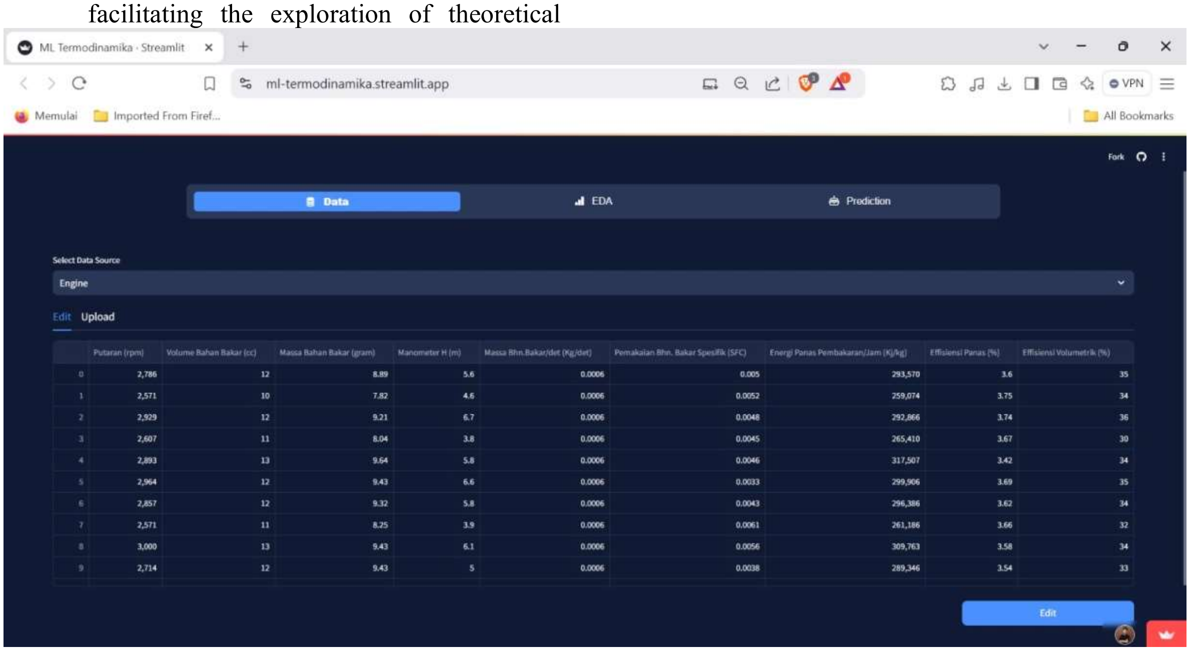

Figure 1: Student Experimental Data Storage Database

The database is equipped with CRUD (Create, Read, Update, Delete) functionality, built using the NoSQL database framework Pymongo, allowing students to efficiently store, access, and manage experimental data from practical sessions. This feature provides flexibility in storing data derived from various experimental apparatus, such as steam power plants, internal combustion engines, wind turbines, and crank mechanisms. With CRUD capabilities, experimental data can be stored in a structured format, enabling students to retrieve data as needed. This CRUD-based database enhances the quality of thermodynamics learning by empowering students to independently manage and organize experimental data. Students are not only presented with data as a final outcome but are also trained to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

appreciate the importance of structured data management. By accessing historical data, students can compare results across testing periods, identify anomalies, and observe long-term trends in experimental data. This fosters a deeper understanding of the variability and complexity of real-world systems while improving their analytical and data management skills.

2. 2D Visualization Feature

The 2D visualization feature in this study was developed using Pygwalker, an interactive Python-based visualization library. One of the main capabilities of Pygwalker is its ability to automatically generate visualizations based on the structure of the provided data. Students only need to select the variables they wish to visualize, and Pygwalker instantly generates relevant graphs. This feature saves time and effort, allowing students to focus on data analysis without worrying about the technicalities of graph creation. Additionally, the tool offers options to customize visualization elements, such as color, opacity, and plot type, making it easy to adjust graphs to meet specific analysis needs. The interface for the 2D visualization feature in the EDA system can be seen in Figure 2.

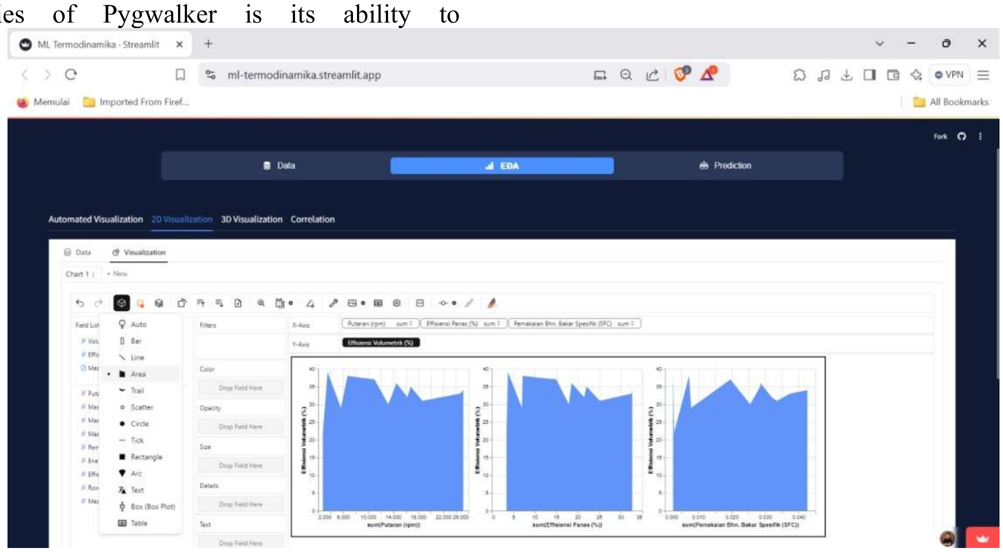

Figure 2: 2D Visualization Using Pygwalker

From Figure 3, it can be observed that the Pygwalker feature supports direct interaction with data through drag-and-drop functionality, filters, and various variable setting options. Students can easily change axis variables, filter data based on specific conditions, and add various visual elements to view data from different perspectives. For instance, in thermodynamics studies, students can explore how input variables like pressure and temperature affect output variables such as thermal efficiency with just a few clicks. This interactive feature enables deeper exploration of data, allowing students not only to view the final results but also to understand the processes within the data. In the digital age and data-driven learning, visualization technology plays a critical role in helping students better understand data. Interactive visualizations allow them to identify patterns and distributions that might be difficult to detect using traditional methods, making abstract concepts easier to grasp. In line with research [31], interactive visualization tools have been shown to support students in drawing more accurate conclusions, as they can explore data from various angles and variables to meet their specific analysis needs. These findings indicate that visualization methods can be utilized as tools to aid in understanding complex concepts, particularly in the field of thermodynamics.

3. 3D Visualization with Contour and Surface Area Diagrams

In thermodynamics education, 3D visualization plays a crucial role in providing a deeper understanding of complex physical phenomena. Representations in the form of contour diagrams and surface area plots allow students to observe variable interactions in three-dimensional space, offering a richer perspective compared to 2D graphs. These visualizations are particularly useful for explaining concepts such as temperature distribution, pressure variations, or other variables affecting thermodynamic systems. Contour diagrams help students visually understand heat transfer and pressure differences, which are critical elements in thermodynamics. These representations enable them to see how variables change spatially, identify relevant patterns, and conduct a deeper analysis of the physical phenomena at play. On the other hand, 3D surface area plots provide an even richer representation by showcasing the relationships between three variables simultaneously. This type of visualization is especially beneficial when students aim to comprehend the complex interactions between variables, such as how pressure and temperature jointly affect the efficiency of a thermal system. In a surface area plot, data is mapped onto a three-dimensional surface, where each point on the surface represents a combination of values for three different variables. For example, in a crank mechanism experiment, students can use a surface area plot to observe how changes in crank angle and mass influence the generated torque. This approach allows students to identify areas on the graph that indicate optimal results. Surface area plots provide a comprehensive view of interdependent variables, helping students understand how changes in one variable affect the overall outcome of the system. The display of the 3D visualization feature integrated into the EDA system can be seen in Figure 3.

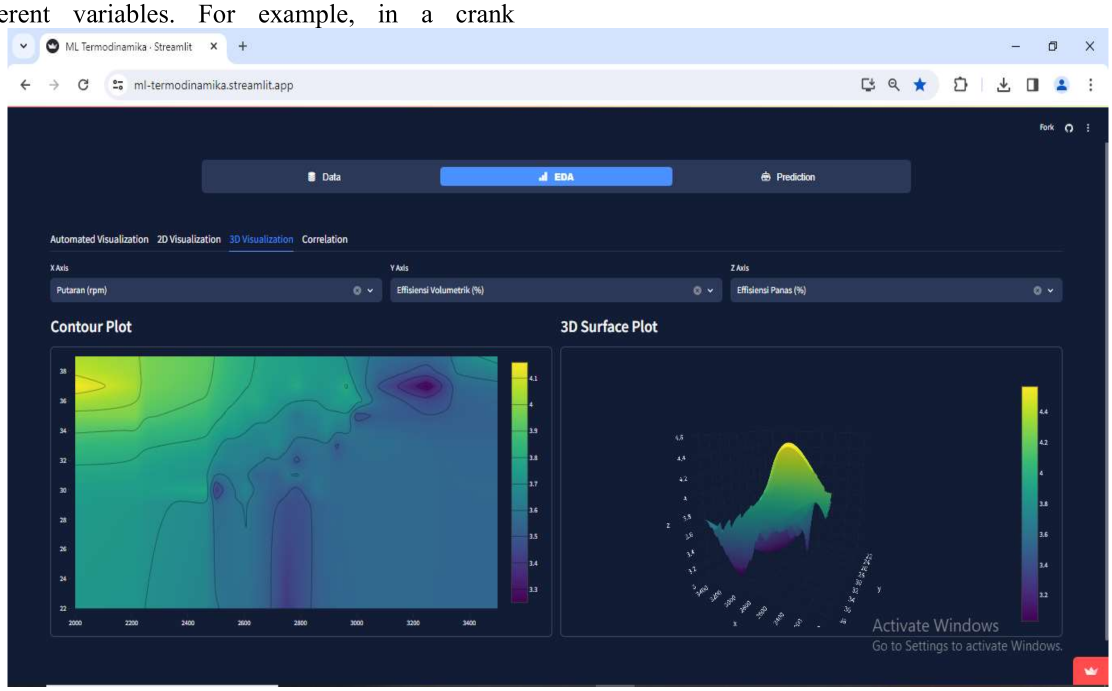

Figure 3: 3D Visualization

Recent studies have shown that 3D visualization significantly enhances students' understanding of complex concepts in science and engineering education, including thermodynamics. 3D visualization allows students to explore variable relationships in depth and helps them grasp phenomena that are challenging to explain using only 2D graphs or theoretical explanations. According to study [32], interactive 3D representations help students visualize abstract concepts and improve their understanding of parameter interactions within a thermal system. Study [33] on the benefits of 3D visualization in fostering active learning highlights how students actively engage with data and explore interrelated variables. The study found that when students manipulate 3D visualizations, they not only find it easier to understand the data but also show improvements in analytical skills and the ability to draw more critical conclusions from experimental data. With 3D representations, thermodynamics students can quickly identify how changes in parameters such as temperature, pressure, or crank angle affect the overall system performance. This aligns with the increasingly emphasized data-driven learning approach in engineering education, where students are encouraged to view processes holistically and develop a deeper understanding through detailed visual exploration.

4. Correlation Analysis Using Heatmap

A correlation heatmap is a visual tool used to display relationships or correlations between various variables in a dataset. In thermodynamics research, correlation heatmaps play a crucial role in helping students understand the interconnections between variables that may influence the performance or outcomes of a system. With a heatmap, students can visually identify which variables are significantly correlated, providing deeper insights into how thermodynamic systems function and how these

Journal of Theoretical and Applied Information Technology
15th January 2025. Vol.103. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

variables interact. In a correlation heatmap, each cell represents the correlation value between two variables, typically illustrated with a color scale to indicate the strength and direction of the relationship, as shown in Figure 4. Strong positive correlations are displayed in yellow, while strong negative correlations are visualized in purple. This allows students to quickly identify which variables have strong correlations, whether positive or negative, with others.

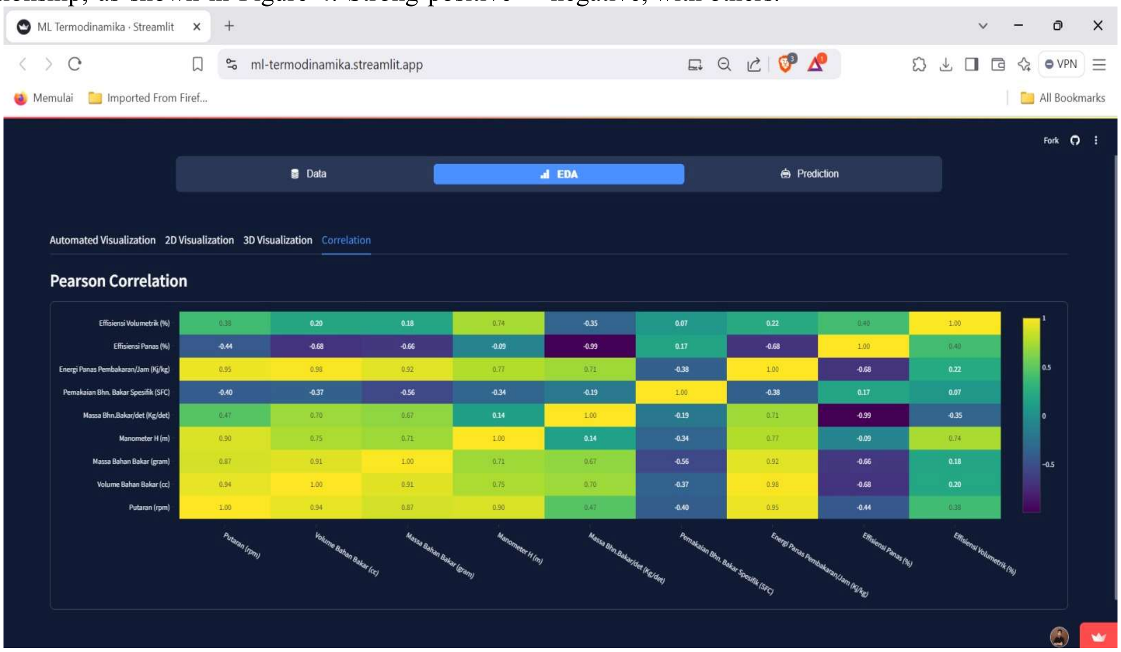

Figure 4: Visualization of Correlation Analysis Using Heatmap

Correlation heatmaps also serve as a tool for identifying key variables that may require further attention in thermodynamic analysis. Variables strongly correlated with output variables or other important system variables can become the focus of deeper analysis. Students can use this information to narrow their analysis to the most relevant variables, saving time and effort in understanding the overall system. In experiments involving crank mechanisms, for instance, a correlation heatmap can reveal relationships between crank angle, piston mass, and generated torque. If students observe a strong correlation between crank angle and actual torque, they can focus on those parameters to optimize the mechanism's performance. Conversely, if a variable shows no significant correlation with others, students may choose to exclude it from further analysis, allowing them to focus on variables with greater impact. Using a correlation heatmap provides students with a highly beneficial visual framework for analyzing and prioritizing variables in thermodynamics studies. By recognizing significant correlation patterns, students can direct their efforts toward key variables that most influence experimental outcomes or system design. This enhances their analytical skills, provides a more structured understanding, and prepares them for challenges in the workforce, where data analysis is increasingly emphasized.

3. Machine Learning Design

The Machine Learning approach used in this research leverages PyCaret, a Python-based library designed to simplify the process of modeling and evaluating Machine Learning. PyCaret enables users to automate various Machine Learning tasks, including model selection, training, evaluation, and optimization [17], [18]. In thermodynamics education, PyCaret assists students in conducting faster and more accurate analysis and predictions without requiring a deep understanding of the technical details of Machine Learning algorithms. Thermodynamic data collected in experiments often involve multiple variables with complex relationships, such as system efficiency. To handle such data, PyCaret offers a variety of Machine Learning algorithms that can be selected based on the analysis needs. Model selection in PyCaret is performed by comparing several suitable algorithms and choosing the one with the best performance based on evaluation metrics. PyCaret makes it easy for students to objectively evaluate each model, providing insights into which model is most suitable for their data. This feature allows students to input their practical experiment data into designated input

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

columns based on the variables measured during the experiments. Using these input variables, students can predict target variables in thermodynamic phenomena, such as thermal efficiency in internal combustion engine performance tests, friction percentage in crank mechanisms, or generator output power in steam power plants. The implementation of the Machine Learning model using PyCaret is displayed in Figure 5.

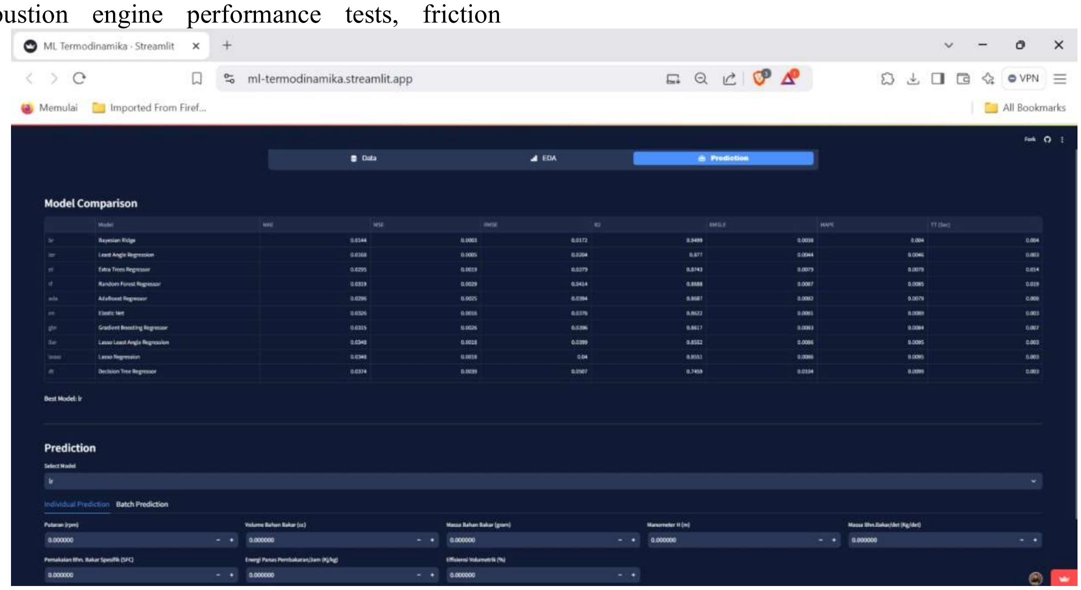

Figure 5: Machine Learning Model Interface Using PyCaret

Through a simple interface, this feature allows students to quickly test various algorithms and compare prediction results with experimental outcomes. Students can evaluate multiple models in a short time, obtaining comprehensive results without the need for complex programming. This accelerates the learning process and enables students to gain insights from their data more efficiently. One example of thermal efficiency prediction results from internal combustion engine performance tests using PyCaret demonstrates that several regression models perform exceptionally well in predicting engine thermal efficiency. Based on evaluation metrics such as MAE, MSE, RMSE, R-squared, RMSLE, MAPE, and execution time (TT), Linear Regression (lr) emerged as the best model with an R-squared value of 0.9779, indicating an almost perfect predictive capability in understanding the relationship between input and output variables. This model also showed low error rates (MAE: 0.0115), making it highly accurate for predicting thermal efficiency. Other models, such as Ridge and Bayesian Ridge (br), also performed well, while the Passive Aggressive Regressor (par) showed poor performance with a negative R-squared value of -25.0259. The results of the Machine Learning models using PyCaret are summarized in Table 5.

Table 5. Prediction Results of Thermal Efficiency Using Machine Learning Models

Model	MAE	MSE	RMSE	R2	RMSLE	MAPE	TT (Sec)
lr	0.0115	0.0002	0.0133	0.9779	0.0029	0.0031	0.282
ridge	0.0164	0.0004	0.0193	0.9572	0.0042	0.0046	0.003
br	0.0144	0.0003	0.0172	0.9499	0.0038	0.004	0.004
lar	0.0168	0.0005	0.0204	0.877	0.0044	0.0046	0.003
et	0.0295	0.0019	0.0379	0.8743	0.0079	0.0078	0.014
rf	0.0319	0.0029	0.0414	0.8688	0.0087	0.0085	0.019
ada	0.0296	0.0025	0.0394	0.8687	0.0082	0.0079	0.008
en	0.0326	0.0016	0.0376	0.8622	0.0081	0.0089	0.003
gbr	0.0315	0.0026	0.0396	0.8617	0.0083	0.0084	0.007

llar	0.0348	0.0018	0.0399	0.8552	0.0086	0.0095	0.003
lasso	0.0348	0.0018	0.04	0.8551	0.0086	0.0095	0.003
dt	0.0374	0.0039	0.0507	0.7459	0.0104	0.0099	0.003
omp	0.0934	0.0141	0.1123	-0.129	0.0242	0.0255	0.003
knn	0.093	0.0142	0.1106	-0.1718	0.0237	0.0252	0.005
lightgbm	0.1137	0.0223	0.1372	-0.447	0.0293	0.0309	0.005
dummy	0.1137	0.0223	0.1372	-0.447	0.0293	0.0309	0.003
huber	0.217	0.1207	0.285	-7.3916	0.064	0.0591	0.004
par	0.3972	0.3169	0.4899	-25.0259	0.1126	0.1062	0.003

With the predictive results from the Machine Learning models, students can gain deeper insights into how variables such as engine speed, fuel volume, and temperature influence the thermal efficiency of an internal combustion engine. In thermodynamics education, this understanding is crucial, as thermal efficiency is one of the key parameters in evaluating the performance of thermal systems. Accurate predictions from the models allow students to explore different scenarios, such as how increasing temperature or altering rotational speed would impact thermal efficiency. This engagement enhances the learning process as students can directly observe the relationships between thermodynamic parameters. The use of Machine Learning provides a more interactive learning experience, enabling students to make predictions based on real experimental data. Students can predict thermal efficiency for conditions that have not been tested in the laboratory, allowing them to explore scenarios without the need for time- and resource-intensive physical experiments. This practical and efficient approach helps students understand the effects of input variable changes on output variables. Additionally, the experience of using Machine Learning prepares students for data-driven decision-making in the professional world. They are not only taught thermodynamics theory but also how to apply it in data analysis and prediction using modern technology. This application of Machine Learning bridges the gap between theory and practice, making learning more relevant and contextual in the digital era. Overall, the Machine Learning approach using PyCaret helps students not only to interpret experimental results but also to perform effective predictive analyses. This enriches thermodynamics education by enabling students to combine technical knowledge with data analysis skills, ultimately providing them with a more comprehensive and profound understanding of the concepts they study.

4. System Testing and Validation

The evaluation of the thermodynamic visualization model was conducted in two main stages: internal testing and expert assessment, involving thermodynamics lecturers from Universitas Negeri Padang. Internal testing aimed to ensure that the visualization model functions as expected. During this phase, each feature of the model was tested to confirm smooth operations, including quick responsiveness to user inputs and seamless transitions between features. This evaluation also included testing the visualization loading times, which needed to remain within ideal limits to ensure users could access the model without difficulty. Based on the results, the model demonstrated stable performance and successfully integrated simulation data and experimental data from students with high accuracy. No technical disruptions or recurring issues were found that could hinder user experience. Next, the model was evaluated by experts using a Likert scale to measure various aspects, including the accuracy of the presented information, visualization quality, the relevance of the model to the learning process, and ease of use. The assessment was conducted by four lecturers experienced in teaching thermodynamics at Universitas Negeri Padang. The results of this evaluation are presented in Figure 6.

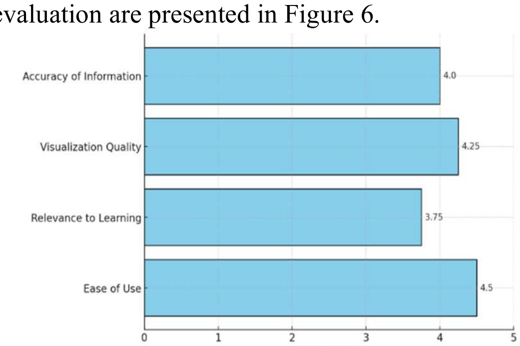

Figure 6: Evaluation Results of the Visualization Model by Experts

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Based on the evaluation results, the accuracy of information aspect received an average score of 4, classified as "Good," indicating that the visualization model effectively conveys thermodynamics concepts accurately. For the visualization quality aspect, experts gave an average score of 4.25, suggesting that the model presents high-quality visualizations, rated between "Good" and "Very Good." The relevance of the model to learning objectives scored an average of 3.75, slightly below "Good," showing that while the model aligns with educational goals, there is room for improvement to better meet learning needs. The ease of use aspect received a high average score of 4.5, indicating that the model is considered user-friendly with a responsive and intuitive user experience. Overall, the evaluation results reflect positive feedback from the experts, demonstrating that the model effectively delivers information and is easy to use. However, there are opportunities to enhance its relevance to learning objectives to better support students in understanding thermodynamics concepts.

This visualization-based learning model in thermodynamics aligns with current educational trends that heavily integrate digital technologies [34], [35], [36]. In an era where traditional approaches often fail to address the needs of interactive and practical learning, the visualization and data analysis-based learning model designed in this study offers a novel approach. By utilizing tools such as EDA visualizations, correlation heatmaps, and Machine Learning predictions through PyCaret, the model enables students to learn through hands-on experience and in-depth analysis—an approach more suited to the digital generation's learning style. The study's findings indicate that a data-driven visualization model can enhance students' understanding of abstract thermodynamics concepts. Through interactive 2D and 3D visualizations, students can directly observe the relationships and interactions between variables in thermodynamic systems. With Machine Learning-based prediction features, students also have the opportunity to test hypothetical scenarios that are difficult to implement physically. They can simulate changes in certain parameters to observe their impact on target variables.This model positively impacts students' learning of thermodynamics. Instead of relying solely on theory or instructor-led instruction, students actively engage in the learning process through data exploration. They can use experimental data to make predictions about thermal efficiency in internal combustion engines or identify the most influential variables in steam power plant systems. This allows students to practice critical thinking, analyze data, and interpret results directly. By bridging the gap between theory and practice, this model equips students with the skills needed in a data-driven world, enabling them to better understand thermodynamic concepts while fostering analytical and problem-solving abilities essential for their future careers.

5. CONCLUSION

This study demonstrates that the visualization-based learning model developed using students' experimental data in thermodynamics education is effective in enhancing understanding of abstract concepts. The model successfully visualizes complex thermodynamic phenomena by transforming practical data into easily comprehensible visual representations. Features such as data storage databases, 2D and 3D visualizations, correlation analysis with heatmaps, and predictions using machine learning significantly enrich the learning process. The visualization model helps students grasp relationships and interactions between variables in thermodynamic systems, which are often difficult to understand theoretically. These visualizations make abstract concepts more tangible, enabling students to connect theory with real-world phenomena in thermal systems and reinforcing their understanding through interactive, data-driven experiences. The evaluation of this model yielded positive results, with high ratings for information accuracy, visualization quality, and ease of use. Although there is room to improve the model's relevance in supporting more specific learning objectives, overall, it is considered effective and appropriate for the learning process. This study demonstrates that the integration of EDA and machine learning offers an innovative approach to science and engineering education, supporting a learning process that is more interactive, practical, and aligned with industry needs in the digital era.

ACKNOWLEDGMENT

The author expresses gratitude to the Directorate of Research, Technology, and Community Service (DRTPM) for funding support through the Research Grant under the Fundamental Research Scheme, as stated in Decision Letter No. 069/E5/PG.02.00.PL/2024 and Contract Agreement No. 2646/UN35.15/LT/2024. This funding has greatly contributed to the smooth execution of the research and the completion of this article.

The author also extends appreciation to the institutions and individuals who provided support,

suggestions, and facilities throughout the research process. It is hoped that the results of this research will offer meaningful benefits and contributions to the relevant field of study.

REFERENCES

[1] J. G. Acevedo, G. Valencia Ochoa, and L. G. Obregon, "Development of a new educational package based on e-learning to study engineering thermodynamics process: combustion, energy and entropy analysis," *Heliyon*, vol. 6, no. 6, p. e04269, Jun. 2020, doi: 10.1016/j.heliyon.2020.e04269.

[2] L. Atarés, M. J. Canet, M. Trujillo, J. Vte. Benlloch-Dualde, J. Paricio Royo, and A. Fernandez-March, "Helping Pregraduate Students Reach Deep Understanding of the Second Law of Thermodynamics," *Educ. Sci.*, vol. 11, no. 9, p. 539, Sep. 2021, doi: 10.3390/educsci11090539.

[3] J.-C. De Hemptinne *et al.*, "A View on the Future of Applied Thermodynamics," *Ind. Eng. Chem. Res.*, vol. 61, no. 39, pp. 14664–14680, Oct. 2022, doi: 10.1021/acs.iecr.2c01906.

[4] S. A. Finkenstaedt-Quinn *et al.*, "Capturing student conceptions of thermodynamics and kinetics using writing," *Chem. Educ. Res. Pract.*, vol. 21, no. 3, pp. 922–939, 2020, doi: 10.1039/C9RP00292H.

[5] S. Husein, Gunawan, A. Harjono, and S. Wahyuni, "Problem-Based Learning with Interactive Multimedia to Improve Students' Understanding of Thermodynamic Concepts," *J. Phys. Conf. Ser.*, vol. 1233, no. 1, p. 012028, Jun. 2019, doi: 10.1088/1742-6596/1233/1/012028.

[6] Dr. T. H. S. Tariq and P. S. Aithal, "Visualization and Explorative Data Analysis," *SSRN Electron. J.*, 2023, doi: 10.2139/ssrn.4400256.

[7] W. Villegas-Ch, M. Román-Cañizares, and X. Palacios-Pacheco, "Improvement of an Online Education Model with the Integration of Machine Learning and Data Analysis in an LMS," *Appl. Sci.*, vol. 10, no. 15, p. 5371, Aug. 2020, doi: 10.3390/app10155371.

[8] A. Karudin, D. Leni, R. Lapisa, Y. P. Kusuma, and M. R. Abbas, "Design of Tools for Visualizing Thermodynamic Concepts in Steam Power Plant Trainer Processes with Web-Based Exploratory Data Analysis (EDA)," *JOIV Int. J. Inform. Vis.*, vol. 8, no. 3, p. 1134, Sep. 2024, doi: 10.62527/joiv.8.3.2139.

[9] F. Jirasek and H. Hasse, "Perspective: Machine Learning of Thermophysical Properties," *Fluid Phase Equilibria*, vol. 549, p. 113206, Dec. 2021, doi: 10.1016/j.fluid.2021.113206.

[10] M. T. Hughes, G. Kini, and S. Garimella, "Status, Challenges, and Potential for Machine Learning in Understanding and Applying Heat Transfer Phenomena," *J. Heat Transf.*, vol. 143, no. 12, p. 120802, Dec. 2021, doi: 10.1115/1.4052510.

[11] S. Amer-Yahia, "Towards AI-powered data-driven education," *Proc. VLDB Endow.*, vol. 15, no. 12, pp. 3798–3806, Aug. 2022, doi: 10.14778/3554821.3554900.

[12] Faculty of Engineering, Universitas Negeri Padang, Indonesia and Waskito, "Integration of Mobile Augmented Reality Applications for Engineering Mechanics Learning with Interacting 3D Objects in Engineering Education," *Int. J. Inf. Educ. Technol.*, vol. 14, no. 3, pp. 354–361, 2024, doi: 10.18178/ijiet.2024.14.3.2057.

[13] D. A. Bhanage, A. V. Pawar, and K. Kotecha, "IT Infrastructure Anomaly Detection and Failure Handling: A Systematic Literature Review Focusing on Datasets, Log Preprocessing, Machine & Deep Learning Approaches and Automated Tool," *IEEE Access*, vol. 9, pp. 156392–156421, 2021, doi: 10.1109/ACCESS.2021.3128283.

[14] Y. Q. A. Syed, G. Hossain, and S. P. Mohammed, "Introducing Data Visualization to High School Students," in *2024 IEEE Integrated STEM Education Conference (ISEC)*, Princeton, NJ, USA: IEEE, Mar. 2024, pp. 1–7. doi: 10.1109/ISEC61299.2024.10664915.

[15] [15] A. Páez and G. Boisjoly, "Exploratory Data Analysis," in *Discrete Choice Analysis with R*, in Use R! , Cham: Springer International Publishing, 2022, pp. 25–64. doi: 10.1007/978-3-031-20719-8_2.

[16] D. Leni, A. Karudin, M. R. Abbas, J. K. Sharma, and A. Adriansyah, "Optimizing stainless steel tensile strength analysis: through data exploration and machine learning design with Streamlit," *EUREKA Phys. Eng.*, no. 5, pp. 73–88, Sep. 2024, doi: 10.21303/2461-4262.2024.003296.

[17] P. Whig, K. Gupta, N. Jiwani, H. Jupalle, S. Kouser, and N. Alam, "A novel method for diabetes classification and prediction with Pycaret," *Microsyst. Technol.*, vol. 29, no. 10, pp. 1479–1487, Oct. 2023, doi: 10.1007/s00542-023-05473-2.

[18] P. R. Pinheiro and T. S. Rodrigues, "Optimization of Regression Models for Rental Price Recommendations: An Integrated Approach Using PyCaret and Gaussian Analytic Hierarchy Process," in *Cybernetics and Control Theory in Systems*, vol. 1119, R. Silhavy and P. Silhavy, Eds., in Lecture Notes in Networks and Systems, vol. 1119. , Cham: Springer Nature Switzerland, 2024, pp. 504–512. doi: 10.1007/978-3-031-70300-3_37.

[19] A. R *et al.*, "Automating Machine Learning Model Development: An OperationalML Approach with PyCARET and Streamlit," in *2023 Innovations in Power and Advanced Computing Technologies (i-PACT)*, Kuala Lumpur, Malaysia: IEEE, Dec. 2023, pp. 1–6. doi: 10.1109/i-PACT58649.2023.10434389.

[20] E. Vivas, H. Allende-Cid, and R. Salas, "A Systematic Review of Statistical and Machine Learning Methods for Electrical Power Forecasting with Reported MAPE Score," *Entropy*, vol. 22, no. 12, p. 1412, Dec. 2020, doi: 10.3390/e22121412.

[21] D. Leni, D. S. Kesuma, Maimuzar, Haris, and S. Afriyani, "Prediction of Mechanical Properties of Austenitic Stainless Steels with the Use of Synthetic Data via Generative Adversarial Networks," in *The 7th Mechanical Engineering, Science and Technology International Conference*, MDPI, Feb. 2024, p. 4. doi: 10.3390/engproc2024063004.

[22] D. Höttecke and D. Allchin, "Reconceptualizing nature-of-science education in the age of social media," *Sci. Educ.*, vol. 104, no. 4, pp. 641–666, Jul. 2020, doi: 10.1002/sce.21575.

[23] Z. Zulyusri, I. Elfira, L. Lufri, and T. A. Santosa, "Literature Study: Utilization of the PjBL Model in Science Education to Improve Creativity and Critical Thinking Skills," *J. Penelit. Pendidik. IPA*, vol. 9, no. 1, pp. 133–143, Jan. 2023, doi: 10.29303/jppipa.v9i1.2555.

[24] A. Alam, "Media Multitasking with M-Learning Technology in Real-Time Classroom Learning: Analysing the Dynamics in Formal Educational Settings for the Future of E-Learning in India," in *2023 2nd International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN)*, Villupuram, India: IEEE, Apr. 2023, pp. 1–6. doi: 10.1109/ICSTSN57873.2023.10151509.

[25] S. L. Franconeri, L. M. Padilla, P. Shah, J. M. Zacks, and J. Hullman, "The Science of Visual Data Communication: What Works," *Psychol. Sci. Public Interest*, vol. 22, no. 3, pp. 110–161, Dec. 2021, doi: 10.1177/15291006211051956.

[26] J. C. Domínguez *et al.*, "MATLAB applications for teaching Applied Thermodynamics: Thermodynamic cycles," *Comput. Appl. Eng. Educ.*, vol. 31, no. 4, pp. 900–915, Jul. 2023, doi: 10.1002/cae.22613.

[27] M. M. El-Awad, "The development of Thermax123: An educational Excel-based modelling platform for thermo-fluid analyses: An educational Excel-based modelling platform for thermo-fluid analyses," Sep. 24, 2024. doi: 10.31224/3949.

[28] M. M. Asad, N. Hussain, M. Wadho, Z. H. Khand, and P. P. Churi, "Integration of e-learning technologies for interactive teaching and learning process: an empirical study on higher education institutes of Pakistan," *J. Appl. Res. High. Educ.*, vol. 13, no. 3, pp. 649–663, Jun. 2021, doi: 10.1108/JARHE-04-2020-0103.

[29] X. Shu and X. Gu, "An Empirical Study of A Smart Education Model Enabled by the Edu-Metaverse to Enhance Better Learning Outcomes for Students," *Systems*, vol. 11, no. 2, p. 75, Feb. 2023, doi: 10.3390/systems11020075.

[30] S. L. Dodman, K. Swalwell, E. K. DeMulder, J. L. View, and S. M. Stribling, "Critical data-driven decision making: A conceptual model of data use for equity," *Teach. Teach. Educ.*, vol. 99, p. 103272, Mar. 2021, doi: 10.1016/j.tate.2020.103272.

[31] L. Perkhofer, C. Walchshofer, and P. Hofer, "Does design matter when visualizing Big Data? An empirical study to investigate the effect of visualization type and interaction use," *J. Manag. Control*, vol. 31, no. 1–2, pp. 55–95, Apr. 2020, doi: 10.1007/s00187-020-00294-0.

[32] K. Seshadri, P. Liu, and D. R. Koes, "The 3Dmol.js Learning Environment: A Classroom Response System for 3D Chemical Structures," *J. Chem. Educ.*, vol. 97, no. 10, pp. 3872–3876, Oct. 2020, doi: 10.1021/acs.jchemed.0c00579.

[33] S. Chatterjee *et al.*, "Click, Zoom, Explore: Interactive 3D (*i* -3D) Figures in Standard Teaching Materials (PDFs)," *J. Chem. Educ.*, vol. 98, no. 11, pp. 3470–3475, Nov. 2021, doi: 10.1021/acs.jchemed.1c00296.

[34] Z. Rahmatika, V. Wiliyanti, R. Diani, E. Daenuri, and C. A. S. A. Putri, "Science, Technology, Engineering, and Mathematics (STEM) approach in physics learning: Meta-analysis study," presented at the PROCEEDINGS OF THE TEGAL INTERNATIONAL CONFERENCE ON APPLIED SCIENCES 2022: Applied Science Research Post-Covid-19 Pandemic, Tegal,

ISSN: **1992-8645** www.jatit.org E-ISSN: **1817-3195**

Indonesia, 2024, p. 020002. doi: 10.1063/5.0200999.

[35] Y. Li *et al.*, "On Computational Thinking and STEM Education," *J. STEM Educ. Res.*, vol. 3, no. 2, pp. 147–166, Jul. 2020, doi: 10.1007/s41979-020-00044-w.

[36] A. W. Prananta, R. R. P. Megahati S, N. Susanto, and J. H. Raule, "Transforming Education and Learning through Chat GPT: A Systematic Literature Review," *J. Penelit. Pendidik. IPA*, vol. 9, no. 11, pp. 1031–1037, Nov. 2023, doi: 10.29303/jppipa.v9i11.5468.