Utilization of Machine Learning for Property Price Segmentation and Prediction

Akbar Andriansyah [1], Mulki Djenfik Dzulkarnain[2], Afni Izzah Afkarinah[3], Fadel Amili[4], Gilang Ramadhika[5], Nurvelly Rosanti[6], Siti Nurbaya Ambo[7], Yana Andharani[8], Mirza Sutrisno[9]

[1,2,3,4,5,6,7,8,9] Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jakarta

✉ akbarandriansyah456 @gmail.com

ARTICLE INFO	ABSTRACT
	Advances in digital technology have encouraged the utilization of artificial intelligence, especially machine learning, in various sectors, including property price analysis. However, there are still many people who do not understand the basic concepts of this technology, so structured and applicable education is needed. To answer this challenge, an activity entitled "Utilization of Machine Learning for Property Price Segmentation and Prediction" was held which aimed to introduce and train participants in the application of machine learning to predict property prices. This activity consists of two main parts, namely webinars and workshops. The webinar focused on introducing the concepts of artificial intelligence, machine learning, and AI Project Cycle as the main method in analyzing house prices. Meanwhile, the workshop provided hands-on training to participants in building prediction models using Google Colab. This activity was carried out through a series of stages, starting from socialization, preparation of materials, pre-test to measure initial understanding, educational and practical sessions, to evaluation through post-test and filling in participant feedback. A total of 39 participants from various backgrounds participated in this activity. The evaluation showed that 38.7% of participants were satisfied, while 51.6% were very satisfied with the program. In addition, the post-test results showed a significant increase in understanding compared to the pre-test results. Based on these results, this activity proved to be successful in providing new insights into the application of machine learning in property price prediction and equipping participants with practical skills that can be applied in the real world. Kemajuan teknologi digital telah mendorong pemanfaatan kecerdasan buatan, khususnya *machine learning,* di berbagai sektor, termasuk analisis harga properti. Namun, masih banyak masyarakat yang belum memahami konsep dasar dari teknologi ini, sehingga diperlukan edukasi yang terstruktur dan aplikatif. Untuk menjawab tantangan tersebut, maka diadakanlah kegiatan bertajuk "Pemanfaatan Machine Learning untuk Segmentasi dan Prediksi Harga Properti" yang bertujuan untuk mengenalkan dan melatih peserta dalam penerapan *machine learning* untuk memprediksi harga properti.

Kegiatan ini terdiri dari dua bagian utama, yaitu *webinar* dan *workshop*. *Webinar* difokuskan pada pengenalan konsep kecerdasan buatan, *machine learning*, dan *AI Project Cycle* sebagai metode utama dalam menganalisa harga rumah. Sementara itu, *workshop* memberikan pelatihan langsung kepada para peserta dalam membangun model prediksi menggunakan Google Colab. Kegiatan ini dilakukan melalui serangkaian tahapan, mulai dari sosialisasi, persiapan materi, *pre-test* untuk mengukur pemahaman awal, sesi edukasi dan praktik, hingga evaluasi melalui *post-test* dan pengisian *feedback* peserta. Sebanyak 39 peserta dari berbagai latar belakang mengikuti kegiatan ini. Hasil evaluasi menunjukkan bahwa 38,7% peserta merasa puas dan 51,6% merasa sangat puas dengan program ini. Selain itu, hasil *post-test* menunjukkan adanya peningkatan pemahaman yang signifikan dibandingkan dengan hasil *pre-test*. Berdasarkan hasil tersebut, kegiatan ini terbukti berhasil memberikan wawasan baru mengenai penerapan *machine learning* dalam prediksi harga properti dan membekali peserta dengan kemampuan praktis yang dapat diaplikasikan di dunia nyata.

A. INTRODUCTION

A house is a basic need that cannot be avoided by people because it serves as a place to shelter and rest after hectic daily activities. Apart from being a place to live, a house can also be a valuable investment asset in the future. The fluctuating movement of house prices, coupled with the high demand for houses, especially those near job centers, offices, shopping centers, and transportation facilities, will certainly affect the price of the house significantly (Saiful, Andryana, & Gunaryati, 2021). As a form of long-term investment, analyzing house prices is necessary to reduce potential losses. This analysis also provides benefits for buyers in predicting house prices according to their preferences. Therefore, the use of machine learning-based systems can be a solution to support this analysis (Putri & Arianto, 2024). Machine learning technology allows computers to learn patterns from historical data and generate predictions without having to be explicitly programmed. One algorithm that is often used in prediction is linear regression, which is known for its simplicity and ability to model the relationship between independent and dependent variables. The results of these algorithms are then evaluated using certain metrics to improve accuracy and optimization (Hallan & Fajri, 2025). In this context, the research falls under the category of supervised learning, where the technique relies on data that is already labeled with a clear target. The goal is to identify causal relationships between input (independent) variables and target (dependent) variables. One of the algorithms used in machine learning is linear regression, which serves to analyze and describe the relationship between a dependent variable, such as house price, and independent variables, such as land area, number of bedrooms, or number of bathrooms. These variables have a significant influence on house prices, and with linear regression, we can explore the patterns and relationships that exist (Nuris, 2024). Machine learning also makes it possible to predict new data by finding patterns in big data and drawing conclusions from those predictions. This technology not only helps in identifying the most influential factors on house

prices, but also provides a stronger basis for making accurate price predictions based on existing house specifications (Hafizh, Subairi, Libriawan, Maulana, & Rizki, 2024). Therefore, the application of this house price prediction model is expected to provide price information in accordance with market conditions and become a reference for property entrepreneurs in offering houses to buyers, as well as helping buyers or investors determine the right price when negotiating (Rahayuningtyas, Rahayu, & Azhar, 2021). The introduction of important concepts in technology and data analysis needs to be socialized more widely to the community. Webinar and Workshop activities with the theme "Utilization of Machine Learning for Property Price Segmentation and Prediction" are part of the Real Work Lecture program and also the result of Internships and Certified Independent Studies. The purpose of this activity is to provide understanding and practical skills to participants, both students and the general public, regarding the basics of the application of machine learning, especially in terms of property price segmentation and prediction. By participating in this activity, participants are expected to understand more about data analysis techniques and the important role of machine learning in predicting property prices, as well as knowing the factors that affect the property market in a more efficient and precise way.

B. METHODS

To carry out the activities described earlier, the author and the team compiled several steps that must be taken. This activity was carried out in two main stages, namely Community Education through Webinars and Training in the form of Workshops. In organizing Webinars and Workshops, the authors and team followed several steps as follows:

1. *Stage 1 (Activity Socialization)*
   At this stage, the author and the team conducted promotions through social media by distributing flyers and registration links at the URL *https://forms.gle/yXjPeqgSb6YAr4Sb9* which contained information about the Webinar and Workshop. Activity posters were published on the Instagram platform, LinkedIn, and distributed through WhatsApp group messages.

2. *Stage 2 (Preparation of Activity Materials)*
   At this stage, the author and team prepare materials for Webinars and Workshops that will be held. The material is compiled in PPT format which will be presented by the presenters during the activity.

3. *Stage 3 (Pre-Test completion by participants)*
   Before the activity begins, participants are asked to fill out a pre-test via the URL *https://bit.ly/TestPengetahuanArtificialnteligence* which contains questions related to the material that will be presented in the Webinar and Workshop. The purpose of this pre-test is to determine the extent of participants' understanding of the material to be discussed, the results of which will later be compared with the post-test filled in after the activity is completed.

4. *Stage 4 (Community Education through Webinar)*
   At this stage, the author and team presented basic material about Machine Learning, especially about modeling for house price prediction. The result of this stage is an introduction to the participants about the basic concepts of Machine Learning, as well as Google Colab tools. The material discussed includes the definition of Artificial Intelligence, Machine Learning, Supervised Learning and Unsupervised Learning, AI Project Cycle, Problem Scoping, Data Acquisition, Data Exploration, Data Pre-processing, Modeling and Evaluation.

5. *Stage 5 (Training through Workshop)*
   At this stage, participants will practice Machine Learning model building, from data preparation, model building, to optimization. Participants will learn to import, clean, and analyze data in Google Colab. Next, they will build and evaluate models using Python and Scikit-Learn. Finally, participants will optimize the model, simulate predictions, and save and

distribute it.At the end of the session, participants are expected to understand the process of building a house price prediction model and be able to apply Machine Learning in their projects.

6. *Stage 6 (Completion of Feedback and Post-Test by Participants)*
   *At the end of the activity, participants were asked to fill out a feedback form to determine their level of satisfaction with the material presented by the speaker. In addition, participants were also asked to fill out a Post-Test which can be accessed viaURL:*https://docs.google.com/forms/d/10nTr0Si4Wb90ucuFMuV7KT3FL-5YmCtwukJSvNW_i9I/edit?ts=678a55ff *The results of the Post-Test will be compared with the results of the Pre-Test to evaluate the extent of participants' understanding of the material taught during the activity.*

C. RESULTS AND DISCUSSION

This Webinar and Workshop activity was carried out by students of the Informatics Engineering study program. Faculty of Engineering, University of Muhammadiyah Jakarta. The result is that this activity is carried out online through the Zoom Meeting Conference with the URL link https://zoom.us/j/98456459230?pwd=Kg9hB9Pk6gPpY02vaaLaFMIqrQoItM.1 *on Thursday, January 16, 2024 at 13.00 - 15.30 WIB. Participants who attended this activity amounted to 39 people from various agencies, the majority of which were attended by students of the Informatics Engineering study program at Muhammadiyah University Jakarta. This activity is carried out online through the Zoom Meeting Conference and there are interactive sessions such as QnA during its implementation. The following is the composition of the Webinar and workshop activities:*

Table 1 Schedule of Activities

Time	Activities	PIC
13.00 – 13.15	Gather the Committee and Spread the Zoom Link and VG Link (pre-test)	Fadel Amili
13.15 – 13.20	Opening by MC	Afni Izzah
13.20 – 13.24	Indonesia Raya	Afni Izzah
13.24 – 13.28	Mars Muhammadiyah	Afni Izzah
13.28 – 13.30	Tilawah	Fadel Amili
	Moderator CV Reading	Afni Izzah
	Presenter CV Reading	Gilang Ramadhika
13.30 – 13.50	Webinar	Mulki Djenfik Dzulkarnain
13.50 – 14.00	Break (post-test)	
	Presenter CV Reading	Gilang Ramadhika
14.00 – 14.40	Workshop	Akbar Andriansyah
14.40 – 14.45	QnA	Gilang Ramadhika
14.45 – 14.55	Quiz	Gilang Ramadhika
14.55 – 15.00	Group photo, attendance link and feedback	Fadel Amili
15.00 - 15.30	Closing	Afni Izzah

Then the dedication stage is carried out as described above:

Stage 1 (Socialization of Activities)

At this stage, the author and team socialize to students and the general public through the distribution of flyers to each WhatsApp group to social media that has been made as shown in Figure 1.

Image 1 Flyer

Stage 2 (Preparation of Activity Materials)

At this stage, the author and team prepare materials for Webinars and Workshops that will be held. The material is compiled in PPT format which will be presented by the presenters during the activity. The prepared materials are Artificial Intelligence, Machine Learning, Supervised Learning and Unsupervised Learning, AI Project Cycle, Problem Scoping, Data Acquisition, Data Exploration, Data Pre-processing, Modeling and Evaluation. The activity materials can be seen in the following figures 2 and 3.
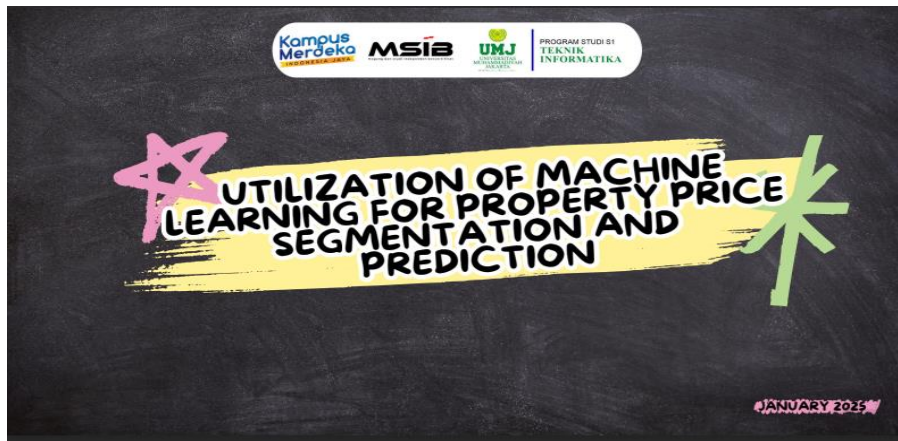
Image 2 Webinar title

Image 3 Workshop title

Stage 3 (Pre-Test completion by participants)

At this stage, participants are asked to take a Pre-Test that has been prepared by the author and the team. This Pre-Test contains a series of questions related to machine learning. The purpose of filling out this Pre-Test is to measure the extent of the participants' understanding before the author and team carry out this activity. As a result, 45 participants have taken the Pre-Test with a fairly good level of understanding.
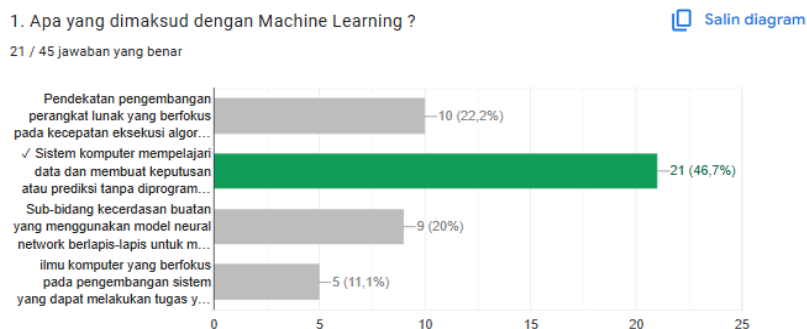
Image 4 Participant Pre-test Questions on Machine Learning

In the picture above you can see that the Pre-test questions were filled in by 45 participants with a percentage of correct answers of 46.7%.
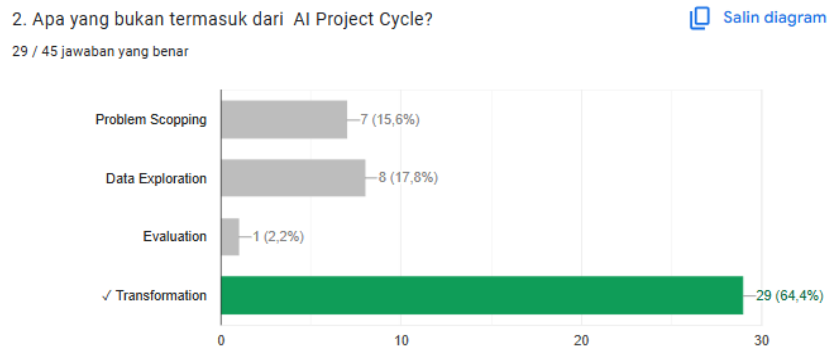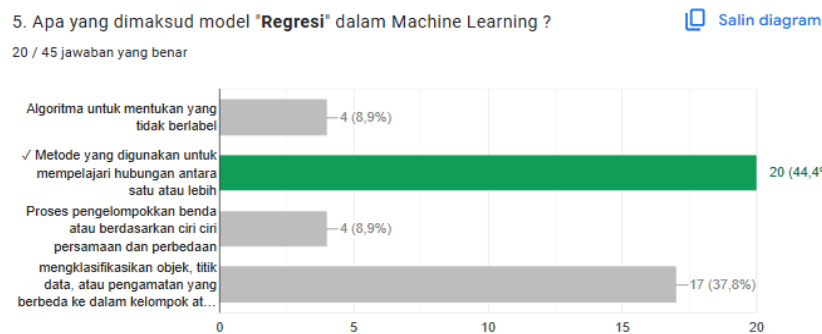

Image 5 Pre-test question on AI Project Cycle

In the picture above you can see that the Pre-test questions were filled in by 45 participants with a percentage of correct answers of 64.4%.
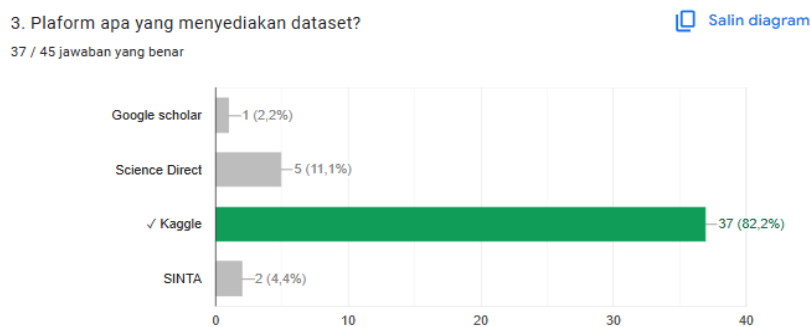
Image 6 Pre-test questions about providing datasets

In the picture above you can see that the Pre-test questions were filled in by 45 participants with a percentage of correct answers of 62.2%.
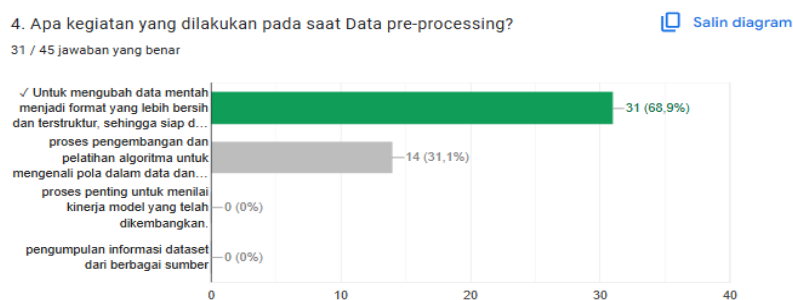
Image 7 Pre-test question about Data pre-processing

In the picture above you can see that the Pre-test questions were filled in by 45 participants with a percentage of correct answers of 68.9%.

Image 8 Pre-test questions about Regression

In the picture above you can see that the Pre-test questions were filled in by 45 participants with a percentage of correct answers of 44.4%.

The pre-test results show that the majority of participants already have a basic understanding of the material that will be discussed in this webinar and workshop.

Stage 4 (Community Education through Webinar)

At this stage, the webinar presenter, Mulki Djenfik Dzulkarnain, delivered pre-prepared material to participants who came from various backgrounds. The material presented included topics on Artificial Intelligence, Machine Learning, and approaches used in the machine learning process, namely Supervised Learning and Unsupervised Learning. The presentation began with an explanation of the basic concepts of Artificial Intelligence, followed by an introduction to Machine Learning as a branch of AI that allows systems to learn from data without being explicitly programmed. The speaker also discussed the two main methods in Machine Learning, namely Supervised Learning, which involves using labeled data to train a model, and Unsupervised Learning, which is used to identify patterns in unlabeled data.

In addition, participants were introduced to the AI Project Cycle, which is the steps that need to be taken in building an artificial intelligence-based project. This process includes several important steps, starting from Problem Scoping which aims to understand the problem to be solved, Data Acquisition to collect relevant information, Data Exploration to understand the characteristics of the data, Data Pre-processing to clean and prepare the data before being used in modeling, to the Modeling and Evaluation stages to assess the performance of the system that has been created.
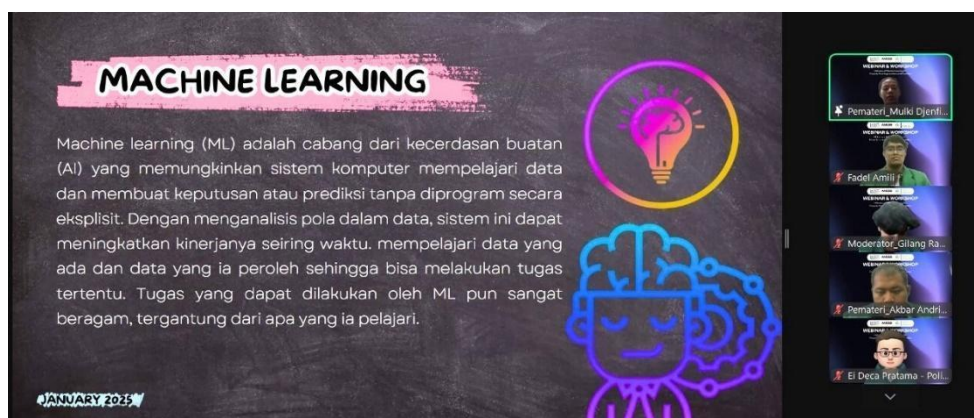

Image 9 Webinar Material Presentation

During the session, participants were encouraged to interact actively through discussions and question and answer sessions, both directly and through the chat column on the webinar platform used. Through this activity, it is hoped that participants will gain a more comprehensive insight into the application of artificial intelligence and the stages required in building an AI-based model.

Stage 5 (Training through Workshop)

At this stage, the Workshop presenter, Akbar Andriansyah, provides direct application of the Webinar material that has been delivered previously. In this workshop session, the speaker utilized Google Colab as a tool. Participants have been invited to access and understand the use of the tool before the activity begins.

The speaker explained the basics of applying Machine Learning in Property Price Segmentation and Prediction, which includes fundamental concepts, data processing, and methods used. In this implementation, Linear Regression is chosen as the main model to build a system that can predict property prices.

Participants were guided to understand the code structure, data pre-processing, and machine learning model building stages. In a systematic way, the speaker explained step by step to produce interactive output, so that participants can interact directly with the prediction model that has been created. The final result of this workshop is a property price prediction model that can be used to estimate prices based on parameters entered by users.
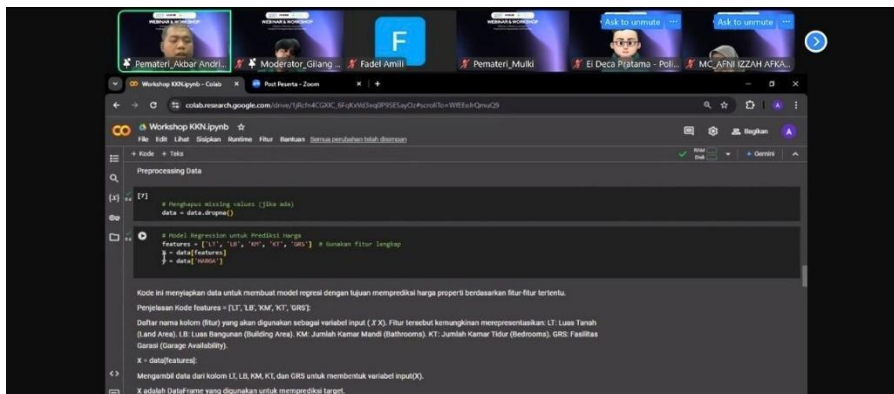

Image 10 Workshop Material Presentation

In the picture above, the speaker explains in detail the steps in the process of making property prediction analysis. Which finally produces output to interact with the predictions that the author describes.

Stage 6 (Completion of Feedback and Post-Test by Participants)

At this stage, participants were asked to fill in the attendance, feedback, and Post-Test forms that had been provided in one form shared via Google Form. To measure participants' answers, a Likert scale was used. Likert scale is a scale consisting of several answer options that describe the level of agreement of participants with existing statements or statements, with customized answer options. In this questionnaire, 4 levels of score assessment were applied with the following conditions: (5) Strongly Agree (4) Agree, (3) Neutral (2) Disagree, and (1) Strongly Disagree. The questions in the questionnaire are as follows:

1. *Does the speaker deliver the material in a structured manner?*

2. *Are Webinar resource persons who provide Webinar material in accordance with their scientific fields?*
3. *Are the Workshop resource persons who provide Workshop material in accordance with their scientific fields?*
4. *Is the Webinar resource person able to explain the material well?*
5. *Is the Workshop resource person able to explain the material well?*
6. *Are audio and visual running well during the session?*
7. *Is the online administration service provided easy to use?*
8. *Overall, how satisfied are you with this activity?*

The post-test consists of questions similar to the pre-test, the results of which can be used to compare participants' understanding before and after participating in this Webinar and Workshop. Below are the results of the feedback and post-test filled out by the participants.

In each Webinar and Workshop session, participants were very enthusiastic about the presentation delivered by the presenters. This can be clearly seen from the feedback questionnaire form given at the end of the session to participants. The average answer in the questionnaire form shows the level of satisfaction of the participants with the material presented.
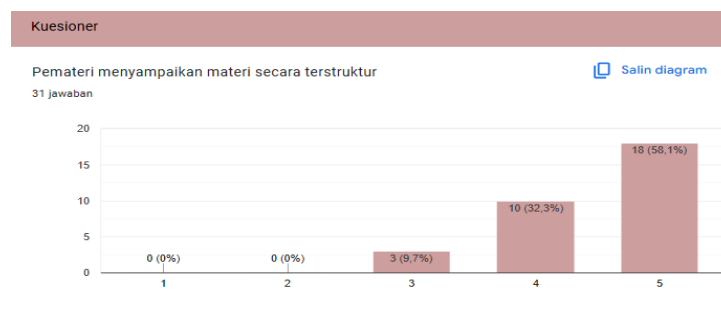
Image 11 Level of Assessment about Presenters delivering material to participants

In Figure 10 above, it can be seen that the feedback given by participants agreed with a percentage of 32.3% and 58.1% strongly agreed. In the next question, the feedback results also show positive numbers as shown in the figure below:
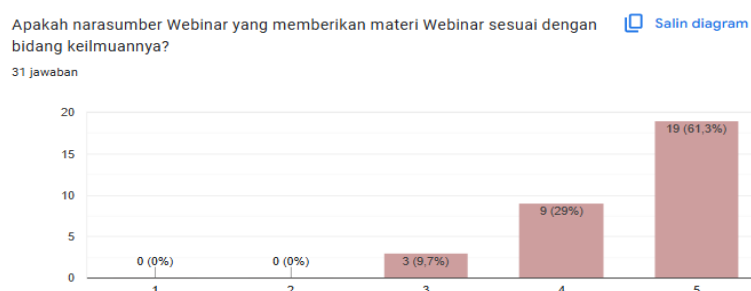

Image 12 Level of Assessment of webinar speakers according to the material in their scientific field

In Figure 12 above, it can be seen that the feedback given by participants agreed with a percentage of 29% and 54.8% strongly agreed. In the next question, the feedback results also show positive numbers as shown in the figure below:
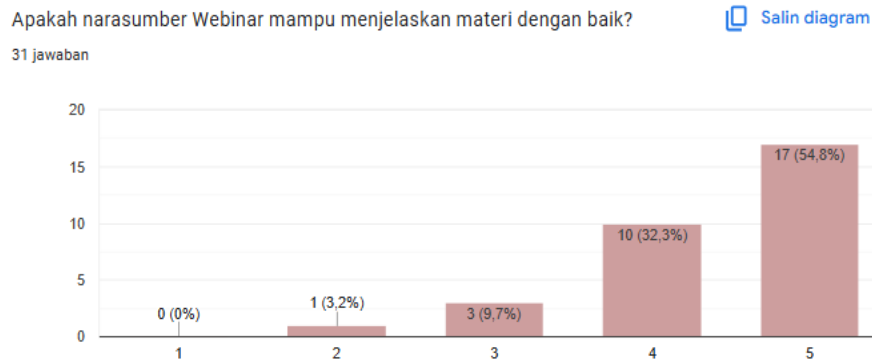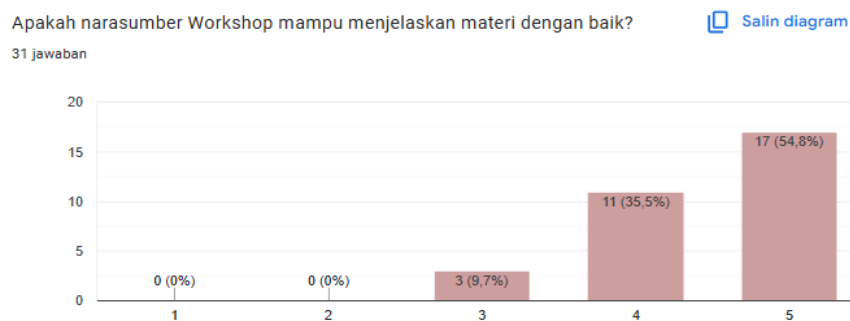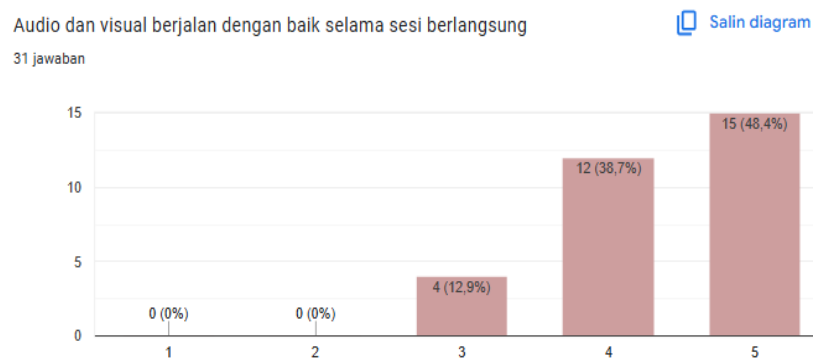
Image 13 Level of assessment of webinar speakers about explaining the material

In Figure 13 above, it can be seen that the feedback given by participants agreed with a percentage of 32.3% and 54.8% strongly agreed. In the next question, the feedback results also show positive numbers as shown in the figure below:

Image 14 Level of assessment of workshop resource persons on explaining the material

In Figure 14 above, it can be seen that the feedback given by participants agreed with a percentage of 35.5% and 54.8% strongly agreed. In the next question, the feedback results also show positive numbers as shown in the figure below:

Image 15 Assessment Level for audio and visual

In Figure 15 above, it can be seen that the feedback given by the participants agreed with a percentage of 38.7% and 48.4% strongly agreed. In the next question, the feedback results also show positive numbers as shown in the figure below:
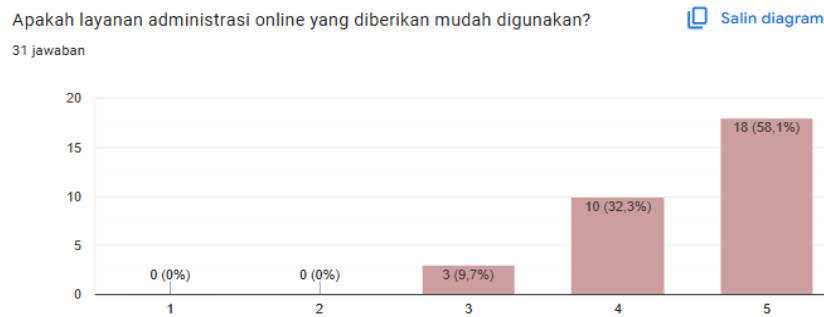
Image 16 Level of administration service assessment

In Figure 16 above, it can be seen that the feedback given by participants agreed with a percentage of 32.3% and 58.1% strongly agreed. In the next question, the feedback results also show positive numbers as shown in the figure below:
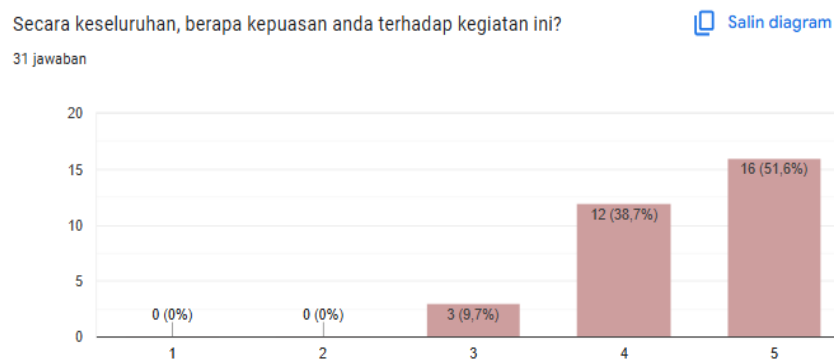

Image 17 Satisfaction rating of webinars and workshops

In Figure 17 above, it can be seen that the feedback given by participants agreed with a percentage of 38.7% and 51.6% strongly agreed. Participants expressed their satisfaction with this activity. Through the questions asked to them and the results of filling out the questionnaire, it can be seen that participants managed to understand the new material according to the theme of the activity. The positive responses from the participants regarding the delivery of the material showed that the presenters had delivered the information well and clearly, so that the participants could learn and understand the new knowledge easily. Thus, this activity was successful and can be understood by the general public.

In addition to the feedback questionnaire, participants were also asked to fill out a post-test after the activity was completed, with the aim of measuring the improvement of participants' understanding. Here are some of the results of the post-test filling by the participants:
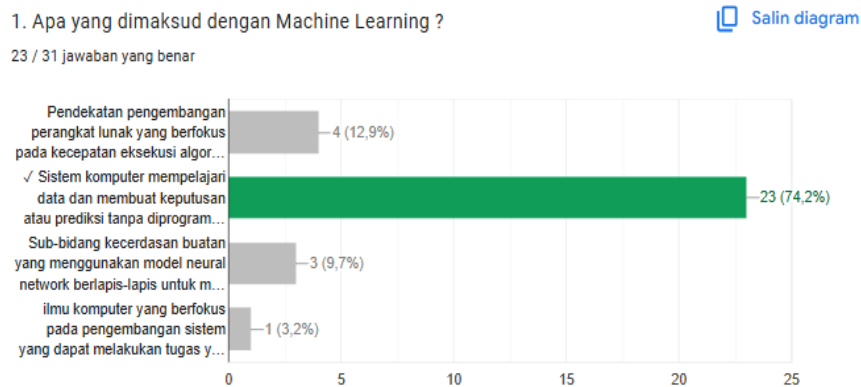
Image 18 Post-test questions about Machine Learning

In Figure 18 above, it can be seen that the post-test results show that 74.2% of participants answered correctly on question number 1, which was completed by 31 participants.
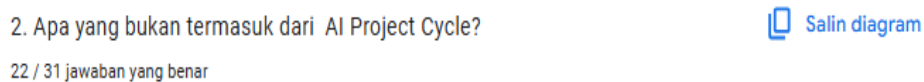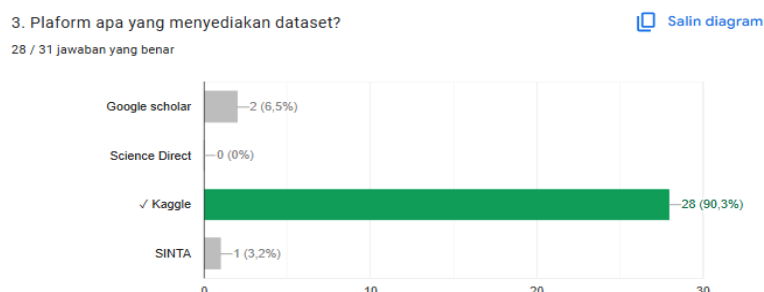

Image 19 Post-test questions about AI Project Cycle

In Figure 19 above, it can be seen that the post-test results show that 71% of participants answered correctly on question number 2, which was completed by 31 participants.

Image 20 Post-test questions about providing datasets

In Figure 20 above, it can be seen that the post-test results show that 90.3% of participants answered correctly on question number 3, which was completed by 31 participants.
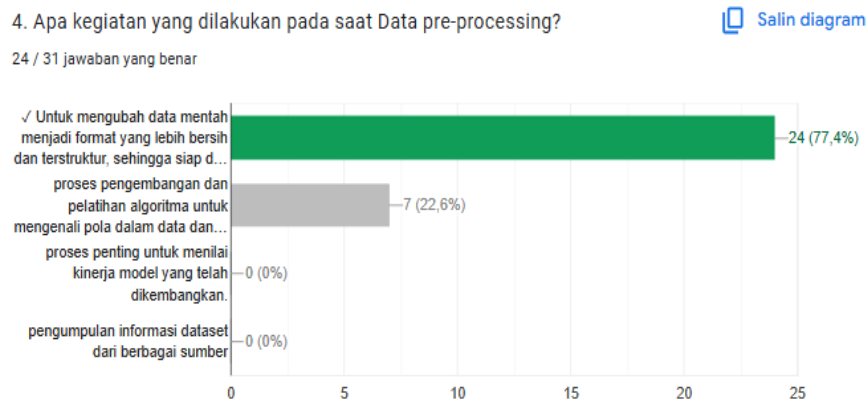
Image 21 Post-test questions about Data Pre-Processing

In Figure 21 above, it can be seen that the post-test results show that 77.4% of participants answered correctly on question number 4, which was completed by 31 participants.
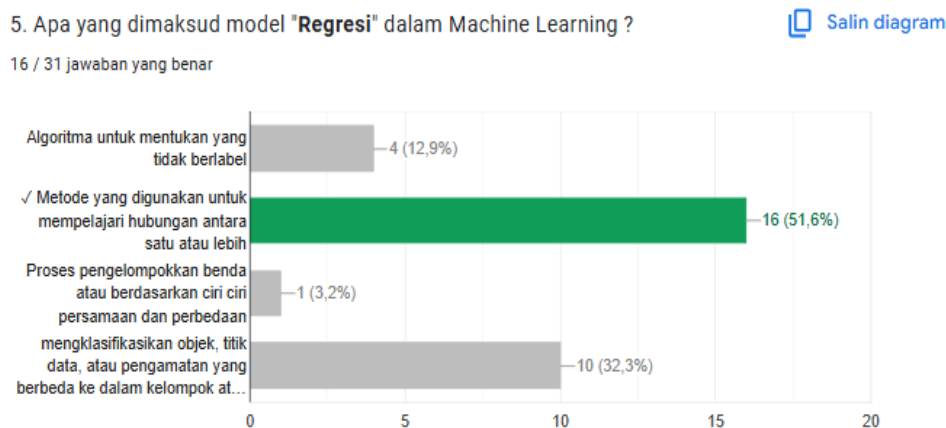

Image 22 Post-test questions about regression models

In Figure 22 above, it can be seen that the post-test results show that 51.6% of participants answered correctly on question number 5, which was completed by 31 participants.

After attending the webinar and workshop, the participants showed a significant improvement in their understanding of machine learning, as seen in the figure above. In addition, they also received certificates that have been endorsed by the Head of the UMJ Informatics Engineering Study Program.

D. CONCLUSION

Based on the results of the webinar and workshop "Utilization of Machine Learning for Property Price Segmentation and Prediction" the author can conclude that this activity provides significant benefits for participants. The pre-test and post-test results show a considerable increase in understanding of machine learning concepts, especially in the application of linear regression to predict property prices.

In addition, in the practical session during the workshop, participants gained hands-on experience in building prediction models using Google Colab and machine learning libraries such as Scikit-Learn. The participants' response to this activity was also very positive, as reflected in the

feedback that showed a high level of satisfaction with the material, the delivery of the presenters, and the technical implementation of the activity.

Therefore, it can be said that this activity succeeded in achieving its objectives, namely providing in-depth understanding and practical skills in data analysis and machine learning. It is hoped that similar activities can continue to be developed in the future to expand the reach of education about machine learning technology and encourage its application in various fields, especially in analyzing and predicting property prices.

E. ACKNOWLEDGEMENTS

F. AUTHOR CONTRIBUTIONS

In the webinar and workshop entitled "Unitilization of Machine Learning for Property Price Segmentation and Prediction," each team member has a clear contribution and responsibility to ensure the smooth implementation until the preparation of scientific articles. Akbar Andriansyah as the chief executive is responsible for the overall course of the activity and oversees the duties of team members, as well as acting as a workshop presenter which is part of the implementation of the webinar. Fadel Amili was in charge of creating promotional content, including flayers, pre-test and post-test forms, and feedback to measure participants' learning outcomes. Afni Izzah Afkarinah acted as Master of Ceremony during the activity, Gilang Ramadhika acted as moderator who conducted the activities during the webinar and workshop sessions. Mulki Dzenfik Dzulkarnain who was in charge of being a webinar presenter by presenting the material discussed including the understanding of Artificial Intelligence, Machine Learning, Supervised Learning and Unsupervised Learning, AI Project Cycle, Problem Scoping, Data Acquisition, Data Exploration, Data Pre-processing, Modeling and Evaluation. Dr. Nurvelly Rosanti acts as a supervisor who provides direction, supervision, and support so that the implementation of webinars and workshops runs smoothly until the checking of scientific articles. All authors work together to ensure the success of this activity and the quality of the resulting article publications.

G. REFERENCES

Hafizh, M. A., Subairi, Libriawan, R. D., Maulana, N. D., & Rizki, A. M. (2024). Prediksi Harga Rumah Di Jabodetabek Menggunakan Metode Artificial Neural Network. *Jurnal Riset Inovasi Bidang Informatika dan Pendidikan Info*, 49.

Hallan, R. R., & Fajri, I. N. (2025). Prediksi Harga Rumah menggunakan Machine Learning Algoritma Regresi Linier. *Jurnal Teknolgi Dan Sitem Informasi Bisnis*, 57.

Nuris, N. (2024). Analisis Prediksi Harga Rumah Pada Machine Learning Menggunakan Metode Regresi Linear. *Jurnal Informatika dan Komputer*, 109.

Putri, N. A., & Arianto, D. B. (2024). Komparasi Penggunaan Information GainPada Machine Learninguntuk Memprediksi Harga Rumah di Jabodetabe. *Jurnal Sains dan Teknologi, 756.*

Rahayuningtyas, E. F., Rahayu, F. N., & Azhar, Y. (2021). Prediksi Harga Rumah Menggunakan General. *Jurnal Informatika,* 60.

Saiful, A., Andryana, S., & Gunaryati, A. (2021). Prediksi Harga Rumah Menggunakan Web ScrappingDan Machine Learning Dengan Algoritma Linear Regression. *Jurnal Teknik Informatika dan Sistem Informasi*, 42.