1. INTRODUCTION

The industrialization of phase 4.0 has transformed modern life, with digital technologies becoming integral to various fields, including education [1]. The use of this technology not only supports the career field, but also serves as a means of learning. Overall, technology as a learning medium can be utilized through various applications that can be downloaded on mobile phones [2]. In the world of education, many utilize technology as a learning tool through applications that can help in learning and solving problems that are not easy to solve. One of them is math [3]. Among the various applications available, photomath is one of the most popular ones used to solve math problems. MicroBlink created Photomath in 2011. It uses advanced text reader technology to scan math problems that users upload [4]. The answer and solution steps will be displayed by this application. This application is very suitable for use as a reference for solving difficult problems [5]. However, in the field of education, this innovation has sparked both support and criticism. A number of user experiences can be seen and written in the photomath application review column [3]. User reviews can include advantages, disadvantages, or problems with the photomath application. This makes important information for user needs increasingly difficult to find because of the many opinions that arise. To address this issue, natural language processing (NLP) techniques enable computers to understand and process human language for tasks such as sentiment analysis. In this research, sentiment analysis is used as a computational study to find opinions, attitudes, and emotions towards an entity in a review column [6].

Sentiment analysis is a technique used to identify how a sentiment is expressed through text and how it can be categorized as positive or negative [7]. Sentiment analysis is usually used to evaluate opinions about an application with the aim of improving its quality in the future. In this context, sentiment analysis can be implemented on photomath application reviews [8]. In this research, there are several references that are relevant to previous research that focuses on sentiment analysis of application reviews. The methods used are SVM, Random Forest and Word2Vec. Some of these studies involve research [9] by M. J. Aufa and A. Qoiriah (2022) focused on sentiment analysis of users of the online learning platform Coursera using the Random Forest algorithm combined with Word2Vec feature extraction. After the filtering process, a total of 706,828 data entries consisting of 5 columns were obtained. The study achieved the highest accuracy with an 80:20 data proportion, a Word2Vec window size of 5, and a feature size of 100, resulting in a train score of 92.7% and a test score of 91.4%. Research [10] by F. A. Larasati et al. (2022) analyzed sentiment in application reviews using the Random Forest method. The study utilized data retrieval through web scraping techniques and Jupyter Notebook tools with the Google-Play-Scrapper API. The data was split into 80% training and 20% testing datasets, achieving an accuracy of 84%, recall of 84%, F1-Score of 84%, and precision of 84% with a tree depth of 65 and 400 trees. This demonstrates the effectiveness of Random Forest in sentiment analysis tasks for application reviews. Research [11] by S. N. Adhan et al. (2024) analyzed the sentiment of Wattpad application reviews on the Google Play Store using the Random Forest method. The study was further optimized using the SMOTE (Synthetic Minority Oversampling Technique) to address class imbalance issues. The findings revealed that unbalanced sentiment classification tends to achieve better performance using the Random Forest method without SMOTE optimization. The reported metrics include an accuracy of 84.05%, precision of 84.71%, recall of 91.60%, F1-Score of 88.02%, FPR of 8.40%, and an AUC of 0.9166, indicating excellent classification performance. Random forest was chosen because it is one of the effective classification algorithms for analyzing sentiment and there are advantages, namely the ability to improve accuracy results if there is missing data, be resistant to outliers, and be effective in data storage [12].

Research [13] by M. A. A. Jihad et al. (2021) focused on sentiment analysis of movie reviews using Word2Vec and Support Vector Machine (SVM). The study applied lemmatization and utilized Word2Vec with 300 dimensions, combined with a linear SVM classification method, achieving the best accuracy of 78.75% and the highest F1-score of 78.74%. This research demonstrated that this approach provided superior performance compared to other tested methods. Research [14] by D. I. Rifai (2024) focused on sentiment analysis of TikTok user reviews using Word2Vec models combined with the Support Vector Machine (SVM) method. The study employed Word2Vec in the forms of CBOW and skip-gram to compare accuracy in sentiment classification. The CBOW model achieved an accuracy of 66%, while the skip-gram model performed slightly better with an accuracy of 68%. The data for this research was sourced from application user reviews on the Google Play Store. Research [15] by P. E. Shopee and S. Watmah (2021) compared the performance of K-Nearest Neighbor (K-NN), Support Vector Machine (SVM), and Random Forest methods in classification tasks. The study focused on evaluating the accuracy, precision, and recall of these classification techniques, finding that the SVM method outperformed the others with an accuracy of 89.4%, precision of 89.5%, and recall of 89.7%. Enquiry [16] by Engineering (2023) analyzed sentiment towards the usage of the Shopee application using the Support Vector Machine (SVM) algorithm. Collected 3000 review data points through a scraping procedure and achieved outstanding performance,

with an accuracy rate of 98% and an F1-score of 98%. This demonstrates the potential of SVM for high-accuracy sentiment classification tasks. The ability of the SVM approach to reliably categorize evaluations into good, negative, and neutral groups led to its selection. Word2Vec was used because it could turn words into vectors. Word2Vec can assist in overcoming difficult sentiment analysis problems by comprehending the meaning of words in user reviews [17]. Although sentiment analysis has been extensively used in earlier research utilizing techniques like Word2Vec, Random Forest, and Support Vector Machine (SVM), the majority of these studies have not examined the effects of various Word2Vec dimensions on performance.

The results of earlier research indicate that the Random Forest and SVM approaches, when paired with Word2Vec feature extraction, exhibit highly effective sentiment analysis performance. Though this parameter can be changed to increase accuracy, prior research has tended to ignore the variance in Word2Vec dimension size. This study aims to analyze the sentiment of user reviews for the Photomath app using Random Forest and SVM methods, optimized with Word2Vec feature extraction. Different dimensions are tested to identify the most effective configuration. This research is expected to identify the best method in classifying review sentiment as positive or negative, as well as provide deeper insights into user experience to support the quality improvement of the Photomath app in the future.