

INTISARI

Angka kematian ibu merupakan salah satu indikator yang dapat menentukan derajat kesehatan masyarakat di suatu negara. Salah satu faktor yang menyebabkan tingginya angka kematian ibu hamil di Indonesia adalah kurangnya pengetahuan ibu tentang kehamilannya. Aplikasi kehamilan berbasis *mobile* merupakan salah satu wujud transformasi digital yang membantu ibu hamil untuk mengakses informasi yang dapat menambah pengetahuan kehamilan secara mudah dan cepat. Meskipun demikian, tingkat adopsi aplikasi kehamilan masih rendah. Analisis sentimen pengguna merupakan cara *stakeholder* dan pengembang untuk mengetahui penerimaan pengguna terhadap aplikasi kehamilan. Permasalahan pada analisis sentimen adalah tidak dapat menangkap aspek dari suatu teks, padahal ada berbagai aspek yang memengaruhi penerimaan pengguna terhadap aplikasi kehamilan, seperti aspek teknis, sosial, pembayaran, dan individu. *Aspect-based sentiment analysis* (ABSA) dapat mengatasi permasalahan tersebut untuk mengklasifikasikan sentimen teks ulasan aplikasi kehamilan berdasarkan aspeknya. Implementasi ABSA memiliki berbagai pilihan metode *feature extraction*, *imbalanced data handling*, dan berbagai model *machine learning* yang dapat digunakan. Namun, belum diketahui performa setiap metode untuk implementasi ABSA pada ulasan aplikasi kehamilan. Oleh karena itu, penelitian ini menguji dan menganalisis metode *feature extraction* TF-IDF dan *bag of words* (BoW), pengaruh *imbalanced data handling* berupa SMOTE, dan beberapa model *machine learning*, yaitu *logistic regression*, *support vector machine* (SVM), *decision tree* (DT), dan *random forest* (RF) untuk implementasi ABSA pada ulasan aplikasi kehamilan. Hasil penelitian menunjukkan kombinasi metode terbaik untuk setiap aspek. Kombinasi metode terbaik untuk aspek teknis adalah model RF dengan *feature extraction* BoW pada data dengan distribusi aslinya yang menghasilkan skor F1 sebesar 88,5%. Kombinasi metode terbaik untuk aspek sosial adalah model SVM dengan *feature extraction* BoW pada data dengan distribusi aslinya yang menghasilkan skor F1 sebesar 87,5%. Kombinasi metode terbaik untuk aspek pembayaran adalah model RF dengan *feature extraction* BoW pada data dengan distribusi aslinya yang menghasilkan skor F1 sebesar 95,3%. Kombinasi metode terbaik untuk aspek individu adalah model RF dengan *feature extraction* BoW pada data dengan distribusi aslinya yang menghasilkan skor F1 sebesar 89,6%. Model yang menggunakan metode BoW memiliki skor *recall* yang lebih tinggi dibandingkan TF-IDF, sedangkan TF-IDF memiliki skor *precision* yang lebih tinggi dibandingkan BoW. Metode *feature extraction* BoW yang dikombinasikan dengan SMOTE menyebabkan performa semua model menjadi buruk. Penerapan SMOTE dapat meningkatkan skor *recall* pada semua model yang menggunakan metode TF-IDF. Di sisi lain, penerapan SMOTE menyebabkan penurunan skor *precision* pada semua model yang menggunakan TF-IDF maupun BoW. Model RF dan SVM menghasilkan performa yang lebih tinggi pada semua aspek. *Logistic regression* mengalami peningkatan performa saat diimplementasikan dengan metode SMOTE. Model DT rawan mengalami *overfitting* dan memiliki performa yang lebih rendah dibandingkan ketiga model lainnya.

Kata kunci: ABSA (*aspect-based sentiment analysis*), aplikasi kehamilan, *feature extraction*, *imbalanced data handling*, *machine learning*.

ABSTRACT

The maternal mortality rate is one of the key indicators used to assess the health status of a population in a country. One of the factors contributing to the high maternal mortality rate in Indonesia is the lack of knowledge among pregnant women about their pregnancy. Mobile-based pregnancy applications represent a form of digital transformation that helps pregnant women access information easily and quickly to enhance their knowledge about pregnancy. However, the adoption rate of pregnancy applications remains low. User sentiment analysis serves as a way for stakeholders and developers to understand user acceptance of pregnancy applications. A major limitation of sentiment analysis is its inability to capture aspects within a text. However, various aspects influence user acceptance of pregnancy applications, such as technical, social, payment, and individual aspects. Aspect-based sentiment analysis (ABSA) addresses this issue by classifying the sentiment of pregnancy application reviews based on their aspects. The implementation of ABSA involves multiple choices for feature extraction methods, imbalanced data handling techniques, and machine learning models. However, the performance of each method for ABSA implementation in pregnancy application reviews remains unknown. Therefore, this study evaluates and analyzes TF-IDF and bag of words (BoW) feature extraction methods, the impact of imbalanced data handling using SMOTE, and multiple machine learning models, including logistic regression, support vector machine (SVM), decision tree (DT), and random forest (RF), for ABSA implementation on pregnancy application reviews. This study reveals the best method combinations for each aspect. The best method combination for the technical aspect is the RF model with BoW feature extraction applied to data with its original distribution, achieving an F1-score of 88.5%. The best method combination for the social aspect is the SVM model with BoW feature extraction applied to data with its original distribution, yielding an F1-score of 87.5%. The best method combination for the payment aspect is the RF model with BoW feature extraction applied to data with its original distribution, achieving an F1-score of 95.3%. The best method combination for the individual aspect is the RF model with BoW feature extraction applied to data with its original distribution, resulting in an F1-score of 89.6%. Models using BoW achieved higher recall scores compared to TF-IDF, while TF-IDF showed slightly higher precision scores than BoW. The BoW feature extraction method combined with SMOTE resulted in poor performance across all models. The implementation of SMOTE improved recall scores for all models using TF-IDF. However, SMOTE led to a decline in precision scores for all models using both TF-IDF and BoW. Among the machine learning models, RF and SVM demonstrated higher performance across all aspects. Logistic regression showed improved performance when implemented with the SMOTE method. Meanwhile, the DT model is prone to overfitting and performs worse than the other three models.

Keywords: ABSA (aspect-based sentiment analysis), pregnancy app, feature extraction, imbalanced data handling, machine learning.