

Al - Zaytoonah University  
of Jordan  
Deanship of Graduate  
Studies



جامعة الزيتونة الاردنية  
عمادة الدراسات العليا

**Sentiment Analysis of Arabic texts on Twitter using Deep and Machine  
Learning Techniques to Extract the impact of the vaccine covid-19 on citizens  
in Dependence on Arabert model**

By

**Omar Albahri**

Supervisor:

**Dr. Mohammed Elbes**

**Al-Zaytoonah University of Jordan,  
2023**

**Abstract**

In the recent period, the use of the Internet on social media has increased, and the world has become an open environment that allows all users of social networking sites to express their opinions, emotions, and stories. One of the most important of these issues is the Coronavirus (COVID-19), Coronaviruses are a family of viruses that can cause diseases such as colds and respiratory infections that began to appear in China in 2019, and in 2020, the World Health Organization declared that the COVID-19 virus had become a global pandemic. Thereafter, companies began trying to produce a vaccine to stop this pandemic. More than one vaccine was produced to stop the spread of this virus, and confusion and questions increased among people on social networking sites, which vaccine is better? Is

there a negative effect of the vaccine? and so on. This research aims to conduct a sentiment analysis on Arabic tweets to extract the impact of the vaccine for COVID-19 on citizens (Positive, Negative, Neutral) using two methods: Machine Learning (ML) and deep learning (DL) algorithms. Many studies revealed the effect of the vaccine using the English language. However, the studies are scarce for the Arabic tweets. In this research, we used arabert model to deal with the Arabic language. We applied four arabert methods (Remove Diacritics, Remove Elongation, Leave the English letters, and Tokenization). The investigated dataset were collected from Twitter with approximately 7186 tweet. We implemented various ML classifiers, namely; Support Vector Machine (SVM), Naive Bayes (NB), Decision Tree (J48), Random Forest (RF), and Decision Tree(REPTree). Furthermore, we used deep learning methods such as Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). The results showed that the RF classifier obtained the highest value on the F1 scale with (91.9%). Using deep learning techniques we obtained better results than ML methods with the RNN/LSTM F1 value of (93.2%).

**Keywords:** Machine Learning, Deep Learning, NLP, Arabic Language, COVID-19.