

Al - Zaytoonah University  
of Jordan  
Deanship of Graduate  
Studies



جامعة الزيتونة الاردنية  
عمادة الدراسات العليا

تحليل المشاعر للنصوص العربية على تويتر باستخدام تقنيات التعلم العميق والآلة

لاستخراج تأثير لقاح كوفيد -19 على المواطنين اعتمادًا على نموذج أرابرت

إعداد

عمر امين البحري

المشرف

الدكتور محمد البس

جامعة الزيتونة الأردنية، 2023

تاريخ المناقشة 2023/01/08م

الملخص

في الفترة الأخيرة ، ازداد استخدام الإنترنت على وسائل التواصل الاجتماعي ،

وأصبح العالم بيئة مفتوحة تتيح لجميع مستخدمي مواقع التواصل الاجتماعي

التعبير عن آرائهم وانفعالاتهم وقضاياهم. ومن أهم هذه القضايا فيروس كورونا ،  
وفيروسات كورونا عائلة من الفيروسات التي يمكن أن تسبب أمراضًا مثل نزلات  
البرد والتهابات الجهاز التنفسي بدأت بالظهور في الصين في عام 2019 ، وفي  
عام 2020 أعلنت منظمة الصحة العالمية أن كوفيد- 19 أصبح فيروس جائحة  
عالمية. وبدأت الشركات تحاول إنتاج لقاح لوقف هذا الفيروس. تم إنتاج أكثر  
من لقاح لمحاولة إيقاف هذا الفيروس ، وتزايد الارتباك والتساؤلات بين  
الأشخاص على مواقع التواصل الاجتماعي ، أي لقاح أفضل؟ هل هناك أثر  
سلبي للقاح؟

تهدف دراستنا إلى تحليل المشاعر للتغريدات العربية لاستخراج تأثير لقاح  
**COVID-19** على المواطنين (إيجابي ، سلبي ، محايد) باستخدام طريقتين:  
التعلم الآلي (ML) وخوارزميات التعلم العميق (DL). تكشف العديد من  
الدراسات عن تأثير اللقاح باستخدام اللغة الإنجليزية. أما هذه الدراسة على اللغة  
العربية. استخدمنا نموذج أرابرت للتعامل مع اللغة العربية. قمنا بتطبيق أربع  
طرق من **arabart** (إزالة علامات التشكيل ، إزالة الاستطالة ، ترك الأحرف

الإنجليزية ، والترميز). كما تم إنشاء مجموعات بيانات وسائل التواصل الاجتماعي العربية من موقع تويتر مع ما يقرب من 7186 تغريدة. استخدمنا مصنفات ML المشهورة . أشهر هذه المصنفات هي (SVM) و (NB) و (J48) و (RF) و (REPTree). قمنا بتقييم دقة المصنف من خلال حساب مقاييس **F1-Measure** والاستدعاء والدقة. من ناحية أخرى ، استخدمنا طرق التعلم العميق ، والتي تحتوي على الشبكات العصبية التلافيفية (CNN) ، والشبكات العصبية المتكررة (RNN). تتضمن طريقة استخراج الميزة المستخدمة تضمين الكلمات (Word2Vector). أظهرت النتائج أن المصنف RF حصل على أعلى قيمة على مقياس **F1** بنسبة (91.9%). باستخدام تقنيات التعلم العميق نحصل على نتائج أفضل من مصنفات ML لأن RNN/LSTM تعطي قيمة أعلى من الشبكات العصبية التلافيفية وتحصل على قيمة **F1** بنسبة (93.2%).

الكلمات المفتاحية: التعلم الآلي ، التعلم العميق ، NLP ، اللغة العربية ،

