نهج ذكي للكشف عن اساءة استخدام الادوية في منشورات وسائل التواصل الاجتماعي العربية

إعداد

أماني خميس محمد عبد الغافر

المشرف

الدكتور بلال الحواشين

المشرف المشارك

الدكتور طارق كنعان

الملخص

تاريخ المناقشة (2022/08/30)

في الآونة الأخيرة ، أدى استخدام الإنترنت بشكل أساسي إلى زيادة مواقع شبكات التواصل الاجتماعية. فأصبح العالم بيئة مفتوحة ، مما أدى إلى اعتماد مستخدمي وسائل التواصل الاجتماعي على هذه المواقع لتبادل الخبرات المختلفة و منها الخبرات الطبية واعتمادها في كثير من الأحيان كمراجع أساسية في الحصول على المشورة الطبية ، مما يؤدي إلى استخدام الأدوية المختلفة بشكل عشوائي و غير آمن وبالتالي إساءة استخدام الأدوية مثل مسكنات الالم والمضادات الحيوية و ادوية الأمراض الجلدية. تؤثر هذه المشكلة المتنامية على جميع الفئات العمرية ؛ و تجعلها عرضة للعديد من المشاكل الصحية ، تهدف دراستنا إلى تصنيف استخدام الادوية في المحتوى العربي في وسائل التواصل الاجتماعي (استخدام صحيح ، استخدام غير صحيح) باستخدام طرق التعلم الآلي و نموذج(AraBERT) مع طرق التعلم العميق, حيث تكشف العديد من الدراسات هذه الظاهرة في اللغة الإنجليزية بينما لا توجد دراسات مماثلة

في اللغة العربية. قمنا بتطبيق المعالجة المسبقة للبيانات التي توفرها البرمجة اللغوية العصبية لتسهيل التعامل مع اللغة العربية. نفذنا ثلاثة طرق في البرمجة اللغوية العصبية (إيقاف إزالة الكلمات ، و التطبيع ، والجذع). وقمنا بجمع مجموعات من البيانات من وسائل التواصل الاجتماعي العربية من الفيس بوك بما يقارب من (7000) منشورًا.

استخدمنا مصنفات التعلم الآلي ؛ الة ناقلات الدعم (SVM) و Naive Bayes (NB) و شجرة القرار (J48) و الغابة العشوائية (RF) و K-Nearest Neighbor (KNN) , طبقنا أيضًا نموذجًا قائمًا على AraBERT ؛ CNN–AraBERT و RNN–AraBERT و LSTM–AraBERT حيث قمنا بتقييم دقة المصنف من خلال حساب قياسات F1–Measure و الدقة والاستدعاء .

وأظهرت النتائج أن مصنف (CNN–AraBERT) حصل على أعلى قيمة لـ F1-Measure بنسبة (98.3٪) في مهام التصنيف الثنائي و (90.99٪) في مهام التصنيف المتعدد.

الكلمات المفتاحية: التعلم الآلي ، البرمجة اللغوية العصبية ، اللغة العربية ، تمثيلات التشفير ثنائي الاتجاه العربية من المحولات (أرابرت) ، وسائل التواصل الاجتماعي.