تحليل المشاعر والتنقيب عن الآراء للمنتجات على تويتر باستخدام تقنيات التعلم العميق و تعلم الآلة

من قبل

مصعب عماد عبدالرحمن الشيخ ظاهر

بأشراف

الدكتور طارق غازي كنعان

جامعة الزيتونة الأردنية، 2022

الملخص

تاريخ المناقشة 5-6-2022

في السنوات الأخيرة ، انتشرت شعبية مواقع التواصل الاجتماعي. يستخدم الأشخاص منصات الوسائط الاجتماعية للتعبير عن آرائهم حول جميع الموضوعات تقريبًا. يتم التعبير عن هذه الآراء في مجموعة متنوعة من التنسيقات ، بما في ذلك المدونات والتغريدات ومنشورات Facebook ولوحات الرسائل عبر الإنترنت ومنشورات Instagram وما إلى ذلك. يتضمن تحليل المشاعر التوصيف الحسابي وتصنيف الآراء المعبر عنها في التعليقات ومنشورات المدونات والتغريدات. عادةً ما يكون الهدف من تحليل المشاعر هو تحديد رأي العميل بشأن منتج أو خدمة. تعد ملاحظات العملاء ضرورية للمؤسسات ، ويمكن الاستفادة من وسائل التواصل الاجتماعي ، وهي منصة قوية ، لتحسين وتوسيع إمكانات الشركة شريطة إجراء تحليل في الوقت المناسب لملاحظات الوسائط الاجتماعية. وبالتالي ، فإننا نركز على تحليل آراء العملاء للعديد من المطاعم في جميع

أنحاء العالم العربي. لتصنيف كل تعليق على أنه إيجابي أو سلبي أو غير ذلك ، استخدمنا أدوات معالجة مسبقة للغة العربية وثلاث تقنيات معالجة لغوية طبيعية معروفة (التطبيع ، وإيقاف إزالة الكلمات ، والتشتت). قمنا أيضًا بتجميع مجموعات بيانات تحتوي على حوالي 1804 منشور من منصات التواصل الاجتماعي العربيةTwitter و Facebook.

في علم الاله ، استخدمنا مصنفات عديدة معروفه ؛ Support Vector Machine (SVM) و Naive Bayes (NB)و Decision Tree (J48) و Random Forest (RF) و K-Nearest Neighbor (KNN) (KNN).حددنا دقة المصنف من خلال حساب F1-Measure و Recall و Precision.في المقابل ، استخدمنا تقنيات التعلم العميق ، والتي تشمل الشبكات العصبية التلافيفية (CNN) والشبكات العصبية العميقة .(DNN) يعد تضمين الكلمات أحد مكونات أسلوب استخراج الميزات الذي تم تنفيذه.(Word2Vector)

أظهرت النتائج أن المصنف SVM حصل على أعلى قيمة لمقياس F1 لكل من مجموعتي بيانات Twitterبنسبة (82.7٪) ومجموعة بيانات Facebook بنسبة (81.5٪). باستخدام تقنيات التعلم العميق ، نحصل على نتائج أفضل من مصنف ML حيث يعطي DNN أعلى قيمة من تقنيات الشبكة العصبية الأخرى في مجموعة بيانات Twitter ، نحصل على قيمة F1 بنسبة (86.8٪) ومجموعة بيانات Facebook بنسبة(85.8) .