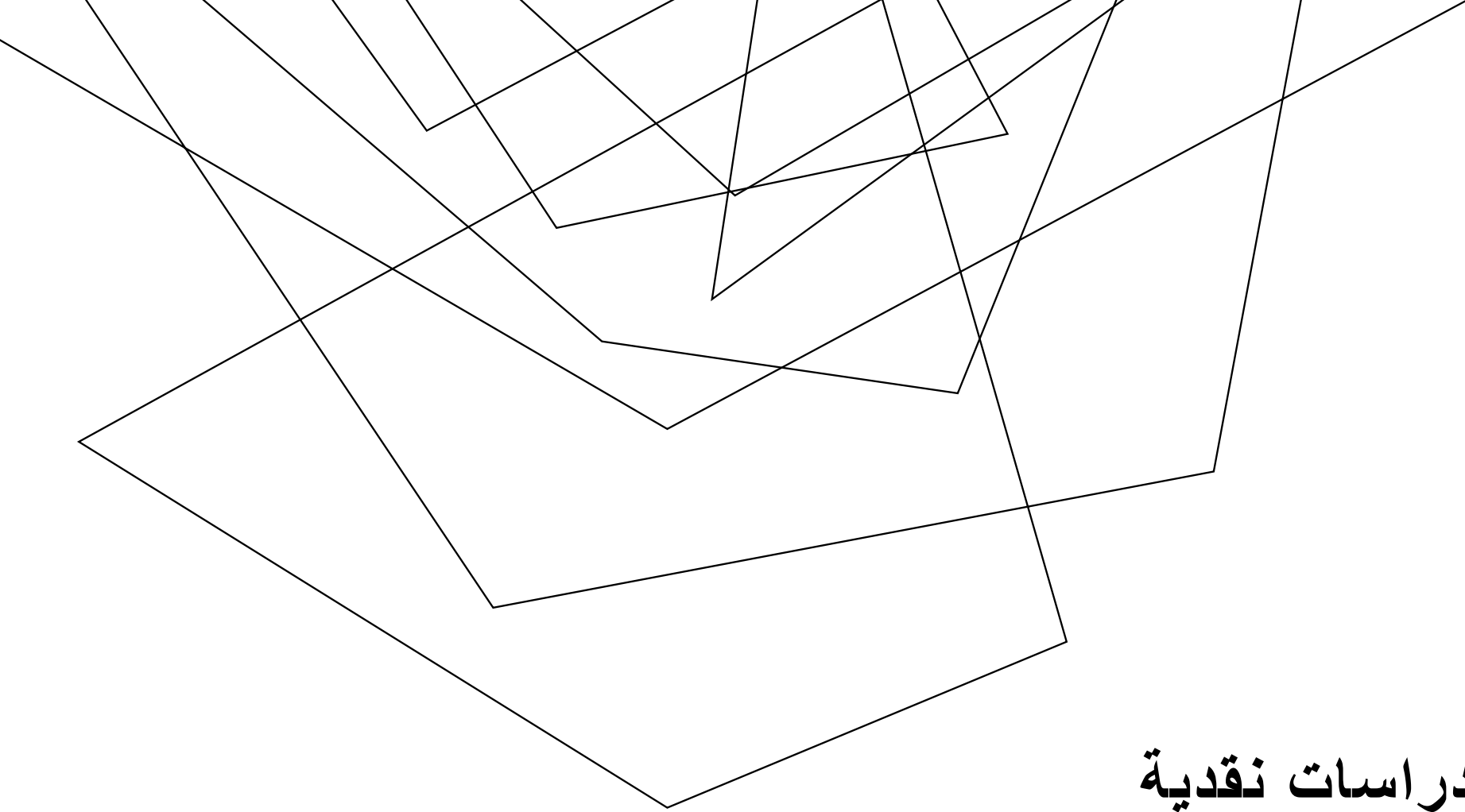




دراسات نقدية

إعداد: د. عبد العزيز الشمران

المهم الخامسة

اختيار رسالة دكتوراه في مجال التخصص، وتحليل وإبداء الملاحظات والتوصيات لتطوير الدراسة
المختارة والاقتراحات لبحوث مستقبلية

1- استخدام خوارزمية: NAÏVE BAYES

• الدراسة اعتمدت على خوارزمية Naïve Bayes لتحليل المشاعر، وهي خوارزمية تعتمد على الاحتمالات لتصنيف النصوص لكنها قد تكون محدودة في دقة النتائج، خاصة مع نصوص اللغة العربية التي تتطلب معالجة أكثر تعقيداً. لذلك، يُوصى بتجربة خوارزميات أكثر تقدماً للتعامل مع البيانات النصية العربية، مثل خوارزميات دعم المتجهات ((SVM، أو الشجرات العشوائية ((Random Forest، أو تقنيات التعلم العميق مثل الشبكات العصبية التلافيفية (CNN أو شبكات LSTM التي تعالج النصوص بتسلسل زمني.

• يمكن أيضاً استخدام نماذج تعتمد على التضمين النصي (Text Embedding Models) مثل AraBERT، وهي نسخة من BERT مصممة خصيصاً للغة العربية، مما قد يُحسن من دقة تصنيف المشاعر عبر فهم السياق النصي بشكل أفضل.

2- محدودية حجم البيانات

أشارت الدراسة إلى جمع البيانات يدويًا، مما أدى إلى وجود عينة صغيرة نسبيًا قد تكون غير كافية لتعميم النتائج. لذا، يمكن أن يفيد توسيع نطاق جمع البيانات ليشمل عددًا أكبر من التعليقات من منصات التواصل الاجتماعي المختلفة، بالإضافة إلى استخدام أدوات التجميع التلقائي مثل واجهة برمجة التطبيقات الخاصة بتويتر (Twitter API) أو فيسبوك لتوفير تدفق مستمر وكبير من البيانات.

كما يمكن تطوير إطار عمل مخصص لتجميع البيانات وتخزينها من مصادر متعددة، لتصبح متاحة لأي تجارب تحليل مستقبلي، ما يعزز من شمولية الدراسة.

3- التحديات في معالجة اللغة العربية

واجهت الدراسة تحديات تتعلق بالموارد والمكتبات الخاصة بمعالجة اللغة العربية، خاصةً فيما يتعلق باللهجات المختلفة التي قد يصعب تصنيفها باستخدام الأدوات التقليدية. توصي الدراسة باستخدام مكتبات أكثر تخصصاً للغة العربية، مثل مكتبة AraVec للتضمين النصي أو farasa لمعالجة اللغة العربية.

للتغلب على مشكلة اللهجات، يمكن إنشاء نماذج تستند إلى تحليل اللهجات بشكل منفصل، أو تدريب نموذج موحد يمكنه التعرف على الاختلافات بين اللهجات، خاصةً اللهجات العامية الشائعة مثل اللهجة المصرية، والشامية، والخليجية.

4- تنظيف البيانات ومعالجة النصوص

شملت الدراسة عمليات تنظيف البيانات من علامات خاصة وتجذيع الكلمات، لكنها لاحظت أن هذه العمليات قد تؤثر على دقة المعنى في النصوص العربية، إذ قد يؤدي التجذيع إلى فقدان دلالات معينة. يمكن تحسين عملية التنظيف عن طريق استخدام أدوات مخصصة لتنظيف النصوص العربية مثل تصفية الكلمات المكررة وتنقية النصوص من الرموز الغير ضرورية، بالإضافة إلى تحسين دقة النموذج باستخدام التضمين السياقي للكلمات، الذي يحافظ على المعنى الكامل للنص.

التوصيات لتطوير الدراسة

1. استخدام نماذج تعلم الآلة المتقدمة: يمكن تحسين النتائج عبر استخدام نماذج أكثر تطورًا، مثل الشبكات العصبية التلافيفية (CNN) أو شبكات LSTM التي تعالج السياق بشكل أفضل، خاصة عند التعامل مع النصوص العربية المعقدة.
2. التوسع في استخدام البيانات: يمكن جمع بيانات من مصادر متعددة مثل فيسبوك وإنستجرام بالإضافة إلى تويتر، لزيادة التنوع في البيانات ورفع دقة التحليل.
3. إعداد قواميس مخصصة: بناء قاموس خاص بالعواطف باللغة العربية يمكن أن يساعد في تحسين دقة التحليل، خاصة إذا تم تخصيصه لقطاع المكتبات والخدمات المعلوماتية.
4. استخدام التحليل المختلط: الاستفادة من تقنيات تحليل المشاعر المعتمد على القواميس إلى جانب تعلم الآلة قد يوفر أداءً متوازنًا، حيث يُجمع بين القاموس اللغوي وخوارزميات تعلم الآلة لتحقيق نتائج أدق.

الاقتراحات لبحوث مستقبلية

1. دراسة أثر اللهجات العربية المختلفة: دراسة كيفية تأثير اللهجات المحلية على نتائج تحليل المشاعر وتطوير نماذج تتعامل مع هذه اللهجات بشكل فعال.
2. تحليل تعليقات المستخدمين في أوقات الأزمات: دراسة استخدام تحليل المشاعر لتقييم ردود فعل المستخدمين في حالات الأزمات والطوارئ، مثل جائحة كورونا، وكيفية استخدام ذلك لتحسين الخدمات.
3. التوسيع في أنواع البيانات المدروسة: مثل استكشاف تحليل المشاعر على مراجعات الكتب الإلكترونية أو التطبيقات الخاصة بالمكتبات وتقييمها.
4. بناء أدوات مفتوحة المصدر: تطوير مكتبات وأدوات تحليل مشاعر خاصة باللغة العربية يمكن لمؤسسات المكتبات استخدامها مباشرة، مما يعزز من تطبيقات الذكاء الاصطناعي في هذا المجال.

المراجع

تحليل المشاعر القائم على الذكاء الاصطناعي كأداة لتقييم فعالية خدمات المعلومات

[https://jslmf.journals.ekb.eg/article_251102_df9102288cd373d8a919c1b70ba3d06c.p
df](https://jslmf.journals.ekb.eg/article_251102_df9102288cd373d8a919c1b70ba3d06c.pdf)

A series of white, overlapping geometric lines and polygons on a black background, located on the left side of the slide.

شكراً