كشف الشائعات العربية باستخدام نقل المعرفة بالاعتماد على السمات النصية و المرئية

الباحثة/ رشا مسلم سالم البلوي

المشرف / د.أماني طارق جمال
المشرف المساعد/ علاء عمر خديدوس

المستخلص

في الآونة الأخيرة ، ازداد استخدام منصات الوسائط الاجتماعية مع سهولة الاستخدام وسهولة الوصول إليها ، مما جعل هذه المنصات مكانًا لانتشار الشائعات بسبب عدم وجود قيود على النشر ومصادقة المحتوى. لذلك ، هناك حاجة للاستفادة من تقنيات الذكاء الاصطناعي للكشف عن الشائعات على منصات التواصل الاجتماعي لمنع آثارها السلبية على المجتمع والأفراد. تستهدف معظم الدراسات الحالية التي تكشف عن الشائعات باللغة العربية نص التغريدة. ومع ذلك ، تحتوي التغريدات على أنواع مختلفة من المحتوى ، بما في ذلك النصوص والصور ومقاطع الفيديو وعناوين URL ، وتلعب الميزات المرئية دورًا أساسيًا في نشر الشائعات.

تقترح هذه الدراسة نموذجًا عربيًا لكشف الشائعات لكشف الشائعات على تويتر باستخدام الميزات النصية والمرئية للتغريدات من خلال نوعين من الاندماج متعدد الوسائط: الاندماج المبكر والمتأخر. علاوة على ذلك ، استفدنا من نقل التعلم Transfer Learning للإصدارات العربية الحديثة من نماذج اللغة المدربة مسبقا BERT ، ونماذج تحويل النص إلى نص T5 ، ونماذج الرؤية الشائعة VGG-19 وResNet50 و InceptionV3 .

تم إجراء تجارب مختلفة لاختيار أفضل مستخرج ميزات نصية من بين ستة عشر نموذجًا للغة ونماذج استخراج الميزات المرئية لبناء النموذج متعدد الوسائط. تم استخدام MARBERTv2 كمستخرج للميزات النصية، في حين تم استخدام ensemble of ResNet50 and VGG-19 كمستخرج للميزات المرئية.

بعد ذلك، تم استخدام نماذج اللغة والرؤية للنماذج الفردية كخط أساس لمقارنة النتائج مع تلك الخاصة بالنماذج متعددة الوسائط. تم تقييم النماذج المقترحة باستخدام مجموعة البيانات المتاحة للجمهور AraFacts، وفيما يتعلق بقضايا الفصول غير المتوازنة، قمنا بجمع مجموعة بيانات الوسائط المتعددة الخاصة بنا ودمجها مع AraFacts . حققت نتائج التجارب لنموذجنا متعدد الوسائط للاندماج المبكر والمتأخر f1-scores 85 ,84 على التوالي. مقارنة بالنماذج أحادية الوسائط، يحقق نموذجا MARBERTv2 وensemble of ResNet50 and VGG-19 90 ,79 على التوالي.

أظهرت نتائج التجارب أن النماذج متعددة الوسائط تعمل بشكل أفضل من نماذج الرؤية القائمة على الصور ولا يمكنها أن تتفوق على النماذج القائمة على النص. توضح النتائج التي توصلنا إليها فعالية الميزات النصية في مهام الكشف عن الشائعات مقارنة بالنماذج متعددة الوسائط.

الكلمات المفتاحية:

معالجة اللغات الطبيعية، الذكاء الصناعي، التعلم العميق، كشف الشائعات ، نقل المعرفة ، الاندماج متعدد الوسائط

Arabic Rumor Detection Using Transfer Learning Based on Textual and Visual Features

by

Rasha Musallam Albalawi

Advisor

Dr. Amani Tariq Jamal

Co-Advisor

Dr. Alaa Omar Khadidos

Abstract

Recently, the use of social media platforms has increased with ease of use and fast accessibility, making them a place of rumor proliferation owing to the lack of posting constraints and content authentication. Therefore, there is a need to leverage artificial intelligence techniques to detect rumors on social media platforms to prevent their adverse effects on society and individuals. Most existing studies that detect rumors in Arabic target the text of a tweet. Nevertheless, tweets contain different types of content, including text, images, videos, and URLs, and their visual features play an essential role in rumor diffusion.

This study proposes an Arabic rumor-detection model to detect rumors on Twitter using textual and visual image features through two types of multimodal fusion: early and late. Furthermore, it incorporated the transfer learning of recent Arabic versions of the pre-trained language model, bidirectional encoder representations from transformers (BERT), text-to-text-transfer-transformers (T5), and popular vision models VGG-19, ResNet50, and InceptionV3. Different experiments were conducted to select the best textual features among sixteen language models and visual feature extractors for building a multimodal model. MARBERTv2 was used as the textual feature extractor, whereas the ensemble of VGG-19 and ResNet50 was used as the visual feature extractor. Subsequently, the language and vision models of the single models were used as a baseline to compare the results with those of the multimodal models. The proposed models were evaluated using the publicly available dataset AraFacts, and regarding unbalanced classes, the multimedia dataset was collected and combined with AraFacts.

The experimental results of our multimodal model for early and late fusion were 85% and 84% f1-scores, respectively. Compared to the unimodal model, the MAR- BERTv2 and ensemble vision models achieve 90% and 79% f1-scores, respectively. The experimental results show that the multimodal models perform better than the vision-based model and cannot outperform the text-based model. Our findings demonstrate the effectiveness of textual features in rumor detection tasks compared with multimodal models.

Keywords: Arabic NLP, Artificial intelligence, Deep learning, Multimodal fusion, Rumor detection, Transfer learning.