Detecting COVID-19 Misinformation in Arabic Tweets

Lama Alsudias

Prof. Paul Rayson

I. Introduction

- People use social media applications such as Twitter to find the news related to COVID-19 and/or express their opinions and feelings about it.
- The Arabic language is spoken by 467 million people in the world and has more than 26 dialects.
- In this poster, we have been Identifying and detecting the **rumours** related to **COVID-19**.

II. Data Collection and Preprocessing

When?	• We analysed the tweets related to COVID-19 from December 2019 to April 2020.
How many?	• We have collected approximately six million tweets in Arabic during this period.
What?	• We obtained the tweets depending on three keywords (كورونا), (كرونا), and (كوفيد-19) , which mean Coronavirus, a misspelling of the name of Coronavirus, and COVID-19 respectively in English.
How?	• We collected the tweets weekly using Twitter API.
Pre-processing	• we pre-processed the tweets through a pipeline of different steps

III. Method

- We applied a top-down strategy, which is where the set of rumours is identified in advance then the data is sampled to extract the posts associated with the previously identified rumours.
- The resulting dataset was 1,048,576 unique tweets from the original 6,578,982 collected.
- In our dataset, out of the one million tweets, we sampled 2,000 tweets to classify them for rumour detection. We manually labelled the tweets based on the list issued by the Ministry of Health to create a gold standard dataset and then applied different machine learning algorithms.

IV. Labelling Guidelines

Rumour in Arabic	Rumour in English
الحيوانات الأليفة تنقل فيروس كورونا.	Pets are transporters of Coronavirus.
البعوض ناقل لكورونا.	Mosquitoes are transporters of Coronavirus.
الأطفال قد لا يصابون بفيروس كورونا.	Children are not infected by Coronavirus.
كبار السن فقط قد يتعرضون لمخاطر سيئة من فيروس كورونا.	Only old people may have a high risk of Coronavirus.
حرارة الطقس أو برودته تقضي على الفيروس.	Hot or cold weather can kill the virus.
الغرغرة بالماء والملح تقضي على الفيروس.	Gargling with water and salt eliminates the virus.
الفيروس لا يبقى على الأسطح.	There are some herbs that protect against from Coronavirus.

V. Example Tweets

Tweet in Arabic	Tweet in English	Label
سيكون هناك انحسار لانتشار فيروس كورونا مع بداية فصل الصيف خصوصا في العالم العربي نظرا لارتفاع درجة الحرارة.	There will be a decrease in the spread of the Corona virus at the beginning of the summer, especially in the Arab world, due to the high temperatures.	1 (false)
الصحة: يعيش الفيروس ويرتكز بالأساس في الجهاز التنفسي لذلك غير وارد انتقاله عن طريق الحشرات أو من خلال لدغة البعوض.	The Ministry of Health: A virus lives and is mainly concentrated in the respiratory system, so it is not likely to be transmitted by insects or by mosquito bites.	-1 (true)
اللهم في هذي الساعة المباركة نسألك أن ترحمنا وتبعد عنا كل داء وبلاء وقنا شر الأمراض والأسقام واحفظ بلادنا وكافة بلاد المسلمين.	Oh God, in this blessed hour, We ask you to have mercy on us and keep away from us all disease and calamity, and protect us from the evil of diseases and sicknesses. Preserve our country and other Muslim countries.	0 (unrelated)

VI. Machine Learning Models

- We applied three different machine learning algorithms: Logistic Regression (LR), Support Vector Classification (SVC), and Naïve Bayes (NB).
- To help the classifier distinguish between the classes more accurately, we extracted further linguistic features.
- The selected features fall into two groups: word frequency, count vector and TF-IDF, and word embedding based (Word2Vec and FastText).
- We used 10-fold cross validation to determine accuracy of the classifiers for this dataset, splitting the entire sample into 90% training and 10% testing for each fold.

VII. Results

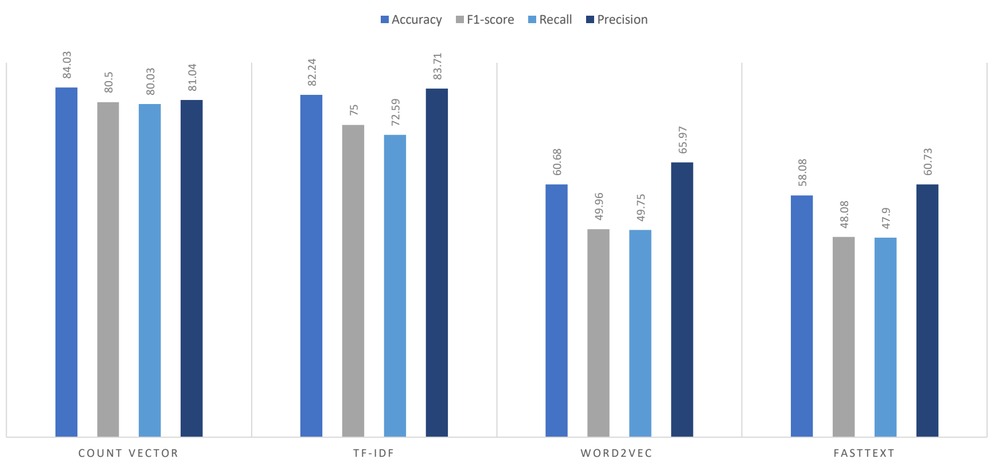

Figure: Results using Logistic Regression

VIII. Conclusion

- In this poster, we identified and analysed one million tweets related to the COVID-19 pandemic in the Arabic language.
- We labelled sample of tweets (2,000 out of 1 million) annotated for false information, correct information, and unrelated.
- We applied multiple Machine Learning Algorithms with different sets of features on the labelled tweets to investigate the replicability and scalability of this annotation.
- The highest accuracy result was 84% achieved by the LR classifier with count vector set of features and SVC with TF-IDF.
- The dataset, including tweet IDs, manually assigned labels for the sampled tweets, and other resources used in this poster are made freely available for academic research purposes.

If you have feedback or comments, please email: l.alsudias@lancaster.ac.uk