A Summary of on the feasibility of predicting fake news appearance – the Spanish case

Luis Ibañez-Lissen
Dept. Computer Science
Computer Security Lab, Universidad Carlos III de Madrid
luibanez@pa.uc3m.es

Lorena González-Manzano
Dept. Computer Science
Computer Security Lab, Universidad Carlos III de Madrid
lgmanzan@inf.uc3m.es

Jose M. de Fuentes
Dept. Computer Science
Computer Security Lab, Universidad Carlos III de Madrid
jfuentes@inf.uc3m.es

Manuel Goyanes
Dept. Media and Communication
Universidad Carlos III de Madrid
mgoyanes@hum.uc3m.es

***Abstract*—The growing amount of news shared on the Internet makes it hard to verify them in real time. An estimation of future fake news would help to focus the detection and verification efforts. Unfortunately, no previous work has addressed this issue yet. Therefore, this work measures the feasibility of predicting the volume of future fake news in a particular context – Spanish contents related to Spain. The approach involves different Artificial Intelligence mechanisms on a dataset of 298k real news and 8.9k fake news in the period 2019-2022. Results show that very accurate predictions can be reached.**

Contribution: *Research already published*

I. Introduction

Fake news are articles that are intentionally and verifiably false. They have been shared and spread as long as the very first newspapers were released. They have been used to destabilize countries, states and manipulate public opinion.

A vast array of research efforts have been focused on automatic detection of fake news. They train machine learning techniques to identify abnormal features in a piece of news. Other works have characterized different aspects of fake news, such as the profile of the victims or their spread pattern.

Despite prior efforts, no comprehensive research currently focuses on predicting the volume of fake news and herein it is where this paper contributes. Mainly, this paper addresses this matter by applying different Artificial Intelligence. (AI) techniques for forecasting the amount of fake news concerning the events happening in a particular case study – news written in Spanish related to Spain. Interestingly, our models are trained only with real news from reputed media, as qualifying news as fake is typically not achievable in real-time.

The research question at stake and related contributions of this work are as follows:

Research question – Is it feasible to predict the future amount of fake news?

- Different AI models are configured and trained, exhibiting relevant differences in their reliability.
- Input data required to perform the prediction is also characterized.
- The reliability of short, medium and long-term predictions is assessed.
- The trained models are publicly released to foster further research.

The structure of this summary is the following. Section II describes de proposal, which is later evaluated in Section III. Finally, Section IV outlines conclusions and future work.

II. System description

A. Overview

The following steps are initially considered in the process of predicting the volumne of fake news (Figure 1).

The process consists of four phases:

1) **Data extraction**. Events from The GDELT (Global Data on Events, Location, and Tone) project [1] are extracted, filtering by the target country, Spain. Moreover, fake news are obtained from Maldita[1] database.
2) **Data pre-processing and feature extraction/generation**. The features at stake are selected and normalized.
3) **Dataset preparation**. Once pre-processed, input data is gathered in groups of different sizes, to understand the most suitable input size. Different features of real news are considered in each experiment. Then, datasets are prepared for each experiment, splitting them in training and testing subsets.
4) **Training and assessment**. The model is trained and performance results are computed.

B. AI models

Three models (plus one variant) have been trained. First, a basic Recurrent Neural network (RNN) Long short-term memory cells (LSTM) model is designed. It consists of a model with a LSTM layer in charge of encoding all the information from the tabular data of GDELT.

A variant of this model is considered by adding an attention layer after the LSTM layer. The idea is to understand whether the basic LSTM improves the quality of the predictions.

Thirdly, an Encoder-decoder is designed. The architecture forces the model to compress the input enforcing a later reconstruction. This may allow the model to extract essential features that help reach better predictions.

[1] https://maldita.es/nosotros-maldita/, last access February, 2024.

ISBN:978-84-09-62140-8

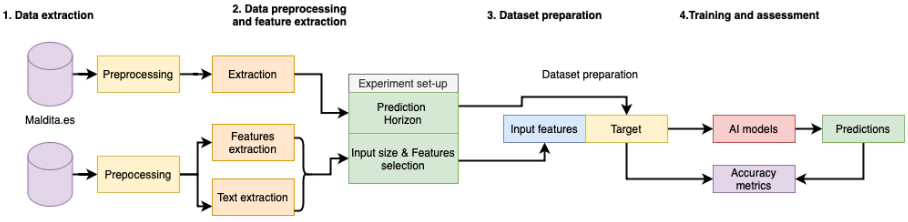

Fig. 1. Approach overview

Moreover, a multi-modal encoder-decoder architecture is designed in order to be able to encode headlines information from the processed URLs. We included a LSTM layer in charge of encoding the headlines after an embedding layer. The idea behind is to encode textual features which may be relevant to the model.

III. Assessment

A. Metrics

Three error functions are considered herein, namely the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE) and the coefficient of determination (R^2). Both MAE and RMSE serve to characterize the mean and standard deviation of the prediction errors and are measured in the predicted unit (i.e., number of fake news). Thus, their optimal value is zero. Finally, the coefficient of determination (R^2) represents the strength of the relationship or the portion of common variation in two time-series or variables. The closer to 1, the better, whereas negative values indicate that a horizontal line would be a better fit than the model at stake.

B. Data preparation

1) GDELT processing: All events available between 2019 and 2022 in GDELT filtering by Spain are collected and filtered by reputed news sources in Spain. A dataset of 298,242 events is produced.

2) Maldita.es processing: The dataset contains 8,990 fake news collected from 4 years, from 2019 until mid 2022. Precisely, they comprise a range of 1,018 days (2019-09-30 to 2022-07-14). For each fake news, timestamps of the first appearances are recorded. All the fake news per day are grouped, computing the total per day.

C. Impact of time horizons and input sizes

Results are depicted in Table I. It shows the prediction errors per model, time horizon and window size. Herein it is analysed the prediction of the amount of fake news in a given period, where 2 to 10 input days were considered for the sake of representativeness. All models report consistent values – when increasing the amount of input days, models learn better the trends. The longer the window sizes, the better the results.

The target period size (up to 7 days) affects the model performance. When predicting the number of fake news in the next 4 and 7 days, models do not have to face the complexity of understanding the possible abrupt changes that may happen from one day to another. The variability of the data is reduced and hence, the model does not have to face that much daily life uncertainty. When predicting the next day, all models show a relatively worse accuracy in direct comparison with the prediction the amount of fake news in the next seven days. Indeed, in terms of MAE, for horizon 1 day, MAE is around 3 and given a mean of 7.72 fake news in the dataset, the error is close to 38%. The situation improves for larger horizons and large windows size, for instance, horizon of 4 days and windows of 8 days, the mean MAE is 6.1 which leads to 21% of error.

TABLE I
PREDICTIONS FOR PERIODS

		LSTM			Encoder-decoder			LSTM + attention		
Horizon	Input Size	MAE	RMSE	R2	MAE	RMSE	R2	MAE	RMSE	R2
Next day	2 Days	4.08	5.78	0.39	4.03	5.61	0.42	3.83	5.45	0.44
	6 Days	3.71	5.58	0.47	3.69	3.58	0.49	3.58	5.43	0.497
	10 Days	3.68	5.59	0.44	3.59	5.29	0.5	3.35	5.06	0.544
Next 2 days	2 Days	6.78	9.43	0.49	6.5	9.25	0.51	6.55	9.2	0.52
	6 Days	5.83	8	0.56	5.96	8.17	0.53	5.51	7.68	0.578
	10 Days	5.41	7.56	0.63	5.26	7.7	0.62	4.88	7	0.681
Next 4 Days	2 Days	8.95	12.3	0.65	8.64	12.1	0.66	8.91	12.3	0.642
	6 Days	7.44	9.96	0.79	6.97	9.51	0.81	6.42	9.29	0.81
	10 Days	6.08	8.49	0.82	5.84	8.29	0.83	5.71	7.68	0.847
Next 7 Days	2 Days	11.7	15.5	0.77	11.4	15.4	0.78	11.3	14.9	0.791
	6 Days	8.5	11.9	0.86	7.92	11	0.88	7.31	11.1	0.878
	10 Days	5.95	8.13	0.93	6.35	8.62	0.93	5.18	7.33	0.947

IV. Conclusions and Future Work

Predicting the amount of expected fake news is a daunting task that has received little attention from the research community. Our results, published in [2], support the feasibility of this approach. The applied techniques have shown their ability to reproduce the fake news trend along a period of 4 years.

A possible future line of research is to extend this work in other contexts and countries to compare results, as well as to deploy more-advanced encoders to deal more deeply with semantic information.

Acknowledgment

This work has been supported by the Madrid Government under the Multiannual Agreement with UC3M (“Fostering Young Doctors Research”, DEPROFAKE-CM-UC3M), and in the context of the V PRICIT, and the Excellence Program for University Professors; by INCIBE grant APAMcyber within the framework of the Recovery, Transformation and Resilience Plan funds, financed by the European Union (Next Generation); and Jose Maria de Fuentes and Lorena Gonzalez have also received support from UC3M’s Requalification programme, funded by the Spanish Ministerio de Ciencia, Innovacion y Universidades with EU recovery funds (Convocatoria de la UC3M de Ayudas para la recualificación del sistema universitario español para 2021-2023).

References

[1] K. Leetaru and P. A. Schrodt, “Gdelt: Global data on events, location, and tone,” in *ISA annual convention*, vol. 2, no. 4. Citeseer, 2013, pp. 1–49.

[2] L. Ibañez-Lissen, L. González-Manzano, J. M. de Fuentes, and M. Goyanes, “On the feasibility of predicting volumes of fake news—the spanish case,” *IEEE Transactions on Computational Social Systems*, 2023.