

Previous research data (anonymous)

This research has been published in relevant journals, or it is now in press but it can not be provided at this point of the reviewing process. Main data of this previous work is extracted to support reviewing. If the paper is accepted the references would be added to clarify the proposal to the readers.

Annotation scheme

For each news item, these three levels of content are annotated:

INVERTED PYRAMID	5W1H	ELEMENTS OF INTEREST
TITLE: attributes ::= <style>, <title_stance>	WHAT attributes ::= <reliability>, <lack_of_information>, <main_event>	QUOTE attribute ::= <author_stance>
SUBTITLE	WHO attributes ::= <reliability>, <lack_of_information>, <role>	KEY EXPRESSION
LEAD	WHEN attributes ::= <reliability>, <lack_of_information>	FIGURE
BODY	WHERE attributes ::= <reliability>, <lack_of_information>	ORTHOTYPOGRAPHY
CONCLUSION	WHY attributes ::= <reliability>, <lack_of_information>	
HOW attributes ::= <reliability>, <lack_of_information>		
Values: style ::= Objective Subjective title_stance ::= Agree Disagree Unrelated	Values: reliability ::= Reliable Unreliable lack_of_information ::= Yes role ::= Subject Target Both main_event ::= n/a	Values: author_stance ::= Agree Disagree Unknowr

Besides, a global reliability and veracity value is assigned to the news item.
Found the complete detailed annotation guidelines at the end of this document.

Evaluation of reliability with news content annotation following the schema:

Experiments	Baseline	model	Model with	
	(TF-IDF)		features	
	Acc	F_1m	Acc	F_1m
SVM	0.662	0.395	0.937	0.925
Random Forest	0.75	0.639	0.912	0.898
Logistic Regression	0.65	0.392	0.912	0.875
Decision Tree	0.737	0.683	0.95	0.948
MLP	0.712	0.57	0.925	0.912
AdaBoost	0.787	0.748	0.95	0.945
GaussianNB	0.612	0.456	0.687	0.57
	Baseline model		Model with	
			features	
BETO	0.85	0.80	0.887	0.854

As can be observed from the previous research results, the approaches trained with a dataset annotated with the schema is very powerful to determine reliability of the documents.

This is the previous experimentation that demonstrates the appropriateness of the annotation.

Evaluation of veracity only with news content annotation following the schema:

Model	True News			Fake News			Acc	Macro F_1
	P	R	F_1	P	R	F_1		
Baseline (Random)	0.551	0.549	0.548	0.498	0.500	0.497	0.526	0.522
Baseline (TF-IDF)	0.609	0.868	0.715	0.726	0.381	0.494	0.637	0.605
<div></div>	0.920	0.550	0.790	0.680	0.950	0.690	0.750	0.740

Using the content annotation features of the schema, the features are improving baselines for veracity detection, and the new approach presented at ICDAR is improving previous results that are not using external evidences.

Conclusion from previous research based only on reliability:

As observed in these results, here fake news items are detected using reliability features only. It can be observed that most of the reliable news items are not true, and this conclusion of previous research leads us to the conclusion that adding external world knowledge could improve coverage detection of fake news, as was demonstrated in our ICDAR manuscript.

Estos datos últimos no están publicados, si nos vale con las tablas que te he puesto antes casi que mejor para justificarlo, lo que está entre las líneas es lo que tengo dudas, el resto al estar publicado o in press no me preocupa

Evaluation results of reliability detection and veracity detection of a subset of Posadas dataset enriched with the proposed annotation schema (name of the annotation schema is hidden):

Experiments	Baseline model NO features		Model with features	
	Acc	F_1m	Acc	F_1m
MLP	49.00	48.58	93.00	92.44
SVM	62.00	38.27	95.00	94.66
LR	38.00	36.76	94.00	93.56
Decision tree	41.00	40.27	87.00	85.97
Random Forest	52.00	51.69	88.00	86.97

Table 8: Experiment results using classical ML approaches for Reliability detection task

Experiments	Baseline model NO features		Model with features	
	Acc	F_1m	Acc	F_1m
MLP	56.00	44.19	74.00	73.62
SVM	52.00	34.21	75.00	74.57
LR	52.00	43.89	77.00	76.60
Decision tree	54.00	52.45	72.00	71.98
RandomForest	60.00	57.55	78.00	77.96

Table 9: Experiment results using classical ML approaches for fake news detection task

Appendix. Annotation schema details.

2. Structure level

The structure level divides a news item into different parts following the Inverted Pyramid journalistic technique. According to this technique, each structure part contains information with different levels of relevance, placing the most important information at the beginning of the news item and the least relevant at the end. Structure labels are described in order of relevance below.

2.1. <TITLE>

The TITLE of the news article provides the main idea of the story. Normally in one sentence, it summarises the basic and essential information about the story and the idea around which the news piece has been created. The main objective of the TITLE is to attract readers.

This label presents two attributes that provide additional information about the TITLE:

2.1.1. style

This attribute allows to mark two values: Objective and Subjective. An objective title will present information in an accurate and informative way, while a subjective title will provide information in an alarmist, connotative or emotional way.

2.1.2. title_stance

This attribute indicates whether the information presented in the BODY is consistent with the information of the TITLE. This consistency is represented by the following values: Agree (information provided is consistent in both parts), Disagree (information provided is inconsistent in one of the parts) or Unrelated (information provided in the TITLE has no relation with the rest of the news item).

2.2. <SUBTITLE>

The SUBTITLE explains the TITLE in more detail. It completes the information by presenting the idea in a very summarised way or can provide additional information not mentioned in the TITLE. The SUBTITLE'S purpose is to keep the reader's attention and encourage him/her to read the whole news article.

2.3. <LEAD>

The LEAD is the paragraph that develops the main information by answering the six key questions allowing to communicate information in an accurate and objective way (the 5W1H journalistic technique).

2.4. <BODY>

The BODY contains all the information developed in the news article. The BODY presents all the background, facts and arguments of the story in detail. The key questions answered in the LEAD are developed in the BODY by explaining all the elements that are involved in the news piece.

2.5. <CONCLUSION>

The main idea of the story can be summarised in a sentence or in a paragraph but, even if the CONCLUSION is considered part of a well-structured article, it does not always appear. It presents the least important information, as it is only a summary of all the important information that has been developed in the previous parts of the news story.

3. Content level

The second level focuses on the essential content elements of news. The approach followed in this level is based on the journalistic technique known as 5W1H, which enables the detection of the key elements needed to accurately communicate a story. The 5W1H questions used for this journalistic method are WHAT, WHO, WHERE, WHEN, WHY and HOW. All the content level labels are marked with the attributes reliability and lack_of_information (if necessary):

□ reliability

Attribute allowing to annotate a news item as Reliable or Unreliable depending on the level of accuracy and objectivity.

3.1. <WHO>

In a sentence, the WHO represents the subject or entity involved or acting in an event. It may usually refer to people, organisations or even personified entities (such a country: e.g., France discovers a vaccine...)

3.2. <WHAT>

The WHAT label refers to the circumstances, events or facts of the action performed by the subject.

3.3. <WHEN>

The WHEN label indicates the time or the moment when the events occurred. It is found in temporary expressions (e.g., on Wednesday, in 2010, last Friday...)

3.4. <WHERE>

This label designates the location where the events occurred. It is found in location expressions, either physical (e.g., in France, in a laboratory) or not (e.g., in Facebook).

3.5. <WHY>

This label refers to the cause of the event. It must not be confused with the purpose.

3.6. <HOW>

This label refers to the way events have developed, the manner or the method in which a given action has been carried out.

4. Elements of Interest

This level allows to mark textual information that could be interesting to differentiate unreliable news from reliable news.

4.1. <KEY_EXPRESSION>

Phraseology that urges readers to share the information or that expresses emotions such as fear, contempt, alarm, hope or economic and ideological purposes.

Examples

ENGLISH:

<KEY_EXPRESSION>SHARE THIS INFORMATION</KEY_EXPRESSION>

<KEY_EXPRESSION>THIS CAN SAVE YOUR LIFE</KEY_EXPRESSION>

<KEY_EXPRESSION>THE LEMON PREVENTS AND CURES
CANCER</KEY_EXPRESSION>

SPANISH:

<KEY_EXPRESSION>COMPARTE ESTA INFORMACIÓN</KEY_EXPRESSION>

<KEY_EXPRESSION>ESTO PUEDE SALVAR TU VIDA</KEY_EXPRESSION>

<KEY_EXPRESSION>EL LIMÓN PREVIENE Y CURA EL CÁNCER</KEY_EXPRESSION>

4.2. <FIGURE>

Label allowing to mark figures in the text, as it is a characteristic that can be verified by fact-checking techniques.

Examples

<FIGURE>8</FIGURE>

<FIGURE>1200</FIGURE>

<FIGURE>560</FIGURE>

4.3. <ORTHOTYPOGRAPHY>

This label is used to mark grammatical, spelling or formatting mistakes that can be found in the text. Some examples of orthotypography are whole sentences in capital letters, suspension points in the middle of the text or incomplete, double spaces, many exclamation marks, grammatical errors, spelling mistakes, lack of cohesion, etc.

Examples

ENGLISH:

<ORTHOTYPOGRAPHY>This will completly change your life</ORTHOTYPOGRAPHY>

<ORTHOTYPOGRAPHY>LEMON IS AN ANTI-CANCER FOOD...that can save your life thanks to...its anti-cancer properties!!!!!!!!</ORTHOTYPOGRAPHY>

<ORTHOTYPOGRAPHY>IF YOU DON'T SHARE THIS INFORMATION, YOU WILL HAVE 5 YEARS OF BAD LUCK</ORTHOTYPOGRAPHY>

SPANISH:

<ORTHOTYPOGRAPHY>Esto camviará completamente tu bida</ORTHOTYPOGRAPHY>

<ORTHOTYPOGRAPHY>EL LIMÓN ES UN ALIMENTO ANTICÁNCER...que puede salvar tu vida gracias a...sus propiedades anticancerígenas!!!!!!!!</ORTHOTYPOGRAPHY>

<ORTHOTYPOGRAPHY>SI NO COMPARTES ESTA INFORMACIÓN, TENDRÁS 5 AÑOS DE MALA SUERTE</ORTHOTYPOGRAPHY>

4.4. <QUOTE>

This label enables the annotation of elements or sentences that textually quote a message or reproduce an already reported idea.

4.4.1. author_stance

The QUOTE label only has an attribute called author_stance represented by the following values: Disagree (to express its disagreement towards the idea), Agree (to share its agreement) or Unknown (just to inform, without showing its stance towards it).

