Automatic Detection of Hungarian Clickbait and Entertaining Fake News

Veronika Vincze[1] and **Martina Katalin Szabó**[2,3]
[1]MTA-SZTE Research Group on Artificial Intelligence, Szeged, Hungary
[2]Computational Social Science Research Center for Educational and Network Studies (CSS-RECENS), Centre for Social Sciences, Budapest, Hungary
[3]Department of Software Engineering, University of Szeged, Szeged, Hungary
{vinczev, martina}@inf.u-szeged.hu

Abstract

Online news do not always come from reliable sources and they are not always even realistic. The constantly growing number of online textual data has raised the need for detecting deception and bias in texts from different domains recently. In this paper, we identify different types of unrealistic news (clickbait and fake news written for entertainment purposes) written in Hungarian on the basis of a rich feature set and with the help of machine learning methods. Our tool achieves competitive scores: it is able to classify clickbait, fake news written for entertainment purposes and real news with an accuracy of over 80%. It is also highlighted that morphological features perform the best in this classification task.

1 Introduction

The growing number of online news and the ability to easily and rapidly distribute information on the internet increasingly stimulate demand for automatic fact checking (Thorne and Vlachos, 2018). As a consequence, linguistic aspects of deception, bias and uncertainty detection have raised worldwide interest recently and have been thoroughly studied in a variety of NLP-applications (Zhou et al., 2004; Mihalcea and Strapparava, 2009; Szarvas et al., 2012; Choi et al., 2012; Girlea et al., 2016; Barrón-Cedeño et al., 2019). However, determining the trustworthiness of news and separating facts from misinformation is still a challenging and often controversial task (Graves, 2018).

There may be several motivations for spreading fake news on the internet. Hoax websites, for instance, publish dubious news (clickbait) in order to spread different kinds of misinformation or make money by spreading commercials. To be more specific, fabricated news draw disproportionate attention on social networks most of the time, outperforming conventional news (Graves, 2018). Consequently, publishing interesting fake news is a great way to spread advertisements on the internet. At the same time, there are pages where the primary purpose is to entertain readers by spreading fake news. In this case the readers are aware of the deceptive nature of the information provided and they read it just for fun.

In this paper we will focus on two types of fake news written in Hungarian, namely clickbait and fake news written for entertainment purposes. Our main goal is to distinguish these two types from real news with machine learning methods. We will also analyse what type of linguistic information can contribute the most to performance.

The paper is structured as follows: First, we will give a short review of the relevant research work and we detail the importance and benefits of the recent analysis. Then, we will present the corpus analysed, along with its basic statistical data and the methods and tools we used in the experiments. Next, we will introduce our rich feature set consisting of statistical, morphological, syntactic, semantic and pragmatic features applied for statistical significance analysis and machine learning experiments. We will discuss the findings of the significance analysis, and then, we will provide a detailed description of the machine learning experiments, along with the results. Finally, we discuss the results and add some ideas for future work.

Proceedings of the 3rd International Workshop on Rumours and Deception in Social Media (RDSM), pages 58–69
Barcelona, Spain (Online), December 12, 2020.

2 Related work

Most of the authors address the issue of automatic fact extraction and binary classification verification task based on machine learning methods and annotated datasets. There are several studies addressing the phenomena of deception and bias in different types of discourse. For instance, most of the authors analyze the phenomena of deception and bias either in speech text (Fetzer, 2008; Fraser, 2010; Scheithauer, 2007; Simon-Vandenbergen et al., 2007) or address the issue of automatic fact extraction and binary classification verification task (Greene and Resnik, 2009; Rubin et al., 2015; Wang, 2017; Thorne and Vlachos, 2018; Thorne et al., 2018; Graves, 2018). In addition, uncertainty detectors have mostly been developed in the biological and medical domains (Szarvas et al., 2012). Also, a few studies seem to address the issue of the systematic analysis of a huge amount of propaganda texts (Propaganda Analysis, 1938; Rashkin et al., 2017; Barrón-Cedeño et al., 2019; Vincze et al., 2019; Kmetty et al., 2020; Szabó et al., 2020).

Recently, several databases have been compiled for the English language which contain both fake and real news. For instance, the dataset described in Vlachos and Riedel (2014) consists of a total of 221 statements, along with their veracity value. The Liar corpus contains approximately 13,000 short political statements (Wang, 2017). The Emergent Corpus consists of 300 statements and 2500 news related to the semantic content of the statements, along with their veracity value (Ferreira and Vlachos, 2016). Our study is most similar in vein to Rubin et al. (2016), which compares satirical news to real news, collected from different websites. Pérez-Rosas et al. (2018) used crowdsourcing to generate fake news on the basis of real news and then carried out machine learning experiments to separate the two types. There are also a few studies that identify clickbaits (see e.g. Biyani et al. (2016) for English and Karadzhov et al. (2017) for Bulgarian).

The above studies focus on the phenomena of deception and bias almost exclusively in sources written in English, so the reported research findings and models are mostly limited to the English language. The same goes for the currently existing datasets. Santos et al. (2020) forms an exception, which, however, aims at the distinction of Brazilian Portuguese fake and real news. To the best of our knowledge, our recent research work is the first attempt at the automatic detection of Hungarian fake news.

The main contributions of our paper are the following:

- We present a novel dataset for detecting fake news in Hungarian;
- We define a rich feature set of linguistic parameters for detecting different characteristics of different types of Hungarian news examined;
- We carry out a detailed statistical analysis of linguistic parameters that may distinguish real news, clickbait and entertaining fake news in Hungarian;
- We perform machine learning experiments with the above mentioned feature set for detecting real news, clickbait and fake news written for entertainment purposes;
- We analyze the effect of each set of features and identify the best combination of these features for the task.

3 The corpus

Our study was conducted on a corpus of 180 online news compiled by us. First, the real news come from several national and regional news portals, e.g. `www.index.hu`, `www.origo.hu` and `www.delmagyar.hu`. Second, the fake news were collected from two sources. On the one hand, we selected some special news published on the 1st of April (Fools' Day) on some of the news portals basically spreading real news, as this day is traditionally seen as an occasion for making pranks. These news were intended to play a trick on readers. On the other hand, we downloaded articles from the Hírcsárda website[1]. These texts are mostly based on parodies of current political and public events, and their

[1] `https://www.hircsarda.hu`

Texts	Number of texts	Number of sentences	Number of tokens
Real news	60	1673	22203
Clickbait	60	1241	13925
Entertaining fake news	60	1310	16193

Table 1: Basic statistical data of the corpus.

explicit purpose is not to spread real or fake news but to entertain the readers. Third, the clickbait news were downloaded from websites that were collected by a Hungarian real news portal and were claimed to be unreliable[2]. Most of the texts were published during the time period between 2017 and 2019. For each type of text, we collected 60 documents. The news were randomly selected but we made sure that a given topic should not be included in the corpus more than one time.

The basic data of our corpus are presented in Table 1.

4 Experiments

In this section, we present our methods for identifying the three types of news with machine learning methods as well as providing statistical significance analysis for the linguistic features defined for the task.

4.1 Feature set

As a first step, we automatically preprocessed the texts with magyarlanc (Zsibrita et al., 2013), a toolkit written in JAVA for the linguistic processing of Hungarian texts. With this tool, the text was first split into sentences, then tokenised and lemmatised, and finally morphologically and syntactically analyzed. Based on the output of magyarlanc, we extracted a high number of linguistic features. Our feature set consists of the following features:

- **Basic statistical features**: the number of sentences; the number of words; the number and rate of lemmas; the average sentence length.
- **Morphological features**: Part-of-speech (or POS) features: the number and frequency of nouns, proper nouns, verbs, adjectives, (demonstrative) pronouns, numerals, adverbs, conjunctions and unanalyzed words (i.e. those with an "unknown" POS tag); the number of punctuation marks; the number and frequency of imperative and conditional verbs; the number and frequency of past and present tense verbs; the number and frequency of first person singular verbs and of first person plural verbs; the number and frequency of frequentative, causative and modal verbs; the number and frequency of comparative and superlative adjectives.
- **Syntactic features**: the number and frequency of subjects and objects as Hungarian is a pro-drop language, meaning that pronominal subjects and objects might not be overt in the clause; the number and frequency of adverbials; the number and frequency of coordinations and subordinations.
- **Semantic features**: the number and frequency of negation words; the number and frequency of content words and function words; the number of frequency of public verbs, private verbs and suasive verbs based on a Hungarian translation of lists found in Quirk et al. (1985). Uncertainty features: the number and frequency of words belonging to several classes of linguistic uncertainty based on Vincze (2014a). Sentiment features: the number and frequency of positive and negative words based on a list of sentiment phrases. We applied two different Hungarian dictionaries for sentiment analysis: one list was a translation of Liu (2012), while the other one contained Hungarian slang words (Szabó, 2015), respectively. Emotion features: the number and frequency of words belonging to the emotions described in Szabó et al. (2016).

  For list-based semantic features we used a simple dictionary-based method. Thus, if a lemma in our corpora matched with any item in our lists, it was counted as a hit.

[2] `https://hvg.hu/tudomany/20150119_atveros_weboldalak`

- **Pragmatic features**: the number and frequency of speech act verbs, based on a list manually constructed by us; the number and frequency of literal quotes and citations, detected on the basis of quotation marks and dashes at the beginning of the sentences. The number and frequency of discourse markers. To find discourse markers in the texts we applied a word list based on Dér and Markó (2007).

4.2 Statistical significance analysis

In order to measure the effect of each feature in distinguishing real news, clickbait and entertaining fake news, we performed statistical significance tests (pairwise t-tests) for all features.

Here we assume that the distinction between different types of fake news and real news is primarily a semantic-pragmatic problem. As a consequence, we also presuppose that morphological and syntactic features of the texts in the three subcorpora will not necessarily differ from each other.

The results of the statistical significance analysis are shown in Table 2. For the sake of simplicity, we provide p-values for features only where a statistically significant difference was found. For comparison, we also report the mean values for each feature for all classes in Table 3.

4.3 Machine learning experiments

In addition to the statistical significance analysis, we seek to automatically discriminate clickbait and fake news from real news. For this purpose, we made use of the above mentioned rich feature set including statistical, morphological, syntactic, semantic and pragmatic characteristics of Hungarian texts.

In our experiments, we used a Random Forest classifier (Breiman, 2001) with ten fold cross validation since it does not easily overfit. . In order to examine which features play the most important role in distinguishing the three groups of news, we divided the features into five main groups based on a linguistic classification, and experimented with all possible combinations of these groups, yielding an extensive ablation analysis. The baseline method was majority classification, which achieved an accuracy of 33.33%.

5 Results

In this section, we present the results of both our statistical analysis and machine learning experiments.

5.1 Results of statistical significance analysis

Table 2 shows the features that exhibit significant differences among three groups of news. Basically, our hypothesis is just partly confirmed: the findings do not show the outstanding role of semantic-pragmatic features because at the same time, morphological characteristics of the tokens proved to be also essential.

Apart from morphological features, the frequency distribution of certain types of semantic contents also seem to be significantly different in the three subcorpora. For instance, there is a significant difference of the frequency of uncertainty markers such as condition, weasel, peacock and hedge between clickbait news and both of the other corpora. While the first type belongs to semantic uncertainty, the latter three are types of uncertainty at the discourse level (Vincze, 2013). In the case of semantic uncertainty, the lexical content (meaning) of the uncertainty marker (cue) is responsible for uncertainty, e.g. *may*, *possible*, *believe* etc. (Vincze, 2014b). In contrast to semantic uncertainty (Szarvas et al., 2012), in the case of discourse level uncertainty, "the missing or intentionally omitted information is not related to the propositional content of the utterance but to other factors", e.g. for cues like *some*, *often*, *much* etc. Bias evoked by discourse-level uncertainty might be viewed as a characteristic feature of clickbait news.

5.2 Results of machine learning experiments

In the best scenario, our machine learning algorithm achieved an accuracy of 82.78%, which was yielded by combining statistical, morphological, semantic and pragmatic features. This proved to be the best combination of features for identifying fake news (with an F-score of 81.4). The combination of morphological and semantic features seemed to be the most effective for identifying real news, obtaining an F-score of 80.3. As for clickbait news, the combination of statistical, syntactic, semantic and pragmatic features yielded the best result (an F-score of 89.1). More detailed results are shown in Table 4.

Feature	click-fake	fake-real	click-real	Feature	click-fake	fake-real	click-real
token #			0.0047	coord. %			0.0306
lemma #	0.0054		0.0013	uncertain %	0.0254		0.0053
lemma %	<0.0001	0.0001		negation %	0.0003		<0.0001
token %			0.0013	invest. #			0.0054
sentence length			0.0007	epistemic %	0.0492		
unknown #	<0.0001	<0.0001		invest. %			0.0094
unknown %	<0.0001	<0.0001		condition %	0.0154		0.0005
noun #	0.0328	0.0379	0.0009	weasel %	0.0178		0.0067
adjective #	0.0137	0.0456	0.0008	peacock %	0.0088		0.0040
pronoun #	0.0213			hedge %	0.0062		<0.0001
numeral #		0.0006	0.0001	joy #	0.0083		0.0102
punct			0.0149	sorrow #	0.0045	0.0253	
proper noun #	<0.0001		<0.0001	love #	0.0159		0.0248
verb %	<0.0001	0.0358	<0.0001	joy %	<0.0001	0.0011	<0.0001
noun %	<0.0001		<0.0001	fear %	0.0310		0.0021
adjective %	0.0001		<0.0001	sorrow %	0.0058	0.0117	
pronoun %	<0.0001		<0.0001	love %	0.0078		0.0007
conjunction %	0.0094		0.0012	anxiety %	0.0288		0.0063
numeral %		0.0005	<0.0001	surprise %		0.0497	0.0030
adverb %	<0.0001	0.0007	<0.0001	positive1 #	0.0273		
proper noun %	<0.0001		<0.0001	negative1 #		0.0221	
comparative #			0.0326	positive2 #		0.0279	0.0280
superlative %	0.0214			positive2 %	<0.0001		<0.0001
past tense #		0.0154		negative1 %	0.0010		
past tense %		0.0018	0.0165	positive2 %	0.0001	0.0011	<0.0001
present tense %		0.0035		emo. negative %	0.0280		0.0125
imperative #	0.0009		0.0286	content %	<0.0001		<0.0001
imperative %	0.0006		0.0008	function %	<0.0001		<0.0001
1Sg verb %		0.0046	0.0255	public verb #		0.0047	0.0002
dem. pronoun #	0.0093	0.0026		private verb %			0.0050
1Pl verb %		0.0357		public verb %		0.0228	0.0003
dem. pronoun %		0.0067	0.0080	suasive verb %	0.0245		
modal verb %		0.0019	0.0440	quote #	<0.0001		0.0206
subject #		0.0418	0.0071	dash #	0.0011		0.0178
subord. #	0.0056	0.0213	0.0002	speech act %	0.0322		0.0179
coord. #			0.0090	quote %	<0.0001	0.0008	
subord. %	<0.0001	0.0005	<0.0001	dash %	0.0058		
adverbial %	0.0259		0.0246	disc. marker %	0.0001	0.0034	<0.0001

Table 2: Statistically significant features. #: number, %: rate.

Feature	click	fake	real	Feature	click	fake	real
token #	232.0833	269.8833	370.0500	uncertain #	2.8000	2.2167	2.8667
sentence #	20.6833	21.8333	27.8833	uncertain %	0.0119	0.0081	0.0073
lemma #	143.0167	188.0333	211.9667	negation #	4.4333	3.6167	4.3500
lemma %	0.6441	0.7334	0.6595	negation %	0.0199	0.0125	0.0104
token %	13.9728	14.8909	15.8212	epistemic #	0.7500	0.6000	0.8333
sentence length	11.6967	12.5064	13.3731	investigation #	0.0500	0.2000	0.4167
unknown #	0.0500	1.3000	0.2333	condition #	2.2333	1.5167	1.9833
unknown %	0.0003	0.0052	0.0005	weasel #	7.2667	7.0000	7.4667
verb #	38.2333	34.9500	44.9167	peacock #	1.8000	1.2000	1.6500
noun #	58.0500	75.7167	107.4667	hedge #	3.6333	3.2667	3.8167
adjective #	24.6000	35.3333	52.7167	doxastic #	2.8000	2.1333	3.2667
pronoun #	17.5667	12.7667	18.0833	epistemic %	0.0035	0.0019	0.0019
conjunction #	6.9333	6.2333	8.3500	investigation %	0.0002	0.0007	0.0016
numeral #	6.4167	7.4000	15.0833	condition %	0.0096	0.0059	0.0044
adverb #	25.6000	24.7667	29.5667	weasel %	0.0311	0.0246	0.0236
punct #	46.1500	53.1500	69.8667	peacock %	0.0077	0.0045	0.0044
proper noun#	4.8667	17.8167	20.9500	hedge %	0.0157	0.0112	0.0091
verb %	0.1682	0.1299	0.1195	doxastic %	0.0117	0.0086	0.0094
noun %	0.2451	0.2784	0.2879	joy #	4.1000	2.2833	2.1333
adjective %	0.1034	0.1269	0.1347	fear #	0.7000	0.4667	0.3833
pronoun %	0.0748	0.0454	0.0465	anger #	0.4000	0.2333	0.5333
conjunction %	0.0304	0.0232	0.0216	sorrow #	1.1000	0.3167	3.0500
numeral %	0.0269	0.0298	0.0467	love #	0.8500	0.3833	0.4000
adverb %	0.1165	0.0888	0.0740	anxiety #	0.5500	0.3333	0.3500
proper noun %	0.0204	0.0706	0.0661	disgust #	0.1833	0.1833	0.2167
superlative #	0.8333	0.5500	0.6667	surprise #	0.3833	0.2500	0.2167
comparative #	0.7167	0.9167	1.3167	joy %	0.0159	0.0076	0.0041
superlative %	0.0316	0.0138	0.0167	fear %	0.0034	0.0017	0.0010
comparative %	0.0328	0.0271	0.0313	anger %	0.0015	0.0007	0.0017
Sg1 verb #	1.4833	1.1500	1.2167	sorrow %	0.0049	0.0014	0.0050
Pl1 verb #	2.6500	2.5000	2.8333	love %	0.0039	0.0014	0.0008
past #	13.8333	11.9167	19.5500	anxiety %	0.0022	0.0009	0.0007
present #	20.1167	19.6333	21.5167	disgust %	0.0007	0.0005	0.0006
past %	0.3758	0.3546	0.4827	surprise %	0.0019	0.0009	0.0003
present %	0.5167	0.5471	0.4364	positive #	11.9167	8.5000	11.7667
imperative #	3.5500	1.5333	2.0667	negative #	8.4167	6.9000	12.4500
cond. verb #	1.1333	1.7333	1.4833	positive2 #	7.4333	7.2833	13.1167
imperative %	0.0890	0.0396	0.0393	negative2 #	16.5833	14.3000	16.5167
cond. verb %	0.0310	0.0506	0.0312	neg. emotive #	0.4667	0.3333	0.6000
Sg1 verb %	0.0317	0.0310	0.0127	positive %	0.0498	0.0303	0.0275
dem. pronoun #	6.4500	4.3000	7.8500	negative %	0.0370	0.0249	0.0303
Pl1 verb %	0.0736	0.0684	0.0401	positive2 %	0.0700	0.0516	0.0395
dem. pron %	0.3606	0.3533	0.4523	negative2 %	0.0328	0.0269	0.0327
noun morph #	0.8986	0.9023	0.9259	neg.emotive %	0.0027	0.0011	0.0008
freq. verb #	0.0833	0.0667	0.0333	content %	0.6806	0.7244	0.7290
modal verb #	1.9667	2.0667	1.4500	function %	0.2566	0.2145	0.2224
caus. verb #	0.2333	0.1333	0.2833	private verb #	4.4000	3.4833	4.1167
freq. verb %	0.0026	0.0017	0.0010	public verb #	1.4500	1.7167	2.8833
modal verb %	0.0546	0.0602	0.0343	suasive verb #	0.6667	0.8333	1.0333
caus. verb %	0.0074	0.0045	0.0070	private verb %	0.1132	0.0941	0.0822
subject #	17.6500	19.9833	28.2000	public verb %	0.0401	0.0524	0.0741
object #	14.1667	14.6333	16.8833	suasive verb %	0.0145	0.0264	0.0233
attributive #	49.3167	70.3333	105.8333	speech act #	4.2333	3.7667	4.9667
adverbial #	15.7167	15.2167	19.6333	quote #	1.9333	5.8667	4.8500
coordination #	16.3667	18.6000	25.9833	dash #	0.0333	0.5333	0.3000
subject %	0.9167	0.9507	0.9927	speech act %	0.0188	0.0147	0.0143
object %	0.7134	0.6927	0.6702	quote %	0.0084	0.0214	0.0109
attributive %	2.4684	3.2076	3.7817	dash %	0.0002	0.0023	0.0008
adverbial %	0.8044	0.6866	0.6747	discourse marker #	10.4833	9.4500	10.7833
coordination %	0.7947	0.8435	0.9329	discourse marker %	0.0456	0.0338	0.0256

Table 3: Mean values for features in each class. #: number, %: rate.

Feature groups	Acc	clickbait			fake news			real news			all		
		P	R	F	P	R	F	P	R	F	P	R	F
stat+morph+synt+sem+prag	79.44	86.9	88.3	87.6	78.2	71.7	74.8	73.4	78.3	75.8	79.5	79.4	79.4
stat+morph+synt+sem	80.56	86.9	88.3	87.6	78.9	75	76.9	75.8	78.3	77	80.5	80.6	80.5
stat+morph+synt+prag	79.44	85	85	85	80.7	76.7	78.6	73	76.7	74.8	79.6	79.4	79.5
stat+morph+sem+prag	**82.78**	88.3	88.3	88.3	**82.8**	**80**	**81.4**	77.4	80	78.7	**82.8**	**82.8**	**82.8**
stat+synt+sem+prag	75	83.8	**95**	**89.1**	67.9	63.3	65.5	71.4	66.7	69	74.4	75	74.5
morph+synt+sem+prag	81.11	85.2	86.7	86	80.4	75	77.6	**77.8**	81.7	79.7	81.1	81.1	81.1
stat+morph+synt	78.33	85	85	85	77.6	75	76.3	72.6	75	73.8	78.4	78.3	78.3
stat+morph+sem	81.11	85.5	88.3	86.9	82.7	71.7	76.8	75.8	83.3	79.4	81.3	81.1	81
stat+synt+sem	72.78	84.6	91.7	88	66	55	60	66.2	71.7	68.8	72.3	72.8	72.3
morph+synt+sem	80	86.7	86.7	86.7	80	73.3	76.5	73.8	80	76.8	80.2	80	80
stat+morph+prag	79.44	83.6	85	84.3	78.6	73.3	75.9	76.2	80	78	76.2	80	78
stat+synt+prag	67.22	77	78.3	77.7	67.9	63.3	65.5	57.1	60	58.5	67.3	67.2	67.2
morph+synt+prag	77.78	82.3	85	83.6	77.2	73.3	75.2	73.8	75	74.4	77.7	77.8	77.7
stat+sem+prag	75.56	84.4	90	87.1	68.5	61.7	64.9	72.6	75	73.8	75.2	75.6	75.3
morph+sem+prag	81.11	**88.1**	86.7	87.4	80	73.3	76.5	75.8	83.3	79.4	81.3	81.1	81.1
synt+sem+prag	71.67	81.5	88.3	84.8	62.5	58.3	60.3	69.5	68.3	68.9	71.2	71.7	71.4
stat+morph	78.89	87.7	83.3	85.5	78.6	73.3	75.9	71.6	80	75.6	79.3	78.9	79
stat+synt	65.56	77	78.3	77.7	60	55	57.4	59.4	63.3	61.3	65.5	65.6	65.5
stat+sem	73.33	83.6	85	84.3	62.1	60	61	73.8	75	74.4	73.1	73.3	73.2
stat+prag	69.44	77.2	73.3	75.2	70.5	71.7	71.1	61.3	63.3	62.3	69.7	69.4	69.5
morph+synt	78.89	82	83.3	82.6	79.7	78.3	79	75	75	75	78.9	78.9	78.9
morph+sem	81.11	84.4	90	87.1	83.7	68.3	75.2	76.1	**85**	**80.3**	81.4	81.1	80.9
morph+prag	78.89	80.6	83.3	82	84.6	73.3	78.6	72.7	80	76.2	79.3	78.9	78.9
synt+sem	72.22	85	85	85	60.9	65	62.9	71.4	66.7	69	72.5	72.2	72.3
synt+prag	66.11	76.3	75	75.6	59.7	61.7	60.7	62.7	61.7	62.2	66.2	66.1	66.2
sem+prag	71.11	80.3	81.7	81	59.3	58.3	58.8	73.3	73.3	73.3	71	71.1	71
stat	63.33	74.6	78.3	76.4	55.9	55	55.5	58.6	56.7	57.6	63.1	63.3	63.2
morph	80.56	86.2	83.3	84	82.1	76.7	79.3	74.2	81.7	77.8	80.9	80.6	80.6
synt	57.78	71.7	71.7	71.7	43.1	41.7	42.4	58.1	60	59	57.6	57.8	57.7
sem	65.56	78.3	78.3	78.3	54.1	55	54.5	64.4	63.3	63.9	65.6	65.6	65.6
prag	58.89	63.8	61.7	62.7	53.4	65	58.6	61.2	50	55	59.5	58.9	58.8

Table 4: Results of machine learning experiments. Acc: accuracy, P: precision, R: recall, F: F-measure. stat: statistical, morph: morphological, synt: syntactic, sem: semantic, prag: pragmatic.

As can be seen, our algorithm is notably more effective than the baseline method: even the least effective combination of features (i.e. syntactic features on their own) obtained an accuracy of 57.78%. The data also show that the system performs best when it comes to the detection of clickbait news: in the best case scenario, the algorithm properly identified 53 texts and only 2 texts were misclassified as fake news and 5 as real news. The performance was a bit weaker in the case of fake and real news. Overall, the method can be considered effective as it classified only 14 unreal news as real news from the 120 clickbait and fake news.

We also examined the efficiency of each feature set in our ablation experiments. As shown in Table 4, morphological features proved to be most effective, but semantic features also contributed to the success of the automatic detection. The results of the significance tests indicated that syntactic and pragmatic features had a somewhat weaker role in distinguishing the classes, which fact was also confirmed by our machine learning experiments.

6 Discussion of the results

Our results showed that our machine learning experiments achieved an accuracy of 82.78% in the classification task of real, clickbait and entertaining fake news. In this case we used all the feature groups with the exception of syntactic features. Here we discuss our results, with special regard to morphological and semantic features. Moreover, we also provide an error analysis.

6.1 The role of morphology

Examining the role of each feature set, our hypothesis is just partly confirmed: the findings do not show the outstanding role of semantic-pragmatic features in our machine learning experiments. Rather, it was morphology that proved to be the most effective.

In addition to the significant differences of frequency distributions of specific morphological features, morphological characteristics of the subcorpora examined played a decisive role in our machine learning experiments. For instance, when we applied morphological features of the subcorpora exclusively, we achieved an F-score of 80.6 that can be considered notably high compared to the best performance of our algorithm (82.8). At the same time, without these features the best performance we achieved was 75.3% (using statistical, semantic and pragmatic features).

In order to further examine the role of morphology, we divided the features into part-of-speech and deeper morphological features and reran the experiments with only parts-of-speech features and deep morphological features, respectively. This analysis showed that as long as the algorithm achieved 78% using part-of-speech features exclusively, the performance was only 58% when we used the deep morphological features, highlighting the outstanding role of POS tags in the task. Therefore, by using only POS-level information (i.e. keeping the feature set very simple), we can achieve an encouraging result for identifying Hungarian fabricated news.

To investigate this further, we analyzed the frequency distribution of parts-of-speech in a meticulous way. There is a significant difference of the frequency distribution of verbs and adverbs among all the subcorpora examined: there are significantly more verbs and adverbs in clickbait news. What is more, there is a significant difference of the noun, adjective and pronoun rates between clickbait and entertaining fake news, as well as clickbait and real news. More specifically, contributors of clickbait news use less nouns and adjectives and more pronouns than contributors of real news and entertaining fake news. From these results we can conclude that the sentence structure of the clickbait news is notably different from that of the real news: by using more verbal elements, clickbait news seem to highlight the events and happenings and pay less attention to the participants of the events (i.e. nominal elements). In other words, clickbait news emphasize what happened and the details behind the act (e.g. actors, objects etc.) are less noteworthy.

As for other morphological features of the subcorpora, authors of clickbait news use more imperative and less conditional verb forms than the other two text types. While the former feature may be considered as a sign of the need for action, the latter might be viewed as a sign of uncertainty. The results also showed that there is a higher frequency of past tense in real news than present tense compared to clickbait and entertaining fake news. In other words, real and fake news appear to concentrate on past events, i.e. they have a descriptive function, whereas clickbait news focus on the "here and now", representing a more "active" and "powerful" discourse, this enticing readers to click on them. It is also worth mentioning that there is a higher occurrence of first person plural verb forms in clickbait and in entertaining fake news than in real news. This feature may be considered as a linguistic strategy for manipulation since in the unreal news, it may evoke a shared feeling of common ground, attempting to deceive the reader (Bártházi, 2008).

6.2 The role of semantics

As for frequency distribution of certain types of semantic contents there is a significantly higher frequency of uncertainty markers such as condition, weasel, peacock and hedge in clickbait news compared to the fake and real news. At the same time, the frequency difference between fake and real news is not significant in this case. The latter feature shows that fake news tend to present information as factual statements, thereby increasing the apparent authenticity and credibility of the content.

With regard to the sentiment and emotional content of the texts, there is a significantly higher frequency of positive sentiment rate in clickbait news. What is more, there is a significantly higher frequency of "love" and "joy" in clickbait news compared to the other two text types. Data also show that these emotions are more frequent in fake news than in real news as well. Based on these results we may conclude that positive attitude characterizes unreliable news more than real news. These results correlate with a previous research finding about the linguistic features of Hungarian communist propaganda texts (Vincze et al., 2019), where it was also proved that propaganda texts bound in positive emotions.

However, we also should mention that there are more words representing anxiety and fear in clickbait news than real and fake news as well. This might reflect the emphatic role of emotions in clickbait news,

which is probably related to the general purpose of these texts, i.e. the more emotional a text is, the more probably it will generate clicks.

6.3 Error analysis

Below we present some examples of news that are misclassified by the system.

Real news classified as fake news:

- *Golden State Warriors are reluctant to celebrate their latest title at the White House*[3]
- *180-year old postcard making company wound up*[4]

Fake news classified as real news:

- *Formula-1 in Szeged, Hungary*[5]
- *Fake Grabovoi numbers appeared*[6]

Clickbait news classified as real news:

- *Scary: Nostradamus's prophecies for 2019*[7]
- *10 shocking photos without an explanation*[8]

Analyzing the incorrectly classified documents, it was revealed that short real news were often misclassified as real news tend to be longer than fake news. On the other hand, clickbait news that contained a lower rate of imperatives and/or a higher rate of conditionals were more prone to be classified as real news. Finally, fake news with lots of negative words were often misclassified as real news.

7 Conclusions and future work

In this work, we reported on the automatic discrimination between Hungarian real news, clickbait news and fake news written for entertaining purposes. Our results confirm that it is possible to successfully detect untrustful news based on a rich – morphological, syntactic, semantic and pragmatic – feature set, and especially morphological features play an important role in the process. Besides morphological features, the added value of semantic features was also apparent, while at the same time, syntactic features did not have a notable effect on our results. Our experiments show the potential for exploiting the morphological, more specifically part-of-speech features for fake news detection.

As a next step of the research, on the basis of the findings of the recent analysis, we would like to apply our methods and tools on texts belonging to other domains. We will attempt to further train our algorithm to discriminate texts containing propagandistic features, bias, misinformation or disinformation. Moreover, we would like to compare these results to previous findings concerning the linguistic features of Hungarian Communist propaganda texts (Vincze et al., 2019). Finally, we would like to compare our results obtained on Hungarian texts with those on English corpora (Barrón-Cedeño et al., 2019) and possibly other languages as well.

Acknowledgements

This work was supported by grantTUDFO/47138-1/2019-ITM of the Ministry for Innovation and Technology, Hungary and by the Hungarian Artificial Intelligence National Laboratory.

[3] http://index.hu/sport/kosarlabda/2017/09/24/donald_trump_sertodes_golden_state_warriors_steph_curry_lebron_james_kobe_bryant/

[4] http://www.origo.hu/gazdasag/20170927-180-eves-kepeslap-keszito-csaladi-vallalkozas-szunik-meg.html

[5] https://szegedpanorama.blogspot.hu/2013/04/forma-1-es-verseny-szegeden.html

[6] http://hircsarda.hu/2015/02/19/mar_hamisitjak_a_grabovoj-szamokat/

[7] http://eztnezdmeg.com/hatborzongato-nostradamus-szerint-ez-var-rank-2019-ben/

[8] http://www.hirvarazs.info/altalanos/tiz-dobbenetes-foto-amire-sose-talaltak-magyarazatot/

References

Alberto Barrón-Cedeño, Israa Jaradat, Giovanni Da San Martino, and Preslav Nakov. 2019. Proppy: Organizing the news based on their propagandistic content. *Information Processing Management*, 56(5):1849–1864.

Eszter Bártházi. 2008. Manipuláció, valamint manipulációra alkalmas nyelvhasználati eszközök a sajtóreklámokban. *Magyar Nyelv*, 104(4):443–463.

P. Biyani, K. Tsioutsiouliklis, and John Blackmer. 2016. "8 amazing secrets for getting more clicks": Detecting clickbaits in news streams using article informality. In *AAAI*.

Leo Breiman. 2001. Random forests. *Machine Learning*, 45(1):5–32, October.

Eunsol Choi, Chenhao Tan, Lillian Lee, Cristian Danescu-Niculescu-Mizil, and Jennifer Spindel. 2012. Hedge detection as a lens on framing in the gmo debates: A position paper. In *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics*, pages 70–79. Association for Computational Linguistics.

Csilla Ilona Dér and Alexandra Markó. 2007. A magyar diskurzusjelölők szupraszegmentális jelöltsége. In *Nyelvelmélet–nyelvhasználat*, pages 61–67. Tinta, Székesfehérvár–Budapest.

William Ferreira and Andreas Vlachos. 2016. Emergent: a novel data-set for stance classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1163–1168, San Diego, California, June. Association for Computational Linguistics.

Anita Fetzer. 2008. "And I Think That Is a Very Straightforward Way of Dealing With It"– The Communicative Function of Cognitive Verbs in Political Discourse. *Journal of Language and Social Psychology*, 27:384–396, 12.

Bruce Fraser. 2010. Chapter 11. hedging in political discourse. In *Perspectives in Politics and Discourse*, pages 201–214. 01.

Codruta Girlea, Roxana Girju, and Eyal Amir. 2016. Psycholinguistic features for deceptive role detection in werewolf. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 417–422.

Lucas Graves. 2018. Understanding the promise and limits of automated fact-checking. Technical report.

Stephan Greene and Philip Resnik. 2009. More than words: Syntactic packaging and implicit sentiment. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 503–511, Boulder, Colorado, June. Association for Computational Linguistics.

Georgi Karadzhov, Pepa Gencheva, Preslav Nakov, and Ivan Koychev. 2017. We built a fake news / click bait filter: What happened next will blow your mind! *ArXiv*, abs/1803.03786.

Zoltán Kmetty, Veronika Vincze, Dorottya Demszky, Orsolya Ring, Balázs Nagy, and Martina Katalin Szabó. 2020. Pártélet: A hungarian corpus of propaganda texts from the hungarian socialist era. In *Proceedings of The 12th Language Resources and Evaluation Conference*, pages 2381–2388.

Bing Liu. 2012. *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.

Rada Mihalcea and Carlo Strapparava. 2009. The lie detector: Explorations in the automatic recognition of deceptive language. In *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, pages 309–312. Association for Computational Linguistics.

Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. 2018. Automatic detection of fake news. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3391–3401, Santa Fe, New Mexico, USA, August. Association for Computational Linguistics.

Institute for Propaganda Analysis. 1938. How to Detect Propaganda. In *Propaganda Analysis. Publications of the Institute for Propaganda Analysis*, volume I, pages 210–218.

Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman, London.

Hannah Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova, and Yejin Choi. 2017. Truth of varying shades: Analyzing language in fake news and political fact-checking. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2931–2937, Copenhagen, Denmark, September. Association for Computational Linguistics.

Victoria L. Rubin, Yimin Chen, and Niall J. Conroy. 2015. Deception detection for news: Three types of fakes. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community*, ASIST '15, pages 83:1–83:4, Silver Springs, MD, USA. American Society for Information Science.

Victoria Rubin, Niall Conroy, Yimin Chen, and Sarah Cornwell. 2016. Fake news or truth? using satirical cues to detect potentially misleading news. In *Proceedings of the Second Workshop on Computational Approaches to Deception Detection*, pages 7–17, San Diego, California, June. Association for Computational Linguistics.

Roney Santos, Gabriela Pedro, Sidney Leal, Oto Vale, Thiago Pardo, Kalina Bontcheva, and Carolina Scarton. 2020. Measuring the impact of readability features in fake news detection. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 1404–1413, Marseille, France, May. European Language Resources Association.

Rut Scheithauer. 2007. Metaphors in election night television coverage in Britain, the United States and Germany. In *Political Discourse in the Media: Cross-cultural perspectives*, pages 75–106. 01.

Anne-Marie Simon-Vandenbergen, Peter White, and Karin Aijmer. 2007. Presupposition and 'taking-for-granted' in mass communicated political argument An illustration from British, Flemish and Swedish political colloquy. In *Political Discourse in the Media*, pages 31–74. 01.

Martina Katalin Szabó, Veronika Vincze, and Gergely Morvay. 2016. Magyar nyelvű szövegek emócióelemzésének elméleti nyelvészeti és nyelvtechnológiai problémái. In *Távlatok a mai magyar alkalmazott nyelvészetben*. Tinta, Budapest.

Martina Katalin Szabó, Orsolya Ring, Balázs Nagy, László Kiss, Júlia Koltai, Gábor Berend, László Vidács, Attila Gulyás, and Zoltán Kmetty. 2020. Exploring the dynamic changes of key concepts of the hungarian socialist era with natural language processing methods. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, pages 1–13.

Martina Katalin Szabó. 2015. Egy magyar nyelvű szentimentlexikon létrehozásának tapasztalatai és dilemmái. In *Segédkönyvek a nyelvészet tanulmányozásához 177*, pages 278–285. Tinta, Budapest.

György Szarvas, Veronika Vincze, Richárd Farkas, György Móra, and Iryna Gurevych. 2012. Cross-Genre and Cross-Domain Detection of Semantic Uncertainty. *Computational Linguistics*, 38:335–367, June.

James Thorne and Andreas Vlachos. 2018. Automated fact checking: Task formulations, methods and future directions. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3346–3359, Santa Fe, New Mexico, USA, August. Association for Computational Linguistics.

James Thorne, Andreas Vlachos, Christos Christodoulopoulos, and Arpit Mittal. 2018. Fever: a large-scale dataset for fact extraction and verification. *arXiv preprint arXiv:1803.05355*.

Veronika Vincze, Martina Katalin Szabó, and Orsolya Ring. 2019. Automatic analysis of linguistic features in Communist propaganda texts. `https://propaganda.qcri.org/bias-misinformation-workshop-socinfo19/paper4_final_Vincze_et_al_propaganda.pdf`.

Veronika Vincze. 2013. Weasels, hedges and peacocks: Discourse-level uncertainty in wikipedia articles. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, pages 383–391.

Veronika Vincze. 2014a. Uncertainty detection in Hungarian texts. In *Proceedings of Coling 2014*.

Veronika Vincze. 2014b. *Uncertainty Detection in Natural Language Texts*. Ph.D. thesis, University of Szeged, Szeged, Hungary.

Andreas Vlachos and Sebastian Riedel. 2014. Fact checking: Task definition and dataset construction. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*, pages 18–22, Baltimore, MD, USA, June. Association for Computational Linguistics.

William Yang Wang. 2017. "Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 422–426, Vancouver, Canada, July. Association for Computational Linguistics.

Lina Zhou, Judee K Burgoon, Jay F Nunamaker, and Doug Twitchell. 2004. Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communications. *Group decision and negotiation*, 13(1):81–106.

János Zsibrita, Veronika Vincze, and Richárd Farkas. 2013. magyarlanc: A toolkit for morphological and dependency parsing of Hungarian. In *Proceedings of RANLP*, pages 763–771.