Automatic Detection of the Bulgarian Evidential Renarrative

Irina Temnikova
Big Data for Smart Society Institute (GATE)
5, James Boucher St., Sofia, 1164, Bulgaria
irina.temnikova@gate-ai.eu

Ruslana Margova
Big Data for Smart Society Institute (GATE)
5, James Boucher St., Sofia, 1164, Bulgaria
ruslana.margova@gate-ai.eu

Stefan Minkov
Big Data for Smart Society Institute (GATE)
5, James Boucher St., Sofia, 1164, Bulgaria
stefan.minkov@gate-ai.eu

Tsvetelina Stefanova
Big Data for Smart Society Institute (GATE)
5, James Boucher St., Sofia, 1164, Bulgaria
tsvetelina.stefanova@gate-ai.eu

Nevena Grigorova
Big Data for Smart Society Institute (GATE)
5, James Boucher St., Sofia, 1164, Bulgaria
freelancer.neva@gmail.com

Silvia Gargova
Big Data for Smart Society Institute (GATE)
5, James Boucher St., Sofia, 1164, Bulgaria
silvia.gargova@gate-ai.eu

Venelin Kovatchev
School of Computer Science, University of Birmingham
University Rd W, Edgbaston, Birmingham, United Kingdom
v.o.kovatchev@bham.ac.uk

Manual and automatic verification of the trustworthiness of information is an important task. Knowing whether the author of a statement was an eyewitness to the reported event(s) is a useful clue. In linguistics, such information is expressed through "evidentiality". Evidentials are especially important in Bulgarian, as Bulgarian journalists often use a specific type of evidential ("renarrative") to report events that they did not directly observe, nor verify. Unfortunately, there are no automatic tools to detect Bulgarian renarrative. This article presents the first two automatic solutions for this task. Specifically - a fine-tuned BERT classifier (renarrative BERT detector, BGRenBERT), achieving 0.98 Accuracy on the test split, and a renarrative rule-based detector (BGRenRules), created with regular expressions, matching a parser's output.

https://doi.org/10.47810/JCLIB.1.2025.04

Both solutions detect Bulgarian texts containing the most frequently encountered forms of renarrative. Additionally, we compare the results of the two detectors with the manual annotation of subsets of two Bulgarian fake text datasets. BGRenRules obtains substantially higher results than BGRenBERT. The error analysis shows that the errors from BGRenRules most frequently correspond to cases in which humans also have doubts. The training dataset (BgRenData), the annotated dataset subsets, and the two detectors are made publicly accessible on Zenodo, GitHub, and HuggingFace. We expect that these new resources will be of invaluable assistance to 1) Bulgarian-language researchers, 2) researchers of other languages with similar phenomena, especially those working on verifying information.

Keywords: *evidentiality, Bulgarian, renarrative, fine-tuned BERT classifier, Python, annotation*

1. Introduction

Verifying the trustworthiness of information and automatically detecting factually incorrect information has become a topic that in the past few years attracted more attention (Guo et al. 2020; Shu et al. 2017; Das et al. 2023). We define "factually incorrect information" as information which contradicts the facts. It is considered that there are different types of factually incorrect information. "Misinformation" refers to factually incorrect information that is not intended to cause harm, while "disinformation" is factually incorrect information that is spread with the intention to deceive, and to cause harm[1]; "malinformation" is called information that stems from the truth, but is often exaggerated in a misleading way (Newman 2021; Wardle et al. 2018). In this article, we address all these three types of factually incorrect information and focus on the information's trustworthiness or reliability.

There are various automatic methods for verifying the trustworthiness of textual information. Depending on where it appears (for example in news media or social media), it may be verified by one or a combination of more than one from the following methods: using specific linguistic features (Dinkov, Koychev, and Nakov 2019; Atanasova et al. 2019; Zhou and Zafarani 2020), matching statements to databases of fact-checked claims (Panchendrarajan and Zubiaga 2024; Hangloo and Arora 2023), using social network-specific features (Rani, Das, and Bhardwaj 2022; Santhosh, Cheriyan, and Nair 2022) such as the popularity of the author, the number of likes, reposts, and comments.

Knowing whether the author of a news article or a social media post has been a witness to the information shared in the text, can help to verify the trustworthiness of the reported information (Grieve and Woodfield 2023). This knowledge is usually expressed linguistically by "evidentiality". The linguistic expressions of evidentiality are called "evidentials". Evidentials have been already used in natural language processing (NLP) for text trustworthiness detection (Su, Huang, and Chen 2010; Su, Chen, and Huang 2011) for other languages, but not for Bulgarian. The evidentials in Bulgarian language

[1] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2020%3A790%3AFIN&qid=1607079662423

are four - indicative, conclusive, renarrative, and dubitative. When using renarrative, the speaker reports something which he did not witness, without assessing whether the source is reliable or not (Nitzolova 2008).

Margova (2023) found that the form of Bulgarian evidential, called "ренаратив" (in English "renarrative") is most frequently used by Bulgarian journalists in one of two situations: 1) when journalists are unsure of the reliability of the information and want to put distance between themselves and their statement; and 2) in news articles that have been flagged as misleading by independent fact-checkers. In fact Renarrative is used in Bulgarian when the speaker/author wants to transmit information provided by a third party without expressing the knowledge and views of the speaker/author. In this way, the speaker/author does not express doubt, nor guarantee the truthfulness of the reported information. Additionally, he/she shows that the responsibility for the reliability of the reported information belongs to its source and depends on the interpretation of the recipient of the information (Nitzolova 2008; Margova 2023).

Due to these useful characteristics, the automatic detection of the Bulgarian "renarrative" can be an important additional linguistic cue for verifying the trustworthiness of information. We hypothesized that since renarrative is expressed morphologically through specific verb forms, it should be detectable by syntactic parsers or Part-of-Speech (POS) taggers. For this reason, we reviewed several publicly available Bulgarian Natural Language Processing (NLP) tools (see Section 2 for details). Unfortunately, we discovered that none of them was explicitly designed for detecting evidentials.

In this article, we propose the first automatic solutions to detect renarrative in Bulgarian texts. Our solutions include a fine-tuned BERT classifier (the renarrative BERT detector, BGRenBERT) and a renarrative rule-based detector (BGRenRules). Both detectors recognize when a short text or a sentence contains at least one form of renarrative. As the forms of renarrative often match the forms of other Bulgarian verb forms, our solutions detect only the renarrative forms which are pointed by an expert to be the most frequent ones, and also those which are the easiest to be automatically distinguished. The article describes the methods followed for fine-tuning the model and adapting the rule-based detector to reach the best results. We also report the performance results and an error analysis of comparing the performance of the two detectors on manually annotated subsets of two Bulgarian fake text datasets (Hardalov, Koychev, and Nakov 2016; Nakov et al. 2022).

In summary, our contributions are the following:

- a new machine learning (ML) text classifier and a rule-based detector for detecting texts with at least one of the most frequent forms of Bulgarian renarrative;
- a new text dataset used for fine-tuning BGRenBERT and adjusting BGRenRules to recognize the correct renarrative forms (BGRenData);
- manually annotated subsets of two Bulgarian fake texts datasets;
- linguistic insights about the presence of renarrative in these datasets.

All these resources are publically shared on Zenodo, GitHub, and HuggingFace to increase their visibility to the research community. We believe that our work will be very useful to Bulgarian linguists and other researchers working on similar topics.

The next sections include: Section 2 presents the linguistic theory about evidentiality and relevant related work, Section 3 explains which forms of renarrative we detect, Section 4 describe all the datasets used, Section 5 provides details about the two detectors, Section 6 contains the results of the comparison of the two methods on the manually annotated dataset subsets and an error analysis. Section 9 contains our planned future work and the conclusions, Section 7 discusses the ethical and legal considerations, Section 8 presents this work's limitations, and finally, Section 10 lists the authors' contributions and the acknowledgements.

2. Linguistic Theory and Related Work

In this section, we provide a short theoretical overview of evidentiality and present the work most relevant to ours. We consider the works most similar to ours those which automatically detect Bulgarian evidentials, and those which use evidentials for verifying texts' trustworthiness.

2.1 Evidentiality

Evidentiality is the grammatical expression of the information source, and usually assists in distinguishing whether the speaker witnessed the reported information, was told about it by somebody else, or, for example, inferred it based on common sense (Aikhenvald 2004, 2015). Evidentiality can express **the source of the knowledge** and **the subjective certainty of the speaker about the truthfulness of the statement** (Ifantidou 2001; Mushin 2000). The following two examples taken from (Su, Huang, and Chen 2010) show some evidentiality differences in English:

- *It must be raining.*
- *I can see it raining.*

According to researchers of this phenomenon (Aikhenvald 2015; Jackobson 1957), evidentiality exists in every language, but it can be expressed differently (for example morphologically or lexically) (Su, Huang, and Chen 2010).

2.2 Evidentiality in Bulgarian

In Bulgarian (Nitzolova 2008) evidentiality expresses the four combinations of the presence or absence of two aspects:

- a subjective view toward the reported information;
- transmitting somebody else's information.

The Bulgarian evidential system includes the following evidentials:

- **Indicative** (the main evidential - when the speaker is a witness of what he or she reports as information);
- **Conclusive** (when the speaker makes conclusions, based on the information);
- **Renarrative** (when the speaker re-transmits the information without saying if it is true or false);
- **Dubitative** (when the speaker expresses doubt about the information);
- **Admirative** (when the speaker admires the information) (Karagjosova 2021). *Some authors don't see admirative as an evidential, but here we accept it as a part of the evidential system in Bulgarian.

Bulgarian evidentials are expressed morphologically through the forms of verb tenses. Bulgarian language has nine tenses. See below examples of the verb "пиша", with their translations into English ("to write"):

Present tense (praesens) "пиша" ("I am writing, I write"), aorist "аз писах" (I wrote), imperfect past tense (imperfectum) "пишеше" ("I was writing"), perfect (perfectum) "писал съм" ("I wrote"), pluperfect (plusquamperfectum) "писал бях" ("I have written"), future tense (futurum) "ще пише" ("I will write"), future perfect tense (futurum exactum) "ще е/бъде писал", ("will have written" and "will have been writing"), future in the past tense (futurum praeteriti) "щях да пиша" ("I was going to write", "I would have written"), future perfect in the past (futurum exactum praeteriti) "щях да съм / да бъда писал" ("I would have written", "I would have been writing", and "I would have had written").

The most important feature of the indicative in modern Bulgarian is the fact that the past tenses (aorist, imperfectum, plusquamperfectum, futurum praeteriti, futurum exactum praeteriti) implicitly indicate that the speaker witnessed the reported event(s). These tenses in indicative can only be used if the speaker is a witness or if the speaker presents himself as a witness (Aleksova 2024).

2.3 Automatic Methods

It was already shown that **evidentiality improves automatic text trustworthiness detection** when used as additional features to an English-language machine learning model (Su, Huang, and Chen 2010).

According to our knowledge, **there are no works which automatically detect Bulgarian evidentials**. As already mentioned in Section 1, we ran an extensive analysis of automatic tools for Bulgarian and none of them was detecting evidentials. The summary of our analysis can be seen in Table 1.

Table 1: Overview of Existing Bulgarian POS Taggers and Parsers and their Support of Evidentials

Name	Supports Evidentials	Comments
SketchEngine (Kilgarriff et al. 2014)	No	Evidentials not included in the tagset
AMontgomerie / bulgarian-nlp (Montgomery 2023)	No	Relying on Universal Dependencies, only Part-of-Speech (POS), no grammatical tags, but supports Named Entity Recognition (NER), sentiment analysis, etc.
polyglot (Al-Rfou 2015)	No	Relying on Universal Dependencies, only POS, no grammatical tags, but supports NER
UDPipe, Universal Dependencies 2.15, Deep Universal Dependencies 2.8, Universal Dependencies 2.5 Models for UDPipe, Universal Dependencies 2.4 Models for UDPipe (Straka 2018)	No	`https://universaldependencies.org/ext-feat-index.html`, `https://ufal.mff.cuni.cz/udpipe`
iarfmoose/roberta-small-bulgarian-pos (Montgomerie 2021b)	No	Relying on Universal Dependencies
iarfmoose/roberta-base-bulgarian-pos (Montgomerie 2021a)	No	Relying on Universal Dependencies
CLASSLA Fork of the Official Stanford NLP Python Library for Many Human Languages (CLASSLA 2021; Ljubešić, Terčon, and Dobrovoljc 2024)	No	Relying on Universal Dependencies
GATE: Universal Dependencies POS Tagger for bg / Bulgarian (Roberts 2020)	No	Relying on Universal Dependencies
LIMA - Libre Multilingual Analyzer (LIMA 2021)	No	Not supporting morphological features

Continued on next page

Table 1 – *Continued from previous page*

Name	Supports Evidentials	Comments
NLP Cube (Boroş, Dumitrescu, and Burtica 2018)	No	Relying on Universal Dependencies
NooJ (Silberztein 2005)	No	Evidentials not included in the tagset
RNNTagger (Schmid 2019)	No	Only Linux supported, only POS tagger
spaCy (Honnibal et al. 2020)	No	No information in the documentation about Bulgarian, but claimed support of Macedonian
Sparv (Borin et al. 2016)	No	TreeTagger integrated for Bulgarian (see row 18) (relying on Universal Dependencies), working with Stanza for POS and lemmatization, best for Swedish and English
Stanza (Qi et al. 2020)	No	Relying on UniversalDependencies
Text Tonsorium (Jongejan 2020)	No	Relying on Universal Dependencies
TreeTagger - a part-of-speech tagger for many languages (Schmid 1994)	No	Based on the BulTreeBank tagset, only finite indicative recognized (`http://bultreebank.org/wp-content/uploads/2017/06/BTB-TR03.pdf`)
UniMorph - Schema and datasets for universal morphological annotation (Sylak-Glassman 2016)	No information[2]	Supporting over 212 features, including evidentiality, need to check for Bulgarian, might rely on UniversalDependencies
melaniab / spacy-pipeline-bg (Berbatova and Ivanov 2023)	No	Supports morphological features, relying on UniversalDependencies

Continued on next page

[2] There is no information in the documentation whether UniMorph supports evidentials for Bulgarian language

Table 1 – *Continued from previous page*

Name	Supports Evidentials	Comments
Bulgarian NLP pipeline in CLaRK System (BTB-Pipe) (BulTreeBank)	No	Based on the BulTreeBank tagset, only finite indicative recognized (`http://bultreebank.org/wp-content/uploads/2017/06/BTB-TR03.pdf`)
The CLASSLA-Stanza model for morphosyntactic annotation of standard Bulgarian 2.1 (Terčon et al. 2023)	No	A version of Classla from 2023 based on the MULTEXT-East tagset
wietsedv/xlm-roberta-base-ft-udpos28-bg (de Vries 2021; de Vries, Wieling, and Nissim 2022)	No	Relying on UniversalDependencies
KoichiYasuoka/bert-base-slavic-cyrillic-upos (Yasuoka 2021a)	No	Relying on UniversalDependencies
KoichiYasuoka/bert-large-slavic-cyrillic-upos (Yasuoka 2021b)	No	Relying on UniversalDependencies
rmihaylov/bert-base-pos-theseus-bg (Mihaylov 2023)	No	Only POS tagger
bgnlp (Fauzi 2021)	No information[3]	Supports lemmatization, NER, POS and morphological features tagging, keyword extraction and commatization; the POS and morphological features model is trained on the Wiki1000+ dataset
AzBuki.ml (Cholakov)	No	Evidentiality not included in the tagset

[3] There is no information in the documentation whether bgnlp supports evidentials for Bulgarian language.

3. Detected Forms of Bulgarian Renarrative

The renarrative in Bulgarian has the peculiarity of having homonymous forms with other evidentials and tenses. Due to this, its recognition often depends only on the context and the interpretation. The present and imperfect forms of renarrative match the forms of the conclusive imperfect, with a difference in the third person. The forms of present and imperfect have the same form as a specific form of the indicative perfect - when the third person form of the auxiliary verb *be* is omitted. The forms of the dubitative aorist are homonymous with renarrative perfectum/plusquamperfectum. The present tense form of the admirative is the same as the present form of the renarrative.

Due to this homonymy, **we focus on the forms of renarrative which can be recognised more easily**, namely those of **the third-person singular and the third-person plural**. To achieve that, some of our datasets are from journalistic headlines where renarrative is usually in the third person. We did not work with the forms of the auxiliary verb *be* (in Bulgarian съм), which has high homonymy with other evidentials.

Table 2 shows the forms which we detect for the Bulgarian verb "нося" (English translation: "to bring, to carry").

Tense	**Forms**
Present/Imperfect	носел,-а,-о **съм** носел,-а,-о [empty]
Perfect/Pluperfect	бил **съм** носил,-а,-о бил [empty] носил,-а,-о
Aorist	носил,-а,-о,**съм** [empty] носил,-а,-о
Renarrative futurum/futurum praeteriti	**щял,-а,-о съм** да нося щял,-а,-о [empty] да носи
Futurum exactum/Futurum exactum praeteriti	щял,-а,-о съм да **съм** носил,-а,-о щял,-а,-о да е носил,-а,-о

Table 2
Automatically detected forms of renarrative (examples, taken from Nitzolova (2008).)

Finally, we made sure that the forms which we aimed to detect are important in the Bulgarian world of news media. We assured this by consulting one of this article's co-authors, who has a long experience in Bulgarian news media.

4. Datasets Used

In this section, we first (in Section 4.1) introduce the text dataset which we built for training BGRenBERT and testing and modifying the BGRenRules. Next, we describe the datasets (Section 4.2) on which we compared the results of the two methods with manual annotation.

4.1 BgRenData - Dataset Used for Preparing the Automatic Solutions

As our work is connected to detecting factually incorrect information, we focus only on the two most frequently considered types of texts in this domain – namely, news articles and social media texts.

For training BGRenBERT and testing and modifying the BGRenRules, we have created a special dataset (BgRenData) of 2891 short texts, half of which contained renarrative, and half did not contain renarrative. The texts were a mixture of news article titles, short social media posts, and ChatGPT-generated sentences. Table 3 shows the contents of this dataset. The news article titles came from two sources: a random selection of funny/fake titles with a variety of topics from Hardalov (Hardalov, Koychev, and Nakov 2016)'s Credible News dataset; and a compilation of news article titles with renarrative taken from (Margova 2023). Originally, the Credible News dataset contained credible and fake news. We selected only the fake news subsets for our analysis. Specifically, they came from:

1. The Bulgarian website with humorous news Ne!Novinite[4] (translation into English: "No!News"), containing topics such as politics, sports, culture, world news, horoscopes, interviews, and user-written articles;
2. The blog website Bazikileaks[5], containing fictitious blogs;
3. The Bulgarian news media bTV humorous (Duplex) Lifestyle subsection[6].

The social media posts were also a random selection of posts with various topics from Temnikova et al. (2023)'s datasets. These datasets were originally selected to contain keywords related to manipulation, lies, and similar topics.

One of our linguists used ChatGPT 3.5 to generate a selection of sentences, containing 3rd person singular and plural forms of renarrative from those in Table 2. We restricted the generation to 3rd person singular and plural because these are the most frequent forms of renarrative in news, according to our colleague Ruslana Margova.

Each text in the dataset was selected in a way to contain at least one form of renarrative from those in Table 2. We did not count their numbers per text, as the task of both our solutions was only to identify the texts which contained at least one renarrative form. BGRenData is available in Zenodo[7], Github[8], and HuggingFace[9] with the license Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0), due to containing tweets and as required by Temnikova's datasets (Temnikova et al. 2023).

[4] https://www.nenovinite.com/
[5] https://neverojatno.wordpress.com/
[6] https://www.btv.bg/lifestyle/all/
[7] https://zenodo.org/records/15871397
[8] https://github.com/silviavg/bg-renarrative
[9] https://huggingface.co/datasets/gate-institute/BGRenData

Class	Sources and Counts					
	All Sources	News Articles Titles		ChatGPT		Social Media
	Counts	Type	Counts	Type	Counts	Counts
Renarr.	1445	Margova (2023)	377	All	998	70
				Sg.	843	
				Pl.	155	
No Renarr.	1446	Hardalov, Koychev, and Nakov (2016)	375	All	1036	35
				Sg.	845	
				Pl.	191	
Both	**2891**					

Table 3
BGRenData - the dataset used for training and testing BGRenBERT and BGRenRules.

4.2 Analyzed Existing Datasets

To compare BGRenBERT with BGRenRules, we used subsets of two publicly accessible datasets with Bulgarian texts, used in NLP methods for fake news detection. The datasets contained the two types of texts of our interest: news article titles and Twitter posts. Specifically, we used the fake news titles subsets of Credible News (Hardalov, Koychev, and Nakov 2016) (already introduced in Section 4.1) and the Twitter posts dataset, provided for the 2022 Conference and Labs of the Evaluation Forum (CLEF) CHECK-THAT! Lab[10], TASK 1D. CLEF 2022 Check-That! Lab Task 1 (Nakov et al. 2022) included Bulgarian and required to predict which Twitter posts are worth fact-checking, with topics related to COVID-19 and politics. We decided to include in the analysis only the tweets from Task 1D “Attention-worthy tweet detection”: Given a tweet, predict whether it should get the attention of policymakers and why. The tweets from Task 1D were annotated with 5 classes: *harmful* (333 tweets), *yes_discusses_cure* (79 tweets), *yes_blame_authorities* (51 tweets), *yes_discusses_action_taken* (25 tweets), and *no_not_interesting* (3186 tweets). To limit manual annotation efforts, we excluded the last category. We refer to the obtained in this way dataset as ***CT1D***.

Before starting manual annotation, we first preprocessed Hardalov’s datasets by:

- removing the titles, used in the training dataset, described in Section 4.1.
- after removing the training titles we were left with only 6 titles from bTV Lifestyle. As they were too few for a meaningful comparison with the Bazikileaks and Ne!Novinite, we removed all bTV titles.
- leaving only one example from two sets of titles, which were almost completely identical, but differing by a date and/or a person’s name. See Examples 1 and 2 below:

[10] https://checkthat.gitlab.io/clef2022/

Example 1 (Highly similar titles 1)
“Петъчен оптимизъм с Росен Плевнелиев - 30.08”;
“Петъчен оптимизъм с Александър Томов - 04.04”
Translations into English:
“Friday optimism with Rosen Plevnevliev - 30.08”;
“Friday optimism with Alexander Tomov - 04.04”

Example 2 (Highly similar titles 2)
“Седмичен не!хороскоп: 21.01-27.01”;
“Седмичен не!хороскоп: 07.01-13.01”
Translations into English:
“Weekly no-horoscope: 21.01-27.01”;
“Weekly no-horoscope: 07.01-13.01”

We refer to the resulting dataset as ***CNClean***, to its Ne!Novinite’s subset as ***CNCleanN*** and to the Bazikileaks subset as ***CNCleanB***.

Next, we manually annotated all the datasets for the presence of the forms of renarrative in Table 2. Annotation was done by two Bulgarian linguists (co-authors of this article). The annotators had to assign the category “renarrative” if the text contained at least one of the forms in Table 2, and “No renarrative” if it did not contain any of these forms.

The cases in which they did not agree or had doubts about were resolved in a discussion with a third Bulgarian linguist (also co-author of this article).

Table 4 summarizes the number of texts per dataset and how many of them were manually annotated as *Renarrative* and *No renarrative*.

As it can be seen, CT1D contains fewer tweets with renarrative (17 texts or 3.48% from 488 texts) than the CNClean (447 texts, or 8.55% from 5230 texts in total). Specifically, CNCleanN contain the highest number of titles with renarratives (438 texts or 9.57% from the total of 4579 texts). The class “yes discusses action taken” from CT1D, instead, contains 0 texts with renarratives.

Datasets	**Subsets**	**Classes and Counts**		
		Ren.	No Ren.	All
CNClean	CNCleanB	9	642	651
	CNCleanN	438	4141	4579
	all the above	447	4783	5230
CT1D	“harmful”	9	324	333
	“yes discusses cure”	2	77	79
	“yes blames authorities”	6	45	51
	“yes discusses action taken”	0	25	25
	all the above	17	471	488

Table 4
Counts of all items in the CNCleanB, CNCleanN, and CT1D datasets, including counts of the texts annotated as containing “renarrative” and “no renarrative”.

Class	Acc.	Recall	Precision	F1-Score
Dev Split Results				
All	0.98			
Renarr.		0.99	0.97	0.98
No Renarr.		0.97	0.99	0.98
Test Split Results				
All	0.98			
Renarr.		1.00	0.97	0.98
No Renarr.		0.97	1.00	0.98

Table 5
Results of BGRenBERT on BGRenData splits.

CNCleanB, CNCleanN, and CT1D are available in Zenodo[11]. The datasets are shared with the same license as the original Hardalov's datasets (Hardalov, Koychev, and Nakov 2016), and due to the fact that they contain tweets (Creative Commons Attribution-NonCommercial 4.0 International – CC BY-NC 4.0).

5. Automatic Detection Methods

5.1 BGRenBERT

To create BGRenBERT, we fine-tuned BERT-WEB-BG-CASED[12] with the BGRenData dataset, described in Section 4.1. We split BGRenData into train, dev, and test sections with the following proportions: 80, 10, 10. The number of training epochs was 5. For the rest of the hyperparameters, we used their default values. We made sure that all targeted forms appeared in all the splits, however they did not have equal distributions. All three splits contained 50% of texts with one or more forms of renarrative and 50% of texts with no renarratives.

Table 5 shows the results of the classifier. As the dataset is balanced for the presence of renarrative or not, but not balanced regarding the number of forms, we report both Accuracy and F1-scores.

5.2 BGRenRules

We compared BGRenBERT with a BGRenRules which used regular expressions on the top of *Classla*'s output (BGRenRules). The regular expressions covered all renarrative forms in Table 2 and were built after consultations with all three Bulgarian linguists, who are co-authors of this article. The dataset used for training the classifier (BGRenData) was also

[11] https://zenodo.org/records/15882529
[12] https://huggingface.co/usmiva/bert-web-bg-cased.

Dataset	Prec.	Rec.	Acc.	F1-Score
BGRenBERT				
CNClean	0.714	0.846	0.958	0.774
CNCleanB	0.127	1.000	0.905	0.225
CNCleanN	0.806	0.842	0.965	0.824
CT1D	0.292	0.824	0.924	0.431
BGRenRules				
CNClean	0.958	0.928	0.990	**0.943**
CNCleanB	0.583	0.778	0.989	0.667
CNCleanN	0.969	0.931	0.991	0.949
CT1D	0.500	0.882	0.965	**0.638**

Table 6
Results of BGRenBERT and BGRenRules on CNClean, CNCleanB, CNCleanN, and CT1D.

used to recursively test and refine BGRenRules. For BGRenRules we used all three splits of BGRenData. We expected that BGRenRules would give better results as it was based on regular expressions, matching the well-structured Bulgarian grammar, compared with the probabilities-based BGRenBERT. Both BGRenRules and BGRenBERT are shared on Zenodo[13], HuggingFace[14] and Github[15] with Creative Commons Attribution 4.0 International license.

6. Comparison of the Methods on the Annotated Subsets of Datasets

To compare the two methods on the same texts, we applied them to CNClean subsets and CT1D. Exactly like human annotators, the methods were assigning the label "renarrative" if the text contained at least 1 form of the renarratives in Table 2 and "no renarrative" if none. We then compared the success of both methods in matching the manual annotations.

Table 6 shows the results of the two methods on these datasets. We would like to clarify that Table 6 shows the evaluation of the two methods on different texts than those used for creating BGRenBERT and BGRenRules.

Due to the higher number of items, we report the results for CNClean, and separately for CNCleanB and CNCleanN. We report only the results for CT1D as a whole, due to the too low numbers of renarratives in its subsets. We report Precision, Recall, Accuracy, and F1-Scores, compared to the manual annotations. As the datasets are not balanced, more attention should be given to the F1-Scores, but we report both. It is clearly visible that the BGRenRules outperforms BGRenBERT. The CT1D results are lower for both methods, probably because of the low number of texts with renarrative.

[13] https://zenodo.org/records/15802264
[14] https://huggingface.co/gate-institute/BGRenBERT
[15] https://github.com/silviavg/bg-renarrative

Figures 1 and 2 show the percentages of False Positives (FP) and False Negatives (FN) from all errors of the two methods in all datasets. It is clear that BGRenBERT makes more FP errors, while BGRenRules – more FN errors. Given this, if a user is interested in having fewer FP (i.e. fewer texts not containing renarrative, but automatically identified as containing renarratives), they should use BgRenRules. If the user, instead, would like to have fewer FN (i.e. texts with renarrative, wrongly recognised to not contain any of its forms) – then BGRenBERT would be a better option.

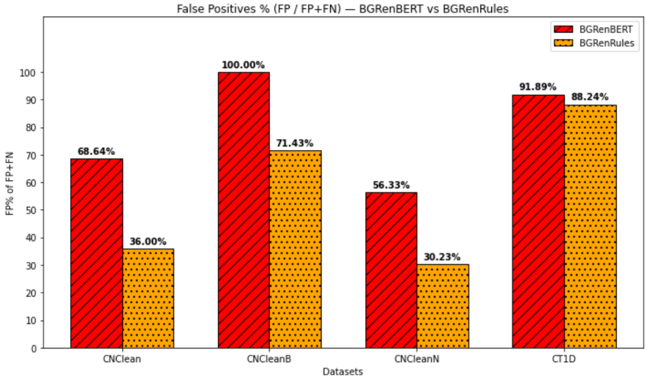

Figure 1
Percentages of False Positives (FP) from all errors of the two methods in all datasets.

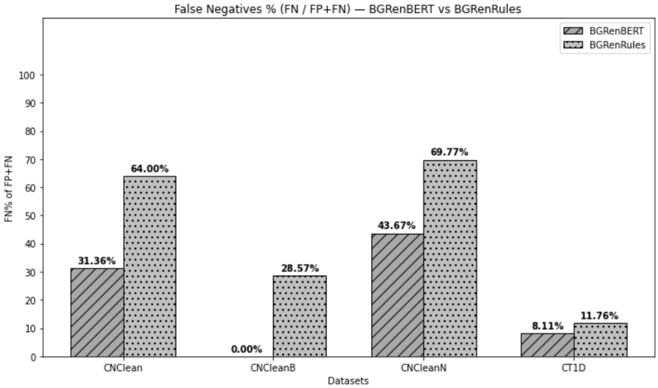

Figure 2
Percentages of False Negatives (FN) from all errors of the two methods in all datasets.

To see a more detailed picture, we manually analysed the False Positives and False Negatives for both datasets and the two methods, comparing them with the manual annotations. The error analysis is presented below.

6.1 Error Analysis of CNClean

As CNCleanB and CNCleanN differ in terms of style and contain enough items, allowing us to build a picture of each dataset's subset, we analyzed them separately.

6.1.1 Analysis of CNCleanB

CNCleanB contains 651 titles, from which 9 (1.38%) were manually annotated as containing renarrative. BGRenBERT made 62 errors, all False Positives (cases in which the model decided that a title contains a renarrative, but there was none). Errors could be grouped as:

- Words ending on '-л'. For example, the model shows the word 'капитал' – in English: 'capital' as a renarrative - obviously because of the suffix '**-ал**';
- Grammatically mistaken absences of comma + past participles ('Доживотен затвор за дядо заклал прасето си!' (in English: 'Life imprisonment for an elderly man who slaughtered his pig!')
- Some quotations headlines, recognised as containing renarrative, like "Марин Райков: 'Бизнесът и медиите трябва да са солидарни с управляващите, а не с народа"', (in English: "Marin Raykov: 'The Business and the media should be in solidarity with the government, not with people."). The possible explanation is that the function of renarrative is to retell a story; thus, the quotation marks are a sign of retelling.
- Linked to the previous case, a lot of punctuation tricks BGRenBERT into deciding that a form is a renarrative, while it actually is not. Only two cases of renarrative were not recognized as such by the model - one is a typical renarrative "Менделеев починал след консумация на български колбас!" (in English: "Mendeleev died after eating Bulgarian sausage!") and the second is difficult to recognise even by humans: "Циганите изпищяха, купували им гласовете с фалшиви банкноти. ЦИК заплаши да касира изборите." (in English: "The gypsies cried out, their votes were bought with fake banknotes. The CEC (the Central Electoral Commission) threatened to cancel the elections.") The second case could be understood either as the evidential conclusive, or as indicative of a statal perfect.

All BGRenRules mistakes instead corresponded to cases in which all human annotators were not sure.

6.1.2 Analysis of CNCleanN

The dataset contains 4579 titles, from which the manually annotated ones with renarrative are 438 (or 9.56%). BGRenBERT tagged 89 titles as containing renarrative, while there was no renarrative. The types of these errors were:

- Words ending in "-л". Examples: "МВР погна похитители на играчки" - translation into English: "The Ministry of Internal Affairs chased kidnappers of toys"; "Протестър №1 роди пудел от Саня Армутлиева" - translation into English: "Protester No. 1 gave birth to a poodle by Sanya Armutlieva"; participle finishing with "-л".
- Present tense, indicative, third person, singular. Examples: "изнасили" (translation into English: "raped"), "уцели" (translation into English: "hit").
- There were also 45 other cases which could not be grouped.

There are also 69 cases of titles, containing renarrative, but not recognised as such by BGRenBERT:

- 12 of them are in a subordinate clause (in the remaining similar cases, the renarrative is correctly recognised). For example: ("Монтират топломери на хората, също излъчвали топлина, приспадат я от сметките." (in English: "They are installing heat meters on people too, since they emit heat — it's deducted from the bills.").
- Elliptical phrases without a personal pronoun (for example: "Ердоган се отказа от мол в Истанбул, прицелил се в Народното събрание в София" (in English: "Erdogan gave up a mall in Istanbul, now he is targeting the Parliament in Sofia").
- There was a case in which the construction was revealed, even though the whole headline was misspelt.

BGRenRules gave better results: it made only 43 errors. From these, the main ones were such in which the renarrative was not recognized:

- Set phrases like "***каквито и да било***" (in English "***any***"), for example, in this already translated sentence: "CSKA insists that the derby with Levski should not be officiated by ***any*** referees"'.
- Renarrative of the auxilary verb "съм" (in English "to be"). For example in "ИЗВЪНРЕДНО!111! РУСИЯ НИ ОБЯВИ ВОЙНА, УДАРЪТ ПО ВОЛЕН БИЛ УДАР ПО ТЯХНАТА ТЕРИТОРИЯ" (translation into English: "BREAKING NEWS!111! RUSSIA DECLARED WAR ON US, THE STRIKE ON VOLEN WAS A STRIKE ON THEIR TERRITORY").
- A reflexive verb, like in this example: "Най-богатият българин във Facebook си платил, за да го следят" (translation into English: "The richest Bulgarian on Facebook paid to be followed.").

6.2 Analysis of the CT1D

CT1D contains small numbers of the different classes, as well as of manually annotated renarratives, and for this reason, we run an error analysis of all classes together. The dataset contains 488 items, out of which 17 were manually annotated as containing renarrative -

the renarrative is 3.48% of all the items. BGRenBERT tagged 34 cases as containing forms of renarrative, but they did not. The errors are related to the following issues:

- Long sentences.
- Insertion of Latin transcription (COVID-19).
- Insertion of special symbols such as #.
- Atypical errors.

BGRenRules worked better than BGRenBERT, as in the previous cases. Specifically, it flagged wrongly as containing renarratives 15 cases and also failed to recognise 2 cases with actual renarratives. All of its errors were hard to resolve even for humans. We give as an example a case which is in general difficult to be recognised as a renarrative: "Три учителки от детска градина в София са с Covid-19, ходили до Сърница" (translation into English: "Three kindergarten teachers in Sofia have Covid-19: they went to Sarnitsa" or "it was said that they went to Sarnitsa"). It is possible that the subordinated clause is in perfect tense, indicative, or that it is in renarrative.

6.3 Overall Observations

We observed that BgRenRules generally made mistakes only in the cases which were doubtful also for humans. This is most probably due to the fact that BGRenBERT is based on probabilities, which may cause errors, while BGRenRules is using regular expressions which closely match the Bulgarian grammar. Such observations, in fact, confirm our initial expectations.

Additionally, we observed that the journalists' poor punctuation skills confuses BGRenBERT. This brings up another factor not previously accounted for in attempts at automatic renarrative detection: punctuation issues. Based on the current analysis, when punctuation is correctly used, BGRenBERT works fine. However, since the lack of commas is a frequent writing mistake, this limitation should be considered in the future. Another main problem for the automatic detection of renarrative is the confusion with past perfect/imperfect participle, as well as with some nouns, finishing with "-л", and verbs in the present tense finishing with "-ели".

7. Ethical and Legal Considerations

The annotators were among the authors of this article. Their annotation work was part of their regular salaries and was decently paid for Bulgaria. The annotated datasets are shared per the requirements and licenses of the datasets the texts were originally part of. We are making BGRenBERT, BGRenRules, and all the annotated datasets open-access with specifically written legal disclaimers. In addition to the license, the disclaimers state that: 1) it should be taken into account that the automatic detection of texts with "renarrative" generates some errors; 2) the presence of a form of "renarrative" should not be taken as a sole indication of the lack of trustworthiness of a statement or a text.

8. Work Limitations

Our work makes an important contribution by presenting the first automatic solutions for detecting Bulgarian texts containing renarrative. However, it also has some limitations:

- We are considering only the most frequent forms of renarrative and do not offer a definitive solution for cases in which humans would also doubt. It would be a better solution to cover all forms of renarrative, including identifying correctly those forms for which humans are also unsure.
- We covered only the 3rd person singular and plural, as these are the most common ones in news articles.
- We recognise only Bulgarian texts, written in Cyrillic alphabet. Social media and forums' posts and answers in Bulgarian are often written with different Latin transliterations.
- Our solutions allow detecting short texts, containing these main forms of renarrative, but not the forms themselves. Future work could include creating a solution which detects the actual beginning and the end of a form of renarrative in the text.

Such limitations could be resolved in our future work or by other researchers. It is also preferable if the automatic detection of renarrative is not a stand-alone solution, but possibly a functionality offered by a Bulgarian syntactic parser.

9. Conclusions and Future Work

In this article, we presented the first automatic solutions for detecting Bulgarian texts or sentences, containing the most frequent forms of the evidential renarrative. After applying them to manually annotated subsets of two datasets with fake Bulgarian texts, we obtained better results from our rule-based solution, BGRenRules. The fine-tuned BGRenBERT, instead, had a worse performance. Such results were justified by the principles on which the classifier and BGRenRules were based. We provided an extensive error analysis. All resources, including the two automatic solutions, the training dataset and the manually annotated subsets of the third-party text datasets, are made publicly available to assist other researchers. In future work, we plan to create a detector which recognizes the exact span of the forms of renarrative, as well as to cover more forms. We also consider adding the indicative present perfect forms as they are often used instead of the renarrative ones. Important questions to consider are whether: 1) the presence of renarrative is statistically meaningful and so frequent; 2) we need automatic methods to protect society from reading fake news containing renarratives. Additionally, while renarrative and evidentiality are important for truthfulness, their combination with other linguistic markers is still not analysed. Finally, a diachronic analysis of the use of renarrative could also be done.

10. Acknowledgments and Authors' Contributions

We would like to thank the "Computational Linguistics in Bulgaria" journal editors and the anonymous reviewers for their suggestions and for signalling imperfections in the article.

This research is part of the GATE project funded by the Horizon 2020 WIDESPREAD-2018-2020 TEAMING Phase 2 Programme under grant agreement no. 857155, the Programme "Research, Innovation and Digitalisation for Smart Transformation" 2021-2027 (PRIDST) under grant agreement no. BG16RFPR002-1.014-0010-C01 and, the project BROD (Bulgarian-Romanian Observatory of Digital Media) funded by the Digital Europe Programme of the European Union under contract number 101083730.

The authors made the following contributions: ***Irina Temnikova*** initiated the research, contributed to most of the datasets collection, and experiments design, calculated the comparative analysis results, generated the graphics, worked on fine-tuning BERT-WEB-BG-Cased, and wrote almost all the article; the focus on renarrative followed ***Ruslana Margova***'s observations during her work as an international news editor in Bulgarian newspapers. She also decided which are the main renarrative forms to detect, supplied part of BGRenData, manually annotated parts of CNCLean and CT1D, contributed to writing the sections on Evidentiality, and on the Detected forms of Bulgarian renarrative, conducted and wrote the manual error analysis, and made changes to the article regarding the linguistic comments of the reviewers. ***Stefan Minkov*** created BGRenRules and provided comments on the final version of the article. ***Tsvetelina Stefanova*** gave linguistic advice, participated in collecting examples with and without renarrative, manually annotated CT1D, checked whether the existing Bulgarian POS taggers and parsers detected evidentials, created Table 1, and made small changes to the final version of the article. ***Nevena Grigorova*** contributed with linguistic advices, generated examples for BGRenData, and gave suggestions for the final version of the paper. ***Silvia Gargova*** created the BGRenData splits for BGRenBERT, fine-tuned one of its previous versions, uploaded BGRenBERT, BGRenRules, and BGRenData to the respective platforms, and made suggestions for the final version of the article. Finally, ***Venelin Kovatchev*** gave suggestions on answering the reviewers' requests and on how to improve the final version of the article, including about the contents of Figures 1 and 2.

References

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. OUP Oxford.

Aikhenvald, Alexandra Y. 2015. Evidentials: Their links with other grammatical categories. *Linguistic Typology*, 19(2):239–277.

Al-Rfou, Rami. 2015. *Polyglot: A massive multilingual natural language processing pipeline*. Ph.D. thesis, State University of New York at Stony Brook.

Aleksova, Krasimira. 2024. Kategoriata evidentsialnost na balgarskiya glagol v obuchenieto po balgarski v srednoto uchilishte [The evidentiality category of Bulgarian verbs in Bulgarian language education at secondary school]. *Balgarski ezik i literatura [Bulgarian Language and Literature]*, 66(1):30–47.

Atanasova, Pepa, Preslav Nakov, Lluís Màrquez, Alberto Barrón-Cedeño, Georgi Karadzhov, Tsvetomila Mihaylova, Mitra Mohtarami, and James Glass. 2019. Automatic fact-checking using context and discourse information. *Journal of Data and Information Quality (JDIQ)*, 11(3):1–27.

Berbatova, Melania and Filip Ivanov. 2023. An improved Bulgarian natural language processing pipeline. *Annual of Sofia University St. Kliment Ohridski. Faculty of Mathematics and Informatics*, 110:37–50.

Borin, Lars, Markus Forsberg, Martin Hammarstedt, Dan Rosén, Roland Schäfer, and Anne Schumacher. 2016. Sparv: Språkbanken's corpus annotation pipeline infrastructure. In *The Sixth Swedish Language Technology Conference (SLTC), Umeå University*, pages 17–18.

Boroş, Tiberiu, Ştefan Daniel Dumitrescu, and Ruxandra Burtica. 2018. NLP-Cube: End-to-end raw text processing with neural networks. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 171–179.

BulTreeBank. Bulgarian NLP Pipeline in CLARK System. http://bultreebank.org/clark/bulgarian-nlp-pipeline-in-clark-system/. BulTreeBank Project website.

Cholakov, Radostin Lozanov. AzBuki.ML - Machine learning platform for natural language processing, implementing recurrent & convolutional neural networks and linguistic algorithms with applications for the Slavic Languages.

CLASSLA. 2021. CLASSLA Fork of the Official Stanford NLP Python Library for Many Human Languages. https://live.european-language-grid.eu/catalogue/tool-service/15639. Software (Tool/Service).

Das, Anubrata, Houjiang Liu, Venelin Kovatchev, and Matthew Lease. 2023. The state of human-centered NLP technology for fact-checking. *Information Processing & Management*, 60(2):103219.

Dinkov, Yoan, Ivan Koychev, and Preslav Nakov. 2019. Detecting Toxicity in News Articles: Application to Bulgarian. *arXiv preprint arXiv:1908.09785*.

Fauzi, Adam. 2021. POS-BERT-bg: BERT model for Bulgarian Part-of-Speech Tagging. https://huggingface.co/auhide/pos-bert-bg. Hugging Face model repository.

Grieve, Jack and Helena Woodfield. 2023. *The Language of Fake News*. Cambridge University Press.

Guo, Bin, Yasan Ding, Lina Yao, Yunji Liang, and Zhiwen Yu. 2020. The future of false information detection on social media: New perspectives and trends. *ACM Computing Surveys (CSUR)*, 53(4):1–36.

Hangloo, Sakshini and Bhavna Arora. 2023. Evidence-aware fake news detection: A review. In *2023 International Conference on Advanced Computing & Communication Technologies (ICACCTech)*, pages 81–86, IEEE.

Hardalov, Momchil, Ivan Koychev, and Preslav Nakov. 2016. In search of credible news. In *Artificial Intelligence: Methodology, Systems, and Applications: 17th International Conference, AIMSA 2016, Varna, Bulgaria, September 7-10, 2016, Proceedings 17*, pages 172–180, Springer.

Honnibal, Matthew, Ines Montani, Sofie Van Landeghem, and Adriane Boyd. 2020. spaCy: Industrial-strength Natural Language Processing in Python.

Ifantidou, Elly. 2001. *Evidentials and Relevance*. John Benjamins Publishing Company.

Jackobson, Roman. 1957. *Shifters, Verbal Categories and the Russian Verb*. Department of Slavic Languages and Literatures, Harvard University.

Jongejan, Bart. 2020. The CLARIN-DK Text Tonsorium. In *CLARIN Annual Conference*, pages 111–121.

Karagjosova, Elena. 2021. Mirativity and the bulgarian evidential system. In *Advances in formal Slavic linguistics 2018*. Language Science Press, pages 133–167.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlỳ, and Vít Suchomel. 2014. The Sketch engine. *Lexicography*, 1(1):7–36.

LIMA. 2021. LIMA - Libre Multilingual Analyzer. https://live.european-language-grid.eu/catalogue/tool-service/15344. Software (Tool/Service).

Ljubešić, Nikola, Luka Terčon, and Kaja Dobrovoljc. 2024. CLASSLA-Stanza: The Next Step for Linguistic Processing of South Slavic Languages. In *Conference on Language Technologies and Digital Humanities (JT-DH-2024)*, Institute of Contemporary History, Ljubljana, Slovenia.

Margova, R. 2023. *Linguistic Features of Fake News*. Phd thesis, Sofia University.

Mihaylov, Rumen. 2023. BERT-base POS model with Theseus training for Bulgarian. https://huggingface.co/rmihaylov/bert-base-pos-theseus-bg. Hugging Face model repository.

Montgomerie, Adam. 2021a. RoBERTa Base Bulgarian POS. https://huggingface.co/iarfmoose/roberta-base-bulgarian-pos. Hugging Face model repository.

Montgomerie, Adam. 2021b. RoBERTa Small Bulgarian POS. https://huggingface.co/iarfmoose/roberta-small-bulgarian-pos. Hugging Face model repository.

Montgomery, Adam. 2023. Bulgarian NLP. https://github.com/AMontgomerie/bulgarian-nlp. GitHub repository.

Mushin, Ilana. 2000. Evidentiality and deixis in retelling. *Journal of Pragmatics*, 32:927–957.

Nakov, Preslav, Alberto Barrón-Cedeño, Giovanni Da San Martino, Firoj Alam, Rubén Míguez, Tommaso Caselli, Mucahid Kutlu, Wajdi Zaghouani, Chengkai Li, Shaden Shaar, et al. 2022. Overview of the CLEF-2022 CheckThat! lab task 1 on identifying relevant claims in tweets. In *2022 Conference and Labs of the Evaluation Forum, CLEF 2022*, pages 368–392, CEUR Workshop Proceedings (CEUR-WS. org).

Newman, Hadley. 2021. Understanding the differences between disinformation, misinformation, malinformation and information: Presenting the dmmi matrix. *Draft Online Safety Bill (Joint Committee).*

Nitzolova, R. 2008. *Balgarska gramatika. Morfologiya [Bulgarian Grammar: Morphology].* UI "St. Kliment Ohridski".

Panchendrarajan, Rrubaa and Arkaitz Zubiaga. 2024. Claim detection for automated fact-checking: A survey on monolingual, multilingual and cross-lingual research. *Natural Language Processing Journal*, 7:100066.

Qi, Peng, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D. Manning. 2020. Stanza: A Python natural language processing toolkit for many human languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*.

Rani, Neetu, Prasenjit Das, and Amit Kumar Bhardwaj. 2022. Rumor, misinformation among web: A contemporary review of rumor detection techniques during different web waves. *Concurrency and Computation: Practice and Experience*, 34(1):e6479.

Roberts, Ian. 2020. GATE: Universal Dependencies POS Tagger for bg / Bulgarian. Software (Tool/Service).

Santhosh, Nikita Mariam, Jo Cheriyan, and Lekshmi S Nair. 2022. A multi-model intelligent approach for rumor detection in social networks. In *2022 International Conference on Computing, Communication, Security and Intelligent Systems (IC3SIS)*, pages 1–5, IEEE.

Schmid, Helmut. 1994. Treetagger-a part-of-speech tagger for many languages. *Ludwig-Maximilians-Universität Munich.*

Schmid, Helmut. 2019. Deep learning-based morphological taggers and lemmatizers for annotating historical texts. In *Proceedings of the 3rd international conference on digital access to textual cultural heritage*, pages 133–137.

Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1):22–36.

Silberztein, Max. 2005. Nooj: a linguistic annotation system for corpus processing. In *Proceedings of HLT/EMNLP 2005 Interactive Demonstrations*, pages 10–11.

Straka, Milan. 2018. UDPipe 2.0 prototype at CoNLL 2018 UD shared task. In *Proceedings of the CoNLL 2018 shared task: Multilingual parsing from raw text to universal dependencies*, pages 197–207.

Su, Qi, Helen Kai-Yun Chen, and Chu-Ren Huang. 2011. A research on the text reliability based on evidentiality. *International Journal of Computer Processing of Languages*, 23(02):201–214.

Su, Qi, Chu-Ren Huang, and Helen Kaiyun Chen. 2010. Evidentiality for text trustworthiness detection. In *Proceedings of the 2010 Workshop on NLP and Linguistics: Finding the Common Ground*, pages 10–17.

Sylak-Glassman, John. 2016. The composition and use of the universal morphological feature schema (unimorph schema). *Johns Hopkins University*, page 6.

Temnikova, Irina, Silvia Gargova, Ruslana Margova, Veneta Kireva, Ivo Dzhumerov, Tsvetelina Stefanova, and Hristiana Krasteva. 2023. New bulgarian resources for studying deception and detecting disinformation. In *Proceedings of the 10th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics*, Poznań, Poland.

Terčon, Luka, Nikola Ljubešić, Petya Osenova, Kiril Simov, and Luka Krsnik. 2023. The CLASSLA-Stanza model for morphosyntactic annotation of standard Bulgarian 2.1. Slovenian language resource repository CLARIN.SI.

de Vries, Wietse. 2021. XLM-RoBERTa base fine-tuned on Bulgarian Universal Dependencies POS. https://huggingface.co/wietsedv/xlm-roberta-base-ft-udpos28-bg. Hugging Face model repository.

de Vries, Wietse, Martijn Wieling, and Malvina Nissim. 2022. Make the best of cross-lingual transfer: Evidence from POS tagging with over 100 languages. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7676–7685, Association for Computational Linguistics, Dublin, Ireland.

Wardle, Claire et al. 2018. Information disorder: The essential glossary. *Harvard, MA: Shorenstein Center on Media, Politics, and Public Policy, Harvard Kennedy School.*

Yasuoka, Koichi. 2021a. BERT-base Slavic Cyrillic model fine-tuned for UPOS. https://huggingface.co/KoichiYasuoka/bert-base-slavic-cyrillic-upos. Hugging Face model repository.

Yasuoka, Koichi. 2021b. BERT-large Slavic Cyrillic model fine-tuned for UPOS. https://huggingface.co/KoichiYasuoka/bert-large-slavic-cyrillic-upos. Hugging Face model repository.

Zhou, Xinyi and Reza Zafarani. 2020. A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR)*, 53(5):1–40.