cepanca_UNAM at CheckThat! 2025: A Language-driven BERT Approach for Detection of Subjectivity in News

Notebook for the CheckThat! Lab at CLEF 2025

Iván Diaz[1,*,†], Jessica Barco[2,†], Joana Hernández[3,†], Edgar Lee-Romero[3,†] and Gemma Bel-Enguix[3]

[1]*Posgrado en Ciencias e Ingeniería de la Computación, Universidad Nacional Autónoma de México, Ciudad de México 04510, México*

[2]*Universidad Autónoma Metropolitana Unidad Iztapalapa, Ciudad de México 09340, México*

[3]*Instituto de Ingeniería, Universidad Nacional Autónoma de México, Ciudad de México 04510, México*

Abstract

Automatic subjectivity detection is a trending issue in Natural Language Processing. Responding to the challenge of finding models that are able to distinguish between objective and subjective segments, the CheckThat! Lab, which is part of CLEF 2025, invites in Task 1 to distinguish whether a sentence from a news article expresses the subjectivity of the author or not. The task has three settings: monolingual, multilingual and zero-shot. Our contribution focuses on monolingual classification in three of the proposed languages: English, Italian and German. The approach used is based on Transformers-based models. We have opted for BERT-base-uncased for English, BERT-base-italian-cased-sentiment for Italian and German BERT large for German. In our work, we have taken into account the lexical features, specifically the distribution of the different grammatical categories in each of the corpora. Despite the simplicity of our models, our results have been competitive, obtaining the second place in German, with a macro-F1 of 0.8280.

Keywords

Subjectivity, Objectivity, Transformer Models, BERT, Binary Classification, LLM, CheckThat! 2025

1. Introduction

The semantic distinction between subjectivity and objectivity has traditionally been understood as a dichotomy that distinguishes whether the author of a sentence appears immersed in the enunciation or not [1]. From the area of NLP, it is seen as the 'aspects of language used to express opinions, evaluations, and speculations' [2, 3]. However, it is clear that the way humans communicate, talk about events and experiences, needs to be unique and arises from our own experiences [4]. Therefore, subjectivity seems unavoidable in human language communication.

For the objective of the present study, we understand subjectivity as statements that rely on personal opinions or emotions noticeable by the grammatical presence of an enunciator. The objective shall be to explore the application of some methods of advanced natural language processing that are still being developed in an attempt of improving the results of its functionality. This paper is an approximation to the tools that attempt to serve in the mentioned improvement of the techniques.

This work originates from the 2025 edition of the CheckThat! Lab [5], held ad CLEF 2025 [6]. CheckThat! 2025 is the eighth version of the competition. Task 1 [7] is devoted segment-level subjectivity detection. It consists of a binary classification in which the system developed had to identify a text sequence as subjective or objective. They posed three possible settings: a) monolingual, b) multilingual,

CLEF 2025 Working Notes, 9 – 12 September 2025, Madrid, Spain

*Corresponding author.

†These authors contributed equally.

diazrysivan@gmail.com (I. Diaz); jessbarco13@gmail.com (J. Barco); joana.hernandez.rebollo@gmail.com (J. Hernández); edgar.lee133@gmail.com (E. Lee-Romero); gbele@iingen.unam.mx (G. Bel-Enguix)

https://github.com/JuanIvanDiazReyes (I. Diaz); https://github.com/GIL-UNAM (G. Bel-Enguix)

no orcid provided (I. Diaz); no orcid provided (J. Barco); 0009-0001-2537-9657 (J. Hernández); 0009-0003-8701-341X (E. Lee-Romero); 0000-0002-1411-5736 (G. Bel-Enguix)

and c) zero shot. Our team decided to participate in the monolingual sub-tasks in the following languages: a) German, b) Italian, and c) English. For all training, we utilized the datasets provided by the organizers of the competition, which were composed of news articles.

2. Related work

Automatic subjectivity detection has been a fundamental topic in Natural Language Processing for years. Subjectivity detection, in particular, is an essential subtask of sentiment analysis because most polarity detection tools are optimized to distinguish between positive and negative text. Subjectivity detection, hence, ensures that factual information is filtered out and only opinionated information is passed on to the polarity classifier.

Early attempts to address the problem were based on the use of lexicons, dictionaries, and other lexical resources that could detect words specifically related to subjectivity [8, 9, 10, 11].

The need to create annotated corpora arose, especially with the emergence of machine learning methods that would allow a supervised approach to the subject. Classical corpora include MPQA (Multi-Perspective Question Answering)) [12], and NewsSD-ENG, compiled from English news articles [13], which has been partially used in the task CheckThat!. One of the main problems of annotating something like subjectivity is that it is a very subjective task. This is why some authors have suggested the paradigm of disagreements [14].

With language models and large language models, all NLP tasks have undergone a major development, and subjectivity detection could not be different. Transformers are a technology where the main focus is in providing pre-trained models to reduce computational cost, reduce carbon emissions, and save time from training conventional models. BERT (Bidirectional Encoder Representations from Transformers) is one of these pre-trained models that will provide substantial output with lax parameters to detect subjectivity [15] [16]

In particular, the BERT-based models are being successfully applied to the task. Satapathy et al. [17] present a multi-task model for detecting and mutually supporting polarity and subjectivity detection. In 2024 CheckThat task, the teams that won in English [18], German [19], and Italian [20], applied BERT models to approach the problem.

In 2024 the JK PCIC UNAM [20] utilized Bert-based models focusing in two languages; English and Italian in news articles to explore whether sentences were written with tints of subjectivity or objectivity.

Our methodology does not include the use of LLMs, although the model has been shown to achieve competitive results [21].

3. Methodology

The languages that we selected for this shared task were English, Italian, and German. The following analysis was made using the datasets for these languages.

3.1. Analysis of the dataset

First, we analyzed the label distribution (OBJ, SUBJ) to detect imbalanced classes, which could compromise model training.

As shown in Table 1, the German dataset exhibits a relatively balanced label distribution, whereas the Italian and English datasets display a significant imbalance, with a pronounced bias toward the OBJ class over SUBJ. This disparity may adversely affect model training, potentially leading to biased predictions.

After analyzing label distribution, we conducted a Part-of-Speech (POS) analysis to explore linguistic patterns across the datasets. In all cases, the "Others" category was predominant, accounting for over

Table 1
Data Split and Distribution

Language	Train			Dev		
	Total	OBJ	SUBJ	Total	OBJ	SUBJ
English	830	532	298	462	222	240
Italian	1613	1231	382	667	490	177
German	800	492	308	491	317	174

50% of POS tags. This indicates a higher frequency of grammatical function words compared to lexical content (e.g., nouns, verbs).

Additionally, we observed a disparity in adverb usage: posts labeled as subjective contained a higher proportion of adverbs. This aligns with the linguistic function of adverbs, which often serve to express opinions, emotions, or personal perspectives.

Table 2
Analysis of Lexical and Syntactic Features in Training Data

Feature	English		Italian		German	
	OBJ	SUBJ	OBJ	SUBJ	OBJ	SUBJ
Word Count	11 708	7253	40 551	13 797	8658	5429
Unique Lemmas	2847	1813	6600	3341	2513	1771
POS Distribution (%)						
Nouns	23.31	22.63	20.54	19.93	23.27	19.16
Adjectives	9.03	9.83	6.49	7.45	6.19	5.58
Verbs	12.12	11.52	9.37	9.89	10.10	10.55
Adverbs	4.57	5.87	4.35	6.4	10.34	14.15
Others	50.96	50.15	61.66	56.3	50.10	50.56

3.2. Machine Learning Models

For this task, we have decided to establish a comparison between the performance of traditional ML and Transformer-based methods. For this purpose, we have performed some tests with the classical classification algorithms and we have taken Logistic Regression, since it is the one that has given the best results. The LR has been used as a baseline to have a benchmark to compare the performance of the fine-tuned BERT-based models we have applied. We vectorized the text with bag-of-words, and used 3-grams and 4-grams. After some evaluations, we did not filter the stopwords

3.3. Transformer Training

We maintain the protocol established last year [20], ensuring comparability between different editions. Our approach employed BERT (Bidirectional Encoder Representations from Transformers) as the primary architecture for subjectivity classification tasks. BERT models, being pre-trained on extensive text corpora, demonstrate particular effectiveness in sequence classification applications. For language-specific implementations, we utilized: BERT-base-uncased for English, BERT-base-italian-cased-sentiment for Italian and German BERT large for German [22].

German BERT large is a language model for German, released in 2020 by the creators of the original German BERT and the dbmdz BERT. Pretrained on approximately 170 GB of text data, it leverages diverse linguistic sources, with the OSCAR corpus being one of its most significant training datasets

[22]. Given OSCAR's broad coverage of web-sourced texts, the model benefits from varied linguistic patterns, making it particularly suitable for this task.

The italian BERT sentiment model was pretrained by Neuraly AI from an instance of bert-base-italian-cased, fine-tuned with a corpora of 45k tweets to perform sentiment analysis in italian. Regardless of the domain of the training dataset which was reported to be football, Neuraly AI claims efficiency in other topics which proved to be an accurate promise when the model was presented with the competition's news datasets. [23]

BERT-base-uncased is also a pretrained model trained in a self-supervised manner with two objectives:; Masked language modeling (MLM) and Next Sentences Prediction (NSP). This way, the model learns an inner representation of the English language that can then be used to extract features useful for downstream tasks [15]. The main objective of our application is to classify pairs of sentences as either positive or negative based on selected parameters. In this case, we focus on subjectivity detection, using words with factual language as a starting point to categorize text as subjective or objective. This approach is also commonly applied in sentiment analysis.

All models were fine-tuned on the provided dataset, with hyperparameter optimization performed exclusively on the training set to maintain evaluation integrity. The fine-tuning process focused on adjusting the supervised classifier's parameters, using: 4 training epochs, Batch size of 16, Maximum input length of 256 tokens.

Model performance was assessed using precision, recall, F1 score, and macro-averaged F1 score for each experimental condition. We prioritize macro-F1 as our primary optimization metric due to its robustness in addressing class imbalance issues inherent in subjectivity classification tasks.

All experiments were conducted in Google Colab using GPU acceleration to handle the computational demands of fine-tuning.

3.3.1. Model Comparison and Hyperparameter Optimization

We conducted a systematic performance comparison between our baseline Logistic Regression (LR) model and BERT-based classifiers across all target languages (English, Italian, German). This dual-model approach served to establish transformer performance gains over classical methods as we can see in Table 3.

Table 3
Performance Metrics by Language and Model

Model	Metric	English	Italian	German
BERT	Precision	0.53	0.59	0.71
	Recall	0.47	0.78	0.82
	F1	0.50	0.67	0.76
	Macro F1	0.62	0.76	0.81
Logistic Regression	Precision	0.67	0.38	0.51
	Recall	0.55	0.38	0.55
	F1	0.61	0.67	0.53
	Macro F1	0.63	0.38	0.63

The Transformer results, which we obtained through systematic hyperparameter tuning on the organizers' development datasets, are presented in Tables 4 (English), 5 (Italian), and 6 (German).

For German and Italian, we observed that reducing the batch size to 16 while increasing the maximum token length to 256 yielded superior performance. This improvement likely stems from the need to preserve complete linguistic structures in longer sentences and the risk of losing critical contextual information (and introducing bias) with shorter sequences.

In contrast, English achieved optimal results with shorter sequences (128 tokens) and smaller batches (16), suggesting different processing requirements for this language.

Table 4
Analysis of Results with Different Settings on the English Development Dataset

Settings	Precision	Recall	F1	Macro F1
Batch size = 32, Max len = 128	0.420	0.798	0.550	0.531
Batch size = 16, Max len = 128	0.536	0.473	0.502	**0.624**
Batch size = 32, Max len = 256	0.482	0.372	0.420	0.575
Batch size = 16, Max len = 256	0.360	0.177	0.177	0.460

Table 5
Analysis of Results with Different Settings on the Italian Development Dataset

Settings	Precision	Recall	F1	Macro F1
Batch size = 32, Max len = 128	0.651	0.632	0.641	0.757
Batch size = 16, Max len = 128	0.711	0.570	0.633	0.758
Batch size = 32, Max len = 256	0.672	0.604	0.636	0.757
Batch size = 16, Max len = 256	0.686	0.593	0.636	**0.758**

Table 6
Analysis of Results with Different Settings on the German Development Dataset

Settings	Precision	Recall	F1	Macro F1
Batch size = 32, Max len = 128	0.702	0.789	0.747	0.796
Batch size = 16, Max len = 128	0.709	0.799	0.751	0.800
Batch size = 32, Max len = 256	0.756	0.678	0.715	0.785
Batch size = 16, Max len = 256	0.707	0.821	0.760	**0.806**

The strong performance of the German model may be attributed to the model's pretrained datasets like OSCAR. While OSCAR's web-sourced texts provide broad linguistic coverage, they may also introduce challenges due to potential misinformation and inaccuracies inherent in internet content.

4. Analysis of the results

We evaluated our models' performance against the official shared task baseline using the organizers' evaluation framework. The standardized scorer provided computes the following classification metrics:

- **Accuracy**:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Macro-Precision (macro-P)**:

$$\text{macro-P} = \frac{1}{N}\sum_{i=1}^{N} P_i, \quad \text{where} \quad P_i = \frac{TP_i}{TP_i + FP_i}$$

- **Macro-Recall (macro-R)**:

$$\text{macro-R} = \frac{1}{N}\sum_{i=1}^{N} R_i, \quad \text{where} \quad R_i = \frac{TP_i}{TP_i + FN_i}$$

- **Macro-F1 (macro-F1)**:

$$\text{macro-F1} = \frac{1}{N}\sum_{i=1}^{N} F1_i, \quad \text{where} \quad F1_i = 2 \times \frac{P_i \times R_i}{P_i + R_i}$$

- **Class-specific metrics (e.g., SUBJ)**:
  - **Precision (SUBJ-P)**:
$$P_{\text{SUBJ}} = \frac{TP_{\text{SUBJ}}}{TP_{\text{SUBJ}} + FP_{\text{SUBJ}}}$$
  - **Recall (SUBJ-R)**:
$$R_{\text{SUBJ}} = \frac{TP_{\text{SUBJ}}}{TP_{\text{SUBJ}} + FN_{\text{SUBJ}}}$$
  - **F1-score (SUBJ-F1)**:
$$F1_{\text{SUBJ}} = 2 \times \frac{P_{\text{SUBJ}} \times R_{\text{SUBJ}}}{P_{\text{SUBJ}} + R_{\text{SUBJ}}}$$

The following tables present our model's performance against the official baseline for each target language. Our results demonstrate that even with a simple approach, we achieved performance above the organizers' baselines. Key improvements are highlighted in the subsequent analysis.

Table 7
Performance comparison on German (Baseline vs. Our Model)

Metric	Baseline	Our Model	Δ
Macro-F1	0.6729	0.8280	+0.1551
Macro-Precision	0.6758	0.8318	+0.1560
Macro-Recall	0.6917	0.8247	+0.1330
SUBJ-F1	0.6176	0.7706	+0.1530
SUBJ-Precision	0.5385	0.7876	+0.2491
SUBJ-Recall	0.7241	0.7542	+0.0301
Accuracy	0.6823	0.8473	+0.1650

As shown in Table 7, our model demonstrates substantial improvements over the baseline across all metrics for German. The most notable gains appear in:

- **Overall performance**: +15.51% macro-F1 (0.8280 vs. 0.6729) and +16.50% accuracy (0.8473 vs. 0.6823), indicating robust generalization.
- **SUBJ-class precision**: A remarkable +24.91% increase (0.7876 vs. 0.5385), suggesting our approach effectively reduces false positives for subjective content while maintaining recall (+3.01%).

The balanced improvement in both precision and recall (macro-P: +15. 60%, macro-R: +13. 30%) implies that our model achieves better classification consistency without sacrificing coverage. The exceptional SUBJ-class performance (F1: +15.30%) particularly highlights the effectiveness of our simple approach to handling subjective German texts.

Our model achieved remarkable German performance, securing second place out of 17 teams in the competition despite employing a relatively simple architecture.

Table 8
Performance comparison on English (Baseline vs. Our Model)

Metric	Baseline	Our Model	Δ
Macro-F1	0.7271	0.7075	-0.0196
Macro-Precision	0.7272	0.7004	-0.0268
Macro-Recall	0.7275	0.7332	+0.0057
SUBJ-F1	0.7331	0.6100	-0.1231
SUBJ-Precision	0.7457	0.5304	-0.2153
SUBJ-Recall	0.7208	0.7176	-0.0032
Accuracy	0.7273	0.7400	+0.0127

For English (Table 8), our model achieved 14th place out of 24 teams, showing modest gains in accuracy (+1.27%) and recall (+0.57%) but significant challenges in subjective content detection (SUBJ-Precision: -21.53%). This performance pattern suggests that while our simple approach generalized well for objective content, it struggled with English-specific subjective constructs, where more complex systems excelled.

Table 9
Performance comparison on Italian (Baseline vs. Our Model)

Metric	Baseline	Our Model	Δ
Macro-F1	0.6528	0.7086	+0.0558
Macro-Precision	0.6503	0.7209	+0.0706
Macro-Recall	0.6844	0.7022	+0.0178
SUBJ-F1	0.5351	0.6091	+0.0740
SUBJ-Precision	0.4470	0.6667	+0.2197
SUBJ-Recall	0.6667	0.5607	-0.1060
Accuracy	0.6927	0.7425	+0.0498

Our Italian results secured us 9th place out of 15 teams, demonstrating remarkable improvements in precision-oriented metrics, particularly for SUBJ classification (+49.15% precision gain). While the recall decrease (-15.90%) indicates our model adopted a more conservative approach to subjective content detection—correctly identifying positives but potentially missing marginal cases—this precision-focused strategy proved effective in the competition context. The balanced performance across metrics (accuracy: +7.19%, macro-F1: +8.55%) suggests our approach successfully negotiated the trade-off between false positives and coverage.

5. Conclusions and future work

In this research, simple models have been used for the task of identifying binary subjectivity in English, Italian, and German. After an exploratory analysis of the data, a logistic regression model was tested as a baseline.

Once the baseline was established, pretrained models such as BERT-base-uncased for English, BERT-base-italian-cased-sentiment for Italian, and German BERT for German were employed for analysis and classification. It was found that the Transformer-based models outperformed the traditional ones. Moreover, despite the simplicity of the technique used, the results obtained were competitive. Our best performance was second place in the German task.

In the future, we are plan to employ other strategies and techniques, including the treatment of corpus imbalance and the incorporation of linguistic elements related to subjectivity, such as sentiment analysis or the use of adjectives.

Additionally, we plan integrating reinforcement learning methods to the models to improve the performance of the algorithms.

Acknowledgments

I. Díaz thanks CONAHCYT scholarship program (CVU: 923309). This research was funded by UNAM, PAPIIT project IG400325.

Declaration on Generative AI

During the preparation of this work, the authors used DeepL and Grammarly in order to: Grammar and spelling check and correct translation to English. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] J.-C. Verstraete, Subjective and objective modality: interpersonal and ideational functions in the english modal auxiliary system, Journal of Pragmatics 33 (2001) 1505–1528. URL: https://www.sciencedirect.com/science/article/pii/S0378216601000297. doi:https://doi.org/10.1016/S0378-2166(01)00029-7.

[2] J. Wiebe, T. Wilson, R. Bruce, M. Bell, M. Martin, Learning subjective language, Computational Linguistics 30 (2004) 277–308. URL: https://aclanthology.org/J04-3002/. doi:10.1162/0891201041850885.

[3] J. Wiebe, Tracking point of view in narrative, Computational Linguistics 20 (1994) 233–287–308.

[4] V. Basile, T. Caselli, A. Balahur, L.-W. Ku, Editorial: Bias, subjectivity and perspectives in natural language processing, Frontiers in Artificial Intelligence Volume 5 - 2022 (2022). URL: https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2022.926435. doi:10.3389/frai.2022.926435.

[5] F. Alam, J. M. Struß, T. Chakraborty, S. Dietze, S. Hafid, K. Korre, A. Muti, P. Nakov, F. Ruggeri, S. Schellhammer, V. Setty, M. Sundriyal, K. Todorov, V. Venktesh, Overview of the CLEF-2025 CheckThat! Lab: Subjectivity, fact-checking, claim normalization, and retrieval, in: J. Carrillo-de Albornoz, J. Gonzalo, L. Plaza, A. García Seco de Herrera, J. Mothe, F. Piroi, P. Rosso, D. Spina, G. Faggioli, N. Ferro (Eds.), Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the Sixteenth International Conference of the CLEF Association (CLEF 2025), 2025.

[6] G. Faggioli, N. Ferro, P. Rosso, D. Spina (Eds.), Working Notes of CLEF 2025 - Conference and Labs of the Evaluation Forum, CLEF 2025, Madrid, Spain, 2025.

[7] F. Ruggeri, A. Muti, K. Korre, J. M. Struß, M. Siegel, M. Wiegand, F. Alam, R. Biswas, W. Zaghouani, M. Nawrocka, B. Ivasiuk, G. Razvan, A. Mihail, Overview of the CLEF-2025 CheckThat! lab task 1 on subjectivity in news article, in: [6], 2025.

[8] I. Maks, P. Vossen, A lexicon model for deep sentiment analysis and opinion mining applications, Decision Support Systems 53 (2012) 680–688. URL: https://www.sciencedirect.com/science/article/pii/S0167923612001364. doi:https://doi.org/10.1016/j.dss.2012.05.025, 1) Computational Approaches to Subjectivity and Sentiment Analysis 2) Service Science in Information Systems Research : Special Issue on PACIS 2010.

[9] J. Steinberger, M. Ebrahim, M. Ehrmann, A. Hurriyetoglu, M. Kabadjov, P. Lenkova, R. Steinberger, H. Tanev, S. Vázquez, V. Zavarella, Creating sentiment dictionaries via triangulation, Decision Support Systems 53 (2012) 689–694. URL: https://www.sciencedirect.com/science/article/pii/S0167923612001406. doi:https://doi.org/10.1016/j.dss.2012.05.029, 1) Computational Approaches to Subjectivity and Sentiment Analysis 2) Service Science in Information Systems Research : Special Issue on PACIS 2010.

[10] E. Riloff, J. Wiebe, Learning extraction patterns for subjective expressions, in: Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing, 2003, pp. 105–112. URL: https://aclanthology.org/W03-1014/.

[11] S.-M. Kim, E. H. Hovy, Automatic detection of opinion bearing words and sentences, in: International Joint Conference on Natural Language Processing, 2005. URL: https://api.semanticscholar.org/CorpusID:2423990.

[12] J. Wiebe, T. Wilson, C. Cardie, Annotating expressions of opinions and emotions in language, Language resources and evaluation 39 (2005) 165–210.

[13] F. Antici, F. Ruggeri, A. Galassi, K. Korre, A. Muti, A. Bardi, A. Fedotova, A. Barrón-Cedeño, A corpus for sentence-level subjectivity detection on English news articles, in: N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, N. Xue (Eds.), Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), ELRA and ICCL, Torino, Italia, 2024, pp. 273–285. URL: https://aclanthology.org/2024.lrec-main.25/.

[14] A. M. Davani, M. Díaz, V. Prabhakaran, Dealing with disagreements: Looking beyond the majority vote in subjective annotations, Transactions of the Association for Computational Linguistics 10 (2022) 92–110. URL: https://doi.org/10.1162/tacl_a_00449. doi:10.1162/tacl_a_00449.

[15] K. L. K. T. Jacob Devlin, Ming-Wei Chang, Bert: Pre-training of deep bidirectional transformers for language understanding, Association for Computational Linguistics (2019).
[16] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, 2017. URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.
[17] R. Satapathy, S. Pardeshi, E. Cambria, Polarity and subjectivity detection with multitask learning and bert embedding, 2022. URL: https://www.mdpi.com/1999-5903/14/7/191. doi:10.3390/fi14070191.
[18] M. Casanova, J. Chanson, B. Icard, G. Faye, G. Gadek, G. Gravier, P. Égré, Hybrinfox at checkthat! 2024–task 2: Enriching bert models with the expert system vago for subjectivity detection, arXiv preprint arXiv:2407.03770 (2024).
[19] M. R. Biswas, A. T. Abir, W. Zaghouani, Nullpointer at checkthat! 2024: Identifying subjectivity from multilingual text sequence, CEUR Workshop Proceedings 3740 (2024) 361–368.
[20] K. Salas-Jimenez, I. Díaz, , H. Gómez-Adorno, G. Bel-Enguix, G. Sierra, Jk_pcic_unam at checkthat! 2024: Analysis of subjectivity in news sentences using transformers-based models, CEUR Workshop Proceedings (2024).
[21] M. Shokri, V. Sharma, E. Filatova, S. Jain, S. Levitan, Subjectivity detection in english news using large language models, in: Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, 2024, pp. 215–226.
[22] B. Chan, S. Schweter, T. Möller, German's next language model, Digital Library, Munich Digitization Center (2020).
[23] NeuralyIA, neuraly/bert-base-italian-cased-sentiment, https://huggingface.co/neuraly/bert-base-italian-cased-sentiment, 2021. Accessed: 2024-05-24.