

Deep Emotions Across Languages: A Novel Approach for Sentiment Propagation in Multilingual WordNets

Jan Kocoń

Department of Artificial Intelligence

Wrocław University of Science and Technology, Poland

jan.kocon@pwr.edu.pl

Abstract—Sentiment analysis involves using WordNets enriched with emotional metadata, which are valuable resources. However, manual annotation is time-consuming and expensive, resulting in only a few WordNet Lexical Units being annotated. This paper introduces two new techniques for automatically propagating sentiment annotations from a partially annotated WordNet to its entirety and to a WordNet in a different language: Multilingual Structured Synset Embeddings (MSSE) and Cross-Lingual Deep Neural Sentiment Propagation (CLDNS). We evaluated the proposed MSSE+CLDNS method extensively using Princeton WordNet and Polish WordNet, which have many inter-lingual relations. Our results show that the MSSE+CLDNS method outperforms existing propagation methods, indicating its effectiveness in enriching WordNets with emotional metadata across multiple languages. This work provides a solid foundation for large-scale, multilingual sentiment analysis and is valuable for academic research and practical applications.

Index Terms—Sentiment Analysis, Multilingual WordNets, Multilingual Structured Synset Embeddings, Cross-Lingual Sentiment Propagation, Deep Neural Networks

I. INTRODUCTION

Sentiment analysis, the process of analyzing human emotions through text, has become increasingly popular in the age of widespread online communication. Whether it's consumer reviews, social media conversations, or news articles, recognizing and interpreting emotional tone automatically is no longer just a technological feat but a necessity for society [1]–[12]. Although deep learning techniques such as BERT [13]–[16] have significantly improved the field, achieving state-of-the-art results in various NLP tasks, the need for high-quality, annotated resources remains a crucial factor for performance [17]–[20]. This is especially challenging for under-resourced languages and cross-lingual sentiment analysis tasks.

Big Data has raised concerns about methodology and scalability. While large language models like GPT-3 and GPT-4 are impressive, they are not practical for specialized tasks like sentiment analysis. These models require significant resources and cannot replace high-quality, domain-specific training data [21]. Additionally, their monolithic size and one-size-fits-all approach are not efficient for customized solutions like sentiment propagation in multilingual WordNets.

In this paper, we introduce two new techniques, Multilingual Structured Synset Embeddings (MSSE) and Cross-Lingual Deep Neural Sentiment Propagation (CLDNS), which aim to

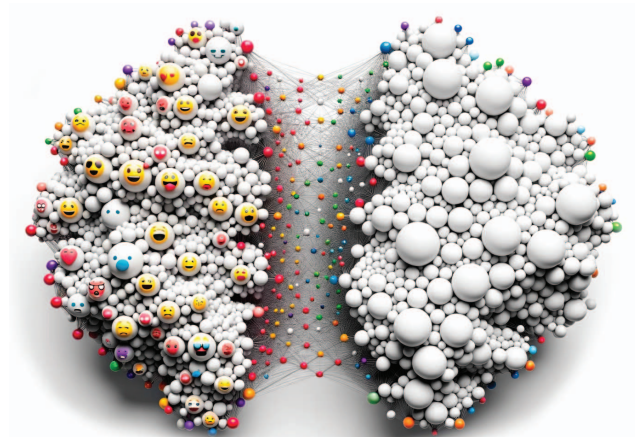


Fig. 1. Propagation of the emotional annotation from one WordNet to another. Example WordNets are connected at the level of synsets and lexical units with inter-lingual relations (Fig. generated with the assistance of DALL-E 3).

address existing gaps in sentiment annotations. These techniques help propagate sentiment annotations from a partially annotated WordNet to its entirely unannotated counterpart and a WordNet in a different language. We thoroughly evaluated the techniques using Princeton WordNet and Polish WordNet, which share a significant set of inter-lingual relations. The results showed that MSSE+CLDNS outperforms current sentiment propagation methods significantly. This improvement leads to enriching WordNets across multiple languages with high-quality emotional metadata.

Our study offers an overview of sentiment analysis resources for the Polish language. We assessed current WordNet sentiment methods and contrasted them with our new approaches. Our main contributions are: (1) introducing a new embedding technique, MSSE, capturing human emotions in graph-based databases; (2) presenting CLDNS, a method for emotion propagation in multilingual WordNets using MSSE; (3) aiming to harmonize modern NLP computational needs with the detailed demands of sentiment analysis in different languages.

II. RELATED WORK

Our work builds upon previous research in sentiment analysis, cross-lingual sentiment propagation, and knowledge base embeddings [17], [22]–[25]. Traditionally, there have been three approaches for creating sentiment lexicons: corpus-based, dictionary-based, and manual annotation [26]. However, manual annotation tends to be expensive, which has led to the development of automatic techniques to expand lexicons.

Recent work has specifically targeted the propagation of sentiment information in WordNet-like structures [27]–[31]. It’s worth noting that current efforts in the field of sentiment analysis have primarily focused on Lexical Units (LUs) or synsets, and have been limited to a single language. In contrast, our research aims to achieve cross-lingual sentiment propagation through the use of Multilingual Structured Synset Embeddings (MSSE) and Cross-Lingual Deep Neural Sentiment Propagation (CLDNS). These methods are designed to work across multiple languages and more accurately capture the broader semantic context of the text.

Embedding techniques are crucial for NLP tasks like sentiment analysis [13], [32], [33]. The latest advances in NLP, such as Transformer models and large pre-trained models like BERT [13], GPT-2, GPT-3, ChatGPT, and GPT-4 [21], [34]–[40], have set new benchmarks for a variety of tasks. Although large language models have proven to be highly efficient, their size and computational demands make them less appropriate for certain tasks, such as propagating sentiment across WordNets. Therefore, our approach seeks to overcome the limitations of traditional WordNet-based methods and large pre-trained models by combining the benefits of structured embeddings and deep learning techniques.

With the advent of large language models, there is now an opportunity to reconsider the way we utilize and expand knowledge bases. These models can be used to generate fresh knowledge and supplement existing resources, which makes it possible to automatically enhance WordNet-like lexicons [41], [42]. Typically, the techniques employed to develop models tend to be very resource-intensive, which makes them challenging to fine-tune for specific tasks like spreading sentiment across different languages [21]. Our MSSE+CLDNS approach aims to bridge this gap in the literature by combining the power of deep models with the efficiency and specificity of structured embeddings.

An increasing trend in sentiment analysis is to move beyond simplistic sentiment polarity and instead analyze the underlying emotions. [43]. The work on EmoReact [44] and GoEmotions [45] extends this further by mapping text to multiple dimensions of emotions, much like our work, which considers not just polarity but also dimensions of arousal and Plutchik’s eight basic emotions [46].

The continuous need for computational power of deep learning models is a major concern, as discussed in [47]. This issue becomes especially challenging in a cross-lingual setting, as each additional language requires a model of similar size to be trained, making the scaling problem more severe. Our

approach aims to address this by utilizing efficient graph-based techniques to propagate sentiment information across different languages, without the need for complex models.

Our work combines different research areas to advance sentiment analysis across multiple languages. We use structured embeddings and neural network architectures designed for this task. Our approach introduces innovative methods for cross-lingual sentiment propagation, providing a more efficient alternative to large language models.

III. MULTILINGUAL STRUCTURED SYNSET EMBEDDINGS

This paper proposes a solution to the problem of limited emotional data in WordNet LUs by introducing Multilingual Structured Synset Embeddings (MSSE). MSSE is an advanced embedding technique that not only captures lexical semantics but also incorporates structural information of the WordNet graph (see Algorithm 1). Our methodology is an extension and refinement of prior methods [6], [48], enriched to handle multilingual WordNets, specifically Princeton WordNet [49] and Polish WordNet [50].

Algorithm 1 MSSE

Require: WordNet graph G , Number of dimensions d , Random walk length l , Number of random walks n , Skip-gram window size w

Ensure: Embeddings for WordNet nodes and edges

```
1: function GENERATECORPUS( $G, I$ )
2:   Initialize an empty corpus  $C$ 
3:   for  $i$  in range( $n$ ) do
4:     Randomly select a starting node  $n$  in  $G$ 
5:     Perform a random walk of fixed length  $l$ , recording
       a sequence of all visited node IDs and edge IDs, ensuring
       each node in a sequence is visited only once
6:     Add the sequence to  $C$ 
7:   end for
8:   return  $C$ 
9: end function
10: function TRAINEMBEDDINGS( $C, d$ )
11:   Initialize empty embedding model
12:   for each node sequence in  $C$  do
13:     Update the embedding model using Skip-gram
       with context size  $w$ 
14:   end for
15:   return Embeddings for WordNet nodes and edges
16: end function
```

MSSE stands out from other approaches for its unique training methodology. Unlike other models, it doesn’t rely on pre-existing models like cross-lingual FastText embeddings [51]. Instead, it is trained from scratch, using only the textual representations of the WordNet nodes and edges. To achieve multilingual alignment, we leverage the dense inter-lingual relations between Princeton WordNet and Polish WordNet [52]–[54]. To conduct random walks across both WordNets, we utilize a Skip-gram algorithm [55]. This neural architecture typically generates word embeddings by predicting local

context based on a large textual corpus. The model assumes that semantically related words often occur in similar contexts, allowing the embeddings to capture such relationships [56]. However, in contrast to the standard Skip-gram approach, we create an artificial corpus from WordNet, based on Goikoetxea et al. work [48]. This corpus is formed through random walks on the WordNet graph, which captures the subtle interplay between synsets (groups of cognitive synonyms), lexical units, and relation types. The embedding of each node is determined by the context along the random walk path. Nodes that are inter-lingual synonyms of each other often appear in close proximity within the context window, which leads to the creation of similar embeddings. This illustrates the potential of MSSE to uncover meaningful multilingual parallels.

A major innovation in MSSE is incorporating inter-WordNet relations, a feature largely overlooked in previous studies [48]. The dense links between Princeton WordNet and Polish WordNet, which were developed under the CLARIN-PL project [50], are crucial. Their extensive inter-connectedness allows us to perform random walks within a single WordNet and across different language-specific WordNets. Furthermore, MSSE includes the type of relation that connects different units or synsets, adding another layer of semantic information to the embeddings.

Our artificial corpus includes identifiers for synsets, lexical units, and various relation types. For example, relation types are labeled with prefixes such as "rSS" for synset-to-synset relations or "rSL" for synset-to-lexical unit relations, similarly as in [31]. Additional markers indicate the specific relation type, such as hyponymy or hypernymy. This approach allows for a more detailed analysis of embeddings, which become semantically rich, capturing the concepts and their intricate relationships within and across languages.

We use the FastText algorithm [32] to create embeddings. This algorithm is known for its effectiveness in various NLP tasks, including sentiment analysis. Our setup generates 300-dimensional embeddings that capture various lexical and emotional nuances. This high-dimensional representation is more informative and versatile, making it suitable for complex tasks like sentiment propagation across multilingual WordNets.

We chose FastText over other options like Word2Vec [57] and GloVe [58] due to its unique ability to capture subword information. FastText represents words using character n-grams, allowing it to handle out-of-vocabulary words and effectively capture variations in word structure. This feature is especially useful in multilingual contexts where languages have diverse morphological structures. Furthermore, FastText has demonstrated strong performance in various NLP tasks [51], [59]–[61], making it a reliable choice for our multilingual sentiment propagation framework.

MSSE is a reliable and efficient multilingual tool for large-scale sentiment analysis. It captures words' emotional undertones and lexical semantics, making it an ideal option for sentiment analysis. When combined with our CLDNS, MSSE far outperforms other sentiment propagation methods.

IV. CROSS-LINGUAL DEEP NEURAL SENTIMENT PROPAGATION

In this section, we will discuss Cross-Lingual Deep Neural Sentiment Propagation (CLDNSP), a mechanism designed to automatically propagate sentiment annotations across different WordNets using the framework of Multilingual Structured Synset Embeddings (MSSE). We will explore the intricacies of CLDNSP, a method for automating the propagation of sentiment annotations across WordNets of different languages.

We address the limitations of previous methods that utilized a simple logistic regression [31] by introducing a deep neural network-based multilabel classifier. Specifically, we propose two configurations:

- 1) *Base Configuration* is based on [31] and consists of a two-layer neural network: an input layer with 300 neurons and an output layer comprising 26 neurons to capture the multiple dimensions of sentiment.
- 2) *Deep Configuration* builds on the Base by adding multiple layers and dropout for regularization, configured as follows: an initial dense layer with 300 inputs and 4096 outputs, followed by a 20% dropout layer; a subsequent dense layer with 1024 neurons and another 20% dropout layer; a penultimate dense layer with 256 neurons; and finally, an output layer with 26 neurons.

In both configurations, we employ the ReLU activation function for hidden layers and a linear activation function for the output layer. We use the coefficient of determination as a loss function and early stopping with patience equal to 30.

The proposed methodology advances multilingual sentiment analysis research by efficiently transferring sentiment annotation between WordNets. It eliminates the need for manual annotations and provides a scalable approach to enrich WordNets with emotional metadata across languages. The CLDNSP method leverages the inter-lingual richness offered by MSSE and adopts a deep neural architecture for sentiment propagation, taking us closer to achieving automated, fine-grained emotional understanding across languages. Our preliminary results suggest that we are on a promising trajectory toward mapping the complex landscape of human emotions across languages.

V. EMOTIONAL ANNOTATION IN POLISH WORDNET

For the experiments, we decided to use the plWordNet 3.0 dataset [62]. The emotional annotation process was conducted to enhance the lexical resource with emotive lexicons, which included sentiment polarities, fundamental human values, and basic emotions. This was done to improve its application in emotional and sentiment analysis. The annotations were manually done by linguists and psychologists while following consistent guidelines. These guidelines included the usage of corpora, substitution tests, and lexical unit properties, as presented in [62]. The annotation dimensions that were used included:

- **Sentiment Polarity (pol):** Encoded on a 6-grade scale distinguishing between strong and weak polarisation:

Strong Positive, Weak Positive, Strong Negative, Weak Negative, Ambivalent, Neutral

- **Basic Emotions (emo):** Based on Plutchik’s framework [46], including Joy, Fear, Surprise, Sadness, Disgust, Anger, Trust, and Anticipation.
- **Fundamental Human Values (val):** Influenced by Puzynina’s work [63] and consists of: Non-usefulness, Mistake, Ugliness, Goodness, Harm, Ignorance, Unhappiness, Beauty, Truth, Happiness, Usefulness, and Knowledge.

Over 31,000 LUs were annotated in plWordNet 4.0 emo. The inter-annotator agreement for sentiment polarity was notably high, encompassing negative (strong and weak), neutral, and positive (strong and weak) polarities. Negative emotions and values were predominant, aligning with trends in other sentiment lexicons. Notably, there are almost 300k manually annotated interlingual relations between Princeton WordNet and plWordNet [54].

VI. EVALUATION

Our evaluation framework is unique in its approach, focusing on multiple levels of WordNet concepts and employing a robust and comprehensive MSSE embedding method to obtain common embeddings for densely connected plWordNet and Princeton WordNet. As a reference embedding method, we used Heterogeneous Structured Synset Embeddings (HSSE) described in [31] with embeddings trained using plWordNet only (with no connections to Princeton WordNet). The propagation procedure is the same as in [31], but instead of logistic regression, we used CLDNSP in two variants (Base and Deep, see Section IV).

We utilize 10-fold cross-validation and propagate emotional metadata from an initial seed (train set) that contains 80% of annotated lexical units (LUs) in plWordNet; a validation set containing 10% of LUs (for early stopping) and 10% in the test set. Our propagation method adapts the approach from [30]. At first, we train the CLDNSP classifier on an initial seed of LUs, where the input for the model is the LU embedding (obtained via HSSE or MSSE), and the output is the multidimensional emotional annotation of the LU. Then, we apply the classifier to all nearest neighbors of seed LUs and repeat the procedure until we reach all LUs in the test set. Finally, we calculate the precision, recall, and F1 score on a test set and repeat the whole procedure on another set of folds. In the end, we have a sample of 10 results. The code is available in GitHub repository¹, and the data can be obtained here² upon request.

A. Statistical Significance Tests

We use the Shapiro-Wilk test to determine if our results are normally distributed. Through this test, we observe p-values higher than the level of significance ($\alpha = 0.05$) for each set of folds in our 10-fold cross-validation. As a result, we fail to reject the null hypothesis that our results are normally

¹<https://github.com/KoconJan/deeprop>

²<http://plwordnet.pwr.edu.pl/wordnet>

distributed. We then use the paired-differences Student’s t-test to check the statistical significance of the observed differences, using a significance level of $\alpha = 0.05$ [64].

VII. RESULTS

A. Comparison of Configurations

Table I presents a comparative analysis of four different configurations employing combinations of embeddings (MSSE or HSSE) and classifiers (Base or Deep). These configurations are denoted as HB, HD, MB, and MD:

- HB – HSSE with Base classifier [31],
- HD – HSSE with Deep classifier (CLDNSP),
- MB – MSSE with Base classifier [31],
- MD – MSSE with Deep classifier (CLDNSP).

Our statistical tests show that there is no significant difference between the HD and MB configurations. However, both of these configurations are significantly different from the HB configuration. This implies that either the deep classifier or the MSSE method (or both) contribute to enhanced performance. The MD configuration yields better F1-score values, for almost all emotional dimensions.

TABLE I
COMPARISON OF THE RESULTS BETWEEN DIFFERENT CONFIGURATIONS (H - HSSE, M - MSSE, B - BASE, D - DEEP (CLDNSP)). MEASURES USED ARE R AND R-SQUARED ($R^2 = 1 - FVU$).

Measure	HB	HD	MB	MD
R	0.628±0.008	0.641±0.015	0.651±0.007	0.693±0.007
R^2	0.394±0.010	0.412±0.019	0.423±0.009	0.480±0.010

B. Impact of Cross-Lingual Information

In Table II, we compare the HD and MD configurations to assess the impact of cross-lingual information. Emotional dimensions are described in Section V. The results show a consistent improvement in performance across all sentiment, emotional, and valuation categories for the MD setup. This highlights the significance of interlingual relations between the Polish and English synsets, emphasizing the importance of cross-lingual information.

C. Non-propagated vs Propagated Configurations

Table III compares non-propagated (HSSE/MSSE) and propagated (HD/MD) configurations. Utilizing the deep classifier (CLDNSP) leads to a significant performance improvement, regardless of the type of embeddings used (monolingual or cross-lingual).

VIII. CONCLUSIONS

Our study highlights the effectiveness of MSSE and CLDNSP, our proposed methods for sentiment and emotion propagation in multilingual WordNets. We found that the MD configuration, which uses both MSSE and CLDNSP, outperforms other configurations in terms of performance metrics. This emphasizes the significance of integrating cross-lingual information and deep neural networks in sentiment

TABLE II
COMPARISON OF THE RESULTS BETWEEN HD AND MD. **BOLD** VALUES INDICATE THE BEST PERFORMANCE AMONG ALL ARCHITECTURES.

Emotion	HD			MD		
	P	R	F1	P	R	F1
val_nonusefulness	67.6±0.9	70.0±1.2	68.8±0.8	68.0±1.0	72.2±1.9	69.8±0.4
val_mistake	62.8±2.6	64.8±1.5	63.6±1.1	64.2±1.6	64.8±3.3	64.4±1.1
val_ugliness	65.4±3.7	47.6±3.2	55.0±3.5	67.2±3.4	45.6±3.0	54.4±3.0
emo_anticipation	42.0±2.7	31.0±3.2	35.2±1.8	37.2±2.3	35.2±3.1	36.0±1.2
val_goodness	47.0±2.3	46.0±2.2	46.2±1.3	48.2±6.9	48.2±3.3	48.0±3.7
val_harm	59.8±2.0	64.6±1.1	62.0±1.4	61.0±3.1	66.2±3.0	63.4±1.5
val_ignorance	59.2±5.8	37.0±4.8	45.4±5.0	60.4±3.6	36.6±6.2	45.4±5.0
val_unhappiness	55.4±1.3	60.2±1.9	57.6±0.9	55.6±2.6	62.2±3.1	58.8±1.8
val_beauty	63.4±2.3	46.2±3.5	53.4±2.1	63.4±5.8	48.0±4.7	54.4±3.6
val_truth	51.0±2.8	28.6±3.1	36.6±2.9	51.6±6.7	33.8±4.9	40.4±3.5
emo_joy	67.4±0.5	60.6±1.7	64.0±0.7	67.2±1.8	63.8±1.8	65.4±1.1
emo_sadness	62.4±1.5	69.0±0.7	65.2±0.8	63.4±1.1	70.2±1.5	66.6±0.9
val_happiness	58.0±1.7	58.4±2.3	58.0±1.4	57.6±3.0	60.6±2.7	59.0±1.6
emo_fear	52.8±3.6	40.4±1.9	46.0±1.6	52.8±5.0	43.8±4.4	47.8±1.1
val_usefulness	59.8±2.5	59.2±2.7	59.4±1.7	61.2±2.2	61.2±1.9	61.2±0.8
val_knowledge	59.6±0.9	44.4±4.3	50.8±2.6	60.4±4.0	44.4±3.8	51.0±2.2
emo_disgust	60.6±2.9	62.8±2.0	61.6±1.7	63.2±2.2	61.2±2.6	62.2±1.3
emo_trust	50.2±2.8	50.2±3.0	50.2±2.3	53.2±3.8	50.8±2.9	52.0±1.6
emo_surprise	45.2±8.1	22.2±4.4	29.4±4.8	40.2±8.9	21.8±4.1	28.4±5.7
emo_anger	68.8±0.8	72.2±2.3	70.6±0.9	69.0±0.7	73.2±1.6	71.2±0.8
pol_strong_positive	59.2±3.3	52.8±2.8	55.8±2.0	61.0±3.8	52.0±3.5	56.0±2.5
pol_weak_positive	54.8±1.3	55.4±3.4	55.2±1.8	57.4±2.5	56.8±2.5	57.4±0.9
pol_strong_negative	56.0±4.1	64.2±2.0	60.0±2.0	57.2±1.5	64.2±2.8	60.4±1.1
pol_weak_negative	54.8±0.8	65.0±2.5	59.2±0.8	56.0±1.9	67.0±2.2	61.0±0.7
pol_ambivalent	53.0±1.2	36.2±2.5	43.0±1.2	50.8±3.8	38.0±2.6	43.8±1.5
pol_neutral	88.2±0.4	95.2±0.8	91.6±0.5	89.4±0.5	94.4±0.9	91.8±0.4
micro	67.4±1.1	68.4±0.9	67.8±0.8	67.6±1.3	69.2±1.5	68.4±0.5
macro	58.6±1.3	54.2±1.1	55.8±0.5	58.8±1.3	55.2±1.5	56.4±0.5
weighted	66.8±1.1	68.4±0.9	67.0±0.7	67.4±1.1	69.2±1.5	68.2±0.4

TABLE III
COMPARISON OF HSSE VS. HD AND MSSE VS. MD CONFIGURATIONS. **BOLD** VALUES INDICATE THE BEST PERFORMANCE AMONG ALL ARCHITECTURES.

Avg.type	HSSE			HD		
	P	R	F1	P	R	F1
micro	66.8±1.1	67.8±0.8	67.4±0.5	67.4±1.1	68.4±0.9	67.8±0.8
macro	57.4±1.1	53.4±1.1	54.8±0.8	58.6±1.3	54.2±1.1	55.6±0.5
weighted	66.2±0.8	67.8±0.8	66.8±0.8	66.8±1.1	68.4±0.9	67.0±0.7
	MSSE			MD		
	P	R	F1	P	R	F1
micro	67.6±1.9	68.2±1.3	68.0±0.7	67.6±1.3	69.2±1.5	68.4±0.5
macro	59.0±1.9	54.4±1.5	55.6±0.5	58.8±1.3	55.2±1.5	56.4±0.5
weighted	67.4±1.7	68.2±1.3	67.6±0.5	67.4±1.1	69.2±1.5	68.2±0.4

propagation tasks. Furthermore, our results indicate that our propagation method is highly effective, as it delivers considerable gains regardless of the embeddings used.

As part of further research, we recommend two steps. Firstly, we suggest extending the MSSE embedding in a similar manner as the HSSE method, but this time using multilingual aligned embeddings [65]. Secondly, we recommend testing the impact of the propagated dataset on the quality of sentiment recognition in text.

ACKNOWLEDGEMENTS

This work was financed by (1) contribution to the European Research Infrastructure "CLARIN ERIC - European Research Infrastructure Consortium: Common Language Resources and Technology Infrastructure", 2022-23 (CLARIN Q); (2) the European Regional Development Fund, as a part of the 2014-2020 Smart Growth Operational Programme, projects no. POIR.04.02.00-00C002/19, POIR.01.01.01-00-0923/20, POIR.01.01.01-00-0615/21, and POIR.01.01.01-00-0288/22; (3) the statutory funds of the Department of Artificial

Intelligence, Wrocław University of Science and Technology; (4) the Polish Ministry of Education and Science within the programme "International Projects Co-Funded"; (5) the European Union under the Horizon Europe, grant no. 101086321 (OMINO). However, the views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor European Research Executive Agency can be held responsible for them.

REFERENCES

- [1] J. Ramteke, S. Shah, D. Godhia, and A. Shaikh, "Election result prediction using twitter sentiment analysis," in *International conference on inventive computation technologies (ICICT)*, vol. 1. IEEE, 2016.
- [2] E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco, "Affective computing and sentiment analysis," *A practical guide to sentiment analysis*, pp. 1–10, 2017.
- [3] J. Kocoń, P. Miłkowski, M. Wierzba, B. Konat *et al.*, "Multilingual and language-agnostic recognition of emotions, valence and arousal in large-scale multi-domain text reviews," in *Language and Technology Conference*. Springer, 2019, pp. 214–231.
- [4] J. Kocoń, M. Zaśko-Zielińska, and P. Miłkowski, "Multi-level analysis and recognition of the text sentiment on the example of consumer opinions," in *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*, 2019.
- [5] J. Kocoń, P. Miłkowski, and M. Zaśko-Zielińska, "Multi-level sentiment analysis of polemo 2.0: Extended corpus of multi-domain consumer reviews," in *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, 2019, pp. 980–991.
- [6] J. Kocoń, A. Janz *et al.*, "Recognition of emotions, valence and arousal in large-scale multi-domain text reviews," *Human language technologies as a challenge for computer science and linguistics*, 2019.
- [7] K. Kanclerz, P. Miłkowski, and J. Kocoń, "Cross-lingual deep neural transfer learning in sentiment analysis," *Procedia Computer Science*, vol. 176, pp. 128–137, 2020.
- [8] M. Wierzba, M. Riegel, J. Kocoń, P. Miłkowski *et al.*, "Emotion norms for 6000 polish word meanings with a direct mapping to the polish wordnet," *Behavior Research Methods*, pp. 1–16, 2021.
- [9] J. Kocoń, J. Radom, E. Kaczmarz-Wawryk, K. Wabnic *et al.*, "Aspectemo: multi-domain corpus of consumer reviews for aspect-based sentiment analysis," in *2021 International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2021, pp. 166–173.
- [10] J. Kocoń, P. Miłkowski, and K. Kanclerz, "Multiemo: Multilingual, multilevel, multidomain sentiment analysis corpus of consumer reviews," in *International Conference on Computational Science*. Springer, 2021.
- [11] E. Cambria, Q. Liu, S. Decherchi, F. Xing, and K. Kwok, "Senticnet 7: A commonsense-based neurosymbolic ai framework for explainable sentiment analysis," in *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 2022, pp. 3829–3839.
- [12] J. Baran and J. Kocoń, "Linguistic knowledge application to neuro-symbolic transformers in sentiment analysis," in *2022 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2022.
- [13] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.
- [14] W. Korczyński and J. Kocoń, "Compression methods for transformers in multidomain sentiment analysis," in *2022 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2022, pp. 419–426.
- [15] J. Kocoń, J. Baran, M. Gruza, A. Janz, M. Kajstura, P. Kazienko, W. Korczyński, P. Miłkowski, M. Piasecki, and J. Szolomicka, "Neuro-symbolic models for sentiment analysis," in *International conference on computational science*. Springer, 2022, pp. 667–681.
- [16] A. Srivastava, A. Rastogi, A. Rao, A. A. M. Shobh, A. Abid, A. Fisch, A. R. Brown, A. Santoro, A. Gupta, A. Garriga-Alonso *et al.*, "Beyond the imitation game: Quantifying and extrapolating the capabilities of language models," *Transactions on Machine Learning Research*, 2023.
- [17] K. Dashtipour, S. Poria, A. Hussain, E. Cambria *et al.*, "Multilingual sentiment analysis: state of the art and independent comparison of techniques," *Cognitive computation*, vol. 8, pp. 757–771, 2016.

- [18] J. Kocoń and M. Maziarz, "Mapping wordnet onto human brain connectome in emotion processing and semantic similarity recognition," *Information Processing & Management*, vol. 58, no. 3, p. 102530, 2021.
- [19] P. Miłkowski, M. Gruza, K. Kanclerz, P. Kazienko, D. Grimling, and J. Kocoń, "Personal bias in prediction of emotions elicited by textual opinions," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics: SRW*, 2021.
- [20] P. Miłkowski, S. Saganowski, M. Gruza, P. Kazienko, M. Piasecki, and J. Kocoń, "Multitask personalized recognition of emotions evoked by textual content," in *2022 IEEE International Conference on Pervasive Computing and Communications*. IEEE, 2022.
- [21] J. Kocoń, I. Cichecki, O. Kaszyca, M. Kochanek, D. Szydło, J. Baran, J. Bielaniewicz, M. Gruza, A. Janz, K. Kanclerz *et al.*, "Chatgpt: Jack of all trades, master of none," *Information Fusion*, p. 101861, 2023.
- [22] T. Brychcín and I. Habernal, "Unsupervised improving of sentiment analysis using global target context," in *Proceedings of the international conference recent advances in natural language processing*, 2013.
- [23] R. He and J. McAuley, "Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering," in *proceedings of the 25th international conference on world wide web*, 2016.
- [24] J. Bučar, M. Žnidarič, and J. Povh, "Annotated news corpora and a lexicon for sentiment analysis in slovene," *Language Resources and Evaluation*, vol. 52, pp. 895–919, 2018.
- [25] A. Rogers, A. Romanov, A. Rumshisky, S. Volkova, M. Gronas, and A. Gribov, "Rusentiment: An enriched sentiment analysis dataset for social media in russian," in *Proceedings of the 27th international conference on computational linguistics*, 2018, pp. 755–763.
- [26] J. Zhao, K. Liu, and L. Xu, "Sentiment analysis: mining opinions, sentiments, and emotions," 2016.
- [27] I. Maks and P. Vossen, "Different approaches to automatic polarity annotation at synset level," in *Proceedings of the First International Workshop on Lexical Resources*, 2011, pp. 62–69.
- [28] F. Sebastiani and A. Esuli, "Sentiwordnet: A publicly available lexical resource for opinion mining," in *Proceedings of the 5th international conference on language resources and evaluation*. European Language Resources Association (ELRA) Genoa, Italy, 2006, pp. 417–422.
- [29] A. Janz, J. Kocon, M. Piasecki, and M. Zasko-Zielinska, "plwordnet as a basis for large emotive lexicons of polish," *Proceedings of Human Language Technologies as a Challenge for Computer Science and Linguistics*, pp. 189–193, 2017.
- [30] J. Kocoń, A. Janz, and M. Piasecki, "Classifier-based Polarity Propagation in a Wordnet," in *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC'18)*, 2018.
- [31] J. Kocoń and A. Janz, "Propagation of emotions, arousal and polarity in wordnet using heterogeneous structured synset embeddings," in *Proceedings of the 10th Global Wordnet Conference*, 2019.
- [32] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," *Transactions of the association for computational linguistics*, vol. 5, pp. 135–146, 2017.
- [33] P. Miłkowski, M. Gruza, P. Kazienko, J. Szołomicka, S. Woźniak, and J. Kocoń, "Multi-model analysis of language-agnostic sentiment classification on multitemo data," in *International Conference on Computational Collective Intelligence*. Springer, 2022, pp. 163–175.
- [34] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever *et al.*, "Language models are unsupervised multitask learners," *OpenAI blog*, vol. 1, no. 8, p. 9, 2019.
- [35] T. Brown, B. Mann, N. Ryder *et al.*, "Language models are few-shot learners," *Advances in neural information processing systems*, 2020.
- [36] P. Kazienko, J. Bielaniewicz, M. Gruza, K. Kanclerz, K. Karanowski, P. Miłkowski, and J. Kocoń, "Human-centered neural reasoning for subjective content processing: Hate speech, emotions, and humor," *Information Fusion*, vol. 94, pp. 43–65, 2023.
- [37] B. Koptyra, A. Ngo, Ł. Radliński, and J. Kocoń, "Clarín-emo: Training emotion recognition models using human annotation and chatgpt," in *International Conference on Computational Science*. Springer, 2023.
- [38] J. Kocoń, J. Baran, K. Kanclerz, M. Kajstura, and P. Kazienko, "Differential dataset cartography: Explainable artificial intelligence in comparative personalized sentiment analysis," in *International Conference on Computational Science*. Springer, 2023, pp. 148–162.
- [39] T. Ferdinan and J. Kocoń, "Personalized models resistant to malicious attacks for human-centered trusted ai," in *The AAAI-23 Workshop on Artificial Intelligence Safety (SafeAI 2023)*. CEUR Workshop Proceedings, 2023.
- [40] W. Mielezszchenko-Kowszewicz, K. Kanclerz, J. Bielaniewicz, M. Oleksy, M. Gruza, S. Woźniak, E. Dzieciół, P. Kazienko, and J. Kocoń, "Capturing human perspectives in nlp: Questionnaires, annotations, and biases," in *ECAI 2023 2nd Workshop on Perspectivist Approaches to NLP*. CEUR Workshop Proceedings, 2023.
- [41] F. Petroni, T. Rocktäschel, P. Lewis, A. Bakhtin, Y. Wu, A. H. Miller, and S. Riedel, "Language models as knowledge bases?" *arXiv preprint arXiv:1909.01066*, 2019.
- [42] D. Kiela, M. Bartolo, Y. Nie, D. Kaushik, A. Geiger, Z. Wu, B. Vidgen, G. Prasad, A. Singh, P. Ringshia *et al.*, "Dynabench: Rethinking benchmarking in nlp," *arXiv preprint arXiv:2104.14337*, 2021.
- [43] S. Mohammad, F. Bravo-Marquez, M. Salameh, and S. Kiritchenko, "Semeval-2018 task 1: Affect in tweets," in *Proceedings of the 12th international workshop on semantic evaluation*, 2018, pp. 1–17.
- [44] B. Felbo, A. Misllove, A. Søgaard, I. Rahwan, and S. Lehmann, "Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm," *arXiv preprint arXiv:1708.00524*, 2017.
- [45] D. Demszky, D. Movshovitz-Attias, J. Ko, A. Cowen, G. Nemade, and S. Ravi, "Goemotions: A dataset of fine-grained emotions," *arXiv preprint arXiv:2005.00547*, 2020.
- [46] R. Plutchik, "A psychoevolutionary theory of emotions," 1982.
- [47] E. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in nlp," *arXiv:1906.02243*, 2019.
- [48] J. Goikoetxea, A. Soroa, and E. Agirre, "Random walks and neural network language models on knowledge bases," in *Proceedings of the 2015 conference of the North American Chapter of the Association for Computational Linguistics: Human language technologies*, 2015.
- [49] G. A. Miller, "Wordnet: a lexical database for english," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.
- [50] M. Piasecki, B. Broda, and S. Szpakowicz, *A wordnet from the ground up*. Oficyna Wydawnicza Politechniki Wrocławskiej Wrocław, 2009.
- [51] E. Grave, P. Bojanowski, P. Gupta, A. Joulin, and T. Mikolov, "Learning word vectors for 157 languages," in *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018)*, 2018.
- [52] E. K. Rudnicka, M. T. Piasecki, T. Piotrowski, Ł. Grabowski, and F. Bond, "Mapping wordnets from the perspective of inter-lingual equivalence," *Cognitive Studies*, no. 17, 2017.
- [53] E. Rudnicka, M. Piasecki, F. Bond, Ł. Grabowski, and T. Piotrowski, "Sense equivalence in plwordnet to princeton wordnet mapping," *International Journal of Lexicography*, vol. 32, no. 3, pp. 296–325, 2019.
- [54] E. Rudnicka, W. Witkowski, and M. Piasecki, "A (non)-perfect match: Mapping plwordnet onto princetonwordnet," in *Proceedings of the 11th Global Wordnet Conference*, 2021, pp. 137–146.
- [55] D. Guthrie, B. Allison, W. Liu, L. Guthrie, and Y. Wilks, "A closer look at skip-gram modelling," in *LREC*, vol. 6, 2006, pp. 1222–1225.
- [56] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," *Advances in neural information processing systems*, 2013.
- [57] K. W. Church, "Word2vec," *Natural Language Engineering*, vol. 23, no. 1, pp. 155–162, 2017.
- [58] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014.
- [59] J. Kocoń and M. Gawor, "Evaluating kgr10 polish word embeddings in the recognition of temporal expressions using bilstm-crf," *arXiv preprint arXiv:1904.04055*, 2019.
- [60] V. Laippala, R. Kyllönen, J. Egbert, D. Biber, and S. Pyysalo, "Toward multilingual identification of online registers," in *Proceedings of the 22nd Nordic Conference on Computational Linguistics*, 2019.
- [61] M. M. Agüero-Torales, J. I. A. Salas, and A. G. López-Herrera, "Deep learning and multilingual sentiment analysis on social media data: An overview," *Applied Soft Computing*, vol. 107, p. 107373, 2021.
- [62] M. Zaško-Zielińska and M. Piasecki, "Towards emotive annotation in plwordnet 4.0," in *Proceedings of the Global Wordnet Conference*, 2018.
- [63] J. Puzynina, "Język wartości [eng. the language of values]," *Scientific Publishers PWN*, 1992.
- [64] T. G. Dietterich, "Approximate statistical tests for comparing supervised classification learning algorithms," *Neural computation*, 1998.
- [65] A. Joulin, P. Bojanowski, T. Mikolov, H. Jégou, and E. Grave, "Loss in translation: Learning bilingual word mapping with a retrieval criterion," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018.