An Exploration of LLM-Guided Detection of Discursive Patterns in Dutch Social Media

Swarupa Hardikar[1,2,*,†], Maya Sappelli[1,*,†] and Annette Klarenbeek[3]

[1]HAN University of Applied Sciences, Ruitenberglaan 26, 6826 CC Arnhem, The Netherlands
[2]Radboud University, Toernooiveld 212, 6525 EC, Nijmegen, The Netherlands
[3]Hogeschool Utrecht, Utrecht, The Netherlands

Abstract

This paper presents a generative large language model (LLM)- guided approach to detect discursive patterns in Dutch social media. Newsrooms of municipalities and public organizations follow public debate on social media to be aware of and prepare for local and global issues and rumours. The onset of these issues and rumours are now detected by communication specialists in newsrooms. Using discourse analysis, we can ground their findings in theory. Devices from discursive psychology such as emotional evaluations are the lowest level components that can help spot and understand issues[1]. Thus, a rule-based NLP approach was developed to highlight these devices in a learning environment[1]. As a next step, we compare the rule-based approach to a large language modeling approach in order to assess the risks and benefits of both methods. We analyze the detection of two discursive patterns in Dutch tweets: magnifying (exaggerated) language use and assigning negative labels to persons or organizations. We compare the generated responses from two Dutch conversational LLMs finetuned on the Dutch language - Geitje-7B-Ultra and Fietje-2-Chat - and a rule-based NLP baseline using a two-fold evaluation process. The results show mixed performance, with the highest performing LLM setups yielding an accuracy of 64% for the maximizing language category and 73% for the negative labels to organizations/persons category. In comparison, the rule-based algorithm achieves an accuracy of 68% for both categories. Although the LLMs perform well in precision, they frequently find patterns in examples where no discursive markers were annotated. Moreover, the rationale analysis shows relatively poor results, pertaining to multiple factors including model size and interpretation of instructions. The results indicate that although there is merit in conducting discursive analysis using generative language models, it comes with the above risks. Recommendations for future work include combining the usage of language models with the rule-based setup for more robust detection as well as further indication of guidelines to improve upon the reasoning process.

Keywords

Discursive pattern detection, Large Language Models, social media analysis

1. Introduction

Social media is an important source of information for newsrooms, not only for news production and dissemination, but also for managing the effects of potential rumors and issues locally. Therefore, newsrooms of municipalities and public organizations follow public debate to be aware of and prepare for local and global issues and rumors. The onset of these issues and rumors are now detected by communication specialists in newsrooms. They can be supported by making sense of what role each message plays in the entire conversation using discourse analysis. This can help spot changes in the discourse, such as changes in topic, relevance or effects of the online discussion in the offline world.

Devices from discursive psychology such as emotional evaluations are the lowest level components that can help spot and understand issues [1]. The identification of these components can be facilitated through the training of communication specialists. In [2], a rule-based NLP approach was developed which highlights a subset of psychological devices in Dutch social media messages. They have developed

[1]https://husite.nl/bep/

EKAW 2024: EKAW 2024 Workshops, Tutorials, Posters and Demos, 24th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2024), November 26-28, 2024, Amsterdam, The Netherlands.

*Corresponding author.
†These authors contributed equally.
swarupa.hardikar@han.nl (S. Hardikar); maya.sappelli@han.nl (M. Sappelli); annette.klarenbeek@hu.nl (A. Klarenbeek)

the approach with the purpose of making a learning environment[1] and the learning experience more interactive.

Some aspects of detecting psychological devices are particularly challenging because of required knowledge of changing contexts that are difficult to capture in a rule-based approach. Generative large language models (LLMs) have recently gained prominence in their use for various NLP tasks, especially because of their reasoning capabilities in application to specific use cases and show promise with analyzing changing contexts. On the other hand, their generative nature means that they are not suited for all analysis tasks.

In this explorative experiment, we aim to compare the use of LLMs for detecting psychological devices to the use of the rule-based NLP algorithm. Moreover, we explore the reasoning capabilities of small open source LLMs, especially in regards to Dutch text. Finally, we also look at the effect of prompt engineering techniques towards the reasoning process in such a task.

To this effect, we can construct the primary research question: How do open source small Dutch LLMs compare against rule-based systems in detecting discursive patterns in Dutch social media?

The research question is supplemented by the following sub-questions:

1. How well can small LLMs reason when given psychologically rich input?
2. Can innate bias hamper the decision-making processes of LLMs?
3. How do prompt engineering techniques influence small LLM reasoning?

2. Relevant Literature

There is ample research on analyzing social media messages from various perspectives. These analyses and methods can be used to tag and enrich messages in order to improve the sense-making of these messages down the line. In this paper, we focus on elements that are related to the onset or development of rumors and societal issues.

One such aspect is the the organization of an event such as a demonstration. Although event detection has been researched extensively [3, 4, 5], they are primarily focused on big or planned events rather than spontaneous emergence of demonstrations.

Another relevant aspect is the detection of whether a tweet is a rumor or not, which helps to understand the truth value of a message. In this regard, Zhao et al. [6] use regular expressions to detect tweets containing rumors based on linguistic expressions such as ”is this true?”, ”really?”. In contrast, Ma et al. [7] use a recurrent neural network approach on groups of tweets using aggregated information across different time intervals to detect rumors. Finally, Tian et al. [8] focus on the early detection of rumors, using stance learning as a measure for user attitude in comments in combination content analysis. Their CNN (convolutional neural network) and BERT neural networks outperform RNN (recurrent neural network) models and simpler models such as SVM (support vector machines).

More recently, LLMs have been used in social media analysis as well: [9] use LLMs to identify issues raised by citizens based on training data obtained from both human annotators and data augmentation using GPT (generative pre-trained transformer). The focus of the issues is on the topical identification of challenges faced by cyclists. The GPT is presented with a prompt to classify the tweet which includes the categories to choose from. Prompts with more examples outperformed the zero shot setting.

While the above methods are focused at messages as a whole, we are interested in communication strategies used within a messages - these can reveal specific intentions of authors, which can support sense-making from an intent-centric perspective. Many strategies can be derived from the discursive framework, but certain strategies such as assigning a negative label to a person or organization, emotional evaluation, factual language use or intensifying language, as used in the rule-based algorithm [2], have not been investigated in NLP research before. However, there are many aspects in social media that have been researched that may be related.

[1]https://husite.nl/bep/

An interesting aspect is cyber-bullying, which could be related to the strategy of 'assigning a negative label to a person or organization'. Sainju et al. [10] take a keyword-based approach to detect bullying messages, achieving a performance of 70% accuracy.

Gopalakrishna et al. [11] provide a method for the detection of stress and relaxation magnitude in social media messages, which may be relevant for emotional evaluations. Another approach related to emotional evaluations comes from [12]. They fine-tune an LLM-based sentiment analyzer on US-based tweets on public support for nuclear power. Their classifier was less likely to overfit compared to other machine learning approaches. Moreover, they used an LLM to segment tweets into topics and found a relation between topic and sentiment, where some topics conveyed a negative sentiment more often.

3. Data and Methodology

3.1. Data

In [2], annotated data was collected in collaboration with 3 linguistic and communication specialists. They annotated a total of 917 posts from 5 cases using a discursive psychological framework [13].

In this experiment, we focus on a primary use case consisting of 98 annotated tweets (542 total) published on the social media platform X concerning the planned cutting of trees in an area of the Netherlands that caused some outrage on X in 2021.

In total, 24 different strategies were annotated. Of these, we only consider the following devices: maximizing/exaggerating language use *("uitvergrotend taalgebruik")* (UT) and negative label(s) to an organization or individual (*"negatief label aan organisatie of persoon"*) (NLO). These two devices are among the more straightforward and well-defined ones. Moreover, they occur relatively frequently and require a range of aspects to consider such as sentiment and syntactical and orthographic features. In this use case, the NLO category occurs in 13.4% of the annotated messages and UT in 7.2%.

3.1.1. Example Selection

Due to compute constraints, the class imbalance in gold labels as well as the qualitative requirement of manually evaluating LLM reasoning per instance, we compiled a small data set of 22 instances per discursive pattern task. These were designed as follows:

- 6 true examples containing the discursive pattern in question.
- 5 false examples manufactured from the above five, with all discursive markers removed.
- 11 false examples not containing the respective discursive pattern. These examples are the same for both categories.

Example Tweets

- **Maximizing (Exaggerating) Language Usage/UT:** "Nederland gaat vlot in de vernieling met grootste catastrofe is de gigantische immigratie. Weet u wat gezinshereniging betekent voor uw belastinggeld. En gehele familieleden clans die hier kunnen komen." [14]
- **Negative Label(s) to Organizations/NLO:** "Opnieuw zijn Romeo's ingezet, ze veroorzaken opstootjes zodat de politie reden heeft om te arresteren" [14]

3.2. Methodology

3.2.1. Language Models

To conduct the experiment for each discursive pattern, we use two open-source Dutch conversational LLMs: Fietje-2-chat [2] (2 billion parameters) and GEITje-7b-ultra [3] (7 billion parameters)

[2]https://huggingface.co/BramVanroy/fietje-2-chat
[3]https://huggingface.co/BramVanroy/GEITje-7B-ultra

Further, for each model, we test 3 prompt setups in total: a 0-shot prompt, a 1-shot prompt and a 1-shot prompt containing clear instructions derived from the annotation guidelines. Each prompt expects a JSON object containing the answer as well as a rationale behind the answer. As the models are finetuned for the Dutch language, the prompts are also fully in Dutch [4].

3.2.2. Rule-based NLP Algorithm

In the context of a learning environment, it was deemed important to have a transparent algorithm for detecting the devices. Therefore, a rule-based algorithm was developed together with the annotators to capture the strategies they use for annotation. The algorithms contain semantic, syntactical, orthographic and sentiment features that are determined using existing and well-known python libraries and models, paticularly SpaCy [15] for syntactic analysis and entity recognition and Pattern.nl [16] for sentiment detection.

The **maximizing language** (UT) algorithm consists of the following aspects:

- use of any term in a word list curated by the communication specialists. This list contains 399 terms consisting of one or more words that usually have a semantic interpretation of being maximizing, such as the word "geweld" (violence)
- use of superlatives ("*grootste* catastrofe")
- use of adverbial modifiers ("dit is *heel* erg")
- use of more than 2 exclamation marks
- use of shouting (e.g. "*WAT EEN ONZIN*")
- use of elongations (e.g. "nooooooit")

The **assigning a negative label to a person or organization** (NLO) is a simple algorithm consisting of the following aspects:

- polarity score < 0 for sentence (pattern.nl 3.6) that contains an entity (spaCy 3.7.5 *nl_core_news_lg* model) OR
- polarity score < 0 for a QUOTE and any entity in the message.

3.2.3. Evaluation

While LLMs are being increasingly used for classification tasks, their black-box properties often limit the user from understanding how the LLM generates the answer. For the current psychologically rich task, this requirement is extremely relevant. Thus, to analyze the reasoning process of the LLM in generating a classification label for this task, we implement a 2-step evaluation procedure for the setups defined in the previous section.

Binary Classification Report The first step solely concerns the yes/no answer identifying the presence of the required discursive pattern generated by the system. It calculates the precision, recall, f1-score and accuracy of the predicted answers using the scikit-learn Python library [17].

As our experimental examples are imbalanced (6 "Yes" and 16 "No" instances per category, we focus primarily on the macro average scores of precision, recall and f1-score for each setup per task. Precision is the ratio of true positives to the sum of true positives and false positives, while recall is the ratio of true positives to the sum of true positives and false negatives. An f-1 score is the harmonic mean of precision and recall [5].

---

[4] The prompts can be found here: https://github.com/swrp-h/discursive-pattern-detection

[5] https://developers.google.com/machine-learning/crash-course/classification/accuracy-precision-recall

Qualitative Assessment of LLM-generated Rationales For the second step, we manually assess the rationales generated for each instance for their correctness. This assessment includes factors such as adherence to the guidelines, presumptions and outright hallucinations. The rationale evaluation is divided into 3 categories in total: ”Correct”, ”Partially Correct” and ”Wrong”, and is conducted by a native speaker of the Dutch language.

4. Results

As previously highlighted, we first indicate the results for the quantitative classification conducted by the LLM- and rule-based systems, followed by a qualitative analysis of the rationales provided for the correct classifications made by the LLM systems.

4.1. Step 1: Classification Results

4.1.1. Magnifying/Exaggerating Language Use (UT)

Overall, we see that for UT, the best-performing setup is that of the rule-based system with a macro f1-score as well as accuracy of 0.68. This is followed by the Fietje setup using the zero-shot prompt (macro f1 = 0.58). The Geitje model shows, counterintuitively, the lowest performance with a macro f1-score of 0.50. On the other hand, its prompt 1 (one-shot examples) and prompt 2 setups (one-shot examples and guidelines) do indicate higher precision and recall compared to the Fietje setups.

Prompt Performances For the Geitje setups, the introduction of a one-shot example (Prompt 2) to the base prompt (Prompt 1) drastically improves performance (with the macro f1-score increasing from 0.19 to 0.50; prompt 1 < prompt 2), but the addition of further guidelines hampers the score (macro f1 = 0.40; prompt 2 > prompt 3). It is also noticed that for Fietje, the addition of further examples and guidelines increasingly worsens its performance instead of improving it (prompt 1 > prompt 2 > prompt 3), with the corresponding macro f1-scores being 0.58, 0.54 and 0.25 respectively.

4.1.2. Negative Labels to Organizations/Persons (NLO)

For the NLO category, the results are more scattered. Although the rule-based system has the highest macro f1-score among all the setups, the zero- and one-shot (prompt 2) Geitje setups show the highest precision and recall (0.70 and 0.72 respectively), which are only slightly higher than those of its prompt 3 (one-shot with guidelines) setup. The highest accuracy is yielded by the zero-shot Fietje setup, but the same setup shows the lowest macro f1-score for the category (0.42).

Prompt Performances For Fietje in this category, we see a continuous increase in scores with further adaptation of the base prompt (prompt 1 < prompt 2 < prompt 3; macro f1: 0.42, 0.50, 0.55), a trend which is not observed for the previous category. On the other hand, for Geitje, there is only a slight increase in performance with the inclusion of a one-shot example (prompt 1 < prompt 2; macro f1: 0.54, 0.59), and the addition of guidelines does not further improve performance. In contrast, the final prompt performs slightly worse than the one-shot-only approach (prompt 2 > prompt 3; macro f1: 0.59, 0.55). This is in trend with the performance of the Geitje setups in the previous category.

4.2. Step 2: Qualitative analysis of LLM-generated rationales

The qualitative analysis is naturally exclusive to the LLM setups, as we attempt to gauge the reasoning process of the chosen models while making their judgment.

4.2.1. Magnifying/Exaggerating Language Use (UT)

Overall, it can be observed that the rationales generated by both models are qualitatively weaker for this category. We see this most distinctly for the Geitje setups, where prompt 1 yields 50% partially correct rationales and 50% completely wrong rationales and no correct rationales. In line with the classification results, there is improvement with the addition of a 1-shot example (prompt 2) with 33% correct and 44% partially correct rationales, but adding further guidelines (prompt 3) considerably hampers the reasoning process, with 62.5% of the rationales being completely incorrect.

Interestingly, the Fietje setups show better reasoning, with the percentage of correct rationales ranging between 28.5% and 42.9%. However, the percentage of incorrect rationales also remains high at 35.7% to 50%. Interestingly, rationales seem to be more correct without the addition of one-shot examples with or without additional guidelines. Similar to the Geitje prompt 2 setup, there is also a higher prevalence of partially correct rationales for the prompt 2 setups (one-shot example) for Fietje at 28.6%.

4.2.2. Negative Labels to Organizations/Persons (NLO)

For this category, LLM rationales are relatively more correct. The Geitje setups indicate the highest percentage of correct rationales among the correctly classified results, with the prompt 1 approach having 75% correct answers and prompts 2 and 3 having 92.3% and 91.7% correct rationales respectively. There are no wrong rationales for the prompt 1 and 2 setups, while the prompt 3 setup indicates 8%. Further, there are fewer partially correct rationales (25% and lower).

The Fietje setups show weaker reasoning capabilities compared to the Geitje ones when classifying negative labels to organizations/persons. In direct contrast to the classification results, the quality of rationales drops consistently from prompt 1 (75%) to prompt 3 (41.7%). On the other hand, the percentages of wrong rationales remain at 18% or lower. In fact, the prompt 2 setup results indicate no wrong rationales, although there is a prevalence of partially correct rationales (45.5%). This prevalence is also present for the one-shot with guidelines (prompt 3) setup (41.7%).

Model	Prompt	Prec.	Recall	F1	Acc.
Geitje	P1	0.12	0.42	0.19	0.23
	P2	0.68	0.66	0.50	0.52
	P3	0.66	0.59	0.40	0.41
Fietje	P1	0.58	0.59	0.58	0.64
	P2	0.54	0.54	0.54	0.64
	P3	0.38	0.45	0.25	0.27
RB Algo	–	0.73	0.78	0.68	0.68

Table 1
Macro-Avg. Results: Magnifying/Exaggerated Language

Model	Prompt	Prec.	Recall	F1	Acc.
Geitje	P1	0.62	0.64	0.54	0.55
	P2	0.70	0.72	0.59	0.59
	P3	0.69	0.69	0.55	0.55
Fietje	P1	0.36	0.50	0.42	0.73
	P2	0.68	0.66	0.50	0.50
	P3	0.69	0.69	0.55	0.55
RB Algo	–	0.61	0.62	0.62	0.68

Table 2
Macro-Avg. Results: Negative Labels to Organizations/Persons

5. Error Analysis and Discussion

5.1. Error Analysis for LLMs

Below, we highlight the most common error types encountered while using the two chosen LLMs for our categories.

- **Echoing of 1-shot example within the output**: This type of error is seen in both models, but predominantly in the Fietje-2-chat outputs. For the latter two prompts, it is seen that the models sometimes simply return the one-shot example provided in the input as output. Thus, although the model scores higher for the classification, it is an incorrect judgment.

- **Contradictory answer-rationale pairs**: In some cases, the models correctly returned a rationale indicating why the tweet can be classified as containing magnifying language, but the corresponding answer label was a "no", indicating disagreement between the two.
- **Inclusion of quotations for judgment**: In a few instances, the model took into account the quote a tweet was referring to, thereby impairing judgment. e.g. In the tweet "Nederland presenteert plannen voor onze natuur. Rutte gaat vertrekken. [LINK] QUOTE @WerkgroepG: [...]: ze kappen 110.000 bomen voor EU subsidie i.h.k.v. herstel van natuur en vervolgens verkoopt SBB die bomen...", although the main tweet does not indicate negative labels being assigned to an organization/individual, the model instead evaluates (and subsequently determines as NLO) the content within the quoted part.
- **Poor understanding of grammatical devices**: The models misinterpret the meanings of superlatives and adverbial modifiers. e.g.: the phrase ("gezondheidsproblemen en meer ellende") is cited to indicate the usage of superlatives, which is grammatically incorrect.
- **Poor Understanding of negative terms**: In a few cases, the models label a tweet as positive for NLO while citing words that are not necessarily negative within context, such as within the example "De tweet bevat negatieve termen zoals 'kappen'..." - in this case, the word *"kappen"* (English: *"cutting"*) is not negative.

5.2. Error Analysis for the Rule-Based Algorithm

- **Flagging of adverbial modifiers which are not magnifying/exaggerating**: Often, the system identifies tweets as positive for UT while flagging irrelevant adverbial modifiers such as "eigenlijk" or "natuurlijk" (*English: "actually", "natural"*).
- **Flagging of contextually irrelevant words**: As it is context-agnostic, the rule-based algorithm can sometimes identify false positives for UT on the grounds of flagging irrelevant words such as "meer" or "mooi" (*English: "more", "nice"*)
- **Overlooking of multi-word hashtags**: For multi-word hashtags such as "#KlimaatReligie" or "#CorruptFuckers", the algorithm is not able to separate the words, making it overlook a negative label to an organization.
- **Weak Sarcasm detection**: In the example "Twee zorgvuldig formulerende oudere heren en de volkse toon van een @VVD bestuurder die nog geen plant of dier uit de na te streven habitat weet te noemen[...]", the system is not able to identify a sarcastic remark towards an organization, which prevents it from making an NLO detection.
- **Named Entity Detection Issues**: For the snippet "TIS WERKELIJK ONGELOFELIJK !!!" written in full caps, the system incorrectly identifies the words as named entities.

5.3. Discussion

The results indicate similar scores for the best-performing LLM setup and the rule-based algorithm, with the latter narrowly demonstrating higher performance.

The chosen LLMs performed sufficiently in terms of surface-level classification, demonstrating capabilities such as: (a) an implicit understanding of exaggerations or negative sentiment being directed at a certain entity and (b) (mostly) the capability of telling quotes apart from the actual tweet. However, the rationale analysis indicates that their reasoning capability is prone to erroneous judgment. Although the size of the selected models - 2 and 7 billion parameters respectively - contributes to such judgment, it is recommended not to take, for our task, LLM-based classifications at face value. The models also often show weakness in identifying the correct grammatical devices and, especially for Fietje-2-chat, are prone to echoing errors.

Interestingly, the prompt experimentation yielded unexpected results. The intuitive assumption was that the addition of a one-shot example and detection guidelines would improve performance; however, our setups showed that the addition of a one-shot example generally improved the performance of only one model and the addition of further guidelines mostly yielded lower scores compared to the one-shot-only setups. This might indicate that detection guidelines that may contradict a model's innate bias may confuse its judgment process.

The rule-based algorithm returns nearly equivalent scores as the LLM setups, even outperforming most setups for the UT category. On the other hand, for the NLO category, the algorithm shows some lacunae, such as the presence of a named entity and a negative sentiment always yielding a positive result or context-blindness. There is also a higher chance of misinterpretation of sentiment or the selection of irrelevant adverbial modifiers with this approach.

Overall, both methods can be improved, and in particular, more research should be done on how the LLM can follow instructions more reliably.

6. Conclusion

When given the task of detecting discursive patterns in tweets, smaller LLMs fall short of optimal scores, performing along similar lines to a straightforward rule-based algorithm.

We saw that introducing guidelines to the prompt negatively affected performance, suggesting that there may be information in the prompt that contradicts the innate information in the model. This is counterintuitive given our assumption that the addition of detection guidelines would mostly better aid the reasoning process. Thus, this warrants further discussion of the interaction between model biases and external instructions, especially in the context of psychologically rich tasks, as well as a re-visitation of the annotation guidelines used to create our prompts.

Although our chosen LLMs are smaller and can be run locally on a personal computer, they utilize more resources compared to the rule-based algorithm, which is computationally efficient as well as transparent. However, the rule-based algorithm shows weaknesses in regards to missing subtext and hyperbole, causing incorrect judgment. It may also return false positives for examples where negative sentiment and named entities coexist but are not linked.

Some limitations of the experiment included the skewed distribution of labels in the dataset and LLM sizes. For the latter, an interesting direction would consist of further expanding this experiment while using state-of-the-art LLMs such as GPT-4 or Claude 3.

We thus conclude that for the task, both approaches - LLM-guided and rule-based - have their merits and weaknesses. Instead of comparing one with the other, it would be beneficial to combine certain aspects of each to construct a system more traceable than a simple LLM classification and better at detecting subtext and implicit devices than a rule-based system.

7. Acknowledgments

This publication is part of the project 360 Graden Newsroom of the research programme RAAK Publiek.

References

[1] B. Stinesen, P. Sneijder, A. Klarenbeek, Geruchtvorming op social media: Een discursief psychologische analyse van twitterberichten over een zoektocht naar twee vermiste kinderen, Tijdschrift voor Communicatiewetenschap 44 (2016).

[2] M. V. D. T. A. K. M Sappelli, L Meijer, Poster: Automatic detection of discursive patterns in dutch social media, in: Computational Linguistics in the Netherlands (CLIN), 2024.

[3] Z. Saeed, R. A. Abbasi, O. Maqbool, A. Sadaf, I. Razzak, A. Daud, N. R. Aljohani, G. Xu, What's happening around the world? a survey and framework on event detection techniques on twitter, Journal of Grid Computing 17 (2019) 279–312.

[4] F. Atefeh, W. Khreich, A survey of techniques for event detection in twitter, Computational Intelligence 31 (2015) 132–164.

[5] A. Hürriyetoğlu, H. Tanev, V. Zavarella, J. Piskorski, R. Yeniterzi, E. Yörük, Challenges and applications of automated extraction of socio-political events from text (case 2021): Workshop and shared task report, arXiv preprint arXiv:2108.07865 (2021).
[6] Z. Zhao, P. Resnick, Q. Mei, Enquiring minds: Early detection of rumors in social media from enquiry posts, in: Proceedings of the 24th international conference on world wide web, 2015, pp. 1395–1405.
[7] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K.-F. Wong, M. Cha, Detecting rumors from microblogs with recurrent neural networks.(2016), in: Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI 2016), ????, pp. 3818–3824.
[8] L. Tian, X. Zhang, Y. Wang, H. Liu, Early detection of rumours on twitter via stance transfer learning, in: Advances in Information Retrieval: 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14–17, 2020, Proceedings, Part I 42, Springer, 2020, pp. 575–588.
[9] V. G. dos Santos, G. L. Santos, T. Lynn, B. Benatallah, Identifying citizen-related issues from social media using llm-based data augmentation, in: International Conference on Advanced Information Systems Engineering, Springer, 2024, pp. 531–546.
[10] K. D. Sainju, N. Mishra, A. Kuffour, L. Young, Bullying discourse on twitter: An examination of bully-related tweets using supervised machine learning, Computers in human behavior 120 (2021) 106735.
[11] R. Gopalakrishna Pillai, M. Thelwall, C. Orasan, Detection of stress and relaxation magnitudes for tweets, in: Companion Proceedings of the The Web Conference 2018, 2018, pp. 1677–1684.
[12] O. H. Kwon, K. Vu, N. Bhargava, M. I. Radaideh, J. Cooper, V. Joynt, M. I. Radaideh, Sentiment analysis of the united states public support of nuclear power on social media using large language models, Renewable and Sustainable Energy Reviews 200 (2024) 114570. URL: https://www.sciencedirect.com/science/article/pii/S136403212400296X. doi:https://doi.org/10.1016/j.rser.2024.114570.
[13] J. Potter, Attitudes, social representations and discursive psychology. (1996).
[14] ???? URL: https://husite.nl/bep/birds-eye-perspective/.
[15] M. Honnibal, I. Montani, spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing, 2017. To appear.
[16] T. D. Smedt, W. Daelemans, Pattern for python, Journal of Machine Learning Research 13 (2012) 2063–2067. URL: http://jmlr.org/papers/v13/desmedt12a.html.
[17] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830.