Analyzing Politeness in Arabic Tweets: A Preliminary Study

Hend Al-Khalifa[1*], Nadia GHEZAIEL HAMMOUDA[2*] and Maria BOUNNIT[3*]

[1]Information Technology Department, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia
[2]Department of Software Engineering, College of Computer Science and Engineering University of Hail, Hail, Saudi Arabia
[3]College of Arabic Language, Cadi Ayyad University, Marrakesh, Morocco
[*]iWAN Research Group, King Saud University, Riyadh, Saudi Arabia

[1]hendk@ksu.edu.sa, [2]n.hamouda@uoh.edu.sa, [3]m.bounnit.ced@uca.ac.ma

Abstract

This paper explores the application of computational methods, specifically transfer learning and large language models (LLMs), for analyzing politeness in Arabic text. The study employs a subset of an Arabic dataset sourced from X (formerly Twitter), focusing on expressions in Modern Standard Arabic (MSA). The proposed approach involves fine-tuning pre-trained Arabic language models and applying zero-shot and few-shot learning methods using various LLMs. The results demonstrate the potential of these techniques for politeness analysis in Arabic social media content, with the fine-tuned models and LLMs achieving varying levels of performance across different evaluation metrics. The study highlights the need for further research to refine methodologies, expand datasets, and incorporate cultural nuances specific to Arabic communication contexts, particularly in the realm of social media interactions.

1 Introduction

Politeness is a fundamental aspect of human communication that helps maintain social harmony and facilitate effective interactions. In the Arabic language, politeness is deeply rooted in cultural norms and social structures, with various linguistic strategies employed to convey respect and courtesy (Ameri et al., 2023). With the increasing prevalence of online communication, particularly on social media platforms like Twitter (X), understanding and analyzing politeness in Arabic text has become a crucial area of research in natural language processing (NLP).

Recent advancements in NLP, such as transfer learning and large language models (LLMs), have shown promising results in various language understanding tasks. However, their application to Arabic politeness analysis remains largely unexplored. This paper aims to bridge this gap by investigating the effectiveness of transfer learning and LLMs in detecting politeness in Arabic social media text.

The main contributions of this paper are as follows:

1. We present a comparative study of transfer learning and LLMs for analyzing politeness in Arabic text, specifically focusing on social media content from Twitter (X). To the best of our knowledge, this is the first study to apply these techniques to Arabic politeness detection in the context of social media.
2. We evaluate the performance of fine-tuned Arabic language models (MarBERT and CamelBERT) and LLMs (GPT-4o-mini, Cohere Command, and JAIS 30B Chat) using various evaluation metrics, providing insights into their strengths and limitations for this task.

The rest of the paper is organized as follows: Section 2 provides a linguistic background on politeness in the Arabic language, discussing its cultural roots and various linguistic strategies. Section 3 reviews related work in computational politeness analysis and Arabic NLP. Section 4 describes the dataset used in this study and the annotation process. Section 5 outlines the experimental design, including the transfer learning and LLM approaches employed. Section 6 presents the results and discusses the performance of the different models. Section 7 conducts an error analysis to identify common challenges and limitations of the proposed approaches. Finally, Section 8 concludes the paper by summarizing the main findings and suggesting directions for future work.

2 Background

The concept of politeness has been extensively studied in pragmatics, particularly through the works of Brown and Levinson (Brown & Levinson, 1987). Their model introduces the idea of a face divided into positive and negative face needs. A positive face refers to the desire to be liked and appreciated, while a negative face involves the need for autonomy and freedom from imposition. Their strategies for maintaining face include:

- Bald on-record: Direct communication without any politeness strategies.
- Positive politeness: Strategies that emphasize camaraderie and mutual respect.
- Negative politeness: Strategies that emphasize the addressee's right to freedom and non-imposition.
- Off-record: Indirect communication that allows for plausible deniability.
- Withholding: Choosing not to say anything at all.

Robin Lakoff adds another dimension to this framework by emphasizing the importance of making the interlocutor feel good and joyful. Her principles of politeness include not imposing, giving options, and making the addressee feel good, highlighting the interpersonal and affective aspects of polite communication (Alaearji & Buraihi, 2021).

Politeness in the Arabic language is deeply rooted in cultural norms and social structures. The language employs various linguistic strategies to convey respect and maintain social harmony. These strategies may include:

- Forms of address: Using plural forms for singular addresses in formal situations (e.g., "هل يمكنكم" instead of "هل يمكنك"), and employing titles and honorifics (e.g., "دكتور", "أستاذ", "سعادة").
- Kinship terms: Using words like "عم" (uncle) or "خالة" (aunt) to address non-relatives respectfully in daily interactions.
- Teknonyms (كنية): Using "أبو" (father of) or "أم" (mother of) followed by a child's name as a form of respect.

As noted by (Ameri et al., 2023) many of these politeness features developed when Arab societies encountered other civilizations, particularly Persian culture. They also mentioned that politeness norms in Arabic are not static but evolve with societal changes.

3 Related Work

This section outlines the methods employed for computational politeness, which are broadly classified into statistical methods, deep learning-based approaches, and Large Language Models. While state-of-the-art research covers various languages, it is notable that politeness features have not been extensively explored in the context of Arabic.

Starting with statistical computational politeness, we found several studies that have utilized SVM methods. (Danescu-Niculescu-Mizil et al., 2013) developed two classifiers to predict politeness in English requests: a Bag of Words (BOW) classifier with unigram features and a linguistically informed classifier incorporating additional linguistic features. They used logistic regression on the SVM outputs to score politeness. However, (Hoffman et al., 2017) contested their approach, revalidating the politeness-labeling tool for broader application in social computing. (Kumar, 2021) worked on an SVM-based classifier for identifying politeness in Hindi. Kumar's approach includes classifiers using unigram and bigram features, as well as additional linguistic features like formulaic expressions and honorifics. (Fu et al., 2020) proposed a three-step pipeline: Plan, Delete, and Generate, for paraphrasing messages to adjust politeness. Their model uses Integer Linear Programming (ILP) to plan appropriate politeness strategies, deletes irrelevant markers, and integrates new strategies to produce a polite paraphrased message.

Recent advances in deep learning for computational politeness include various approaches for both identification and generation. For Politeness Identification, (Aubakirova & Bansal, 2016) used a Convolutional Neural Network (CNN) with the Stanford Politeness

Corpus to predict politeness, outperforming previous methods and identifying new politeness markers. Also, (Mishra et al., 2023) employed a hierarchical transformer network to accurately predict politeness by leveraging contextual information from previous dialogue utterances. In (Dasgupta et al., 2023), they proposed a graph-induced transformer network (GiTN) combining Graph Convolution Networks (GCN) and BERT for detecting formality and politeness in text.

Concerning the Politeness Generation, (Sennrich et al., 2016) introduced an attention-based encoder-decoder NMT system for translating English to German with controllable politeness via target-side T-V annotations. (Feely et al., 2019) proposed a formality-aware NMT system for English to Japanese translation that enhances translation quality, particularly for formal and polite sentences. (Niu & Bansal, 2018) developed three weakly supervised models—Fusion, Label-Fine-Tuning (LFT), and Polite Reinforcement Learning (Polite-RL)—for generating contextually consistent polite responses in open-domain dialogues without parallel data. (Firdaus et al., 2020), (Golchha et al., 2019), and (Wang et al., 2020) highlighted the importance of politeness in user satisfaction and retention. Golchha et al. proposed a reinforced pointer-generator model for courteous responses in customer-care dialogues. Firdaus et al. used a pointer-generator network to produce courteous responses in Hindi and English. While Wang et al. suggested a sequence-to-sequence framework to add politeness and positivity in customer support responses. (Mishra et al., 2022) introduced the Politeness Adaptive Dialogue System (PADS), which uses reinforcement learning to incorporate politeness into dialogue management based on user satisfaction feedback.

The last approach is based on Large Language Models (LLMs), such as GPT-3, LLaMA2, and ChatGPT, which are advanced generative AI systems with billions of parameters trained on extensive textual data. These models have significantly advanced various NLP tasks, including emotion recognition and dialogue. However, their ability to adhere to politeness norms remains a crucial issue. Research by (Li et al., 2023) and (Ziems et al., 2024) showed that state-of-the-art LLMs, like ChatGPT and GPT-3, perform reasonably well in predicting politeness and classifying utterances into polite, neutral, or impolite categories.

Despite significant advancements in computational politeness research across various languages, there remains a notable gap in addressing the unique linguistic and cultural aspects of politeness in Arabic. The intricate interplay between language, culture, and social norms in Arabic presents distinct challenges for computational models aiming to analyze and generate polite language. To bridge this gap, our paper focuses on analyzing politeness in Arabic, by applying transfer learning techniques.

4 Dataset

Our dataset consists of 500 tweets selected from the ArSarcasm database[1]. We specifically chose tweets labeled as "MSA" because it is considered formal, and we aimed to focus on the challenges presented by this linguistic variety. We also filtered the tweets based on the sentiment labeled as positive to examine whether a positively sentiment expression could help detect polite lexical markers more easily.

The tweets present a variety of themes and tones, ranging from political and religious subjects to cultural events and personal expressions. They are primarily written in MSA with occasional English phrases and dialectic words. The tweets presented entail a variety of subjects and moods from political or religious discussions to cultural events, and personal opinions. The Arabic employed in these tweets is split between MSA used in formal statements and quotes of religious nature on one hand and the vernacular languages used in less formal and more intimate interactions on the other. This dualism in language choices characterizes the richness and flexibility of Arabic which allows speakers to adjust their speech according to context as well as their audience.

In our work, we aim to detect politeness in the selected expressions. We have chosen to start with manual annotation, meaning we annotated the expressions without using available systems and software, adopting a linguistic approach. Based on the non-exhaustive analysis conducted by (Alaearji & Buraihi, 2021), we utilize the following tools:

- Politeness markers: A lexical tool examining the words used in each expression. If the expression contrains one or more words from

[1] https://github.com/iabufarha/ArSarcasm

the category of politeness, it is labeled as polite.

- Intention and Purpose of the Expression: A pragmatic tool arising from understanding the finality of the formulated expression. We opted for: appreciation, respect, love, admiration.

labels. The process will be our baseline and it involves tokenizing the input text, formatting it appropriately, and adding task-specific layers to the pre-trained model, followed by training with adjusted hyper parameters to optimize performance. The model's effectiveness is evaluated using metrics like accuracy, precision, and F1-score on split datasets.

Tweet	**politeness markers(lexical)**	**intention / finality (pragmatic)**	**Politeness?**
خلق_جميل احترام اي انسان فقط #__ لأنه انسان مهما كانت ديانته	جميل / احترام	**Appreciation**	**POLITE**
القرار صهيوني"... #حماس "" ترفض منع استخدام مكبرات الصوت في المساجدhttps://t.co/jTAAch4M2w https://t.co/NjjE1gKjF4"	-	-	**NEUTRAL**
الهلال الاحمر الكويتي»: وقعنا »" اتفاقية مع «الهلال القطري» لتوزيع المساعدات الإغاثية على النازحين السوريين من حلب"	-	-	**NEUTRAL**
لله درك #قطر وقادتها#قطر تلغي " الاحتفال بـ #اليوم_الوطني تضامناً مع #حلب... #سوريا #ستبقي_حلب https://t.co/wV2pFqGjdD"	لله درك	**Blessing**	**POLITE**

Table 1: Examples of Annotated Tweets

We must nonetheless point out that our analysis is subject to a certain degree of subjectivity, considering the nature of the expressions analyzed (tweets), the absence of some elements that might be considered essential such as the knowledge of the speaker and the interlocutor, as well as the exact context and conditions of formulation. Additionally, some tweets may not be in pure MSA, which can affect the consistency of our analysis.

Table 1 shows sample tweets annotated based on their politeness markers and intention.

5 Experimental design

Our approach (as shown in Figure 1) is based on transfer learning, where we adapt finetuning pre-trained language models such as CamelBERT[2] and MarBERT[3] models to the task of politeness analysis using a dataset annotated with politeness

Additionally, we employed zero-shot and few-shot learning methods with GPT-4o-mini[4], Cohere Command[5] and JAIS 30B Chat[6] for politeness analysis, constructing prompts to guide the model and making predictions based on their pre-existing knowledge without fine-tuning. The selection of these models is justified by their unique strengths and relevance to the task. GPT-4o-mini's multimodal approach and superior non-English performance, Cohere Command efficiency and robustness, and JAIS 30B Chat's specific design for Arabic and English make them well-suited for a comprehensive evaluation of Arabic language processing capabilities. Note that Few-shot learning incorporates a small-annotated dataset (between 7 and 20 in our case) to provide context in the prompts.

2 https://github.com/CAMeL-Lab/CAMeLBERT
3 https://huggingface.co/UBC-NLP/MARBERT
4 https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/
5 https://cohere.com/command
6 https://www.core42.ai/jais.html

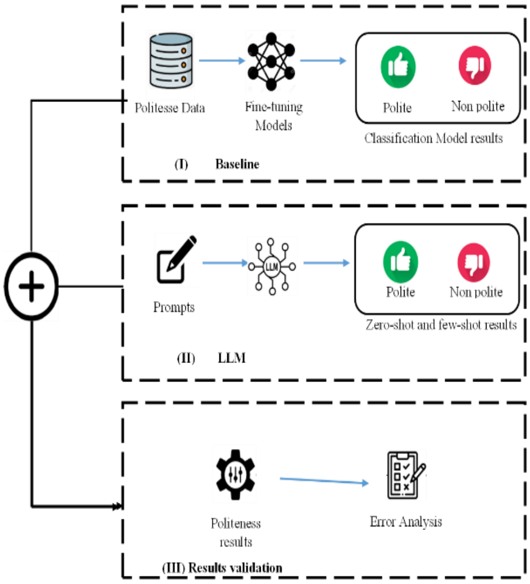

Figure 1: Overall summary of the proposed experiments.

6 Results and Discussion

In this section, we present the various outcomes of our experiments.

Model	Zero-Shot	Few-Shot (7 shots)	Few-Shot (20 shots)
GPT-4o-mini	65.13%	70.73%	66.39%
Cohere Command	54.31%	59.15%	58.87%
JAIS 30B Chat	**65.33%**	**70.87%**	**67.59%**

Table 3: Zero/Few shot learning Evaluation Results.

6.1 Baseline

Table 2 provides the results of MarBERT and CamelBERT models evaluated on accuracy, precision, and F1-score. MarBERT shows higher accuracy (54%) compared to CamelBERT (46%), indicating it classifies more instances correctly overall. However, CamelBERT excels in precision (37% vs. 27%) and F1-score (41.8% vs. 35%), demonstrating it is more reliable in positive predictions and better balanced between precision and recall. Overall, while MarBERT has better accuracy, CamelBERT offers superior performance in terms of precision and F1-score, suggesting it might be more effective for tasks where precision and recall are crucial.

Model	Accuracy	Precision	F1-score
MarBERT	0.54	0.27	0.35
CamelBERT	0.46	0.37	0.418

Table 2: Baseline Evaluation Results.

6.2 Zero/Few shot learning

Table 3 compares the performance of three language models: GPT-4o-mini, Cohere Command, and JAIS 30B Chat on the task of politeness classification in the Arabic language. The evaluation is conducted under three different settings: zero-shot learning, few-shot learning with 7 examples, and few-shot learning with 20 examples.

In the zero-shot setting, where the models are tested without any training examples, GPT-4o-mini and JAIS 30B Chat achieve similar accuracy scores of 65.13% and 65.33%, respectively. Cohere Command lags behind with an accuracy of 54.31%.

With few-shot learning using 7 examples, all three models show improvement. JAIS 30B Chat takes the lead with 70.87% accuracy, closely followed by GPT-4o-mini at 70.73%. Cohere Command also improves but remains in third place with 59.15% accuracy.

When the number of few-shot examples is increased to 20, there is a slight decrease in performance for all models compared to the 7-shot setting. JAIS 30B Chat maintains its lead with 67.59% accuracy, followed by GPT-4o-mini at 66.39% and Cohere Command at 58.87%.

Overall, JAIS 30B Chat and GPT-4o-mini demonstrate superior performance in both zero-shot and few-shot settings compared to Cohere Command for the task of Arabic politeness classification. The results also suggest that increasing the number of few-shot examples from 7 to 20 does not necessarily lead to improved performance for this particular task and dataset.

7 Error Analysis

The baseline models, MarBERT and CamelBERT, yielded divergent results in the politeness analysis of 500 manually annotated tweets. MarBERT classified all tweets as polite, indicating a biased output that exceeded linguistic explanations. In contrast, CamelBERT's predictions aligned well with the manual annotations, with only 12% of expressions remaining incompatible. This discrepancy can be attributed to two main factors: linguistic complexity and data bias. The complex structure and nuances of tweets, including ellipses, wordplay, and implicit references, pose a significant challenge for CamelBERT, requiring deep contextual understanding. Additionally, CamelBERT may have inherited and amplified biases present in its training data, particularly in the context of polite language in various contexts, including religious ones. This bias led to a loss of accuracy in predictions and limited the model's efficiency in politeness analysis.

On the other hand, the error analysis of the politeness prediction models, GPT-4o-Mini, Cohere Command, and JAIS, reveals several factors contributing to the discrepancies between manual annotations and model predictions. GPT-4o-Mini shows a binary approach to politeness, often missing the subtleties that might classify an expression as slightly polite or contextually polite, while Cohere Command exhibits a significant bias towards classifying tweets as polite, lacking a nuanced understanding of politeness gradations. JAIS performs better than the other models but still struggles with contextual understanding of polite content. The main issues identified include confusion between language and social practice, where the models fail to capture the intention and purpose of the entire sentence, as well as errors in analyzing certain sentences, possibly due to the presence of complex linguistic structures or ambiguous expressions. Additionally, the reliance on emojis, particularly heart-shaped ones, can disrupt the politeness prediction, as the models may interpret their presence as a sign of politeness, even when the rest of the tweet's content does not include lexical or pragmatic markers of politeness. These findings highlight the need for more sophisticated models that can better capture the nuances of politeness in Arabic text, taking into account both lexical markers and the overall intention and context of the communication.

8 Conclusion and Future work

In this paper, we presented a comparative study of transfer learning and large language models (LLMs) for analyzing politeness in Arabic text sourced from X (formerly Twitter). Our approach involved fine-tuning pre-trained Arabic language models, specifically MarBERT and CamelBERT, and applying zero-shot and few-shot learning methods using various LLMs, including GPT-4o-mini, Cohere Command, and JAIS 30B Chat. The results demonstrated the potential of these techniques for politeness analysis in Arabic social media content, with the fine-tuned models and LLMs achieving varying levels of performance across different evaluation metrics. The error analysis revealed several factors contributing to the discrepancies between manual annotations and model predictions, including confusion between language and social practice, errors in analyzing certain sentences, and reliance on emojis. Despite the challenges and limitations identified, this study highlights the potential of transfer learning and LLMs for analyzing politeness in Arabic text and underscores the importance of developing language-specific resources and incorporating cultural knowledge and pragmatic understanding into computational models.

Future research should focus on expanding the dataset, incorporating cultural and pragmatic knowledge, investigating advanced architectures, addressing data bias, and extending the insights gained from this study to other NLP tasks in Arabic. By addressing these research directions, future work can contribute to the development of more effective and culturally-aware politeness analysis models for Arabic text, ultimately leading to better communication and social interactions in the digital sphere.

Limitations

This study has several limitations that should be acknowledged. First, the dataset used in this research is relatively small, consisting of only 500 tweets, and focuses solely on Modern Standard Arabic (MSA). This limited dataset may not fully capture the diversity of Arabic dialects and the wide range of politeness expressions used in social media. Second, although the manual annotation was performed by a linguistic expert in the field, the process is still subject to a degree of subjectivity, as it relies on the annotator's understanding of politeness markers and pragmatic intentions, which may be influenced by their cultural background and individual perception. Third, the study focuses on a binary classification of politeness (polite or neutral), which may oversimplify the nuances of politeness in Arabic communication. Finally, the performance of the models may be affected by the limited size of the dataset and the potential biases present in the pre-trained language models used for transfer learning and few-shot learning.

References

Alaearji, A. A., & Buraihi, R. H. (2021). The Principle of Kindness and Politeness in Arabic Linguistic Thought. *International Journal on Humanities and Social Sciences*, *27*, 42–50. https://doi.org/10.33193/IJoHSS.27.2021.336

Ameri, M., Zeighami, A., & Mirahmadi, S. R. (2023). A study of the polite method in the Arabic language according to Brown and Levinson’s theory of politeness. *Studies on Arabic Language and Literature*, *13 (36)*, 61–88. https://doi.org/10.22075/lasem.2022.23600.1286

Aubakirova, M., & Bansal, M. (2016). Interpreting Neural Networks to Improve Politeness Comprehension. In J. Su, K. Duh, & X. Carreras (Eds.), *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing* (pp. 2035–2041). Association for Computational Linguistics. https://doi.org/10.18653/v1/D16-1216

Brown, P., & Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge University Press.

Danescu-Niculescu-Mizil, C., Sudhof, M., Jurafsky, D., Leskovec, J., & Potts, C. (2013). *A computational approach to politeness with application to social factors*. the 51st Annual Meeting of the Association for Computational Linguistics.

Dasgupta, T., Sinha, M., & Geetha Praveen, C. (2023). Graph Induced Transformer Network for Detection of Politeness and Formality in Text. *Companion Proceedings of the ACM Web Conference 2023*, 221–224. https://doi.org/10.1145/3543873.3587352

Feely, W., Hasler, E., & de Gispert, A. (2019). Controlling Japanese Honorifics in English-to-Japanese Neural Machine Translation. In T. Nakazawa, C. Ding, R. Dabre, A. Kunchukuttan, N. Doi, Y. Oda, O. Bojar, S. Parida, I. Goto, & H. Mino (Eds.), *Proceedings of the 6th Workshop on Asian Translation* (pp. 45–53). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-5203

Firdaus, M., Ekbal, A., & Bhattacharyya, P. (2020). Incorporating Politeness across Languages in Customer Care Responses: Towards building a Multi-lingual Empathetic Dialogue Agent. In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 4172–4182). European Language Resources Association. https://aclanthology.org/2020.lrec-1.514

Fu, L., Fussell, S., & Danescu-Niculescu-Mizil, C. (2020). Facilitating the Communication of Politeness through Fine-Grained Paraphrasing. In B. Webber, T. Cohn, Y. He, & Y. Liu (Eds.), *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 5127–5140). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-main.416

Golchha, H., Firdaus, M., Ekbal, A., & Bhattacharyya, P. (2019). Courteously Yours: Inducing courteous behavior in Customer Care responses using Reinforced Pointer Generator Network. In J. Burstein, C. Doran, & T. Solorio (Eds.), *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 851–860). Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1091

Hoffman, E. R., McDonald, D. W., & Zachry, M. (2017). Evaluating a Computational Approach to Labeling Politeness: Challenges for the Application of Machine Classification to Social Computing Data. *Proc. ACM Hum.-Comput. Interact.*, *1*(CSCW), 52:1-52:14. https://doi.org/10.1145/3134687

Kumar, R. (2021). *Towards automatic identification of linguistic politeness in Hindi texts* (arXiv:2111.15268). arXiv. https://doi.org/10.48550/arXiv.2111.15268

Li, C., Pang, B., Wang, W., Hu, L., Gordon, M., Marinova, D., Balducci, B., & Shang, Y. (2023). How Well Can Language Models Understand Politeness? *2023 IEEE Conference on Artificial Intelligence (CAI)*, 230–231. https://doi.org/10.1109/CAI54212.2023.00106

Mishra, K., Firdaus, M., & Ekbal, A. (2022). Please be polite: Towards building a politeness adaptive dialogue system for goal-oriented conversations. *Neurocomputing*, *494*, 242–254. https://doi.org/10.1016/j.neucom.2022.04.029

Mishra, K., Firdaus, M., & Ekbal, A. (2023). Predicting Politeness Variations in Goal-Oriented Conversations. *IEEE Transactions on Computational Social Systems*, *10*(3), 1095–1104. IEEE Transactions on Computational Social Systems. https://doi.org/10.1109/TCSS.2022.3156580

Niu, T., & Bansal, M. (2018). Polite Dialogue Generation Without Parallel Data. *Transactions of the Association for Computational Linguistics*, *6*, 373–389. https://doi.org/10.1162/tacl_a_00027

Sennrich, R., Haddow, B., & Birch, A. (2016). Controlling Politeness in Neural Machine Translation via Side Constraints. In K. Knight, A. Nenkova, & O. Rambow (Eds.), *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 35–40). Association for Computational Linguistics. https://doi.org/10.18653/v1/N16-1005

Wang, Y.-C., Papangelis, A., Wang, R., Feizollahi, Z., Tur, G., & Kraut, R. (2020). *Can You be More Social? Injecting Politeness and Positivity into Task-Oriented Conversational Agents* (arXiv:2012.14653). arXiv. https://doi.org/10.48550/arXiv.2012.14653

Ziems, C., Held, W., Shaikh, O., Chen, J., Zhang, Z., & Yang, D. (2024). Can Large Language Models Transform Computational Social Science? *Computational Linguistics*, *50*(1), 237–291. https://doi.org/10.1162/coli_a_00502