Emotion Analysis in Texts

Sanja Štajner
Karlsruhe, Germany
stajner.sanja@gmail.com

Roman Klinger
Institut für Maschinelle Sprachverarbeitung
University of Stuttgart, Germany
roman.klinger@ims.uni-stuttgart.de

Abstract

Emotion analysis in text is an area of research that encompasses a set of various natural language processing (NLP) tasks, including classification and regression settings, as well as structured prediction tasks like role labeling or stimulus detection. In this tutorial, we provide an overview of research from emotion psychology which sets the ground for choosing adequate NLP methodology, and present existing resources and classification methods used for emotion analysis in texts. We further discuss appraisal theories and how events can be interpreted regarding their presumably caused emotion and briefly introduce emotion role labeling. In addition to these technical topics, we discuss the use cases of emotion analysis in text, their societal impact, ethical considerations, as well as the main challenges in the field.

1 Description and Relevance

Automatic emotion detection in texts has been gaining popularity since 2010's (Acheampong et al., 2020). The systems for automatic emotion detection are often used on social media posts for public opinion analysis, e.g. with respect to climate change (Loureiro and Alló, 2020), to obtain better consumer insights (Sykora et al., 2022), enhance prediction of corporate financial performance (Wang et al., 2023), or predict outcome of elections (Srinivasan et al., 2019). Automatic emotion detection systems are also envisioned to have an important role in building empathetic chatbots and virtual agents (Paiva et al., 2017; Rashkin et al., 2019; Lin et al., 2019b; Shin et al., 2019; Lin et al., 2019a; Ma et al., 2020). More importantly, emotion analysis could be used to aid suicide prevention (Pestian et al., 2012; Desmet and Hoste, 2013), and depression detection (Deshpande and Rao, 2017; Shanthi et al., 2022).

In the computational linguistics (CL) research community, the most commonly used emotion models are Ekman's model (Ekman and Friesen, 1981) consisting of six basic emotions (*anger*, *disgust*, *fear*, *joy*, *sadness*, and *surprise*), and Plutchik's model (Plutchik, 1982), which is commonly used focusing on eight primary emotions (*anger*, *anticipation*, *disgust*, *fear*, *joy*, *sadness*, *surprise*, and *trust*). However, some studies opt for different emotion frameworks or customized emotion sets. For example, Brynielsson et al. (2014), Mohammad et al. (2015), Demszky et al. (2020), Bostan et al. (2020), and Huguet Cabot et al. (2021) use customized emotion sets, Neviarouskaya et al. (2010) use attitudes, and Troiano et al. (2023) use appraisals. Since 2005, over 15 datasets manually annotated for emotions has been compiled and made freely available. The majority of datasets is in English, and they cover a variety of domains and text types: Twitter data (Schuff et al., 2017; Mohammad et al., 2015); personal reports on emotional events (Scherer and Wallbott, 1994; Troiano et al., 2019); sentences from fairy tales (Alm et al., 2005); daily dialogs from websites for English language learners (Li et al., 2017); dialog utterances from the television sitcom Friends (Hsu et al., 2018); movie subtitles (Öhman et al., 2020); news headlines (Bostan et al., 2020; Strapparava and Mihalcea, 2007); and Reddit comments (Demszky et al., 2020; Huguet Cabot et al., 2021). The XED dataset (Öhman et al., 2020), a manually annotated dataset of movies subtitles in English and Finish has been extended to 35 further languages by annotation projection to the parallel sentences in those languages.

From the computational perspective, the research community has used a wide range of approaches for emotion detection and classification, e.g., traditional machine learning approaches that use emotion dictionaries (Mohammad et al., 2015), linear classifiers with various lexical, syntactic, semantic, and structural features (Alm et al., 2005), maximum entropy classifiers with bag-of-words as features (Bostan and Klinger, 2018), support vector machines and naïve Bayes classifiers with

Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics
Tutorial Abstracts, pages 7–12

various lexical, syntactic, and semantic features (Brynielsson et al., 2014), CNN-based classifiers (Hsu et al., 2018), BERT-based classifiers (Demszky et al., 2020; Öhman et al., 2020), multi-task learning (Huguet Cabot et al., 2021), zero-shot learning (Plaza-del Arco et al., 2022; Gebremichael Tesfagergish et al., 2022), and few-shot learning (Guibon et al., 2021). Given that different architectures were tested on different domains, text types, and class types and distributions, it is not clear which models should be considered state of the art. Commercial emotion analysis models commonly use either dictionary-based approaches (due to their domain customisation capabilities which do not require large amounts of labelled training data) or BERT-based models (due to their domain-agnostic adaptation capabilities in the case of sufficient amounts of labelled training data).

Since 2010's, CL research community has been exponentially increasing the effort in building models for recognising and discerning among Ekman's or Plutchik's basic emotions in texts (Acheampong et al., 2020), and building manually annotated datasets, despite of studies in emotion psychology which suggested that detecting emotions in text is difficult and unreliable (Plutchik, 2001; Lang, 2010). The CL studies have pointed out several challenges in emotion annotation in texts: missing context in short utterances (Öhman et al., 2020; Mohammad, 2012), non-literal meaning (Mohammad, 2012), different perspectives one may take, i.e., the reader's, writer's, or text's (Buechel and Hahn, 2017; Alm et al., 2005), and high subjectivity of the task (low inter-annotator agreements were found even among trained annotators (Alm et al., 2005; Schuff et al., 2017; Štajner, 2021)).

Despite the various challenges in emotion analysis from texts, which were reported by researches in emotion psychology or natural language processing (NLP), many tools for emotion analysis are available without a thorough description of challenges and failure modes, e.g. Text2emotion[1] and NRCLex[2] Python libraries. A large number of for-profit companies offer emotion analysis from texts, either using pre-trained models, or customised models trained on clients' data, e.g. BytesView[3], Komprehend[4], IBM Watson Natural Language Understanding.[5] When using the paid emotion analysis APIs, the identification of failure modes on specific datasets or in specific applications, the risk of unintended harms and other ethical considerations are usually shifted to the user of APIs. Those tasks then become extremely difficult given that companies that offer paid APIs often do not disclose the model specifications and datasets the models were trained on.

This tutorial has several goals. First, it provides an overview of most commonly used emotion models and their grounding in emotion psychology, their limitation and challenges from a psychological perspective as well as from NLP perspective. Second, it provides an extensive overview of freely available emotion analysis datasets, their annotation strategies and limitations. Third, it provides an extensive overview and critical comparison of NLP models used for emotion analysis in texts, ranging from traditional machine learning classifiers based on emotion dictionaries to transformer-based classification systems and zero-shot and few-shot learning models. Finally, this tutorial aims at raising awareness about various ethical issues concerning emotion analysis and the still present challenges in emotion analysis in texts (the absence of standardized annotation and evaluation procedures, common failure modes, etc.) which need to be considered when using emotion analysis in real-world applications to avoid unintended harms.

To provide the tutorial participants with a better understanding of the challenges in emotion analysis and help them get started with developing novel models for emotion analysis, we will implement (at the end of the second part of the tutorial) a small annotation exercise.

2 Type: cutting-edge

The first part of the tutorial is an introduction to emotion psychology and the use cases of emotion analysis. The second and third part of the tutorial present cutting-edge NLP research on emotion analysis in texts.

[1] https://pypi.org/project/text2emotion/

[2] https://pypi.org/project/NRCLex/

[3] https://www.bytesview.com/emotion-analysis

[4] https://komprehend.io/emotion-analysis

[5] https://www.ibm.com/cloud/watson-natural-language-understanding

3 Target Audience

This tutorial is well-suited for various audiences: junior and senior researches working on emotion annotation and evaluation of emotion detection models; junior and senior researches working on novel models for emotion analysis, especially those using deep-learning paradigms; industry practitioners who wish to better understand limitations of publicly available emotion analysis tools and models. There are no prerequisites for attending. However, to fully understand the discussion about strengths and limitations of different computational models, a basic knowledge of commonly used non-neural and neural classifiers is recommended.

4 Tutorial Structure

This tutorial contains three thematic parts, each to be covered in a one-hour time slot. The first part introduces emotion models, findings of relevant psychological studies, and use cases. The second part focuses on existing datasets for emotion analysis in texts, and strengths and weaknesses of the computational models which have been proposed so far. The third part covers the fine-grained emotion analysis tasks such as emotion role labeling and stimulus detection, as well as the interpretation of events with appraisal theories. In this part, we also discuss the main challenges in emotion analysis in texts, and ethical considerations for its real-world applications.

Part 1: Foundations

- Emotion theories in psychology
- Emotion recognition reliability in vision and language and what we can expect in NLP
- Use cases and social impact

Part 2: Resources and Computational Models

- Resources for emotion classification
- Resources for emotion intensity prediction
- Non-neural models
- Multi-task and transfer-based models
- Zero-shot and few-shot learning
- Interactive annotation exercise

Part 3: Further Topics

- Event evaluation-based approaches (OCC model and appraisals)
- Emotion role labeling and stimulus/cause detection
- Open challenges in emotion analysis
- Ethical Considerations

5 Reading List

Although no particular prior knowledge is necessary for attending the tutorial, we recommend the attendees which are new to the emotion analysis to read the following works from the references section:

- Peter J. Lang. 2010. Emotion and motivation: Toward consensus definitions and a common research purpose. Emotion review 2, 3:229–233.
- Robert Plutchik. 2001. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. American scientist 89, 4:344–350.
- Laura Ana Maria Bostan and Roman Klinger. 2018. An analysis of annotated corpora for emotion classification in text. In Proceedings of the 27th International Conference on Computational Linguistics, pages 2104–2119.
- Emily Öhman, Marc Pàmies, Kaisla Kajava, and Jörg Tiedemann. 2020. XED: A multilingual dataset for sentiment analysis and emotion detection. In Proceedings of the 28th International Conference on Computational Linguistics, pages 6542–6552.
- Enrica Troiano, Laura Oberländer, and Roman Klinger. 2023. Dimensional modeling of emotions in text with appraisal theories: Corpus creation, annotation reliability, and prediction. Computational Linguistics, 49(1).

6 Instructors' Research Interests and Areas of Expertise

Sanja Štajner has over 14 years of research experience across academia and industry on various psycholinguistic topics in NLP. The last four years, she has led and participated in industry-oriented projects that combined psychology and NLP focusing on sentiment analysis, emotion detection, personality modelling, and mental health assessment. Sanja served as a COLING 2018 area chair for psycholinguistics and cognitive modelling track, and an ACL 2022 demo chair. She has experience as tutorial presenter (COLING 2018, AIST 2018, RANLP 2017) for international audiences and as a lecturer at Masters and PhD levels.

Roman Klinger is senior lecturer at Stuttgart University, where he teaches courses on Emotion Analysis since 2016 (see `https://www.emotionanalysis.de/`). He has been principal investigator on several externally funded projects with focus on emotion analysis. Roman served as senior area chair for sentiment analysis and argumentation mining at ACL 2022 and EACL 2021 and for evaluation and resources at EACL 2023. He was organizer of the WASSA workshop (on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis) in 2018, 2019, 2022, and 2023.

7 Tutorial Materials

All tutorial materials will be made publicly available at: `eacl2023tutorial.github.io`.

8 Ethics Statement

One of the main goals of the tutorial is to raise awareness about open challenges in emotion analysis which can lead to possible unintended harms and ethical issues with models commonly used for emotion analysis in real-world applications.

Acknowledgements

Roman Klinger's work is partially funded by the German Research Council (DFG), project "Computational Event Analysis based on Appraisal Theories for Emotion Analysis" (CEAT, project number KL 2869/1-2).

References

Francisca Adoma Acheampong, Chen Wenyu, and Henry Nunoo-Mensah. 2020. Text-based emotion detection: Advances, challenges, and opportunities. *Engineering Reports*.

Cecilia Ovesdotter Alm, Dan Roth, and Richard Sproat. 2005. Emotions from text: Machine learning for text-based emotion prediction. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 579–586, Vancouver, British Columbia, Canada. Association for Computational Linguistics.

Laura Ana Maria Bostan, Evgeny Kim, and Roman Klinger. 2020. GoodNewsEveryone: A corpus of news headlines annotated with emotions, semantic roles, and reader perception. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 1554–1566, Marseille, France.

Laura Ana Maria Bostan and Roman Klinger. 2018. An analysis of annotated corpora for emotion classification in text. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 2104–2119, Santa Fe, New Mexico, USA.

Joel Brynielsson, Fredrik Johansson, Carl Jonsson, and Anders Westling. 2014. Emotion classification of social media posts for estimating people's reactions to communicated alert messages during crises. *Secur. Informatics*, 3(1):7.

Sven Buechel and Udo Hahn. 2017. Readers vs. writers vs. texts: Coping with different perspectives of text understanding in emotion annotation. In *Proceedings of the 11th Linguistic Annotation Workshop*, pages 1–12, Valencia, Spain. Association for Computational Linguistics.

Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade, and Sujith Ravi. 2020. GoEmotions: A dataset of fine-grained emotions. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4040–4054, Online. Association for Computational Linguistics.

Mandar Deshpande and Vignesh Rao. 2017. Depression detection using emotion artificial intelligence. In *2017 International Conference on Intelligent Sustainable Systems (ICISS)*, pages 858–862.

Bart Desmet and Véronique Hoste. 2013. Emotion detection in suicide notes. *Expert Systems with Applications*, 40(16):6351–6358.

Paul Ekman and Wallace V. Friesen. 1981. *The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding*, pages 57–106. De Gruyter Mouton, Berlin, Boston.

Senait Gebremichael Tesfagergish, Jurgita Kapočiūtė-Dzikienė, and Robertas Damaševičius. 2022. Zero-shot emotion detection for semi-supervised sentiment analysis using sentence transformers and ensemble learning. *Applied Sciences*, 12:8662.

Gaël Guibon, Matthieu Labeau, Hélène Flamein, Luce Lefeuvre, and Chloé Clavel. 2021. Few-shot emotion recognition in conversation with sequential prototypical networks. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 6858–6870, Online and Punta Cana, Dominican Republic.

Chao-Chun Hsu, Sheng-Yeh Chen, Chuan-Chun Kuo, Ting-Hao Huang, and Lun-Wei Ku. 2018. EmotionLines: An emotion corpus of multi-party conversations. In *Proceedings of the Eleventh International*

Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan.

Pere-Lluís Huguet Cabot, David Abadi, Agneta Fischer, and Ekaterina Shutova. 2021. Us vs. them: A dataset of populist attitudes, news bias and emotions. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 1921–1945, Online.

Peter J. Lang. 2010. Emotion and motivation: Toward consensus definitions and a common research purpose. *Emotion review 2*, 3:229–233.

Yanran Li, Hui Su, Xiaoyu Shen, Wenjie Li, Ziqiang Cao, and Shuzi Niu. 2017. DailyDialog: A manually labelled multi-turn dialogue dataset. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 986–995, Taipei, Taiwan. Asian Federation of Natural Language Processing.

Zhaojiang Lin, Andrea Madotto, Jamin Shin, Peng Xu, and Pascale Fung. 2019a. Moel: Mixture of empathetic listeners. *CoRR*, abs/1908.07687.

Zhaojiang Lin, Peng Xu, Genta Indra Winata, Zihan Liu, and Pascale Fung. 2019b. Caire: An end-to-end empathetic chatbot. *CoRR*, abs/1907.12108.

Maria L. Loureiro and Maria Alló. 2020. Sensing climate change and energy issues: Sentiment and emotion analysis with social media in the u.k. and spain. *Energy Policy*, 143:111490.

Yukun Ma, Khanh Linh Nguyen, Frank Z. Xing, and Erik Cambria. 2020. A survey on empathetic dialogue systems. *Information Fusion*, 64:50–70.

Saif Mohammad. 2012. #emotional tweets. In **SEM 2012: The First Joint Conference on Lexical and Computational Semantics – Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation (SemEval 2012)*, pages 246–255, Montréal, Canada. Association for Computational Linguistics.

Saif M. Mohammad, Xiao-Dan Zhu, Svetlana Kiritchenko, and Joel D. Martin. 2015. Sentiment, emotion, purpose, and style in electoral tweets. *Information Processing and Management*, 51:480–499.

Alena Neviarouskaya, Helmut Prendinger, and Mitsuru Ishizuka. 2010. @AM: Textual attitude analysis model. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 80–88, Los Angeles, CA.

Emily Öhman, Marc Pàmies, Kaisla Kajava, and Jörg Tiedemann. 2020. XED: A multilingual dataset for sentiment analysis and emotion detection. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 6542–6552, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Ana Paiva, Iolanda Leite, Hana Boukricha, and Ipke Wachsmuth. 2017. Empathy in virtual agents and robots: A survey. *ACM Transactions on Interactive Intelligent Systems*, 7(3).

John P. Pestian, Pawel Matykiewicz, Michelle Linngust, Brett South, Ozlem Uzuner, Jan Wiebe, Kevin Bretonnel Cohen, John Hurdle, and Christopher Brew. 2012. Sentiment analysis of suicide notes: A shared task. *Biomedical Informatics Insights*, 5:3–16.

Flor Miriam Plaza-del Arco, María-Teresa Martín-Valdivia, and Roman Klinger. 2022. Natural language inference prompts for zero-shot emotion classification in text across corpora. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 6805–6817, Gyeongju, Republic of Korea.

Robert Plutchik. 1982. A psychoevolutionary theory of emotions. *Social Science Information 21*, pages 529–553.

Robert Plutchik. 2001. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist 89*, 4:344–350.

Hannah Rashkin, Eric Michael Smith, Margaret Li, and Y-Lan Boureau. 2019. Towards empathetic open-domain conversation models: A new benchmark and dataset. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5370–5381, Florence, Italy. Association for Computational Linguistics.

K. R. Scherer and H. G. Wallbott. 1994. Evidence for universality and cultural variation of differential emotion response patterning. *Journal of personality and social psychology*, 66 2:310–28.

Hendrik Schuff, Jeremy Barnes, Julian Mohme, Sebastian Padó, and Roman Klinger. 2017. Annotation, modelling and analysis of fine-grained emotions on a stance and sentiment detection corpus. In *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 13–23, Copenhagen, Denmark. Association for Computational Linguistics.

N. Shanthi, Albert Alexander Stonier, Anli Sherine, T. Devaraju, S. Abinash, R. Ajay, V. Arul Prasath, and Vivekananda Ganji. 2022. An integrated approach for mental health assessment using emotion analysis and scales. *Healthcare Technology Letters*, n/a(n/a).

Jamin Shin, Peng Xu, Andrea Madotto, and Pascale Fung. 2019. Happybot: Generating empathetic dialogue responses by improving user experience lookahead. *CoRR*, abs/1906.08487.

Satish Mahadevan Srinivasan, Raghvinder S. Sangwan, Colin J. Neill, and Tianhai Zu. 2019. Power of predictive analytics: Using emotion classification of twitter data for predicting 2016 us presidential elections. *Social media and society*, 8:211–230.

Carlo Strapparava and Rada Mihalcea. 2007. SemEval-2007 task 14: Affective text. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 70–74, Prague, Czech Republic.

Martin Sykora, Suzanne Elayan, Ian R. Hodgkinson, Thomas W. Jackson, and Andrew West. 2022. The power of emotions: Leveraging user generated content for customer experience management. *Journal of Business Research*, 144:997–1006.

Enrica Troiano, Laura Oberländer, and Roman Klinger. 2023. Dimensional modeling of emotions in text with appraisal theories: Corpus creation, annotation reliability, and prediction. *Computational Linguistics*, 49(1).

Enrica Troiano, Sebastian Padó, and Roman Klinger. 2019. Crowdsourcing and validating event-focused emotion corpora for German and English. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4005–4011, Florence, Italy. Association for Computational Linguistics.

Sanja Štajner. 2021. Exploring Reliability of Gold Labels for Emotion Detection in Twitter. In *Proceedings of the 13th international conference on Recent Advances in Natural Language Processing (RANLP)*, pages 1350–1359.

Qiping Wang, Tingxuan Su, Raymond Yiu Keung Lau, and Haoran Xie. 2023. Deepemotionnet: Emotion mining for corporate performance analysis and prediction. *Information Processing & Management*, 60(3):103151.