Editorial for the 4th AAAI-21 Workshop on Affective Content Analysis

Kokil Jaidka[1], Niyati Chhaya[2], Lyle Ungar[3], Jennifer Healey[2] and Atanu Sinha[2]

[1]National University of Singapore, Singapore
[2]Adobe Research
[3]University of Pennsylvania, USA

Abstract

The fourth AAAI Workshop on Affective Content Analysis @ AAAI-21, focused on affect in collaborative creation. A new dataset, called the CLAff-Diplomacy dataset, was released.

Introduction

AffCon-2021 is the fourth Affective Content Analysis workshop @ AAAI. The workshop series (i) builds upon the state of the art in neural and AI methods, for modeling affect in interpersonal interactions and behaviors and (ii) brings a confluence of research viewpoints representing several disciplines. The field of affective content analysis refers to the interdisciplinary research space of Computational Linguistics, Psycholinguists, Consumer psychology, and HCI looking at online communication, its intentions, and the reactions it evokes. The purpose of the workshop was to bring together cross–disciplinary research and mechanisms for affect analysis, as well as to pool together resources for further research and development. The workshop is supported by a committee of keen and experienced researchers in the field of AI.[1]

A large share of content created are outcomes of collaboration. Among others, a basic question worth examining is whether and how collaboration among creatives impact the affective characteristics of the content. A follow up question then is how to model and computationally measure affect in collaborative creation. These are difficult questions especially because the process of exchanges among collaborators is opaque, only the outcome is transparent.

In 2021, collaboration took on an extra meaning in a physically distanced world, with even more reliance on computer-mediated cooperation. Understanding the dynamics of affect in collaborative content is more topical. Therefore, the theme of AffCon-2021 was "Affect for Collaborative Creation". The theme is relevant for increasingly decentralized workplaces, asynchronous collaborations, and computer-mediated communication. Studying and codifying user reactions in this setup can help understand the society

and aid towards better tools for content analysis. The computer mediated collaboration may also produce data to examine these phenomena.

Workshop Topics and Format

Presentations at the workshop featured HCI researchers, computer science researchers, and experts in marketing science. Topics included new approaches that address new directions (e.g., the affect analysis of music), new applications (e.g., affect analysis to help children with autism), and ongoing challenges in collaborative creation (e.g., network dissemination of hate speech). These fall under the broad topics of interest of the workshop:

- Affect in Collaborative Content
- Affect in Communication co-creation
- Affective Reactions in Co-creation and collaboration
- Affectively responsive interfaces
- Deep learning-based models for affect modeling in content (image, audio, and video)
- Mirroring affect
- Psycho–demographic Profiling
- Affect–based Text Generation
- Multi-modal Affect
- Stylometrics, Typographics, and Psycho-linguistics
- Cognitive and psychological computational models of creativity
- Affective needs and Firm-Consumer co-creation Behavior
- Computational models for Consumer Behavior theories of innovation
- Affective Lexica for Online Marketing Communication
- Affective human-agent, human-computer, and human-robot interaction

Overview of the papers

The workshop featured four keynote talks and six paper presentations. The following sections briefly describe the keynote and sessions.

[1]For the full Program Committee list, see https://sites.google.com/view/affcon2021/committees

Keynotes

The title of the first keynote by Dr. Rosalind Picard[2] from the Massachusetts Institute of Technology was "Best practices in automating affect recognition". In the talk, Dr. Picard first reviewed how AI researchers have traditionally framed affect recognition as a pattern recognition problem, and the pitfalls posed by this framing. This was followed by suggestions for best practices, with applications to health monitoring through fitness devices. Dr. Picard's first recommendation was that scholars need to be more careful about how they frame the outcome of predictive AI as a reality vs. a likelihood. She observed that researchers are more answerable to the general public about the interpretation of their findings than ever before. The second takeaway was that, like computers, even humans are not great at recognizing affect. Therefore, artificial intelligence that relies on human labels too may be fooled. Finally, in addressing the impasse about the use of facial recognition technologies in affect computing, Dr. Picard highlighted that intent matters – such data and research is important when it is used to facilitate research that can empower less-abled or vulnerable populations.

The next keynote was "What Text Analysis Cannot Tell Us: The Importance of Observation in Understanding Creative Teams" by Dr. Page Moreau[3] from the University of Wisconsin-Madison. Dr. Moreau used examples from previous work (0) to highlight the importance of cohesion and social sensitivity (reading affect in the eyes) in creative collaboration. Dr. Moreau discussed a conceptual framework to integrate visual cues into recognizing affect, and her recommendation for computer scientists was to consider the signals in turn-taking behavior (whether in actions or speech) to measure dynamic trust.

The talk by Dr. Cristian Danescu-Niculescu-Mizil[4] was titled "Towards an artificial intuition: Conversational markers of (anti)social dynamics" in which he discussed whether conversational dynamics can predict outcomes of social interactions. Dr. Danescu-Niculescu-Mizil provided an overview of a decade of his work studing online group and conversation dynamics. Observing the lack of meta-conversational channels online, he discussed a framework for modeling the subtle pragmatic and rhetorical choices of participants in a conversation, that can point to the nature of the social relation between interlocutors, as well as to the future trajectory of this relation.

Last but not the least, the talk by Dr. Devi Parikh[5] was titled "AI-assisted Human Creativity" wherein she presented several of her projects that have explored how AI can inspire human creativity through the media of sketches, typography, dance, and generative art. Interactive demos explored during the talk[6] allow users to play with or against AI to build new doodles, dances, and other art forms.

---

[2]https://web.media.mit.edu/ picard/

[3]https://wsb.wisc.edu/directory/faculty/page-moreau

[4]www.cs.cornell.edu/ cristian/

[5]https://www.cc.gatech.edu/ parikh/

[6]http://doodlergan.cloudcv.org

Papers:

The first paper by Li, Bhat, and Barmaki (0) presented an open-source multimodal affect analysis framework designed to help children with autism. Their work addressed the challenge of transferring existing methods to a new domain with insufficient labeled training data, and where individuals may show less evidence of positive facial expression. The authors reported an overall accuracy of 72.4% in predicting three main affect states (positive, negative, and neutral) of children with autism. The authors highlighted that speech features predictive of negative states in speech emotion recognition are more distinguishable than facial features, which can even be a challenge for human experts.

The second paper of the workshop won the Best Paper award for the workshop and was titled "Comparison and Analysis of Deep Audio Embeddings for Music Emotion Recognition" (0). The authors validated different audio embedding methods for music emotion inference over four music datasets, demonstrating that solutions with deep audio embeddings (L3-net) improve over the state of the art. The authors demonstrate that complex hand-crafted features offer an improvement for music emotion recognition over simpler features. Visualizing embeddings suggests that L3-net embeddings are able to distinguish timbre through gender and genre, which may possibly explain its superior performance.

The third paper, titled "ABL-MICRO: Opportunities for Affective AI Built Using a Multimodal Microaggression Dataset" (0) highlighted how Affective AI can help to expose and encourage difficult conversations. The paper releases a dataset of over 3000 text and sound instances of microaggressions built from listening and annotating speech from popular American television shows and mining text from websites. The characteristics of the microaggressions included the dataset include racist, homophobic, and sexist remarks, mostly geared towards people of color and women.

The fourth paper titled "Empirical Assessment and Characterization of Homophily in Classes of Hate Speeches" (0) offer a unique approach to involving social network features in the study and prediction of hate speech. The authors apply their methods on an existing dataset (0) which they additionally annotated for hate labels. Their results suggest higher homophily in users associating with topics of racism and nationalism.

Next, the paper "Towards A Six-Level Framework of Emotional Intelligence for Customer Service Chatbots" (0) offers a conceptual framework for evaluating the emotional intelligence of conversational agents in terms of emotional understanding, and emotional strategy. The framework can inspire further research towards an evaluation toolkit for chatbot performance and an outcome-oriented approach to surmounting the challenges associated with building empathetic bots.

Last but not the least, Fong and Kumar (0) report the predictive performance of a deep neural network emotion classifier that uses different audio transformations (spectrograms) designed to capture specific music concepts. The authors report that their Mel square DNN classifier outperforms the SVM classifier developed by the creators of the

CAL500exp dataset (0). The authors suggest that domain knowledge can enable the development of better performing emotion classifiers. They also recommend the use of filters designed to capture specific concepts (time vs. frequency) to understand what black box filters are and are not capturing at a high level.

CL-Aff Shared Task

A new dataset, titled the CL-Aff Diplomacy dataset, was released.[7] The dataset comprises utterances by players in an online game called Diplomacy, labeled for their rapport characteristics. Five annotators were asked to indicate the overall presence of rapport, and then the presence of its subcategories. The original dataset, collected and released by (0) included utterance-level annotations about whether the receiver trusted the speaker, and whether the speaker was lying to the speaker. For the auxiliary rapport labels provided in the dataset, the overall percentage agreement was 75.8%. Annotations for the subcategories were used as an additional filter to denoise the annotations and identify false positives. The lack of participation in 2021 gave the organizers an opportunity to check annotation quality and revise the annotation task. Subsequently, efforts are underway to get the data re-annotated by experts in time for the CLAff Shared Task in 2022.

The possible applications of this dataset would be to examine the role of affect in building task-based collaboration and trust. A call for participation for anew challenge will be announced in late 2021.

Outlook

Many workshops and conferences are now exploring the problems around affective computing which indicates its importance and relevance for AI researchers and practitioners. However, a drop in the overall submission was observed this year. As compared to last year, where the workshop saw 38 papers submitted, this year only 12 papers were submitted to the workshop, of which 5 full papers and one extended abstract were accepted. A competitive acceptance rate upheld a high standard of research quality; however the drop in overall submissions was likely due to the challenges involved in a COVID-stricken year, including but not limited to the expense of participating in a remote AAAI workshop. Sixty-five attendees signed up to the conference. Furthermore, free access to the workshop was made available to 24 applicants who requested for a fee waiver due to different reasons. During the workshop, each session was typically attend by forty or more participants, who were highly engaged in the workshop and interacted with the speakers and presenters in the Q&A sessions.

In the coming years, we as organizers will have to try harder to have wider outreach and diversity in its participation, and to devise new ways to overcoming We will continue to attempt a hybrid workshop in order to foster greater participation and build a more inclusive community

that steers the emerging space of computational approaches for affective content analysis.

Acknowledgments

References

Ding, Q., Carmona, V.I.S., Liu, M., Peng, F., Zhang, Y., Hu, C.: Towards A Six-Level Framework of Emotional Intelligence for Customer Service Chatbots. In: Proceedings of the AAAI-21 Workshop on Affective Content Analysis, New York, USA, AAAI (2021)

Fong, H., Kumar, V.: Music Speaks in Emotions. In: Proceedings of the AAAI-21 Workshop on Affective Content Analysis, New York, USA, AAAI (2021)

Jaidka, K., Singh, I., Lu, J., Chhaya, N., Ungar, L.: A report of the CL-Aff OffMyChest Shared Task: Modeling Supportiveness and Disclosure. In: Proceedings of the AAAI-20 Workshop on Affective Content Analysis, New York, USA, AAAI (2020)

Koh, E., Dubnov, S.: Comparison and Analysis of Deep Audio Embeddings for Music Emotion Recognition. In: Proceedings of the AAAI-21 Workshop on Affective Content Analysis, New York, USA, AAAI (2021)

Li, J., Bhat, A., Barmaki, R.: A Two-stage Multi-modal Affect Analysis Framework for Children with Autism Spectrum Disorder. In: Proceedings of the AAAI-21 Workshop on Affective Content Analysis, New York, USA, AAAI (2021)

Nagar, S., Gupta, S., Bahushruth, C., Barbhuiya, F., Dey, K.: Empirical Assessment and Characterization of Homophily in Classes of Hate Speeches. In: Proceedings of the AAAI-21 Workshop on Affective Content Analysis, New York, USA, AAAI (2021)

Peskov, D., Cheng, B.: It takes two to lie: One to lie, and one to listen. In: Proceedings of ACL. (2020)

Ribeiro, M.H., Calais, P.H., Santos, Y.A., Almeida, V.A., Meira Jr, W.: "like sheep among wolves": Characterizing hateful users on twitter. arXiv preprint arXiv:1801.00317 (2017)

Wang, S.Y., Wang, J.C., Yang, Y.H., Wang, H.M.: Towards time-varying music auto-tagging based on cal500 expansion. In: 2014 IEEE International Conference on Multimedia and Expo (ICME), IEEE (2014) 1–6

Washington, G., Mance, G., Aryal, S., Kengni, M.: ABL-MICRO: Opportunities for Affective AI Built Using a Multimodal Microaggression Dataset. In: Proceedings of the AAAI-21 Workshop on Affective Content Analysis, New York, USA, AAAI (2021)

---

[7]The dataset is available at https://github.com/kj2013/claff-diplomacy

[8]https://sites.google.com/view/affcon2021/committees

Woolley, A.W., Chabris, C.F., Pentland, A., Hashmi, N., Malone, T.W.: Evidence for a collective intelligence factor in the performance of human groups. science **330**(6004) (2010) 686–688