MULTIMODAL PROBABILISTIC LEARNING OF HUMAN COMMUNICATION

Time: Wednesday 5:00-8:20pm

Classroom: In-person SLH 200

Piazza: all communications are through Piazza (https://piazza.com/usc/spring2024/csci535)

Instructor:

Prof. Mohammad Soleymani (soleymani@ict.usc.edu)

TA: TBD

Office: ICT 338 (by appointment only)

Office hours: Friday 3:30 on Zoom

Living document; Due to change

Recommended preparation: CSCI 567 CSCI 561 or CSCI 566 or equivalent. Students should have a proper academic background in probability, statistics and linear algebra. Previous experience in machine learning is required as the assignments will require using basic machine learning methods. This course is not a replacement for the Machine Learning course (CSCI 567).

Academic integrity: Students are expected to abide by the academic integrity rules including submitting individual work for assignments, avoiding plagiarism and citing other works in project reports. Any violation of these rules will result in a 0 in the affected work and a report to the school. For more information, please consult (https://viterbischool.usc.edu/academic-integrity/).

Introduction and Purposes

Human face-to-face communication is a little like a dance, in that participants continuously adjust their behaviors based on verbal and nonverbal displays and signals. Human interpersonal behaviors have long been studied in linguistic, communication, sociology and psychology. The recent advances in machine learning, pattern recognition and signal processing enabled a new generation of computational tools to analyze, recognize and predict human communication behaviors during social interactions. This new research direction has broad applicability, including the improvement of human behavior recognition, the synthesis of natural animations for robots and virtual humans, the development of intelligent tutoring systems, and the diagnoses of social disorders (e.g., autism spectrum disorder).

The objectives of this course are:

1. To give a general overview of human communicative and social behaviors (language, vocal and nonverbal) and show a parallel with computer science (natural language processing, speech processing and computer vision);

2. To understand the multimodal challenge of human communication (e.g. speech and gesture synchrony) and learn about multimodal signal processing;
3. To understand the social aspect of human communication and its implication on computational modeling;
4. To learn about recent advances in machine learning and pattern recognition to analyze, recognize and predict human social and communicative behaviors;
5. To give students practical experience in the computational study of human social communication through a course project.

Course format

Course includes lectures, discussion, assignments and a project. There will be a total of four assignments which will be done individually. Each student will also present one relevant paper and lead the discussion on the topic. Students will be asked to form teams (3-4 people) and present their project advancement for mid-term and at the end of the semester.

Reverse classes (flipped) will be used for lectures. Starting from **week 2,** you will receive a recorded lecture for the following week every Thursday. Every class will start with a Quiz on the lecture and paper reading assignment. We will use the class time for discussions, paper presentations and practicums.

Course Material

Required:

Reading material will be based on published technical papers available via the ACM/IEEE/Springer digital libraries or freely available online. All USC students have automatic access to these digital archives.

Basic knowledge of programming in Python is required.

Optional:

- Burgoon, J. K., Magnenat-Thalmann, N., Pantic, M., & Vinciarelli, A. (Eds.). (2017). Social signal processing. Cambridge University Press.
- *Multimodal Processing and Interaction,* Gros, Potamianos and Maragos, SpringerLink, 2008, DOI: 10.1007/978-0-387-76316-3 (freely available on SpringerLink for USC students)
- *Nonverbal Communication in Human Interaction (7th edition), Mark Knapp and Judith Hall, Wadsworth, 2010*
- *Speech and Language Processing (2nd edition), Daniel Jurafsky and James Martin, Pearson, 2008*
- *Machine Learning for Audio, Image and Video Analysis: Theory and Applications*, Francesco Camastra and Alessandro Vinciarelli, Springer, 2008, DOI: 10.1007/978-1-84800-007-0 (freely available on SpringerLink for USC students)

Course Topics and Readings

** Topics and readings may change based on student interest **

Lectures
(Wed 5-8:20pm)
Week 1 (10 January)
Introduction and communication models

- Human communication dynamics
- Applications and domains
- Communication models
- Mid-term and final projects
- Datasets and sensing tool

Logistics and course format
Reading material

- Vinciarelli et al. (2012), Social Signal Processing
- Krauss et al. (2002), The psychology of Verbal Communication
- (optional) Pentland (2008), Honest Signals, Ch. 1
- (optional) Morency et al. (2010), Human Communication dynamics

Class activity: Project topics and team building
Week 2 (17 January)
Machine Learning reminder/recap: basic concepts

- Classification and Evaluation methods
- Training/validation and testing
- Decision models
- Support Vector Machine
- Neural networks/deep learning
- CNNs and RNNs
- ML toolbox (scikit-learn)

Reading material

- (optional) https://github.com/jakevdp/sklearn_tutorial
- (optional) Courville et al., 2015 deep learning
- (optional) Langley and Kibler (1991), The Experimental Study of Machine Learning

Class activity: Google Colab; USC HPC; choosing papers; project teams

Week 3 (24 January) **[Wednesday, 24 Jan, Assignment 1 release]**
Study Design, Evaluation and Analysis
User studies (guest lecture TBD - January 24)

- Coder agreement, kappa
- Statistical analysis
- Student t-test, effect-size

Class activity: Human subject study group exercise; paper presentations
Reading material

- Lucas et al. (2014), It's only a computer: Virtual humans increase willingness to disclose
- (optional) Costa et al., CHI 2018. Regulating Feelings During Interpersonal Conflicts by Changing Voice Self-perception

- (optional) Wainer (1984) - How to display data badly - The American Statistician
- (optional) Leroy (2011), Designing User Studies in Informatics

Week 4 (31 January)
***** Draft project proposals due **(February 1)***
Project proposal presentations (Flash talks – 1 slide overview with the main points - 5 minutes)
Affective messages and personality traits

- Emotion theory
- Social emotions
- Social behaviors
- Big five personality dimensions

Reading material

- McDuff, et al. Longitudinal Observational Evidence of the Impact of Emotion Regulation Strategies on Affective Expression. 2019.
- de Melo et al. (2013). Reading People's Minds: From Emotion Expressions in Interdependent Decision Making
- (optional) Gratch et al. (2013), Felt emotion and social context determine the intensity of smiles in a competitive video game
- (optional) Gratch and Marsella (2005), Emotion Psychology
- (optional) Barrick and Mount (1991), Big Five personality

Week 5 (7 February)
Vocal messages

- Phonetics and phonology
- Prosody and voice quality
- Vocal expressions
- Audio representation and basic feature extraction
- Wav2Vec, HuBERT, Whisper

Class activity: Audio analysis practicum (Librosa); paper presentations
Reading material

- Bachorowiski et al. Sounds of emotion: production and perception of affect-related vocal acoustics 2003
- (Optional) Schuller et al., (2011), Recognising realistic emotions and affect in speech: State of the art and lessons ..
- (Optional) Eyben et al. (2015), The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research …
- (optional) Bachorowsky et al. (2001), The acoustic features of laughter
- (optional) Ladefoged (2004), A course in phonetics

Week 6 (14 February)

Visual messages (facial expression)

- Facial expressions
- Facial action units
- Image and video representation learning
- FaceAPI, OpenFace, OpenPose
- Visual messages (gaze, gesture)
- Gesture, gaze, posture and proxemics
- Face encoders (CNN, ViT)

Class activity: Visual analysis practicum (MediaPipe/LibreFace); paper presentations

Reading material

- Chu et al., 2017
- Cowen et al, Sixteen facial expressions occur, 2020.
- Kramer (2008) Nonverbal communication
- OpenSense
- PML
- (optional) Argyle and Dean (1965) Eye-Contact, Distance and Affiliation (optional) Kendon (1995) Gesture

Week 7 (21 February)

Verbal and Conversational messages

- Discourse analysis
- Turn-taking and backchannel
- Semantics and pragmatics
- Speech and dialogue acts
- Language models and N-grams
- Boundaries, fillers and disfluencies
- Syntax and part-of-speech tagging
- Word and sentence embedding (GloVe, BERT, etc.)

Class activity: NLP practicum - NLTK, spaCy; paper presentations

Reading material

- DeMasi et al, A Multi-Persona Chatbot for Hotline Counselor Training, EMNLP 2020
- (optional) Zhao et al., Joint dialogue act segmentation and recognition, 2019
- Ghosh et al. Affect-LM:, ACL 2017
- (optional) Hoegen et al. Conversational style, 2019
- (optional) Stolcke et al. CL 2000: Dialogue act modeling for automatic tagging and..
- (optional) Bohus and Horvitz (2010), Computational Turn-taking
- (optional) Jurafsky and Martin (2008), Speech and Language Processing, 4.1-4.4, 5.1-5.3 and 12.1-12.2
- (optional) Liu et al. (2004) Metadata extraction
- (optional) Clark and Brennan (1991)

Week 8 (28 February)
Multimodal behavior recognition

- Multimodal fusion
- Multimodal alignment,
- Audio-visual recognition
- Long Short-term Memory Networks
- Transformers

Class activity: Multimodal fusion practicum; paper presentations
Reading material - fusion

- Tao & Busso, Bimodal LSTM 2017.
- (optional) Baltrušaitis et al., 2018
- (optional) Kang et al, 2010
- (optional) Soleymani et al, 2019

Week 9 (6 March)

- **Mid-term project presentations (mid-term reports are due) (report deadline March 9th)**

Week 10 (13 March)

- **No class - spring break**

Week 11 (20 March)
Multimodal learning for emotion recognition
Human behavior understanding for social robotics
Class activity: Multimodal emotion recognition - OpenSense; paper presentations
Reading material

- Deepali et al. Conversational Error Analysis in Human-Agent Interaction. 2020.
- Celiktutan et al. Computational Analysis of Affect,
- Personality, and Engagement in Human–Robot Interactions, 2018
- Dong et al., Embodiment in socially interactive robots, 2019
- (optional) Soleymani et al, 2019

Week 12 (27 March)
Virtual Humans

- VHToolkit, guest talk TBD
    - Behavior planning and generation
- Multimodal behavior generation
- Designing ECAs
- Guest lecture by on gesture generation TBS

Class activity: Virtual Human Toolkit

Reading material

- Kucherenko et al., Gesticulator, ICMI 2020.
- (optional) Devault et al. Simsensei 2014
- (optional) Ding et al. (2014) Laughter Animation Synthesis
- (optional) Cafaro et al (2016) The Effects of Interpersonal Attitude of a Group of Agents on User's Presence and Proxemics Behavior

Week 13 (3 April)
Speech recognition - guest lecture TBD
Dialogue management - guest lectures TBD

Class activity: Whisper, PyAnnote, LLM; paper presentations

Reading material

- Speech and Language Processing, Jurafsky & Martin, Chapter 16 Speech recognition
- Speech and Language Processing, Jurafsky & Martin, Chapter 15 Chabot & Dialogue

Week 14 (10 April)
Ethics, bias and fairness
Bias and ethics
- Bias
- Fairness
- Trade-off and design considerations
- Methods for improving fairness in machine learning

Reading material - bias and ethics
- McDuff et al., 2019
- Raghavan et al., 2019
- Yan et al., 2020

Multimodal sentiment analysis
- Sentiment vs emotions
- Text-based sentiment analysis
- Multimodal sentiment analysis
- Speech activity detection

Class activity: Fairness evaluation, preparing an IRB protocol; paper presentations

Reading material - sentiment
- Soujanya et al., Beneath the Tip of the Iceberg: Current Challenges and New Directions in Sentiment Analysis Research. 2020.

- Pang et al. EMNLP 2002: Thumbs up? Sentiment classification using machine learning techniques
- Soleymani et al. 2017, definition
- Tavabi et al, 2019

Week 15 (17 April)

Health

- Multimodal behavior analysis in Health
- Mental health assessment 19 April
- Neurodevelopmental disorders
- Neurodegenerative disorders

Reading material - health

- Tavabi et al., 2020
- Cummings et al, 2015

Dyadic and Multiparty Interactions

- Dyadic Modeling
- Recurrence plot analysis
- Synchrony and entrainment
- Team processes and assessment
- Climate and cohesion
- Knowledge formation

Class activity: Measuring dyadic synchrony/entrainement; paper presentations

Reading material

- Tsai, T.J., Stolcke, A. and Slaney, M., 2015. A study of multimodal addressee detection in human-human-computer interaction. IEEE Transactions on Multimedia, 17(9), pp.1550-1561.
- Kozlowski, 2015
- Neubauer et al. 2016
- Zhang et al, 2018

Week 16 (24 April)

Final project presentations

Final reports are due May 6 2024

Bibliography

Introduction and communication models

1. Vinciarelli, A., Pantic, M., Heylen, D., Pelachaud, C., Poggi, I., D'Errico, F., & Schroeder, M. (2011). Bridging the gap between social animal and unsocial machine: A survey of social signal processing. *IEEE Transactions on Affective Computing*, *3*(1), 69-87.
2. Krauss, R.M. (2002). The psychology of verbal communication. In, N. Smelser & P. Baltes (eds.), *International Encyclopedia of the Social and Behavioral Sciences.* London: Elsevier.

3. (optional) Morency, L.-P., Modeling Human Communication Dynamics, IEEE Signal Processing Magazine, September 2010
4. (optional) Pentland, Honest Signals, Chapter 1

Machine Learning: Basic Concepts

5. Tom Fawcett, An introduction to ROC analysis, *Pattern Recognition Letters,* Volume 27, Issue 8, June 2006, Pages 861–874
6. (optional) Langley and Kibler (1991), The Experimental Study of Machine Learning, Unpublished paper.

Study Design, Evaluation and Analysis

8. Gale M. Lucas, Jonathan Gratch, Aisha King, Louis-Philippe Morency, It's only a computer: Virtual humans increase willingness to disclose, Computers in Human Behavior, Volume 37, August 2014, Pages 94-100.
9. Costa, J., Jung, M. F., Czerwinski, M., Guimbretière, F., Le, T., & Choudhury, T. (2018, April). Regulating feelings during interpersonal conflicts by changing voice self-perception. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-13)
10. (optional) H Wainer - How to display data badly - The American Statistician, 1984
11. (optional) Leroy (2011), Designing User Studies in Informatics, Springer.
12. (optional) http://www.sparknotes.com/psychology/psych101/researchmethods/

Affective messages and personality traits

13. McDuff, Daniel, Eunice Jun, Kael Rowan, and Mary Czerwinski. "Longitudinal Observational Evidence of the Impact of Emotion Regulation Strategies on Affective Expression." IEEE Transactions on Affective Computing (2019).
14. de Melo, C. M., Carnevale, P. J., Read, S. J., & Gratch, J. (2013, September 30). ReadingPeople's Minds From Emotion Expressions in Interdependent Decision Making. Journal of Personality and Social Psychology. Advance online publication. doi: 10.1037/a0034251
15. Jonathan Gratch, Lin Cheng, Stacy Marsella and Jill Boberg, Felt emotion and social context determine the intensity of smiles in a competitive video game, Face and Gesture 2013
16. (optional) Gratch & Marsella, 2005 Lessons from Emotion Psychology for the Design of Lifelike Characters
17. (optional) Mr Barrick, Mk Mount (1991) The Big Five Personality Dimensions And Job Performance: A Meta-Analysis - Personnel Psychology

Vocal messages

17. Bachorowiski, Jo-Anne, and Michael J. Owren. "Sounds of emotion: production and perception of affect-related vocal acoustics." (2003).
18. Schuller et al., (2011), Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge, Volume 53, Issues 9–10, 2011, Pages 1062–1087

19. (Optional) Eyben, Florian, et al. "The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing." *IEEE transactions on affective computing* 7.2 (2015): 190-202.
20. (Optional) Bachorowski et al. (2001), The acoustic features of human laughter, *Journal of the Acoustic Society of America*, 110 (3), pp. 1581-1597
21. (Optional) Ladefoged (2004), A course in phonetics

Visual messages

21. Cowen, A. S., Keltner, D., Schroff, F., Jou, B., Adam, H., & Prasad, G. (2020). Sixteen facial expressions occur in similar contexts worldwide. Nature, 1-7.
22. Chu, Wen-Sheng, Fernando De la Torre, and Jeffrey F. Cohn. "Learning spatial and temporal cues for multi-label facial action unit detection." *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*. IEEE, 2017.
23. (optional) Krämer, N. C. (2008). Nonverbal Communication. In J. Blascovich & C. Hartel (eds.), Human behavior in military contexts (pp. 150-188). Washington: The National Academies Press.
24. (optional) Fernando de la Torre and Jeffrey F. Cohn, Facial Expression Analysis, *Visual Analysis of Humans*, 2011, 377-409
25. (optional) Adam Kendon, An Agenda for Gesture Studies, This article appeared in Volume 7 (3) of the Semiotic Review of Books.
26. (optional) Michael Argyle and Janet Dean, Eye-contact, distance and Affiliation, Sociometry, Vol. 28, No. 3, pp. 289-304, 1965

Virtual humans

26. Kucherenko, T., Jonell, P., van Waveren, S., Henter, G. E., Alexandersson, S., Leite, I., & Kjellström, H. (2020, October). Gesticulator: A framework for semantically-aware speech-driven gesture generation. In *Proceedings of the 2020 International Conference on Multimodal Interaction* (pp. 242-250).
27. DeVault, David, Ron Artstein, Grace Benn, Teresa Dey, Ed Fast, Alesia Gainer, Kallirroi Georgila et al. "SimSensei Kiosk: A virtual human interviewer for healthcare decision support." In Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems, pp. 1061-1068. 2014.
28. (optional) Ding, Yu, Ken Prepin, Jing Huang, Catherine Pelachaud, and Thierry Artières. "Laughter animation synthesis." In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pp. 773-780. 2014.
29. Ahuja, Chaitanya, Dong Won Lee, Ryo Ishii, and Louis-Philippe Morency. "No Gestures Left Behind: Learning Relationships between Spoken Language and Freeform Gestures." In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings, pp. 1884-1895. 2020.

Verbal messages

30. Zhao, Tianyu, and Tatsuya Kawahara. "Joint dialog act segmentation and recognition in human conversations using attention to dialog context." Computer Speech & Language 57 (2019): 108-127.

31. Ghosh, Sayan, et al. "Affect-LM: A Neural Language Model for Customizable Affective Text Generation." *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2017.
32. (optional) Pang et al. EMNLP 2002: Thumbs up? Sentiment classification using machine learning techniques, http://www.aclweb.org/anthology/W02-1011
33. (optional) Andreas Stolcke , Noah Coccaro , Rebecca Bates , Paul Taylor , Carol Van Ess-Dykema , Klaus Ries , Elizabeth Shriberg , Daniel Jurafsky , Rachel Martin , Marie Meteer, Dialogue act modeling for automatic tagging and recognition of conversational speech, Computational Linguistics, v.26 n.3, p.339-373, September 2000
34. (optional) Jurafsky and Martin (2008), Speech and Language Processing, Sections 4.1-4.4, 5.1-5.3 and 12.1-12.2
35. (optional) Soo-Min Kim and Eduard Hovy (2004) Determining the Sentiment of Opinions, Proceedings of the COLING conference, Geneva
36. (optional) Yang Liu, Elizabeth Shriberg, Andreas Stolcke, Dustin Hillard, Mari Ostendorf, Barbara Peskin, and Mary Harper. 2004. The ICSI-SRI-UW Metadata Extraction System, ICSLP 2004

Conversational messages

37. Hoegen, R., Aneja, D., McDuff, D., & Czerwinski, M. (2019, July). An end-to-end conversational style matching agent. In Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents (pp. 111-118).
38. Duncan (1974) Some Signals and Rules for Taking Speaking Turns in Conversations
39. Bohus, D., Horvitz, E., (2010) - Computational Models for Multiparty Turn-Taking, Microsoft Technical Report MSR-TR-2010-115
40. (optional) Jurafsky and Martin (2008), Speech and Language Processing, Sections 17.1-17.4 and 21.1-21.4

Multimodal behavior recognition (1/3)

41. Tao, F., & Busso, C. (2017, August). Bimodal Recurrent Neural Network for Audiovisual Voice Activity Detection. In *INTERSPEECH* (pp. 1938-1942).
42. Baltrušaitis, Tadas, Chaitanya Ahuja, and Louis-Philippe Morency. "Multimodal machine learning: A survey and taxonomy." IEEE Transactions on Pattern Analysis and Machine Intelligence 41.2 (2018): 423-443.C. Christoudias, K. Saenko, L.-P. Morency, and T. Darrell (2006) Co-Adaptation of Audio-Visual Speech and Gesture Classifiers, International Conference on Multimodal Interactions (ICMI 2006)
43. I. McCowan, D. Gatica-Perez, S. Bengio, G. Lathoud, M. Barnard, M., D. Zhang, "Automatic analysis of multimodal group actions in meetings", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 27, pp. 305–317, 2005
44. (optional) Gros, Potamianos and Maragos (2008) Multimodal Processing and Interaction, SpringerLink, Chapter 1 [SpringerLink or USC blackboard]
45. (optional) Pradeep K. Atrey, M. Anwar Hossain, Abdulmotaleb El Saddik and Mohan S. Kankanhalli, Multimodal fusion for multimedia analysis: a survey, Multimedia Systems, Volume 16, Number 6 (2010), 345-379 (optional) A. Nefian, L. Liang, X. Pi, X. Liu and K.

Murphy, (2002) Dynamic Bayesian networks for audio-visual speech recognition, EURASIP Journal on Applied Signal Processing, Volume 2002, Issue 1

Social robotics

46. Aneja, Deepali, Daniel McDuff, and Mary Czerwinski. "Conversational Error Analysis in Human-Agent Interaction." In Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, pp. 1-8. 2020.
47. Celiktutan, Oya, Evangelos Sariyanidi, and Hatice Gunes. "Computational Analysis of Affect, Personality, and Engagement in Human–Robot Interactions." Computer Vision for Assistive Healthcare. Academic Press, 2018. 283-318.
48. (optional) Deng, Eric, Bilge Mutlu, and Maja J. Mataric. "Embodiment in socially interactive robots." Foundations and Trends® in Robotics 7.4 (2019): 251-356.

Multimodal behavior recognition (gaze, gesture and speech activity)

49. (optional) K. Stefanov and J. Beskow, "A real-time gesture recognition system for isolated swedish sign language signs," in Linköping University Electronic Press, 2017.
50. (optional) K. Stefanov, G. Salvi, D. Kontogiorgos, H. Kjellström and J. Beskow, "Modeling of human visual attention in multiparty open-world dialogues," in ACM Transactions on Human-Robot Interaction, 2019.
51. (optional) K. Stefanov, "Webcam-based eye gaze tracking under natural head movement," in arXiv preprint arXiv:1803.11088, 2010.
52. (optional) K. Stefanov, J. Beskow and G. Salvi, "Self-supervised vision-based detection of the active speaker as support for socially-aware language acquisition," in IEEE Transactions on Cognitive and Developmental Systems, 2019.
53. (optional) X. Anguera, S. Bozonnet, N. Evans, C. Fredouille, G. Friedland and O. Vinyals, "Speaker diarization: a review of recent research," in IEEE Transactions on Audio, Speech, and Language Processing, 2012.
54. H. Admoni and B. Scassellati, "Social eye gaze in human-robot interaction: a review," in ACM Journal of Human-Robot Interaction, 2017.
55. (optional) S. Rautaray and A. Anupam, "Vision based hand gesture recognition for human computer interaction: a survey," in Springer Artificial Intelligence Review, 2015.

Sentiment

56. Soleymani, M., Garcia, D., Jou, B., Schuller, B., Chang, S. F., & Pantic, M. (2017). A survey of multimodal sentiment analysis. *Image and Vision Computing*, *65*, 3-14.
57. Poria, Soujanya, Devamanyu Hazarika, Navonil Majumder, and Rada Mihalcea. "Beneath the Tip of the Iceberg: Current Challenges and New Directions in Sentiment Analysis Research." IEEE Trans. Affective Computing 2020.

Dyadic and team processes

53. Tsai, T.J., Stolcke, A. and Slaney, M., 2015. A study of multimodal addressee detection in human-human-computer interaction. IEEE Transactions on Multimedia, 17(9), pp.1550-1561.
54. Neubauer, C., Woolley, J., Khooshabeh, P., & Scherer, S. (2016, October). Getting to know you: a multimodal investigation of team behavior and resilience to stress. In

Proceedings of the 18th ACM International Conference on Multimodal Interaction (pp. 193-200). ACM.

55. Zhang, Yanxia, et al. "TeamSense: assessing personal affect and group cohesion in small teams through dyadic interaction and behavior analysis with wearable sensors." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2.3 (2018): 150.
56. Scherer, Stefan, Zakia Hammal, Ying Yang, Louis-Philippe Morency, and Jeffrey F. Cohn. "Dyadic behavior analysis in depression severity assessment interviews." In *Proceedings of the 16th International Conference on Multimodal Interaction*, pp. 112-119. 2014.

Bias and ethics

55. McDuff, Daniel, et al. "Characterizing bias in classifiers using generative models." *Advances in Neural Information Processing Systems*. 2019.
56. Yan, Shen, Di Huang, and Mohammad Soleymani. "Mitigating Biases in Multimodal Personality Assessment." *Proceedings of the 2020 International Conference on Multimodal Interaction*. 2020.

Health

57. Cummins, Nicholas, Stefan Scherer, Jarek Krajewski, Sebastian Schnieder, Julien Epps, and Thomas F. Quatieri. "A review of depression and suicide risk assessment using speech analysis." *Speech Communication* 71 (2015): 10-49.
58. Tavabi, Leili, Anna Poon, Albert Skip Rizzo, and Mohammad Soleymani. "Computer-Based PTSD Assessment in VR Exposure Therapy." In *International Conference on Human-Computer Interaction*, pp. 440-449. Springer, 2020.
59. Cohen, A. S., Mitchell, K. R., & Elvevåg, B. (2014). What do we really know about blunted vocal affect and alogia? A meta-analysis of objective assessments. In *Schizophrenia Research* (Vol. 159, Issues 2–3, pp. 533–538). Elsevier. https://doi.org/10.1016/j.schres.2014.09.013
60. Jaiswal, Shashank, Michel F. Valstar, Alinda Gillott, and David Daley. "Automatic detection of ADHD and ASD from expressive behaviour in RGBD data." In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pp. 762-769. IEEE, 2017.

Grades

Grading breakdown

- Quizzes (10%) (the lowest grade will be discarded to allow for an absence)
- Three assignments 30% (10% each)
- One paper discussion moderation (10%)
- Course project:
    - § Mid-term project report and presentation 15%
    - § Final report and presentation 35%

Grading scale

At the end of the semester, the following chart will be used for converting the grades.

94 - 100 = A	70 – 73.99 = C
90 – 93.99 = A-	66 – 69.99 = C-
86 – 89.99 = B+	62 – 65.99 = D+
82 – 85.99 = B	58 – 61.99 = D
78 – 81.99 = B-	54 – 57.99 = D-
74 – 77.99 =C+	53.99 and below = F

Attendance

- Students are expected to attend every class (1 free absence allowed) and participate actively during the discussions.

Lecture

From week 2, lectures will be released a week in advance. Students are expected to watch the lecture prior to the class. We will use the class time for questions, discussion on paper reading and practicums.

Paper discussion

Each week, we give a paper to read and students are expected to read it carefully and answer a few questions before the class.

Paper presentation and discussion moderation

Students are also expected to choose one paper (a non-exclusive list will be given) to be presented to their peers during the class.

Assignments

- These four assignments will be designed to give hands-on experience with multimodal data analysis and machine learning (e.g., SVM, neural networks) for multimodal behavior recognition.
- Students will need to submit their code (zip files) with their answer to each practical exercise.

Course project:

- The goal of this course is to analyze human communicative behaviors in social settings using state-of-the-art statistical and probabilistic models. The course project is specifically designed to give students practical experience in computational study of human social communication.
- Students should team up (minimum three; max four). The mid-term and final report will need to outline the tasks of each participant. Team projects will be expected to include a deeper analysis than individual projects.
- **Mid-term report:** The mid-term report will present a qualitative analysis of the selected dataset and communicative behaviors. The report should include correct

transcription and annotations of the language, vocal and nonverbal behaviors. Using standard statistical tools and qualitative observations, the students should highlight the challenges with this dataset (and communicative behaviors) and suggest an approach to solve them.

- **Final report and presentation:** Using the same dataset as the mid-term report, the final report will include a quantitative analysis of the human communicative behaviors. The final report should be phrase as a research paper describing either a comparative study of different statistical and probabilistic approaches or a new technique for behavior modeling.

Statement for Students with Disabilities

Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to TA) as early in the semester as possible. DSP is located in STU 301 and is open 8:30 a.m.–5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776.

Statement on Academic Integrity

USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect one's own academic work from misuse by others as well as to avoid using another's work as one's own. All students are expected to understand and abide by these principles. *Scampus,* the Student Guidebook, contains the Student Conduct Code in Section 11.00, while the recommended sanctions are located in Appendix A: http://www.usc.edu/dept/publications/SCAMPUS/gov/. Students will be referred to the Office of Student Judicial Affairs and Community Standards for further review, should there be any suspicion of academic dishonesty. The Review process can be found at: http://www.usc.edu/student-affairs/SJACS/.