

Education

- 2019 – 2023 **University of Cambridge**, Cambridge, UK.
Ph.D., Engineering (Machine Learning) – Supervisor: Adrian Weller
Affiliations: Machine Learning Group, Computational and Biological Learning Lab
- 2015 – 2019 **Carnegie Mellon University**, Pittsburgh, PA.
B.S. and M.S., Electrical and Computer Engineering – Advisor: José Moura

Fellowships and Positions

- 2019 – now **Leverhulme Center for the Future of Intelligence**, *Student Fellow*, Cambridge, UK.
- 2019 – now **Partnership on AI**, *Research Fellow*, San Francisco, CA.
- 2017 – 2019 **Carnegie Mellon University**, *Research Assistant*, Pittsburgh, PA.
Worked with José Moura (ECE), Pradeep Ravikumar (MLD), Radu Marculescu (ECE), Fei Fang (ISR), Zico Kolter (CSD), and Manuela Veloso (MLD)

Publications

Conference Papers

- [C7] **Umang Bhatt**, Adrian Weller, & José Moura. Evaluating and Aggregating Feature-based Model Explanations. To appear at the International Joint Conference on Artificial Intelligence (IJCAI) 2020. Yokohama, Japan.
- [C6] **Umang Bhatt**, Alice Xiang, Shubham Sharma, Adrian Weller, Ankur Taly, Yunhan Jia, Joydeep Ghosh, Ruchir Puri, José Moura, & Peter Eckersley. Explainable Machine Learning in Deployment. ACM Conference on Fairness, Accountability, and Transparency (FAT*) 2020. Barcelona, Spain.
- [C5] Boty Dimanov, **Umang Bhatt**, Mateja Jamnik, & Adrian Weller. You Shouldn't Trust Me: Learning Models Which Conceal Unfairness From Multiple Explanation Methods. European Conference on Artificial Intelligence (ECAI) 2020. Santiago de Compostela, Spain.
- [C4] Brian Davis*, **Umang Bhatt***, Kartikeya Bhardwaj*, Radu Marculescu, & José Moura. On Network Science and Mutual Information for Explaining Deep Neural Networks. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2020. Barcelona, Spain.
- [C3] Aaron Roth, Samantha Reig, **Umang Bhatt**, Johnathan Schulgach, Tamara Amin, Afsaneh Doryab, Fei Fang, & Manuela Veloso. A Robot's Expressive Language Affects Human Strategy and Perceptions in a Competitive Game. IEEE International Conference on Robot and Human Interactive Communication (IEEE-ROMAN) 2019. New Delhi, India.
- [C2] **Umang Bhatt**, Pradeep Ravikumar, & José Moura. Building Human-Machine Trust via Interpretability. Extended Abstract. AAAI Conference on Artificial Intelligence (AAAI) 2019. Honolulu, HI.

- [C1] **Umang Bhatt**, Edgar Xi, Shouvik Mani, & J. Zico Kolter. Intelligent Pothole Detection and Road Condition Assessment. Bloomberg Data for Good Exchange (D4GX) 2017. New York, NY.

Selected Workshop Papers

- [W4] Javier Antorán, **Umang Bhatt**, Tameem Adel, Adrian Weller, & José Miguel Hernández-Lobato. Getting a CLUE: A Method for Explaining Uncertainty Estimates. ICLR Workshop on Machine Learning in Real Life (ML-IRL) 2020. Addis Ababa, Ethiopia. **(Selected Oral Presentation)**
- [W3] **Umang Bhatt**, Adrian Weller, Muhammad Bilal Zafar, & Krishna Gummadi. Counterfactual Accuracies for Alternative Models. ICLR Workshop on Machine Learning in Real Life (ML-IRL) 2020. Addis Ababa, Ethiopia.
- [W2] **Umang Bhatt**, Brian Davis, & José Moura. Diagnostic Model Explanations: A Medical Narrative. AAAI Spring Symposium on Interpretable AI for Well-Being 2019. Stanford, CA.
- [W1] **Umang Bhatt**. Maintaining the Humanity of Our Models. AAAI Spring Symposium on AI, Society, and Ethics 2018. Stanford, CA.

Scholarships and Awards

- 2019 – 2022 Leverhulme Center for the Future of Intelligence PhD Scholarship
- 2019 Lovett Family Endowed Scholarship
- 2019 Best Presentation Award at The AAAI Spring Symposium on Interpretable AI for Well-Being
- 2017 Carnegie Mellon Undergraduate Research Presentation Award
- 2017 NSF I-Corps Site Award – for research commercialization
- 2016 H. F. McCullough Memorial Scholarship

Teaching Experience

- 2017 – 2019 **Carnegie Mellon University**, *Teaching Assistant*.
- S19 18-661 (Machine Learning for Engineers - Masters) taught by Gauri Joshi
- F18 10-701 (Machine Learning - PhD) taught by Ziv Bar-Joseph and Pradeep Ravikumar
- S18 15-388/15-688 (Practical Data Science) taught by Zico Kolter
- F17 15-122 (Principles of Imperative Computation) taught by Illiano Cervesato
- S17 15-110 (Principles of Computing) taught by Margret Reid-Miller

Service

- Organizer ICML Workshop on Human Interpretability (2020)
- Reviewer NeurIPS (2019-2020), ICML (2019-2020)
- PC Member ACM Conference on AI and Finance (2020), ICLR Workshop on Debugging Machine Learning Models (2019), AAAI/ACM Conference on AI, Ethics and Society Student Programs (2019)

Other Experience

- 2020 – now **Credo AI**, *Consultant*, Palo Alto, CA.
Creating an AI auditing platform - in stealth

- 2018 – 2020 **.406 Ventures**, *Student Fellow*, Boston, MA.
Sourced startups and performed first-round due diligence on ventures
- 2018 – 2020 **MERAT**, *Co-Founder*, Seattle, WA.
Started a non-profit to help corporations adopt AI responsibly
- 2017 – 2019 **NavTalent (now HumanCapital)**, *Campus Director*, San Francisco, CA.
Connected top-tier engineers with high-growth, impact-driven startups
- 2011 – 2019 **BAPS Children's Activities**, *National Development Committee*, USA.
Planned and oversaw nation-wide conventions and weekly assemblies on Hinduism for over 5,000 children, ages 6-14, with the team of fifteen other volunteers
- 2017 – 2018 **Perceptsense**, *Co-Founder*, Pittsburgh, PA.
Built products to harvest vehicular telematics data – pipeline now used by Honda Motors
- 2018 **Microsoft**, *Program Management Intern*, Redmond, WA.
Project: explainable conversational agents for technical hardware documentation
- 2017 **Groupon**, *Product Management Intern*, Chicago, IL.
Project: personalized and targeted pricing algorithms to drive purchase frequency
- 2016 **InquisitHealth**, *Software Engineering Intern*, New York, NY.
Project: speaker diarization in medical conversations and adverse event detection in patient SMS

Skills

Programming Python, Java, C, C#, R, Javascript, Tensorflow, PyTorch, Matlab, HTML/CSS

Courses Advanced Probabilistic Machine Learning (CAM), Machine Learning (CMU), Advances in AI for Social Good (CMU), Network Science (CMU), Pattern Recognition Theory (CMU)

Languages English (native), Gujarati (fluent), and Spanish (proficient)

References

Adrian Weller, Director of AI, The Alan Turing Institute and Principal Research Fellow, Department of Engineering, University of Cambridge: aw665@cam.ac.uk

José Moura, Philip L. and Marsha Dowd University Professor, Department of Electrical and Computer Engineering, College of Engineering, Carnegie Mellon University: moura@cmu.edu

Alice Xiang, Research Lead, Partnership on AI: alice@partnershiponai.org