

Umang Bhatt

CONTACT usb20@cam.ac.uk
INFORMATION umangsrbhatt.github.io

EDUCATION **University of Cambridge**, Cambridge, UK
Ph.D. in Engineering (Machine Learning) Sept 2019 – Present
Advisor: [Adrian Weller](#)
Affiliations: [Machine Learning Group](#), [Computation and Biological Learning Lab](#)

Carnegie Mellon University, Pittsburgh, PA
M.S. in Electrical and Computer Engineering Aug 2017 – May 2019
Advisor: [José M.F. Moura](#)
B.S. in Electrical and Computer Engineering Aug 2015 – May 2019

POSITIONS Fellow, **Mozilla Foundation** Oct 2020 – Present
Student Fellow, **Leverhulme Center for the Future of Intelligence** Sept 2019 – Present
Research Fellow, **Partnership on AI** June 2019 – Nov 2020
Research Assistant, **Carnegie Mellon University** Jan 2017 – Sept 2019
Mentors: Pradeep Ravikumar (MLD), Zico Kolter (CSD), Fei Fang (ISR), and Radu Marculescu (ECE)

- PEER-REVIEWED CONFERENCE PUBLICATIONS
- [1] **Getting a CLUE: A Method for Explaining Uncertainty Estimates**
International Conference on Learning Representations (ICLR) 2021 (Oral)
Javier Antorán, **Umang Bhatt**, Tameem Adel, Adrian Weller, José Miguel Hernández-Lobato
 - [2] **FIMAP: Feature Importance by Minimal Adversarial Perturbation**
AAAI Conference on Artificial Intelligence (AAAI) 2021
Matt Chapman-Rounds, **Umang Bhatt**, Erik Pazos, Marc-Andre Schulz, Kostas Georgatzis
 - [3] **Evaluating and Aggregating Feature-based Explanations**
International Joint Conference on Artificial Intelligence (IJCAI) 2020
Umang Bhatt, Adrian Weller, José M.F. Moura
 - [4] **Explainable Machine Learning in Deployment**
ACM Conference on Fairness, Accountability, and Transparency (FAccT) 2020
Umang Bhatt, Alice Xiang, Shubham Sharma, Adrian Weller, Ankur Taly, Yunhan Jia, Joydeep Ghosh, Ruchir Puri, José M.F. Moura, Peter Eckersley
 - [5] **You Shouldn't Trust Me: Learning Models Which Conceal Unfairness From Multiple Explanation Methods**
European Conference on Artificial Intelligence (ECAI) 2020
Botty Dimanov, **Umang Bhatt**, Mateja Jamnik, Adrian Weller
 - [6] **On Network Science and Mutual Information for Explaining Deep Neural Networks**
IEEE Conference on Acoustics, Speech and Signal Processing (ICASSP) 2020
Brian Davis*, **Umang Bhatt***, Kartikeya Bhardwaj*, Radu Marculescu, José M.F. Moura
 - [7] **A Robot's Expressive Language Affects Human Strategy and Perceptions in a Competitive Game**
IEEE Conference on Robot and Human Interactive Communication (ROMAN) 2019
Aaron Roth, Samantha Reig, **Umang Bhatt**, Johnathan Schulgach, Tamara Amin, Afsaneh Doryab, Fei Fang, Manuela Veloso
 - [8] **Intelligent Pothole Detection and Road Condition Assessment**
Bloomberg Data for Good Exchange (D4GX) 2017
Umang Bhatt, Edgar Xi, Shouvik Mani, J. Zico Kolter

| | |
|------------------------------------|---|
| SELECT WORKSHOP PUBLICATIONS | <p>[9] Effects of Uncertainty on the Quality of Feature Importance Estimates <i>AAAI Workshop on Explainable Agency in Artificial Intelligence 2021</i> Torgyn Shaikhina*, Umang Bhatt*, Roxanne Zhang, Kostas Georgatzis, Alice Xiang, Adrian Weller</p> <p>[10] On the Fairness of Causal Algorithmic Recourse <i>NeurIPS Workshop on Algorithmic Fairness through Causality and Interpretability 2020</i> Julius von Kügelgen, Umang Bhatt, Amir-Hossein Karimi, Isabel Valera, Adrian Weller, Bernhard Schölkopf</p> <p>[11] Machine Learning Explainability for External Stakeholders <i>IJCAI Workshop on Explainable AI 2020</i> Umang Bhatt, McKane Andrus, Alice Xiang, Adrian Weller</p> <p>[12] Diagnostic Model Explanations: A Medical Narrative <i>AAAI Spring Symposium on Interpretable AI for Well-Being 2019 (Best Presentation Award)</i> Umang Bhatt, Brian Davis, José M.F. Moura</p> <p>[13] Building Human-Machine Trust via Interpretability (Extended Abstract) <i>AAAI Conference on Artificial Intelligence (AAAI) 2019</i> Umang Bhatt, Pradeep Ravikumar, José M.F. Moura</p> |
| BOOK CHAPTERS | <p>[14] On Trustworthy Artificial Intelligence for Healthcare <i>Healthcare Information Technology for Cardiovascular Medicine</i>. Ed: A.B. Bhatt. Springer 2021 Umang Bhatt, Zohreh Shams</p> |
| PREPRINTS | <p>[15] Uncertainty as a Form of Transparency: Measuring, Communicating, and Using Uncertainty Umang Bhatt, Yunfeng Zhang, Javier Antorán, Vera Liao, Gabrielle Melançon, Prasanna Sattigeri, Ranganath Krishnan, Omesh Tickoo, Riccardo Fogliato, Jason Stanley, Lama Nachman, Rumi Chunnara, Adrian Weller, Alice Xiang</p> |
| INVITED TALKS | <p><i>Challenges in Deploying Explainable Machine Learning</i></p> <ul style="list-style-type: none"> • Robust and Responsible AI (RsqrD) Developers Meetup July 2020 • Keynote: ICML Workshop on Extending Explainable AI July 2020 • QuantumBlack (McKinsey) AI seminar May 2020 <p><i>Explainable Machine Learning in Deployment</i></p> <ul style="list-style-type: none"> • Keynote: Mozilla All-Hands June 2020 • ACM Conference on Fairness, Accountability, and Transparency (FAccT) Jan 2020 <p><i>Aggregating Feature-based Model Explanations</i></p> <ul style="list-style-type: none"> • International Joint Conference on Artificial Intelligence (IJCAI) July 2020 • Fiddler Labs May 2019 • AAAI Spring Symposium on Interpretable AI for Well-being Mar 2019 <p><i>Maintaining the Humanity of Our Models</i></p> <ul style="list-style-type: none"> • AILA Symposium on Creating a Fair and Ethical Future Oct 2018 • AAAI Spring Symposium on AI and Society Mar 2018 <p><i>Intelligent Pothole Detection</i></p> <ul style="list-style-type: none"> • University of Chicago's Data Science for Social Good Conference Oct 2017 • Bloomberg Data for Good Exchange Sept 2017 |
| TEACHING EXPERIENCE | <p>Thesis Co-Supervisor, University of Cambridge</p> <ul style="list-style-type: none"> • Daniel Ley, MEng in Information Engineering 2021 <p>Teaching Assistant (Supervisor/Grader), University of Cambridge</p> <ul style="list-style-type: none"> • <i>Probabilistic ML</i> (4F13) for Zoubin Ghahramani and J.M. Hernández-Lobato Michaelmas 2020 <p>Teaching Assistant, Carnegie Mellon University</p> <ul style="list-style-type: none"> • <i>Machine Learning for Engineers - Masters</i> (18-661) for Gauri Joshi Spring 2019 • <i>Machine Learning - PhD</i> (10-701) for Ziv Bar-Joseph and Pradeep Ravikumar Fall 2018 |

| | | |
|---------------------------------|--|--|
| | <ul style="list-style-type: none"> • <i>Practical Data Science</i> (15-688) for Zico Kolter • <i>Principles of Imperative Computation</i> (15-122) for Illiano Cervesato • <i>Principles of Computing</i> (15-110) for Margret Reid-Miller | Spring 2018 Fall 2017 Spring 2017 |
| PROFESSIONAL SERVICE | Co-Organizer <ul style="list-style-type: none"> • ELLIS Workshop on Human-Centric Machine Learning • ICML Workshop on Human Interpretability in Machine Learning • IBM + Partnership on AI Workshop on Explainable AI Program Committee <ul style="list-style-type: none"> • International Conference on Machine Learning (ICML) • ACM Conference on Knowledge Discovery and Data Mining (KDD) • ACM Conference on Fairness, Accountability, and Transparency (FAccT) • Conference on Artificial Intelligence and Statistics (AISTATS) • Conference on Artificial Intelligence (AAAI) • International Conference on Learning Representations (ICLR) • ACM Conference on Artificial Intelligence and Finance (ICAIF) • Conference on Neural Information Processing Systems (NeurIPS) | May 2021 July 2020 Feb 2020 2021 2021 2021 2021 2021 2021 2020 2020 |
| FELLOWSHIPS, AWARDS, AND HONORS | Mozilla Fellowship Leverhulme Center for the Future of Intelligence PhD Scholarship Fully funded by DeepMind and the Leverhulme Trust Partnership on AI Research Fellowship Best Presentation Award , AAAI Spring Symposium on Interpretable AI for Well-Being Family Endowed Scholarship , The Andrew Carnegie Society Undergraduate Research Presentation Award , CMU NSF I-Corps Site Award , for research commercialization H. F. McCullough Memorial Scholarship , CMU | 2020 – 2021 2019 – 2023 2019 – 2020 2019 Lovett 2019 2017 2017 2016 |
| PROFESSIONAL EXPERIENCE | Advisor, Credo AI , Palo Alto, CA • Creating an AI auditing platform - in stealth Student Fellow, .406 Ventures , Boston, MA • Sourced startups and performed first-round due diligence on ventures Program Management Intern, Microsoft , Redmond, WA • Project: explainable conversational agents for technical hardware documentation Co-Founder, Perceptsense , Pittsburgh, PA • Built products to harvest vehicular telematics data; pipeline acquired by Honda Motors Product Management Intern, Groupon , Chicago, IL • Project: personalized and targeted pricing algorithms to drive purchase frequency Software Engineering Intern, InquisitHealth , River Edge, NJ • Project: speaker diarization and adverse event detection in doctor-patient conversations | June 2020 – Present July 2018 – June 2020 May 2018 – Aug 2018 Jan 2017 – May 2018 May 2017 – Aug 2017 June 2016 – Sept 2016 |