Umang Bhatt

CONTACT usb20@cam.ac.uk Information umangsbhatt.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 Enrichment Student, The Alan Turing Institute

Oct 2021 - Present Student Fellow, Leverhulme Center for the Future of Intelligence Sept 2019 - Present Fellow, Mozilla Foundation Oct 2020 - Sept 2021 Research Fellow, Partnership on AI June 2019 - Sept 2020 Jan 2017 - Sept 2019

Research Assistant, Carnegie Mellon University Mentors: Pradeep Ravikumar (MLD), Zico Kolter (CSD), Fei Fang (ISR), and Radu Marculescu (ECE)

PEER-REVIEWED CONFERENCE **PUBLICATIONS**

[1] Uncertainty as a Form of Transparency: Measuring and Communicating Uncertainty

AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society (AIES) 2021

Umang Bhatt, Javier Antorán, Yunfeng Zhang, Q. Vera Liao, Prasanna Sattigeri, Riccardo Fogliato, Gabrielle Melançon, Ranganath Krishnan, Jason Stanley, Omesh Tickoo, Lama Nachman, Rumi Chunara, Madhulika Srikumar, Adrian Weller, Alice Xiang

[2] 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

[3] FIMAP: Feature Importance by Minimal Adversarial Perturbation

AAAI International Conference on Artificial Intelligence (AAAI) 2021 Matt Chapman-Rounds, *Umang Bhatt*, Erik Pazos, Marc-Andre Schulz, Kostas Georgatzis

[4] Evaluating and Aggregating Feature-based Explanations

International Joint Conference on Artificial Intelligence (IJCAI) 2020 *Umang Bhatt*, Adrian Weller, José M.F. Moura

[5] 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

[6] 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

[7] 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

[8] Building Human-Machine Trust via Interpretability

AAAI International Conference on Artificial Intelligence (AAAI) 2019 (Student Abstract) Umang Bhatt, Pradeep Ravikumar, José M.F. Moura

[9] 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

SELECT WORKSHOP PUBLICATIONS

[10] Do Concept Bottleneck Models Learn As Intended?

ICLR Workshop on Responsible AI 2021

Andrei Margeloiu*, Matth Ashman*, *Umang Bhatt**, Yanzhi Chen, Mateja Jamnik, Adrian Weller

[11] Fast Conformal Classification using Influence Functions

Symposium on Conformal and Probabilistic Prediction and Applications (COPA) 2021 Umang Bhatt, Adrian Weller, Giovanni Cherubin

[12] δ -CLUE: Diverse Sets of Explanations for Uncertainty Estimates

ICLR Workshop on Security and Safety in Machine Learning Systems 2021 Dan Ley, **Umang Bhatt**, Adrian Weller

[13] On the Fairness of Causal Algorithmic Recourse

NeurIPS Workshop on Algorithmic Fairness through Causality and Interpretability 2020 Julius von Kügelgen, **Umang Bhatt**, Amir Karimi, Isabel Valera, Adrian Weller, Bernhard Schölkopf

[14] Counterfactual Accuracies for Alternative Models

ICLR Workshop on Machine Learning in Real Life 2020 Umang Bhatt, Adrian Weller, Muhammad Bilal Zafar, Krishna Gummadi

[15] Diagnostic Model Explanations: A Medical Narrative

AAAI Spring Symposium on Interpretable AI for Well-Being 2019 (Best Presentation Award) Umang Bhatt, Brian Davis, José M.F. Moura

[16] Intelligent Pothole Detection and Road Condition Assessment

Bloomberg Data for Good Exchange (D4GX) 2017 Umang Bhatt, Edgar Xi, Shouvik Mani, J. Zico Kolter

Book Chapters

· Harvard SEAS

[17] Trust in Artificial Intelligence: Clinicians are Essential

Healthcare Information Technology for Cardiovascular Medicine. Springer 2021 Umang Bhatt, Zohreh Shams

SELECT FELLOWSHIPS AND AWARDS	The Alan Turing Institute Enrichment Studentship Mozilla Fellowship Partnership on AI Research Fellowship Leverhulme Center for the Future of Intelligence PhD Scholarship Fully funded by DeepMind and the Leverhulme Trust Best Presentation Award, AAAI Spring Symposium on Interpretable AI for Well-Being Lovett 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	2021 - 2022 2020 - 2021 2019 - 2020 2019 - 2023 2019 - 2019 2019 2017 2017 2016
Grants	Co-Investigator, "Social Explainability (SOXAI) for Trustworthy AI" £107,000 from EPSRC via Research Centre on Privacy, Harm Reduction, and Adversarial Influence PI: Frens Kroeger (Coventry); CI: James Hancock (Stanford) and Beate Grawemeyer (Coventry)	2021-2022
SELECT INVITED TALKS	 Beyond Feature Importance: Explaining Uncertainty Estimates Huawei Strategy & Technology Workshop International Conference on Learning Representations (ICLR) Technical University of Denmark 	Oct 2021 May 2021 Apr 2021

Apr 2021

	Challenges in Deploying Explainable Machine Learning	
	Imperial College Explainable AI Seminar	Feb 2021
	Robust and Responsible AI (Rsqrd) Developers Meetup	July 2020
	Keynote: ICML Workshop on Extending Explainable AI	July 2020
	Keynote: Mozilla All-Hands Meeting	June 2020
	QuantumBlack (McKinsey) AI Seminar	May 2020
	 ACM Conference on Fairness, Accountability, and Transparency (FAccT) 	Jan 2020
	Aggregating Feature-based Model Explanations	
	International Joint Conference on Artificial Intelligence (IJCAI)	July 2020
	• Fiddler Labs	May 2019
	AAAI Spring Symposium on Interpretable AI for Well-being	Mar 2019
	• Element AI	Dec 2018
TEACHING	Thesis Co-Supervisor/Mentor, University of Cambridge	
Experience	Varun Babbar, MEng in Information Engineering	2021 – Present
	Javier Abad Martinez, Research Assistantship	2021 – Present
	• Dan Ley, MEng in Information Engineering (Next: Data Scientist at JP Morgan)	2020 - 2021
	• Charlie Rogers Smith, Research Assistantship (Next: Engineer at Future of Humanity Institut	
	Teaching Assistant (Supervisor/Grader), University of Cambridge	-,
	• Probabilistic ML (4F13) for Zoubin Gharamani and J.M. Hernández-Lobato	Michaelmas 2020
	Teaching Assistant, Carnegie Mellon University	
	Machine Learning for Engineers - Masters (18-661) for Gauri Joshi	Spring 2019
	• <i>Machine Learning - PhD</i> (10-701) for Ziv Bar-Joseph and Pradeep Ravikumar	Fall 2018
	• Practical Data Science (15-688) for Zico Kolter	Spring 2018
	Principles of Imperative Computation (15-122) for Illiano Cervesato	Fall 2017
	• Principles of Computing (15-110) for Margret Reid-Miller	Spring 2017
		1 0
Professional	Co-Organizer	
SERVICE	ELLIS Workshop on Human-Centric Machine Learning	May 2021
CERVICE	ICML Workshop on Human Interpretability in Machine Learning	July 2020
	• IBM + Partnership on AI Workshop on Explainable AI	Feb 2020
		100 2020
	Program Committee	
	AAAI – International Conference on Artificial Intelligence	2021, 2022
	AISTATS – Conference on Artificial Intelligence and Statistics	2021, 2022
	ICLR – International Conference on Learning Representations	2021, 2022
	NeurIPS – Conference on Neural Information Processing Systems A COMP – A COMP	2020, 2021
	ICAIF – ACM Conference on Artificial Intelligence and Finance	2020, 2021
	UAI – Conference on Uncertainty in Artificial Intelligence Output Description: O	2021
	ICML – International Conference on Machine Learning ICML – International Conference on Machine Lea	2021
	KDD – ACM Conference on Knowledge Discovery and Data Mining FA - T - ACM Conference on Friends Associated Hilliam and Transport and Tran	2021
	 FAccT – ACM Conference on Fairness, Accountability, and Transparency Reviewer 	2021
	Journal of Artificial Intelligence Research (JAIR)	
	ACM Transactions on Computer-Human Interaction (TOCHI)	
	Artificial Intelligence Journal (AIJ)	
	ACM Transactions on Interactive Intelligent Systems (TiiS)	
	8,	
Professional	Advisor, Responsible AI Institute , Austin, TX	Oct 2021 – Present
Experience	Building tools and certification programs for Responsible AI	1100011
		e 2020 – June 2021
	Created an AI auditing platform; backed by AI Fund and Andrew Ng	
		y 2018 – June 2020
	Sourced startups and performed first-round due diligence on ventures	, ,
		n 2017 – May 2018
	Built products to harvest vehicular telematics data; pipeline acquired by <i>Honda Motors</i>	,
	* * * * * * * * * * * * * * * * * * * *	