HUA SHEN

Ph.D Candidate in Informatics - The PennState University - PA, USA

✉ huashen218@psu.edu; 🌐 http://hua-shen.org; 🎓 Google Scholar

EDUCATION

Ph.D. Candidate, Informatics   *2019.08 - Present*
College of IST, Penn State University.
Advisor: Dr. Ting-Hao (Kenneth) Huang

Ph.D. Student, Computer Science   *2018.08 - 2019.08*
Department of CSE , Lehigh University. *Graduate Fellowship*
Advisors: Dr. Ting Wang, Brian D. Davison (transferred)

M.S., Management Science and Engineering   *2013.09 - 2016.06*
School of Information, Renmin University of China. *Exempt from Admission Exam*
Advisor: Dr. Xun Liang

B.S., Information Security   *2009.09 - 2013.06*
Department of CSE, University of Science and Technology Beijing. *Ranking: 2/35*

RESEARCH INTERESTS

My research interests lie in general area of **AI**, **HCI**. My current research focuses on **human-centric machine learning interpretability**, particularly on improving AI interpretability for human in **natural language processing** and **computer vision** tasks. I am also interested in exploring **Human-AI Interaction** through **crowdsourcing** to better interpret deep learning models.

PUBLICATIONS

Conference Papers

○ **Hua Shen**, Ting-hao (Kenneth) Huang. How Useful Are the Machine-Generated Interpretations? A Human Evaluation on Guessing the Wrongly Predicted Labels. The 8th AAAI Conference on Human Computation and Crowdsourcing (HCOMP '20).

○ Xinyang Zhang, Ningfei Wang, **Hua Shen**, Shouling Ji, Ting Wang. Interpretable Deep Learning under Fire. Proceedings of the 29th USENIX Security Symposium (USENIX'20).

○ Ren Pang, **Hua Shen**, Xinyang Zhang, Shouling Ji, Yevgeniy Vorobeychik, Xiapu Luo, Alex X. Liu, Ting Wang. The Tale of Evil Twins: Adversarial Inputs versus Poisoned Models. The 27th ACM Conference on Computer and Communications Security (CCS'20).

○ **Hua Shen**, Xun Liang. A Quantitative Analysis Decision System Based on Deep Learning and NSGA-II for FX Portfolio Prediction". The 31st International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (IEA-AIE'18).

○ **Hua Shen**, Xun Liang. A Time Series Forecasting Model Based on Deep Learning Integrated Algorithm with Stacked Autoencoders and SVR for FX Prediction. The 25th International Conference on Artificial Neural Networks (ICANN'16).

○ **Hua Shen**, Xun Liang, Mingming Wang. Emergency Decision Support Architectures for Bus Hijacking Based on Massive Image Anomaly Detection in Social Networks. The 2015 IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC'15).

Book

○ Xun Liang, Xiaoping Yang, **Hua Shen**. Social Commerce Theory and Practice[M], Tsinghua University Press, 2014, ISBN No. 9787302381129.

National Patent

○ Xun Liang, **Hua Shen**, Run Cao. An innovative Emergency Discovery Method for Micro-blogging, Archived by State Intellectual Property Office of The P.R.C (SIPO), Patent Publication No. CN103577404A.

○ Xiaofei Li, Xun Liang, Xiaoping Zhou, Xiaojing Shi, **Hua Shen**, Haiyan Zhang. A Cross-Platform Microblog Community Account Matching Method, Archived by State Intellectual Property Office of The P.R.C (SIPO), Patent Publication No. CN104765729A.

HIGHLIGHTED PROJECTS

Usefulness Evaluation for AI Interpretation by General Users   *May 2020 - June 2020*

○ We investigate whether showing machine-generated interpretations helps MTurk crowdworkers understand the incorrectly predicted labels by image classifiers. It is published at HCOMP'20 conference.

○ I generate a set of saliency maps (e.g., SmoothGrad, ExtremalPerturb, GradCAM) for ResNet model on ImageNet dataset. The human evaluation is finished on Amazon MTurk crowdsourcing platform.

Crowd-Powered Conversational AI System Explanation   *May 2020 - Present*

○ I am developing a dialogue system with frontend as Facebook Messenger and backend powered by Amazon Mechanical Turk. I will conduct interpretation related analysis on the dialogue system.

○ I use full-stack techniques including *Python*, *Pytorch* for deep learning models, *SQLAlchemy* for database, *HTML, CSS, JavaScript, ReactJS* for website development.

iAlgebra: Towards Interactive Interpretability of Neural Networks   *Sep 2019-Sep 2020*

○ I independently implement the first-of-its-kind interactive framework for interpreting deep neural networks. The framework includes a *Pytorch-iAlgebra* package to fulfill the models interpretation as Pytorch backend and a user interface to interact with human as frontend.

○ I conducted all the user studies independently using MTurk crowdsourcing in representative tasks such as inspecting adversarial inputs, resolving model inconsistency and cleansing contaminated data.

Testing Robustness by Attacking Interpretation and DNN Models   *Sep 2018-May 2019*

○ We propose a new class of attacks that generate adversarial inputs misleading both interpretations and classifiers. The work is published in USENIX'20.

○ We attacked different types of saliency maps, including GradSaliency, CAM, RTS, MASK. The evaluation is tested on real applications like skin cancer and ImageNet datasets.

Analyzing Robustness by Explaining Question Answering Models   *Sep 2018-May 2019*

○ We aim to analyze QA model robustness by interpreting adversarial attack. I designed an interpretation method for QA model to interpret layer-wise and token-wise attention-based importance scores.

○ I analyzed SQuAD datasets (with adversaries as ADDSent, ADDAny, etc.) on QA models (BiDAF, Match-LSTM) using Pytorch.

TEACHING ASSISTANT EXPERIENCE

The Pennsylvania State University

○ Application Development Studio I	Fall 2020
○ Algorithmic Methods	Fall 2020

Renmin University of China

○ Calculus I and Calculus II	Spring 2015 and Fall 2015
○ Linear Algebra	Fall 2014
○ Probability Theory and Mathematical Statistics	Fall 2013 and Spring 2014

SELECTED HONORS

Graduate Student Travel Grant, IST, PennState,	2020
Travel Grant Award, CRA-W Grad Cohort Workshop	2019
Graduate Fellowship, Lehigh University	2018-2019
Outstanding Master Degree Candidate, RenMin University of China	2015
Outstanding Merit Student, Beijing Municipal	2013
Excellent Graduate, Beijing Municipal	2013
Outstanding-Dissertation Award, USTB	2013
National Endeavor Fellowship, Nationwide *(twice)*	2011 and 2012
Vice Chairman, The 22nd Student Union of USTB	2011-2012
Alumni Funding Scholarship, USTB	2012