Texas A&M University
https://mengnandu.com
Phone: +1 979 267 5303
Email: dumengnan@tamu.edu

Mengnan Du

Education

2017.8 - Now **Ph.D Computer Science**, *Texas A&M University*.
Advisor: Dr. Xia Ben Hu

2014.9 - 2017.3 **M.S. Electrical Engineering**, *Beihang University*, Beijing, *GPA: 3.80/4.00*.
Advisor: Prof. Xingming Wu, Prof. Weihai Chen and Prof. Zhengguo Li

2010.9 - 2014.6 **B.E. Electrical Engineering**, *North China Electric Power University*, Beijing, *GPA: 3.91/4.00*.
Advisor: Associate Prof. Guangsheng Liang

Industry Experience

2020.5 - 2020.8 **Machine Learning Research Intern**, *Adobe*, San Jose (Remotely in College Station, Texas).
○ Bias detection and mitigation for few shot learning applications.
○ Mentors: Varun Manjunatha, Tong Sun, Rajiv Jain, Jiuxiang Gu, Ruchi Deshpande, Franck Dernoncourt

2017.3 - 2017.8 **Software Engineer Intern**, *Intel Asia-Pacific Research & Development Ltd*, Shanghai.
○ Worked on an AR (Augmented reality) project for Smart glasses.
○ Developed image classification and speech recognition algorithms which can be deployed on Android system.

Research Interests

2017 -Till now Natural Language Understanding and Explainable AI

Publications (Google Scholar) Citations: 380+, h-index: 9, as of November 2020

AAAI Huiqi Deng, Na Zou, **Mengnan Du**, Weifu Chen, Guocan Feng, Xia Hu, "A Unified Taylor Framework for Revisiting Attribution Methods", Thirty-Fourth AAAI Conference on Artificial Intelligence, 2021.

ICWSM Sina Mohseni, Fan Yang, Shiva Pentyala, **Mengnan Du**, Yi Liu, Nic Lupfer, Xia Hu, Shuiwang Ji, and Eric Ragan, "Machine Learning Explanations to Prevent Overtrust in Fake News Detection", International AAAI Conference on Web and Social Media, 2021.

CACM **Mengnan Du**, Ninghao Liu, and Xia Hu, "Techniques for Interpretable Machine Learning", Communications of the ACM, 2020. Highlighted Article on the cover page of the January 2020 issue

IEEE IS **Mengnan Du**, Fan Yang, Na Zou, and Xia Hu, "Fairness in Deep Learning: A Computational Perspective", IEEE Intelligent Systems, 2020.

CIKM **Mengnan Du***, Shiva Pentyala*, Yuening Li, Xia Hu, "Towards Generalizable Deepfake Detection with Locality-aware AutoEncoder", The 29th ACM International Conference on Information and Knowledge Management, 2020. (*Equal contribution)

KAIS **Mengnan Du**, Ninghao Liu, Fan Yang, and Xia Hu, "Learning Credible DNNs via Incorporating Prior Knowledge and Model Local Explanation", Knowledge and Information Systems, 2020.

KDD Ruixiang Tang, **Mengnan Du**, Ninghao Liu, Fan Yang, and Xia Hu, "An Embarrassingly Simple Approach for Trojan Attack in Deep Neural Networks", The 26rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2020.

SDM Fan Yang, Ninghao Liu, **Mengnan Du**, Kaixiong Zhou, Shuiwang Ji and Xia Hu, "Deep Neural Networks with Knowledge Instillation", SIAM International Conference on Data Mining, 2020.

BMC Nur Hafieza Ismail, Ninghao Liu, **Mengnan Du**, Zhe He and Xia Hu, "A Deep Learning Approach for Identifying Cancer Survivors Living with Post-Traumatic Stress Disorder on Twitter", BMC Medical Informatics and Decision Making, 2020.

ICDM **Mengnan Du**, Ninghao Liu, Fan Yang, and Xia Hu, "Learning Credible Deep Neural Networks with Rationale Regularization", IEEE International Conference on Data Mining, 2019. Best Paper Candidate

WWW **Mengnan Du**, Ninghao Liu, Fan Yang, Shuiwang Ji, and Xia Hu, "On Attribution of Recurrent Neural Network Predictions via Additive Decomposition", The Web Conference, 2019. Best Paper Candidate

WWW Fan Yang*, Shiva Pentyala*, Sina Mohseni*, **Mengnan Du***, Hao Yuan*, Rhema Linder, Eric Ragan, Shuiwang Ji and Xia Hu, "XFake: Explainable Fake News Detector with Visualizations", The Web Conference, demo track, 2019. (*Equal contribution)

WSDM Ninghao Liu, **Mengnan Du** and Xia Hu, "Representation Interpretation with Spatial Encoding and Multimodal Analytics", ACM International Conference on Web Search and Data Mining, 2019.

BCB Nur Hafieza Ismail, **Mengnan Du**, Diego Martinez and Zhe He, "Multivariate Multi-step Deep Learning Time Series Approach in Forecasting Parkinson's Disease Future Severity Progression", ACM BCB 2019.

MEDINFO Nur Hafieza Ismail, Ninghao Liu, **Mengnan Du**, Zhe He, and Xia Hu, "Identification of Cancer Survivors Living with Post-Traumatic Stress Disorder (PTSD) on Social Media", The World Congress on Medical and Health Informatics, poster, 2019.

CIKM Yuening Li, Xiao Huang, Jundong Li, **Mengnan Du**, and Na Zou, "SpecAE: Spectral Autoencoder for Anomaly Detection in Attributed Networks", ACM International Conference on Information and Knowledge Management, 2019. INFORMS 2019 Best Refereed Paper Finalists

IJCNN Yuening Li, Ninghao Liu, Jundong Li, **Mengnan Du**, and Xia Hu, "Deep Structured Cross-Modal Anomaly Detection", International Joint Conference on Neural Networks, 2019.

KDD **Mengnan Du**, Ninghao Liu, Qingquan Song, and Xia Hu, "Towards Explanation of DNN-based Prediction with Guided Feature Inversion", The 24rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2018.

DSP **Mengnan Du**, Xingming Wu, Weihai Chen, Zhengguo Li, "Supervised Training and Contextually Guided Salient Object Detection", Digital signal processing, 2017.

JEI **Mengnan Du**, Xingming Wu, Weihai Chen, Jianhua Wang, "Exploiting multiple contexts for saliency detection", Journal of Electronic Imaging, 2016.

SCIENCE CHINA Xingming Wu, **Mengnan Du**, Weihai Chen, Jianhua Wang, "Salient object detection via region contrast and graph regularization", SCIENCE CHINA Information Sciences, 2016.

ICIEA Xingming Wu, **Mengnan Du**, Weihai Chen, Zhengguo Li, "Exploiting Deep Convolutional Neural Network and Patch-level CRFs for Indoor Semantic Segmentation", in IEEE Conference on Industrial Electronics and Applications, 2016.

CCDC **Mengnan Du**, Xingming Wu, Weihai Chen, Jianhua Wang, "Fusing region contrast and graph regularization for saliency detection", Chinese Control and Decision Conference, 2015.

Measurement Control Technology Guangsheng Liang, **Mengnan Du**, Zihao Zhou, et al. "Attitude control based on complementary filter for two-wheel self-balance car", Measurement Control Technology, 2015. (**in chinese**)

Workshop Papers

CHI workshop Sina Mohseni, Fan Yang, Shiva Pentyala, **Mengnan Du**, Yi Liu, Nic Lupfer, Xia Hu, Shuiwang Ji, Eric Ragan, "Trust Evolution Over Time in Explainable AI for Fake News Detection", Fair & Responsible AI Workshop at CHI, 2020.

CVPR workshop Haofan Wang, Zifan Wang, **Mengnan Du**, Fan Yang, Zijian Zhang, Sirui Ding, Piotr Mardziel, Xia Hu, "Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks", IEEE CVPR Workshop on Fair, Data-Efficient and Trusted Computer Vision, 2020.

SEPDA Nur Hafieza Ismail, Ninghao Liu, **Mengnan Du**, Zhe He, and Xia Hu, "Using Deep Neural Network to Identify Cancer Survivors Living with Post-Traumatic Stress Disorder on Social Media ", The 4th Interpretable Workshop on Semantics-Powered Data Mining and Analytics, 2019.

Book Chapter

Springer Nur Hafieza Ismail, **Mengnan Du**, Xia Hu, "Social Media and Psychological Disorder", Social Web and Health Research, Springer, 2019. 171-192.

Preprints

Fairness Ruixiang Tang, **Mengnan Du**, Yuening Li, Zirui Liu and Xia Hu, "Mitigating Gender Bias in Captioning Systems", arXiv:2006.08315, 2020.

Watermark	Ruixiang Tang, **Mengnan Du**, and Xia Hu, "Deep Serial Number: Computational Watermarking for DNN Intellectual Property Protection", arXiv:2011.08960, 2020.
XAI	Fan Yang, **Mengnan Du**, and Xia Hu, "Evaluating Explanation Without Ground Truth in Interpretable Machine Learning", arXiv:1907.06831, 2019.
XAI	Ninghao Liu, **Mengnan Du**, and Xia Hu, "Adversarial Machine Learning: An Interpretation Perspective", arXiv:2004.11488, 2020.
XAI	Weijie Fu, Meng Wang, **Mengnan Du**, Ninghao Liu, Shijie Hao and Xia Hu, "Distribution-Guided Local Explanation for Black-Box Classifiers", 2019.
NAS	Yijun Bian, Qingquan Song, **Mengnan Du**, Jun Yao, Huanhuan Chen and Xia Hu, "Sub-Architecture Ensemble Pruning in Neural Architecture Search", arXiv:1910.00370, 2019.

Professional Talks

2020	"Towards Generalizable Deepfake Detection with Locality-aware AutoEncoder", CIKM-20
2020	"Towards Effective Interpretation of Deep Neural Networks: Algorithms and Applications", Invited talk at Bio-Seminar, ECE Department of Texas A&M University
2019	"Learning Credible Deep Neural Networks with Rationale Regularization", ICDM-19
2019	"On Attribution of Recurrent Neural Network Predictions via Additive Decomposition", WWW-19
2018	"Towards Explanation of DNN-based Prediction with Guided Feature Inversion", KDD-18

Services

Conference	Program Committee member of AAAI-2021, NAACL-21, AAAI-21 Workshop on Content Authoring and Design, NeurIPS-2020, ACL-2020, EMNLP-2020, AAAI-2020, KDD-2020, and AMIA-2020. Reviewer of ICML UDL Workshop 2020, and ICRA-2019
Journal	Reviewer of ACM Computing Surveys, IEEE Transactions on Neural Networks and Learning Systems, IEEE Signal Processing Letters, Journal of Electronic Imaging, IEEE Access

Major Honors & Awards

2020	Graduate Research Excellence Award, Texas A&M University
2020	NortonLifeLock Graduate Fellowship 2020 finalist
2020	CIKM Student Travel Award
2019	ICDM 2019 Best Paper Candidate
2019	Texas A&M Institute of Data Science Graduate Travel Grant
2019	ICDM Student Travel Award
2019	INFORMS 2019 Best Refereed Paper Finalists
2019	WWW 2019 Best Paper Shortlist
2019	TAMU CSE Student Travel Award covering the attendance of the WWW 2019
2018	KDD Student Travel Award