Hassan Ali

PhD Candidate, UNSW

hassanalikhatim@gmail.com | +61435077456 | Lahore, Pakistan
0000-0002-1701-0390 | hassanalikhatim | hassanalikhatim.github.io

What do I think of myself?

I am a self-motivated machine learning researcher who strives to enable real-word deployment of machine learning models that people can trust.

Education

Sep 2023 – Ongoing | **University of New South Wales (UNSW), Sydney, Australia**
PhD in Trustworthy Machine Learning

Sep 2017 – Aug 2019 | **National University of Sciences and Technology (NUST), Islamabad, Pakistan**
MS in Electrical Engineering

Sep 2013 – Aug 2017 | **University of Engineering and Technology (UET), Lahore, Pakistan**
BS in Electrical Engineering

Work Experience

Sep 2021 - Sep 2023 | **Information Technology University (ITU)**
Research Assistant
- Human-centric Robust ML-driven IoT Smart Services

Jan 2021 - Nov 2021 | **Information Technology University (ITU)**
Research Assistant
- Mitigating Anti-social Behavior through Beneficial AI

Publications

2023
1. **Ali, H.**, Khan, M. S., AlGhadhban, A., Alazmi, M., Alzamil, A., Al-utaibi, K. & Qadir, J. Con-detect: Detecting adversarially perturbed natural language inputs to deep classifiers through holistic analysis. *Computers & Security* **125,** 103367 (2023).
2. Butt, M. A., Qayyum, A., **Ali, H.**, Al-Fuqaha, A. & Qadir, J. Towards secure private and trustworthy human-centric embedded machine learning: An emotion-aware facial recognition case study. *Computers & Security* **125,** 103058 (2023).
3. Qayyum, A., Butt, M. A., **Ali, H.**, Usman, M., Halabi, O., Al-Fuqaha, A., Abbasi, Q. H., Imran, M. A. & Qadir, J. Secure and Trustworthy Artificial Intelligence-Extended Reality (AI-XR) for Metaverses. *ACM Comput. Surv.* (2023).

2022
4. **Ali, H.**, Khan, M. S., Al-Fuqaha, A. & Qadir, J. Tamp-X: Attacking explainable natural language classifiers through tampered activations. *Computers & Security* **120,** 102791 (2022).

2021
5. **Ali, H.**, Khan, M. S., AlGhadhban, A., Alazmi, M., Alzamil, A., Al-Utaibi, K. & Qadir, J. All your fake detector are belong to us: evaluating adversarial robustness of fake-news detectors under black-box settings. *IEEE Access* **9,** 81678–81692 (2021).
6. Petrick, N., Akbar, S., Cha, K. H., Nofech-Mozes, S., Sahiner, B., Gavrielides, M. A., Kalpathy-Cramer, J., Drukker, K., Martel, A. L. & BreastPathQ Challenge Group, f. t. SPIE-AAPM-NCI BreastPathQ Challenge: an image analysis challenge for quantitative tumor cellularity assessment in breast cancer histology images following neoadjuvant treatment. *Journal of Medical Imaging* **8,** 034501–034501 (2021).

2020

7. Khalid, F., **Ali, H.**, Hanif, M. A., Rehman, S., Ahmed, R. & Shafique, M. *FaDec: A Fast Decision-based Attack for Adversarial Machine Learning* in *2020 International Joint Conference on Neural Networks (IJCNN)* (2020), 1–8.

2019

8. **Ali, H.**, Khalid, F., Tariq, H. A., Hanif, M. A., Ahmed, R. & Rehman, S. SSCNets: Robustifying DNNs using Secure Selective Convolutional Filters. *IEEE Design & Test* **37,** 58–65 (2019).

9. Khalid, F., **Ali, H.**, Tariq, H., Hanif, M. A., Rehman, S., Ahmed, R. & Shafique, M. *QuSecNets: Quantization-based defense mechanism for securing deep neural network against adversarial attacks* in *2019 IEEE 25th International Symposium on On-Line Testing and Robust System Design (IOLTS)* (2019), 182–187.