

# Sandhya Tripathi, Ph.D.

✉ sandhyat@wustl.edu

🌐 <https://www.linkedin.com/in/sandhyatripathi/>

## Summary

Experienced Postdoctoral Research Scholar adept at developing, evaluating and executing high-quality research into the improvement of AI-based clinical decision support systems. Seeking a position where I can contribute my ideas in pursuing AI applications in healthcare.

## Work Experience

Aug'20 – Pres

📌 **Post Doctoral Research Scholar**, Department of Anesthesiology, Washington University School of Medicine (WUSM) St. Louis, USA.

- Developed schema matching technique for Electronic Health Records (EHRs) integration using AutoEncoders that utilizes expert knowledge too.
- Evaluated machine learning models for the problem of algorithmic bias and unfairness using first principles of statistical learning techniques.
- Developed model report card to evaluate a machine learning model's performance in terms of additional utility over conventional risk scores, uncertainty incorporation, fairness across age, sex, race etc. Will be shown to decision makers in addition to AI's decision.

Feb'20 – Jul'20

📌 **Visiting Researcher** Department of Anesthesiology, WUSM.

- Evaluated existing outlier detection techniques for medication and laboratory data from EHRs to be finally used in the machine learning model deployment.

## Education

Jul'15 – Jan'20

📌 **Ph.D. Machine Learning, Indian Institute of Technology Bombay, India .**

Thesis title: *Label Noise Robustness and Feature Subset Selection in Classification: Discriminative and Generative Approaches.*

- Designed algorithms to learn in the presence of label noise (due to crowdsourcing or sensor breakdown) in a cost-sensitive classification setup using GANs & conventional squared loss.
- Developed a theoretical framework using exponential loss for scalable binary classifiers.
- Modelled a novel classification game for explainable & interpretable feature selection.
- Teaching Assistantship: Data Analytics in Operations Research, Online Machine Learning, Computer Programming & Algorithms (Lab component), Decision Analysis and Game Theory, Engineering Statistics, Optimization Techniques.

Jul'12 – Jun'15

📌 **M.Sc. Operations Research , Indian Institute of Technology Bombay, India .**

Jul'09 – Jun'12

📌 **B.Sc. Statistics (H), Lady Sri Ram College for Women, India.**

## Technical skills

Programming skills	📌 Python, Pytorch, AMPL, Docker (Proficient).
Statistical and simulation tools	📌 R, SPSS, Anylogic, Scilab, Mathematica (Intermediate).
Databases	📌 MySQL, PostgreSQL (Beginner).

## Research Publications

- 1 **Tripathi, S.,** Fritz, B. A., Abdelhack, M., Avidan, M. S., Chen, Y., & King, C. R. (2022). Deep learning to jointly schema match, impute, and transform databases. *arXiv preprint arXiv:2207.03536*.
- 2 **Tripathi, S.,** Fritz, B. A., Avidan, M. S., Chen, Y., & King, C. R. (2022). Algorithmic bias in machine learning based delirium prediction. *arXiv preprint arXiv:2211.04442*.

- 3 Abdelhack, M., Zhang, J., **Tripathi, S.**, Fritz, B. A., Avidan, M. S., Chen, Y., & King, C. R. (2021). A modulation layer to increase neural network robustness against data quality issues. *arXiv preprint arXiv:2107.08574*.
- 4 Hemachandra, N., Patil, K., & **Tripathi, S.** (2020). Equilibrium points and equilibrium sets of some  $GI/M/1$  queues. *Queueing Systems*, 96(3), 245–284.
- 5 Petety, A., **Tripathi, S.**, & Hemachandra, N. (2020). Attribute noise robust binary classification (Student Abstract). In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, pp. 13897–13898).
- 6 **Tripathi, S.**, Fritz, B. A., Abdelhack, M., Avidan, M. S., Chen, Y., & King, C. R. (2020). (Un) fairness in post-operative complication prediction models. *arXiv preprint arXiv:2011.02036*.
- 7 **Tripathi, S.**, & Hemachandra, N. (2020). GANs for learning from very high class conditional noisy labels. *arXiv preprint arXiv:2010.09577*.
- 8 **Tripathi, S.**, Hemachandra, N., & Trivedi, P. (2020). Interpretable feature subset selection: A Shapley value based approach. In *IEEE International Conference on Big Data (Big Data)* (pp. 5463–5472). IEEE.
- 9 **Tripathi, S.**, & Hemachandra, N. (2019). Cost sensitive learning in the presence of symmetric label noise. In *Advances in Knowledge Discovery and Data Mining* (pp. 15–28). Springer International Publishing.
- 10 **Tripathi, S.**, & Hemachandra, N. (2018). Scalable linear classifiers based on exponential loss function. In *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data* (pp. 190–200).

## Presentations and posters

- Oct'20     ■ **IEEE DSAA 2020** Conducted a tutorial on 'Label noise: Problems and solutions'.
- Mar'19     ■ **IRCIN, University of Tokyo** Attended and presented a poster titled "GANs based label noise robust classification" in Neuro-inspired Computation Course.
- Jul'16     ■ **ECQT 2016, Toulouse** Presented the work titled "Equilibrium sets of some  $GI/M/1$  queues" in European Conference on Queueing Theory.
- Aug'15     ■ **MLSS 2015, University of Kyoto** Attended and presented a poster titled "Some Equilibrium Prediction Problems in  $M/M/1$  Queues" in International Machine Learning Summer School.

## Miscellaneous Experience

### Biology course projects (research)

- Jul–Nov'17     ■ **BB553 Bio-informatics**, Identification of protein-protein interface residues using ExpERM
- Explored methods for predicting protein-protein interface residues using both supervised and unsupervised machine learning approaches.
- **BB607 Proteomics**, Human Proteome Reference Maps
- Presented a survey of pioneering research papers on human proteome map.

### Academic services

- **PC member:** ECML2020, IEEE BigData'20 Special Track-XAI, IJCAI-21, ML4H-21,-22
- **Reviewer:** SADHANA IAS journal, Journal of Biomedical Informatics, Neural Networks

### Certification

- Sep'20     ■ **AI for Medicine Specialization.** Awarded by Coursera.

## Non-academic achievements

- Jun'16     ■ Member of **Indian Youth Delegation** to China organized by Government of India.
- Jun'11     ■ Awarded **NCC 'C'** grade senior division certificate.