Mehak Gupta

✉ mehakg@smu.edu
sites/mehak25/
in,

Current Position

2023 **Assistant Professor**, *Southern Methodist University*.
- **Instructor** Operating Systems

Research Interest

My research interest is in Artificial Intelligence (AI) and its interdisciplinary applications. Within AI, I am interested in problems related to natural language processing, representation learning, predictive modeling with supervised and semi-supervised deep learning models, and digital transformation with applications in healthcare and other domains. I am also interested in improving the fairness and explainability of AI systems to enhance their usability and adaptability.

Education

2018 – 2023 **PhD., Computer Science**, *University of Delaware (4.0 GPA)*.
Thesis Title: Deep Learning Predictive Modelling for Electronic Health Record (*Completed*)
Advisor: Prof. Rahmatollah Beheshti

2012–2014 **Masters in Engineering, Software Engineering**, *Thapar University, India*.
Graduation grade: 8.85 GPA

2006–2010 **Bachelors of Engineering, Computer Engineering** , *Punjabi University, India* .
Graduation grade: 73%

Research Experience (4+ years) [Google Scholar Profile]

Jun'19–Jul'23 **Graduate Research Assistant**, *Healthy IAife Lab, University of Delaware*.
Cleaning and pre-processing real-world EHR data for downstream analysis tasks, developing and evaluating deep learning prediction models (such as RNN-LSTM, Seq2Seq models, attention models, and GANs) and interpreting results for clinical outcome prediction.
- **Influence of social determinants on BMI trajectories of children:** Investigate the influence of social environmental community-level variables on children with different BMI trajectories. *(Submitted for Review)*
- **An Extensive Data Preprocessing Pipeline for MIMIC-IV:** Developed open-source repository to clean, pre-process, model, and evaluate open-source MIMIC-IV EHR data for downstream prediction modelling using machine learning and deep learning models. *(accepted)*
- **Long-term and short-term time-series predictions to determine future health trajectory:** Developed semi-supervised generative adversarial model with recurrent layers to predict different lengths of future time-series of health trajectory by learning from labeled and unlabeled samples. *(accepted)*
- **Concurrent Imputation and Prediction in time-series EHR:** Developed semi-supervised generative adversarial model with recurrent layers to impute missing entries in time-series EHR data and predict future values. *(accepted)*
- **Transformer architecture to predict signs of cardiovascular disease:** Implemented a transformer model to learn EHR data representations and predict the values of major modifiable risk factors of cardiovascular disease *(accepted)*
- **Obesity Prediction with EHR Data:** Predict future BMI for children using pediatric EHR data using hybrid LSTM attention-based models. *(accepted)*

Sep' 21–May'22 **Data Science Fellow**, *Delaware Data Innovation Lab, USA*.

- Analyzed DHIN health claims by aggregating individual health claims at census tract level to study the prevalence of 27 chronic disease for Delaware. Published R-shiny report to communicate findings two technical and non-technical audience.
- Collected census level data for Delaware from ACS 5-year estimates and other environmental and social variables with individual-level health claims data to study social determinants of health for 27 chronic diseases. Published API for access by community.
- Developed R shiny webpage to determine eligibility for housing assistance in Delaware.

Publications

Conferences .

ML4H '2022 **Mehak Gupta**, *Brennan Gallamoza, Nicolas Cutrona, Pranjal Dhakal, Raphael Poulain, and Rahmatollah Beheshti.*, **"An Extensive Data Processing Pipeline for MIMIC-IV"**, In Proceedings of the 2nd Machine Learning for Health symposium, volume 193 of Proceedings of Machine Learning Research, pages 311–325. PMLR. Acceptance Rate: 28/85. View.

MLHC '2022 Raphael Poulain**, Mehak Gupta**, *Rahmatollah Beheshti.*, **"Few-Shot Learning with Semi-Supervised Transformers for Electronic Health Records"**, Machine Learning for Healthcare Conference, pages 1–21, 2022. Acceptance Rate: 33/104. View.

IAAI '2021 **Mehak Gupta**, *Raphael Poulain, Thao-Ly T Phan, H Timothy Bunnell and Rahmatollah Beheshti.*, **"Flexible-window Predictions on Electronic Health Records"**, In Proceedings of the AAAI Conference on Artificial Intelligence, 36(11):12510-12516. Acceptance Rate: 22%. View.

IEEE-BIBM '2021 Raphael Poulain**, Mehak Gupta**, *Thao-Ly T Phan, H Timothy Bunnell and Rahmatollah Beheshti.*, **"Transformer-based Multi-target Regression on Electronic Health Records for Primordial Prevention of Cardiovascular Disease"**, In Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Acceptance Rate: 19%. View.

ACM-BCB '2021 **Mehak Gupta**, *Thao-Ly T Phan, H Timothy Bunnell and Rahmatollah Beheshti.*, **"Concurrent Imputation and Prediction on EHR data using Bi-Directional GANs"**, In Proceedings of the 12th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics. Acceptance Rate: 23%. View.

Elsevier '2014 **Mehak Gupta**, *R. Aggarwal*, **"Transforming Relational Database to Graph Database Using Neo4j"**, In Proceedings of the Second International Conference on Emerging Research in Computing, Information, Communication and Applications, Bangalore, India.

Journals .

ACM-Health '2021 **Mehak Gupta**, *Thao-Ly T Phan, H Timothy Bunnell and Rahmatollah Beheshti.*, **"Obesity prediction with ehr data: A deep learning approach with interpretable elements"**, ACM Transactions on Computing for Healthcare (HEALTH), 3(3):1–19, 2022. Impact factor: 2.38. View.

Teaching Experience

2022 **Substitute Instructor**, *Department of Computer and Information Sciences, University of Delaware*, Facilitated classes in absence of the instructor for .

- "INTRODUCTION TO ARTIFICIAL INTELLIGENCE" (Course Code: CISC681)

2019 **Substitute Instructor**, *Department of Computer and Information Sciences, University of Delaware*, Facilitated classes in absence of the instructor for .

- "INTRO TO SOFTWARE ENGINEERING" (Course Code: CISC275)

2018–2019 **Graduate Teaching Assistant**, *Department of Computer and Information Sciences, University of Delaware*, Led weekly labs, guided and provided assistance to students during laboratory/office hours. Explained concepts, graded assignments and final exams for .

- "INTRO TO SOFTWARE ENGINEERING" (Course Code: CISC275)
- "MACH ORGANZTN & ASSEMBLY LANG" (Course Code: CISC260)

2013 **Lab Instructor**, *Department of Computer and Information Sciences, Thapar University*, Led weekly labs, guided and provided assistance to students during laboratory hours. Created lab assignments and exams and conducted final vocabulary tests for.
- "Introduction to Relational Databases"

Honors and Awards

2022 **Frank A. Pehrson Graduate Student Award for Outstanding Computer Science Research**, *Awarded by Graduate College, University of Delaware.*

2022 **Professional Development Travel Award**, *Awarded by Graduate College, University of Delaware to travel to a conference for poster presentation.*

2022 **Travel Award**, *Awarded by Machine learning for Health Conference..*

2022 **Dissertation Fellowship**, *Awarded by Graduate College, University of Delaware in recognition of the top doctoral students across campus based on nominations from graduate programs and recommendations from a faculty review committee. .*

2021, 2022 **CRA-W Travel Scholarship**, *Awarded by Computing Research Association Women to attend Grad Cohort Workshop, 2021, 2022.*

2020 **Distinguished Graduate student Award** , *Awarded by CIS, University of Delaware CIS 2020 in recognition of graduate students who distinguished themselves through exceptional scholarship, research and noteworthy performance in leadership and service..*

Research Presentations

2022 **An Extensive Data Processing Pipeline for MIMIC-IV**, *Machine learning for Health Conference Poster Presentation, In-Person.*

2022 **Deep Learning Predictive Modelling for Electronic Health Records**, *University of Delaware, PhD Proposal Presentation Talk, via Zoom.*

2021 **Flexible-window Predictions on Electronic Health Records**, *AAAI Conference on Artificial Intelligence, Paper Presentation Talk, via Zoom.*

2021 **Concurrent Imputation and Prediction on EHR data using Bi-Directional GANs**, *12th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, Paper Presentation Talk, via Zoom.*

2021 **Deep Learning for Obesity Prediction using Electronic Health Records**, *CRA-W Grad Cohort Workshop, Poster Presentation Talk, via Zoom.*

2020 **Obesity Prediction with EHR Data: A Deep Learning Approach with Interpretable Elements**, *University of Delaware, PhD Prelims Presentation Talk, via Zoom.*

2019 **A Survey of Deep Learning Techniques in EHR Prediction Modelling**, *Series of Special Interest Group in Artificial Intelligence, University of Delaware, Presentation Talk, Newark, USA.*

Undergraduate/Graduate Student Mentoring

Under the supervision and collaboration with my advisor Prof. Rahmat Beheshti, served as graduate research mentor for three undergraduate and three graduate students in performing EHR data analysis and developing models for disease outcome prediction. Collaborated in designing research projects, worked closely with students to guide them through the research process, designing and implementing an approach, evaluating the approach, and presenting their research in paper.

2022 Haritima Manchanda, *Developed in-house pipeline to transform OMOP-EHR data to FHIR and develop proof-of-concept SMART-FHIR app to showcase output of EHR data prediction models.*

2022 Brennan Gallamoza, Pranjal Dhakal, Nicolas, Cutrona, *Co-authored a research paper "An Extensive Data Preprocessing Pipeline for MIMIC-IV" (accepted).*

2021 Raphael Poulain, *Provided guidance on pre-processing techniques and evaluating an approach and presenting in a research paper (accepted)*.

2021 Raphael Poulain, *Introduced to the EHR data and provided guidance on implementing and evaluating an approach and presenting in a research paper. (accepted)*.

Service to Profession

2022 **Reviewer**, *ACM-CHIL 2023 (4 papers)*.

2022 **Reviewer**, *NeurIPS Workshop TS4H 2022 (3 papers)*.

2022 **Reviewer**, *Machine Learning 4 Health 2022 (3 papers)*.

2022 **Senator**, *Graduate Student Government, University of Delaware.* , Served Student Life committee to promote better graduate life for students who are parents..

2022 **Research Grant Contributor**, *Participated in discussion and writing of "Responsible AI to combat childhood obesity", with PI Rahmat Beheshti, "NIH (R01)"*.

2022 **Sub-Reviewer**, *ACM-Conference on Bioinformatics, Computational Biology, and Health Informatics (1 paper)*.

2022 **Reviewer**, *ACM-Conference on Health, Inference and Learning (3 papers)*.

2021 **Lead Mentor**, *EmPOWER University of Delaware*, Provide mentorship to graduate students joining Computer and Information Sciences. Planned events to build community and share resources to ease transition to graduate school.

2020–2023 **Webmaster**, *Women in Engineering, University of Delaware*, Provide supportive community for all women in STEM and planned events and maintained website for the committee.

2020 **Reviewer**, *Journal of Childhood Obesity (1 paper)*.

Industry Experience (4+ years)

2014–2017 **Associate Consultant**, *Software AG, Bangalore*.

Johnson & Johnson (J&J) Consumer R&D, *Developed Drag and Drop plugin using Angular JS for J&J web application*.

Standard Chartered Bank (SCB), *Key developer for transaction banking services*.

Solution Book (SB), *Designed Java/slang interfaces for interaction between various applications. Designed database architecture and JSON files and developed REST services*.

Kier U.K., *Developed SOAP web services and canonical documents for data exchange between systems*.

J&J CAPRI, *Developed the read through and write behind logic for asynchronous data synch using Terracotta caching*.

2010–2012 **System Engineer**, *Infosys Ltd, Pune*.

Aetna Health Insurance, *Worked on client side in Rating Support Application as a developer. Migrated Aetna application from tiered architecture to Service Oriented Architecture. Performed XML migrations for features using XSL. Performed top-k query optimization rating application.* .

Professional Memberships and Affiliations

2022 **Senator**, *Graduate Student Government, University of Delaware*.

2021–2022 **Data Science Institute Fellow**, *University of Delaware*.

2021 **CIS Representative**, *Student Advisory Committee, College of Engineering, University of Delaware*.

2020–2023 **CIS Representative**, *Women in Engineering, University of Delaware*.

2020–2022 **Peer/Lead Mentor**, *EmPOWER CIS, University of Delaware*.

Technical Skills

Programming	Python | R | Keras | PyTorch | tensorflow | Java | R shiny | SQL
Machine Learning	Neural Networks | Recurrent Neural Network | Attention Models | Convolutional Neural Network | Regression | Random Forest | k-means | Gradient Boosting | Generative Adversarial Network
Certifications	Oracle Certified Professional JAVA SE 6 Programmer | Coursera Introduction to Data Science with Python | Coursera Applied Machine Learning with Python | Coursera Text Mining with Python

References

Prof. Rahmatollah Beheshti, *Assistant Professor, University of Delaware*, PhD Advisor.
rbi@udel.edu 302-831-0072

Dr. H. Timothy Bunnell, *Principal Research Scientist and Director, Center for Pediatric Auditory and Speech Sciences, Nemours Children's Health*, PhD Committee Member and Advisor on Clinical Data Modelling.
Tim.Bunnell@nemours.org

Prof. Austin J. Brockmeier, *Assistant Professor, University of Delaware*, PhD Committee.
ajbrock@udel.edu

Héc Maldonado-Reis, *Director, Tech Impact*, Mentor on Data Science for Social Good Fellowship.
Hector@techimpact.org 302-437-5404

Dr. Thao-Ly T. Phan, *Medical Director, Pediatrics, Nemours Children's Health*, Clinical Advisor on AI in health projects.
ThaoLy.Phan@nemours.org