

TANZIRA NAJNIN

Phone: +1 (631) 532-9285 | Email: tanzira.najnin@utsa.edu | [Google Scholar](#) | [LinkedIn](#)

RESEARCH INTERESTS

Data Science, Machine Learning, Artificial Intelligence, Computational Biology

EDUCATION

<u>Institution</u>	<u>Degree</u>	<u>Field</u>	<u>Graduation</u>
The University of Texas at San Antonio, TX	Ph.D. (CGPA: 4.00)	Computer Science	Aug 2018 – July 2024
The University of Texas at San Antonio, TX	M.Sc. (CGPA: 4.00)	Computer Science	Aug 2018 – July 2023
Shahjalal University of Science & Technology, Sylhet	B.Sc. (CGPA: 3.45)	Computer Science & Eng.	Feb 2012 – Dec 2016

PROFESSIONAL EXPERIENCES

<u>Institution</u>	<u>Rank</u>	<u>Dates</u>
The University of Texas at San Antonio, TX	Graduate Teaching Assistant	Aug 2018 – May 2024
Metropolitan University, Sylhet, Bangladesh	Lecturer	Sep 2017 – July 2018
Metropolitan University, Sylhet, Bangladesh	Teaching Assistant	Jan 2017 – Sep 2017

PUBLICATIONS

[Tanzira Najnin](#), Sakhawat Hossain Saimon, Maryam Zand, Nahim Adnan, Zhijie Liu, Tim Hui-Ming Huang, and Jianhua Ruan (2024). A novel approach for predicting breast cancer metastasis through measurement of global transcriptional dysregulation. In review.

Sakhawat Hossain Saimon, [Tanzira Najnin](#), and Jianhua Ruan. A Novel Feature Space Augmentation Method to Improve Classification Performance and Evaluation Reliability. In review.

Li-Ling Lin, [Tanzira Najnin](#), Chun-Wei Chen, Mia S. Ramos, Maricar Galang, Jeffery Chavez, Bijaya Nayak, Chiou-Miin Wang, Sakhawat H. Saimon, Randall D. Robinson, Jianhua Ruan, Bruce J Nicholson, Nameer B Kirma (2024). Single-cell biomarker gap junction gene panel for the non-surgical diagnosis of endometriosis: a proof-of-concept study. In review.

[Tanzira Najnin](#), Sakhawat Hossain Saimon, Garry Sunter, and Jianhua Ruan. "[A Network-Based Approach for Improving Annotation of Transcription Factor Functions and Binding Sites in Arabidopsis thaliana](#)." Genes 14, no. 2 (2023): 282.

Nahim Adnan, [Tanzira Najnin](#), and Jianhua Ruan. "[A Robust Personalized Classification Method for Breast Cancer Metastasis Prediction](#)." Cancers 14, no. 21 (2022): 5327.

MENTORSHIP EXPERIENCES

<u>Institution</u>	<u>Role</u>	<u>People</u>	<u>Dates</u>
Metropolitan University, Bangladesh	Programming Trainer	CS trainees (competitive programming): <ul style="list-style-type: none">o 2 teams of 6 undergraduate students	Jan 2017 – Jul 2018
Metropolitan University, Bangladesh	Research Mentor	CS mentees (research): <ul style="list-style-type: none">o 3 CS undergraduate students	Jan 2017 – Jul 2018

TEACHING EXPERIENCES

<u>Institution</u>	<u>Rank</u>	<u>Courses Taught/Recited</u>	<u>Dates</u>
The University of Texas at San Antonio	Teaching Assistant	<ul style="list-style-type: none">o Data Scienceo Bioinformaticso Data Structureso Intro to Java Programming	Aug 2018 – Present
Metropolitan University, Bangladesh	Lecturer, Teaching Assistant	<ul style="list-style-type: none">o Data Structureso Algorithmso Compiler Constructiono Data Structures	Jan 2017 – Jul 2018

AWARDS, SCHOLARSHIPS & LEADERSHIP

<u>Institution</u>	<u>Award/Leadership</u>	<u>Dates</u>
CS Graduate Student Association, UTSA	Vice-president, Project Chair, Secretary	Fall 2019 – Fall 2023
The University of Texas at San Antonio	COS Alvarez Scholarship	Fall 2021 – Spring 2024
Computing Research Association	Grad Cohort Scholarship for Women	Apr 2022
AnitaB.org	Grace Hopper Scholarship	Sep 2021
Bangladesh Govt.	Secondary School Certificate Scholarship	Jan 2011

PRESENTATIONS

Tanzira Najnin, Sakhawat Hossain Saimon, Garry Sunter, and Jianhua Ruan (2022). A Network-Based Approach for Improving Annotation of Transcription Factor Functions and Binding Sites in *Arabidopsis thaliana*. UTSA CS poster presentation.

Tanzira Najnin, Sakhawat Hossain Saimon, Maryam Zand, Nahim Adnan, Zhijie Liu, Tim Hui-Ming Huang, and Jianhua Ruan (2022). A novel approach for predicting breast cancer metastasis through measurement of global transcriptional dysregulation. CRA-WP conference poster presentation, New Orleans.

Tanzira Najnin, Jianhua Ruan (2024). A Mixture-of-Experts Approach for Pan-Cancer Prognosis Prediction. The Southwest Data Science Conference 2024 at Baylor University.

Tanzira Najnin (2022). Methods for gene regulatory network-based knowledge discovery and cancer outcome prediction. PhD proposal presentation, UTSA.

RESEARCH PROJECTS

Breast Cancer Metastasis Prediction: This study addresses the dual challenges of understanding metastasis in breast cancer and predicting metastatic risk at earlier stages. While previous efforts have made strides in knowledge discovery and prediction accuracy separately, a comprehensive approach remains elusive. We present a novel computational method that simultaneously supports knowledge discovery and metastasis prediction by comparing gene regulatory networks (GRNs) to identify metastasis-associated changes and introducing a dysregulation score for direct prediction. Experimental results demonstrate significant GRN changes and systems-level insights, surpassing complex machine learning models in predictive accuracy. This work bridges the gap between knowledge discovery and accurate, explainable phenotype prediction in breast cancer metastasis.

A Personalized Breast Cancer Classification Method: Predicting breast cancer metastasis early is crucial to reduce deaths. Existing machine-learning models using gene expression often struggle with diverse molecular characteristics and different cancer subtypes. To address this, we developed a method for personalized classifiers trained on patient subsets with similar characteristics. Results across multiple datasets demonstrated significantly improved prediction accuracy compared to models trained on the entire dataset or specific subtypes. Personalized classifiers based on both positively and negatively correlated patients outperformed those based only on positively correlated patients, emphasizing the importance of proper patient subset selection. Moreover, the proposed approach yielded more robust features and identified patient-specific features, making it a promising tool for personalized cancer medicine.

Enhancing Transcription Factor Functions and Annotation in *Arabidopsis thaliana*: Transcription factors play a vital role in cellular regulation, interpreting DNA sequences and signals to control gene expression. Uncovering their functional associations traditionally relies on resource-intensive experiments or yields low-quality results through computational analysis. This paper introduces a data-driven, statistics-based strategy for predicting transcription factor functions in the model plant *Arabidopsis thaliana*. Leveraging a vast gene expression compendium, we construct a genome-wide transcriptional regulatory network, identifying regulatory relationships and downstream targets for each transcription factor. Functionally enriched gene ontology terms are then determined, annotating transcription factors with specific biological processes. DNA binding motif discovery further supports our predictions, aligning well with experimental databases. Statistical analysis of the network reveals intriguing patterns in system-level transcriptional regulation. This methodology holds promise for improving transcription factor annotation and understanding transcriptional regulation in various species.

Novel Data Augmentation Technique: Class imbalance, small sample sizes in high-dimensional spaces, and measurement uncertainty complicate many real-world classification tasks. While prior work has focused on imbalance, the application of data augmentation, specifically for improving overall performance and stability with limited samples, remains uncertain. This paper introduces a novel feature-space augmentation technique designed for high-dimensional data in classification tasks. The method utilizes uniform random sampling and leverages local feature distributions to introduce synthetic instances. The core augmentation algorithm is class-invariant, offering a unique opportunity to simultaneously enhance and stabilize performance by augmenting unlabeled instances. Comprehensive evaluations involving multiple classifiers and metrics demonstrate that the proposed algorithm significantly improves performance and enhances overall reliability compared to existing feature space augmentation methods.

Mixture-of-Experts (MoE) Approach for Pan-cancer Prediction (Ongoing): Cancer manifests in various types with diverse subtypes, some of which exhibit higher aggressiveness. Estimation of the disease trajectory is crucial for advancing precision medicine and elevating overall survival rates. Despite the significance of multi-type cancer studies, there is currently a lack of comprehensive research in this field. This study aims to address this gap by utilizing data from 27 cancer types to predict progression-free interval status. This study introduces an innovative approach utilizing a Mixture of Experts (MoE) method to predict outcome. The core concept involves selecting predominantly similar patients to train personalized machine-learning models for test patients. A thorough evaluation through k-fold and leave-one-out cancer cross-validation underscores the significant performance improvement of our method over conventional machine learning algorithms. This advancement holds promise for enhancing the precision of cancer prognosis and contributing to improved patient outcomes.