

SHANTANU GHOSH

LinkedIn : <https://www.linkedin.com/in/i-am-shantanu-ghosh/>
Github : <https://github.com/Shantanu48114860>

beingshantanu2406@gmail.com
+1 (352)871-3965
3800 SW 34th ST, Apt W220,
Gainesville, FL 32608, USA

OBJECTIVE *To secure a PhD position in the area of Machine Learning/Deep Learning with a focus of Causal Inference and Deep Learning*

RESEARCH INTERESTS *Deep Learning, Causal Inference, Computer Vision*

EDUCATION **University of Florida**, Gainesville, FL, USA
Master of Science, Computer and Information Sciences
Advisor: Dr. Mattia Proserpi
Aug, 2019 - May, 2021 **GPA: 3.88/4**

West Bengal University of Technology, West Bengal, India
Bachelor of Technology from Institute of Engineering and Management
Computer Science and Engineering
Aug, 2008 - June, 2012 **GPA: 8.38/10**

PUBLICATION

- *Deep Propensity Network using a Sparse Autoencoder for Estimation of Treatment Effects* - **Shantanu Ghosh**, Jiang Bian, Yi Guo, Mattia Proserpi. Journal of the American Medical Informatics Association (JAMIA) (**Under review**)
- *Causal AI with Real World Data: Do Statins Protect From Alzheimer's Disease Onset?* - Mattia Proserpi, **Shantanu Ghosh**, Zhaoyi Chen, Marco Salemi, Tianchen Lyu, Jiang Bian. American Medical Informatics Association (AMIA), 2021 (**Under review**)
- *Propensity Score Synthetic Augmentation Matching using Generative Adversarial Networks (PSSAM-GAN)* - **Shantanu Ghosh**, Jiang Bian, Mattia Proserpi. Thirty-Fifth AAAI Conference on Artificial Intelligence (**Under review**)

RESEARCH EXPERIENCE **Data Intelligence Systems Lab (DISL)**, Gainesville, FL, USA
Graduate Student Assistant **March 2020 - Present**
Advisor: Dr. Mattia Proserpi, Dr. Jiang Bian, Dr. Yi Guo

- Currently working on to improve PSSAM-GAN framework by incorporating Adversarial Autoencoder and InfoGAN.

- Using Pytorch, reproduced the model DCN-PD proposed in the paper **[Paper]**(*ICML 2017 - Workshop on Principled Approaches to Deep Learning*), that has been utilized in the research study whether usage of Statins is useful toward the onset of Alzheimer disease. Paper is under review at **AMIA, 2021**. **[Code]**

- Developed Deep Propensity Network - Sparse Autoencoder(**DPN-SA**) - a sparse auto encoder based neural network model to calculate propensity score, outperformed logistic regression and LASSO by **36-63%**, and DCN-PD (baseline models) by **6-10%** across all datasets. Paper is under review at **JAMIA**. **[Code]**

- Developed Propensity Score Synthetic Augmentation Matching using Generative Adversarial Networks (**PSSAM-GAN**) - an algorithm to generate synthetic treated samples to remove imbalance within a observational dataset for Propensity score matching (PSM). The model improved the performance by **38%** and **5%** over DCN-PD and TAR-NET(baseline models) respectively. Paper is under review at **AAAI, 2021**. **[Code]**

Multimedia Communications and Networking Laboratory (MCN), Gainesville, FL, USA
Independent Researcher **Feb 2020 - May 2020**

Advisor: Dr. Dapeng (Oliver) Wu

Developed a Deep Convolutional Multitask Neural Network(**MTL-TCNN**) to classify textures under the supervision of Dr Prof Dapeng Oliver Wu of the Department of Electrical & Computer Engineering in the University of Florida. **[Report]** **[Code]**

PROFESSIONAL EXPERIENCE

Lexmark International India Pvt Ltd, Kolkata, West Bengal, India

Software Engineering Professional II

Oct 2016 - July 2019

Developed the ISP component of the product Publishing Platform for Retail(PPR) with active participation in 2 major releases. Also, worked on Lexmark Digital Media Platform, a multi-tenant enterprise video content management platform hosted in Amazon Web Services.

Cognizant Technology Solutions India Pvt Ltd, Kolkata, West Bengal, India

Associate, Projects

March 2013 - September 2016

Developed WCF web services in the Contract First Approach to provide secure communication between different In-house application using Service Oriented Architecture (SOA), C# .Net 4.5, Oracle Client 11g. Trained C# to new recruits in Cognizant Academy.

COURSE PROJECTS

Classification of Handwritten Characters

Oct 2019 - Dec 2019

Fundamentals of Machine Learning, University of Florida, FL, USA

Inspired by the famous architecture "Lenet-5", developed a deep CNN model to classify Handwritten Characters using a custom Handwritten Character Dataset prepared by Prof Alina Zare by utilizing the Adam Optimizer, Batch Normalization and dropout and achieved a classification accuracy of **97.3%** on a customised data set prepared by Prof Zare **[Code]**

- **Technology/Tools:** Python, Pytorch, Skit-learn

Implementation of P2P network

Nov 2019 - Dec 2019

Computer Networks, University of Florida, FL, USA

Created a peer-to-peer(P2P) network for file downloading. Developed components – peer and file owner. The file owner has a file, and breaks the file into chunks of 100KB. Each peer connects to the file owner to download some chunks with the help of two threads, one acting as a server that uploads the local chunks to another peer (referred to as upload neighbor), and the other acting as a client that downloads chunks from a third peer (referred to as download neighbor). Tested the code with max **5** peers and max file size of **13.3 MB**. **[Code]**

- **Technology/Tools:** Socket Programming, Java

Dataset augmentation using InfoGAN and ConditionalGAN

Oct 2020 - present

Deep Learning for Computer Graphics, University of Florida, FL, USA

Currently implementing MNIST and CIFAR10 dataset augmentation using InfoGAN and ConditionalGAN to improve the classification accuracy of an image classifier.

- **Technology/Tools:** Python, Pytorch, Skit-Learn

Comparative study of the performance of VAE and GAN

Oct 2020 - present

Machine Learning, University of Florida, FL, USA

Currently working to prepare a comparative study of the performance of VAE and GAN on MNIST and CIFAR10 dataset.

- **Technology/Tools:** Python, Pytorch, Skit-Learn

TECHNICAL SKILLS

Languages : Python, C++, C, Java, C#, Javascript/Typescript

Database : MySQL, Oracle 9i/10g, MS SQL Server, DB2

Web Development : Angular, Node.js, WCF

Machine Learning : TensorFlow, PyTorch, Scikit-learn

GRADUATE COURSES

• Fundamentals of Machine Learning • Distributed Operating Systems • Computer Networks • Mathematics for Intelligent Systems • Advanced Data Structures
• Machine Learning • Deep Learning Computer Graphics • Fundamentals of Probability • Analysis of Algorithm (Spring 2021)

ACHIEVEMENTS .

- Recipient of **National Scholarship** Award from **Central Government Human Resource Development Department of Higher Education, India** for excellent result in Higher Secondary Examination in the state of West Bengal, India.
- Topped with **1%** of all candidates appeared in **West Bengal Joint Entrance Examination** in 2008.
- Received **Star Employee** award in Q4, 2013 and Q4, 2015 in Cognizant Technology Solutions.
- Received **Achievement Award** of 4500 USD during the admission of graduate studies in the University of Florida in Fall 2019.
- Received **Research Assistantship** under the supervision of Prof Kevin Butler in FICS Lab, starting from Jan 2021 in the domain of explainable AI.