

# Ashutosh Tiwari

Computational Data Science Grad Student with *six* years of work experience, can join immediately

812-606-5974 | [ashutiwa@iu.edu](mailto:ashutiwa@iu.edu) | [Linkedin @ashutosh-tiwari](#) | [Homepage](#) | [Github @thunderrock](#)

## EDUCATION

---

### Indiana University, Bloomington

Master of Science in Data Science (Computational and Analytical Track) 3.87/4.0

Bloomington, IN

Aug. 2021 – May. 2023

### National Institute of Technology

Bachelor of Science in Computer Science 8.32/10.0

Patna, India

Aug. 2011 – Jun. 2015

## WORK EXPERIENCE

---

### Senior Software Dev Engineer (ML Platform)

Jan. 2019 – Jul. 2021

[SWIGGY](#)

DATA SCIENCE PLATFORM, TIME SERIES FORECASTING

- Was part of team that worked on Feature Store and pipeline which feeds on-demand features to deployed ML models at production scale(4Bn rows, 1Mn QPS / day). Pipeline supported multichannel ingestion, i.e. Spark, Flink and user files etc.
- Founding member of Forecasting and Correlation Platform which was considered by many teams to forecast concerned time series. These forecasts power critical scaling decisions across organizations in real time.
- Led DAQ, a tool used to scrape APIs at scale. Used to collect data for analysis/ model training at a scale of 15 M rows daily.

### Software Development Engineer (NLP)

Sep. 2017 – Jan. 2019

[FLIPKART](#) (a Walmart company)

SEARCH RELEVANCE, QUERY INTENT, NLP

- Was responsible for improvements/inception of search intent models(CRF/Neural Network based), identifying error classes, coming up with solutions, and fixing them. These models power user search and discovery for millions every day.
- Implemented a FastText based query store classifier, which predicts the category of a tail query.
- Implemented the first workflow to automate training and auto-deployment of various search models in Flipkart. First was written using Luigi and later migrated to Airflow.
- Wrote a generic framework using Airflow which at runtime creates generic dags for different ML models and orchestrates their training to deployment flow, including data and model validations.
- Implemented large scale (4Bn+ datapoints) pipelines using Cascading/HDFS to extract data from user events and then transform it to be used for training these models.

### Software Development Engineer

Sep. 2016 – Sep. 2017

[GROUPON](#)

BACKEND ENGINEERING

- Worked on a component called Cyclops, an interface between Customer representatives and internal services.

### Software Engineer

Sep. 2015 – Aug. 2016

[NETSPEED SYSTEMS](#) (Acquired by Intel)

GRAPH ALGORITHMS, NETWORK ON CHIP

- Led engineering efforts on modules like Polarity based Arbitration, Multi-Cast Filtering, Structural Latency Breakdown etc.

## PUBLICATIONS

---

Accepted at [NetSci 2023](#) (Poster Presentation) and [IC2S2 2023](#) (Parallel Talk) as **first author**

- [1] **Ashutosh Tiwari**, Prof. Sadamori Kojaku, Prof. Yong-Yeol Ahn, “Biased Contrastive Learning debiases Graph Neural Networks,” *International Conference on Network Science (NetSci)*, 2023.

In this work, we propose a non-parametric contrastive learning framework to learn debiased graph embeddings with respect to sensitive node attributes and structural homophily. Through empirical evaluations on different datasets, we demonstrate that our method offers a better approach to debiasing compared to existing approaches and thus results in more organic recommendations across different GNN architectures.

## RESEARCH EXPERIENCE / INDEPENDENT STUDY

---

- Implementing novel training frameworks that result in unbiased (or fair) recommendation systems from “biased” datasets at [IUNI](#) with Prof. YY Ahn and Prof. S Kojaku as part of my Independent Study. Summer 2022 - Present
- Working as a paid RA for, “**User Intent as a Network**” with Prof. YY Ahn, P Kantak and FB Yara. Project is funded by Kelly Business School. Goal is to be able to quantify and thus act upon that “intent” network. Summer 2022 - Fall 2022
- Was part of [NLP Lab@IUB](#) with Prof. D Cavar. Contributed to design of [TieML](#) and Events’ Timeline modelling. Fall 2021

## TEACHING EXPERIENCE

---

Teaching Assistant for <u>Network Science</u> (INFO-I 606) with Prof. YY Ahn	Spring 2023
Teaching Assistant for <u>Machine Learning</u> (CSCI-B 555) with Prof. R Khardon	Fall 2022
Teaching Assistant for <u>Network Science</u> (INFO-I 606) with Prof. YY Ahn	Spring 2022

## SELECTED PROJECTS

---

<b>BiasNet</b>   <i>Deep Reinforcement Learning, PyTorch, Actor Critic Algorithm, On-policy Model Free</i>	<a href="#">Code/Report</a>
• Learning to fight in Street Fighter II with induced relational bias from differential game scenes.	
<b>DeepFoodie</b>   <i>Python, Tensorflow, Self Supervised Deep Clustering, Deep Learning, Transfer Learning</i>	<a href="#">Code/Report</a>
• Clustering dishes on basis of their ingredient embeddings. These ingredient embeddings are generated by a NN.	
<b>BlindNet</b>   <i>Python, Pytorch, Deep Learning, Transfer Learning</i>	<a href="#">Code/Report</a>
• Image to vector generation on Coco dataset.	
<b>Continuous Dominant Set Repair</b>   <i>C++, Graph Algorithms, Guha and Khuller's Algo., Greedy</i>	<a href="#">Code</a>
• Repairs a broken link in Continuous Dominant Set in $O(\Delta^2)$ , where $\Delta$ being the avg cardinality of connected graph.	
<b>Humana Mays Healthcare Analytics Case Competition</b>   <i>Boosted Trees, Feature Engineering</i>	<a href="#">Code</a>
• 11th rank on leaderboard. Hosted by TAMU and Humana Mays, 2021.	
<b>AnalyticsVidhya Solutions</b>   <i>Python, Tensorflow, Torch, Time Series, Feature Engineering, Catboost</i>	<a href="#">Code</a>
<b>Investigating Bias Manifolds</b>   <i>Bias Manifolds, Bias Progression, Measuring Bias, Python, Word2vec</i>	<a href="#">Code/Report</a>

## TECHNICAL SKILLS

---

Search Relevance, Recommendations, Graph Neural Networks, Fairness Aware Modeling, Computational ML, NLP, Computer Vision, DL, Spark, Python, Scala, C++, Contrastive Learning  
**Frameworks/Libraries/Tools:** Pytorch Geometric, Pytorch, Tensorflow, Sklearn, Numpy, Pandas, Matplotlib, HDFS, Spark, Kafka, Flink, AWS Sagemaker, DynamoDB, Faiss, Django

## RELEVANT COURSES

---

Computational ML(B555), Deep Learning(E533), Reinforcement Learning(B659), Computer Vision(B657), Statistics(S520), Independent Study, Working on Fairness Aware AI and Bias manifolds in Graph ML (B669), Advanced Database Systems(B561)