

# Ashutosh Tiwari

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](https://www.linkedin.com/in/ashutosh-tiwari) | [Homepage](https://ashutosh-tiwari.github.io) | [Github@thunderrock](https://github.com/ashutosh-tiwari)

## WORK EXPERIENCE

- Senior Software Dev Engineer** Jan. 2019 – Jul. 2021  
*SWIGGY* *MACHINE LEARNING 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(ML)** 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.

## 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

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

## TECHNICAL SKILLS

Graph Neural Networks, Fairness Aware Modeling, Computational ML, NLP, Computer Vision, DL, Spark, Python, Scala, C++  
**frameworks & tools:** Pytorch Geometric, Pytorch, Tensorflow, Sklearn, Numpy, Pandas, Matplotlib, HDFS, Spark, Kafka, Flink, AWS Sagemaker, DynamoDB, FAISS, Django

## EDUCATION

<b>Indiana University, Bloomington (OPT start date July 3, 2023, stem eligible)</b>	Bloomington, IN
<i>Master of Science in Data Science (Computational and Analytical Track) 3.87/4.0</i>	Aug. 2021 – May. 2023
<b>National Institute of Technology</b>	Patna, India
<i>Bachelor of Science in Computer Science 8.32/10.0</i>	Aug. 2011 – Jun. 2015