# Ashutosh Tiwari

**Data Science Grad Student with *six* years of work experience, can join July 3**

812-606-5974 | findashutoshtiwari@gmail.com | Linkedin@ashutosh–tiwari | Homepage | Github@thunderock

## WORK EXPERIENCE

- **Senior Software Dev Engineer** <span style="float:right">Jan. 2019 – Jul. 2021</span>
  *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)** <span style="float:right">Sep. 2017 – Jan. 2019</span>
  *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** <span style="float:right">Sep. 2016 – Sep. 2017</span>
  *GROUPON* — *BACKEND ENGINEERING*
  - Worked on a component called Cyclops, an interface between Customer representatives and internal services

- **Software Engineer** <span style="float:right">Sep. 2015 – Aug. 2016</span>
  *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 Network Science (INFO-I 606) with Prof. YY Ahn | Spring 2023 |
| Teaching Assistant for Machine Learning (CSCI-B 555) with Prof. R Khardon | Fall 2022 |
| Teaching Assistant for Network Science (INFO-I 606) with Prof. YY Ahn | Spring 2022 |

## SELECTED PROJECTS

**Continuous Dominant Set Repair** | *C++, Graph Algorithms, Guha and Khuller's Algo., Greedy* <span style="float:right">Code</span>
- 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* <span style="float:right">Code/Report</span>

**BiasNet** | *Deep Reinforcement Learning, PyTorch, Actor Critic Algorithm, On-policy Model Free* <span style="float:right">Code/Report</span>
- 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* <span style="float:right">Code/Report</span>
- 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* <span style="float:right">Code</span>
- 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

## EDUCATION

**Indiana University, Bloomington (OPT starts July 3, 2023, stem eligible)** <span style="float:right">Bloomington, IN</span>
*Master of Science in Data Science (Computational and Analytical Track) 3.87/4.0* <span style="float:right">Aug. 2021 – May. 2023</span>
**National Institute of Technology** <span style="float:right">Patna, India</span>
*Bachelor of Science in Computer Science 8.32/10.0* <span style="float:right">Aug. 2011 – Jun. 2015</span>