Shrey Gupta

Website: shrey-gupta.github.io
Github: github.com/shrey-gupta

Email: shrey.gupta@emory.edu
Mobile: +1-317-560-6046
Atlanta, US

Education

- **Emory University** | Doctor of Philosophy - Computer Science 2019 - present
  Advisor: Dr. Avani Wildani | **Thesis:** Supervised Learning by applying domain adaptation for knowledge transfer.
- **IIIT-Delhi** | Masters of Technology - Computational Biology 2015 - 2017
  Advisor: Dr. Vinod Scaria | **Thesis:** A Study of Biomedical Information Extraction Systems.
- **Guru Gobind Singh Indraprastha University** | Bachelors of Technology - Computer Science 2011 - 2015
  Undergraduate Thesis: Abstractive Text Summarization using Bag-of-words Encoder.

Skills Summary

- **Languages**: Python, R, MATLAB, C/C++, JavaScript, SQL/Pl-SQL, Java, HTML, CSS, Bash, LaTeX
- **Frameworks**: Scikit, Scipy, Numpy, Pandas, TensorFlow, Keras, Matplotlib, Pip, Conda, Django, NodeJS
- **Tools/Platforms**: GIT, PyCharm, Jupyter, Eclipse, Visual Studio, R-Studio, Heroku, MySQL, Spark/Hadoop, Linux, AWS, Google Cloud, Google Collab.
- **ML Experience**: Deploying ML systems – model analysis, feature engineering, dimensionality reduction, parameter optimization, performance & error analysis.
  Analyzing distribution shifts – co-variate, concept, and using active learning for model fitting.
  Applying bayesian inference for probabilistic modelling and causal analysis.
- **Relevant Courses**: Graduate Algorithms, Advanced Machine Learning, Deep Learning, Data Mining, Statistical Computation, Reinforcement Learning, Big Data Analysis, Applied Mathematics I, II, & III.

Experience

- **Graduate Research Assistant** | Emory University Aug'19 - present
  - **Transfer Learning for $PM_{2.5}$ Prediction**:
    – Developed transfer learning model (ML for few-shot learning) for $PM_{2.5}$ data prediction using *tensorflow* framework.
    – Our algorithm outperformed competitve methodologies 75% of the times (paper submitted).
  - **Persistence in Bayesian Belief Update**:
    – Full stack web development of a tool to measure the persistence of user beliefs using Bayesian Probabilistic Modelling.
    – The tool was deployed on *Heroku* and *Prolific* using *nodeJS* and *mongoDB*.
- **Research Assistant** | Wayne State University Aug'18 - May'19
  - **Detecting Seizures in Neonates using EEG Data**:
    – Implemented Bayesian Switching Factor Analysis (BSFA) in *python* to map the highly-active brain regions in neonatal babies during seizures (Was awarded)
- **Teaching** Aug'15 - Dec'20
  - **Course Instructor (Emory U.)**: Introduction to Python Programming [Fall'20]
  - **Course Instructor (Wayne State U.)**: Fundamental Structures in Computer Science [Fall'18 & Spring'19]
  - **Teaching Assistant (IIIT-Delhi)**: Computer Systems Management, Computer Networks, Computer Architecture, Data Structures and Algorithms, Introduction to Quantitative Biology [Fall 2015 - 2016].
- **Research Intern** | IGIB Delhi, India May'15 - June'18
  - **A Study of Biomedical Information Extraction Systems**:
    – Full stack web development of a tool using *django* to detect & extract disease symptoms from clinical texts.
    – It overcame the limitations of many search based MD tools and was deployed at AIIMS, Delhi.
- **Research Intern** | IIT-Madras, India May'14 - Aug'14
  - **Keyword Search in Databases**:
    – Full stack development of a tool in *Django* which was an improvement upon BANKS, a keyword search tool for databases *sans* SQL queries.
    – It improved the algorithmic (time/space) efficiency of BANKS by 30%.
- **Software Development Intern** | Futor Labs, India May'13 - Aug'13
  - **Database Synchronization**:
    – Developed stored procedures (functional scripts) using *Pl-SQL* and *MySQL workbench* to achieve synchronization in organizational database.
    – Improved organizational database update efficiency by introducing a one-click solution to synchronization.

Additional Projects

- **Mitigating Write/Update Latency in SMR Drives** Aug'19 – Feb'20
  – Reviewed techniques for reducing latency in SMR drives during the write/update operations. Github-link

- **Online Tool for Reporting of Incidental Findings** Aug'16 – Feb'17
  – Full stack web development of a tool using *django* and *mongoDB* to check if the DNA sequencing data consists of matches to a provided variants lists.

- **Causal Inference & Modelling of Datasets** May'16 – Aug'16
  – Implemented Markov Blanket methodology in *python* to determine the causal relationships in the dataset. Github-link

Publications

1. **Gupta, S.**, Bi, J., Liu, Y., & Wildani, A. ISTRBoost: Importance Sampling Transfer Regression using Boosting. (Under Review) (2022) https://doi.org/10.48550/arXiv.2204.12044

2. Vatsyayan⋆, A., Sharma⋆, P., **Gupta, S.**, Sandhu, S., Venu, S. L., Sharma, V.,... Rajab, A. et al. (2021). DALIA - A Comprehensive Resource of Disease Alleles in Arab Population. PloS one, 16(1), e0244567. (2021) doi:10.1371/journal.pone.0244567

Honors and Awards

- **Travel Grant Recipient**, File and Storage Technologies (FAST 2020), organized by USENIX. [2020]
- **Complete Tuition Waiver for Master's Degree**, Department of Biotechnology, India. [2015]
- **Scholarship for Qualifying Graduate Aptitude Test in Engineering (GATE)**, Ministry of Human Resource Development, Government of India. [2015]
- **98.5$^{\text{the}}$ percentile holder in Graduate Aptitude Test in Engineering (GATE)**, Ministry of Human Resource Development, Government of India. [2015]
- **Awarded Certificate of Excellence**, UN Conference on Sustainable Development (Rio+20), India. [2013]

Volunteer Experience

- **Oral Presentation**, "Transfer Learning for Predicting $PM_{2.5}$", GRITS Seminar'22 at Emory U. [2022]
- **Symposium Organizer**, Symposium on Functional Genomics, CCB@IIIT-Delhi, India. [2017]
- **Symposium Organizer**, Symposium on Network Biology, CCB@IIIT-Delhi, India. [2017]