Nilesh Gupta

Research Fellow
Machine Learning and Optimization Group
Microsoft Research India

nileshgupta2797@gmail.com
nilesh2797.github.io
Google Scholar

Research Interests

Machine Learning (Efficient & Large-scale), *Web Search & Recommendation*, *Graph Neural Networks*

Education

Indian Institute of Technology Bombay *2015 - 2019*
B.Tech (Honours) in Computer Science and Engineering *GPA: 8.94/10*
Advisor: Prof. Shivaram Kalyanakrishnan

Work Experience

Microsoft Research India *2019 - Current*
Research Fellow in Machine Learning and Optimization Group Advisor: Dr. Manik Varma
Working on algorithms and applications of Extreme Classification

Tower Research Capital *May 2018 - June 2018*
Core Dev Enginnering Intern
Worked on building modular and efficient exchange recovery client, Awarded Pre-Placement Offer

Publications

- **ZestXML: Zero-Shot Extreme Multi-label Learning for Tagging, Recommendation and Advertising**
  Nilesh Gupta, Sakina Bohra, Yashoteja Prabhu, Saurabh Purohit and Manik Varma
  (under review at WWW The Web Conference), 2021.
- **Extreme Regression for Dynamic Search Advertising**.
  Yashoteja Prabhu, Aditya Kusupati, **Nilesh Gupta** and Manik Varma.
  *International Conference on Web Search and Data Mining (**WSDM**), 2020* **(Long Oral)**.
  *Workshop on eXtreme Classification: Theory and Applications @ **ICML**, 2020.*

Research Experience

Personalized Recommendation for MSN News and Feeds *Ongoing*
Advisor: Dr. Manik Varma, *Microsoft Research India*

- Working on improving personalization in MSN homepage recommendation through extreme classification methods
- Our work proposes to capture user's fine grained interests which most of the existing works fail to address properly

ZestXML: Zero-Shot Extreme Multi-label Learning for Tagging, Recommendation and Advertising
Advisor: Dr. Manik Varma, *Microsoft Research India* *July 2020 - Oct 2020*

- Developed novel ZestXML algorithm which extends extreme multi-label learning to zero-shot scenarios
- Proposed new generative model for retrieval which is significantly more efficient and accurate than traditional IR methods
- Deploying ZestXML in Sponsored Search Advertising on Bing improved Click Yield of IR-based system by 17%

XReg: Extreme Regression for Dynamic Search Advertising *July 2019 - Oct 2019*
Advisor: Dr. Manik Varma, *Microsoft Research India* [Paper]

- Worked on the XReg algorithm, a novel and highly scalable method which efficiently solves the eXtreme Regression task
- Our work proposes new labelwise inference in extreme classifiers which can provide elegant solutions to large scale recommendation problems like DSA
- Deployment of XReg on DSA in Bing resulted in a relative gain of 27% in query coverage

Scalable Parabel: Optimizing Parabel to scale on 50M labels *Aug 2019 - Dec 2019*
Advisor: Dr. Manik Varma, *Microsoft Research India*

- Implemented highly scalable and effecient version of Parabel extreme classifier which got used in multiple applications across Bing and MSN
- The efficient version is upto 10x more RAM efficient and it can train and predict on large scale datasets (∼50M labels) in few hours even on single core commodity desktop

Learning Complex Behaviours and Keepaway in Robocup 3D *July 2018 - May 2019*
Advisor: Prof. Shivaram Kalyanakrishnan, *Undergraduate Thesis, IIT Bombay* [Thesis]

- Developed NEAT based optimization framework for learning high level behaviours of agents in challenging Robocup 3D simulated soccer environment.

- Learnt behaviours consistently outperformed existing hand coded strategies on Keepaway sub task of simulated soccer

Verification of Concurrent Programs under Weak Memory Model *May 2017 - July 2017*
Advisor: Prof. Parosh Aziz Abdulla, *Internship, Uppsala Universitet*

- Implemented modular, efficient and scalable version of stateless DPOR model checking algorithm in C++
- Studied context bound analysis for concurrent programs in Release-Acquire weak memory model

Selected Awards and Honors

- Ranked **4th** in **ACM-ICPC** Asia Regionals and **6th** in **ACM-ICPC** India Online *2017*
- Awarded **AP** Grade for exceptional performance (awarded to less than top 1%), in **Digital Logic Design Lab** and **Computer Programming and Utilisation** *2017*
- **All India Rank 384** in JEE Advanced (IIT-JEE) 2015 among 150,000 candidates *2015*
- Secured **99.97** percentile in JEE Main Paper I amongst 1.3 million student *2015*
- Awarded the prestigious **KVPY** Fellowship from Government of India *2015*
- Ranked **2nd** in **Regional Mathematics Olympiad** (RMO) and among **top 300** students in **Indian National Mathematics Olympiad** (INMO) *2014*

Teaching & Responsibilities

- *Undergraduate Teaching Assistantship* - Computer Science and Engineering, IIT Bombay
  - Computer Programming and Utilisation - *Prof. Ganesh Ramakrishnan* *Autumn 2018*
  - Computer Programming and Utilisation - *Prof. Krishna S.* *Autumn 2017*
  - Basic Calculus - *Prof. Amiya K. Pani* *Autumn 2016*
- *MOOC Teaching Assistantship* - IITBombayX, edX
  - Data Structures and Algorithms - *Prof. Deepak B. Pathak* *Spring & Autumn 2017*
- *Managing Extreme Classification Reading Group* - Microsoft Research India *2020 - 2021*

Relevant Courses & electives

Machine Learning

- **Specialized**: Advanced Machine Learning, Organization of Web Information, Foundations of Intelligent Learning agents, Fundamentals of Image Processing
- **Fundamentals**: Fundamentals of Machine Learning, Artificial Intelligence, Calculus, Linear Algebra, Numerical Analysis

Others

- **Theory**: Applied Algorithms, Data Structures Algorithms, Design Analysis of Algorithms, Logic for Computer Science, Discrete Structures, Automata Theory
- **System**: Digital Logic Design, Computer Networks, Computer Architecture, Operating Systems, Database Systems, Implementation of Programming Languages

Academic Research Projects

Robust Lane Detection for SeDriCa *Spring 2016*
Innovation Cell, IIT Bombay

- Developed and implemented lane detection algorithm for autonomous ground vehicle SeDriCa (Self Driving Car).
- Our submission at Intelligent Ground Vehicle Competition 2016 bagged 5th position in Advanced Auto-Nav Challenge

Automated Music Transcription *Spring 2018*
Advisor: Prof. Ajit Rajwade, *IIT Bombay*

- Developed online tool for real time transcription of music played on piano to MIDI notation.
- Solution based on background subtraction for detection of pressed keys after normalizing the frame for illumination and transformations

Deep Reinforcement Learning in Mario *Autumn 2018*
Foundation of Intelligent Learning Agents, Prof. Shivaram Kalyanakrishnan

- Explored applications of NEAT algorithm on learning controls for agent in MARIO environment having no prior domain knowledge about the environment