

# YASH GARG

yashgarg1232.github.io | yashgarg1232@gmail.com

Machine Learning Researcher, Nokia Bell Labs

Ph.D. in Computer Science

## EDUCATION

---

### Ph.D. in Computer Science

Dec 2020

Arizona State University, Tempe, Arizona

- Recipient of Graduate Collage Doctoral Fellowship
- Recipient of Computer Science Doctoral Fellowship
- Outstanding Mentor Award
- ACM Student Grant

### M.S. in Computer Science

Dec 2015

Arizona State University, Tempe, Arizona

### B.E. in Computer Science & Engineering

June 2013

Rajiv Gandhi Technological University, Bhopal, India

## EXPERIENCE

---

### Nokia Bell Labs

Jan 2021 - Present

Machine Learning Researcher

Murray Hills, NJ

- Automated end-to-end framework for tabular and sequential data processing and modelling

### Arizona State University

May 2015 - Dec 2020

Graduate Research Assistant

Tempe, AZ

- Developed a principled approach to discover insights from the data to perform **single-shot hyperparameter search** and **retraining-free sparsification** of network parameter
- Leveraged **multi-scale patterns** contained in the data to design novel **attention mechanism**, such as localized, cross, and multi-scale multi-head attention, for **multimedia retrieval**

### Nokia Bell Labs

Jun 2019 - Aug 2019

Data Science Intern

Murray Hills, NJ

- Developed an **automated representation** learning framework for rare event detection in streaming time series. Designed a **budgeted** approach to **adaptively learn** the **length of buffer window** for learning representation for streaming time series. Patent in review.

### Eaton Corporation

May 2018 - Aug 2018

Data Science Intern

Menomonee Falls, WI

- Designed a **recurrent ensemble model** for **accurate time series forecasting** in high-dimensional sensory networks. Developed a NodeJS based **spatio-temporal visualization** engine for 62k sensor spread in 85kmi<sup>2</sup>

## TECHNICAL STRENGTHS

---

### AI/Machine Learning

Tensorflow, PyTorch, Keras, PyCaret, Scipy scikit-learn, Pandas

### Languages

MatLab, Python, R, Java

### Databases

MongoDB, MySQL, PostgreSQL

### Web Technologies

JavaScript, jQuery, HTML, CSS

## SELECTED PROJECTS

---

**AutoRep: Automated representation learning for data streaming** Python

- Developed a novel framework for **streaming data** (over cloud) from remote location for **rare event detection**. Designed a **budgeted** approach to **adaptively learn** the **length of buffer window** for learning representation with high variability. Patent in review.

**StormViz: An analytics tool for decision-making in hurricane** HTML, CSS, JS, MongoDB

- Developed a **cloud-based spatio-temporal visualization** engine. Stored unstructured data in MongoDB for scalable data storage and indexing. Implemented variety of visualization using Leaflet Maps API

**QTagger: Real-time tag recommender for Stack Overflow Questions** Python

- Implemented a real-time state-full LSTM-based *top-5* tag recommendation engine for user provide question. Processed 1.6 million questions with 87k potential tags. Performed data wrangling, trained word-2-vec model, and trained a tag deep recommender model

## PEER-REVIEWED PUBLICATIONS

---

★ First Author | ◇ Co Author | + Equal Contribution | <https://yashgarg1232.github.io/research/>

### Journals

- ◇ Selego: Robust Variate Selection for Accurate Time Series Forecasting, ECML-PKDD 2021
- + DataStorm: Coupled, Continuous Simulations for Complex Urban Environment, TDS 2021
- ◇ DataStorm-FE: A Data and Devision-Flow Engine for Coupled Simulation Ensembles, P-VLDB 2018

### Conferences

- ★ XM2A: Multi-Scale Multi-Head Attention with Cross-Talk, MIPR 2021
- ★ SDMA: Saliency-Driven Mutual Cross Attention, ICPR 2021
- ★ SAN: Scale-Space Attention Network, ICDE 2020
- ★ iSparse: Output Informed Sparsification of Neural Networks, ICMR 2020
- ★ RACKNet: Robust Allocation of Convolutional Kernels in Neural Networks, ICMR 2019
- +★ On the Effectiveness of Distance Measures for Similarity Search in Sensory Data, ICMR 2017

### Demonstrations

- + SIMDMS:Data Management and Analysis to Support Decision Making Through Large Simulation Ensembles, EDBT 2017

### Workshop

- + Load-Adaptive Continuous Coupled-Simulation Ensembles with DataStorm and Chameleon, CC 2019
- ◇ NOTES2: Network-of-Traces for Epidemic Spread Simulations, AAAI Workshop, 2015

### Dissertation/Thesis

- ★ On Feature Saliency and Deep Neural Networks, Ph.D. Dissertation, ASU 2020
- ★ Multi-Variate Time Series Similarity Measures and their Robustness Against Temporal Asynchrony, MS Thesis, ASU

## PATENT

---

Akyamac, Ahmet, Lehman, Gerald, and *Garg, Yash*, “Apparatus, Method, and System for Providing a Sample Representation for Event Prediction”. Filed Jan 8, 2020 (FI). In Review.