Rishabh Misra

r1misra@ucsd.edu ☎+1 (650)-686-9605 linkedin.com/in/misrarishabh github.com/rishabhmisra

EDUCATION

M.S. Computer Science	**University of California San Diego**	**GPA**: 3.90/4	**Sep 2016 – Jun 2018**
B.E. Computer Engineering	**Thapar University, India**	**GPA**: 9.88/10	**Jul 2011 – Jul 2015**

INDUSTRY EXPERIENCE

- **SDE Intern @ Amazon.com, Seattle, Washington** **Jun 2017 – Sep 2017**
  - I am interning in the **DataForge** team which provides a **platform for running Big Data operational workloads** consistently within service level agreement. I am working towards **designing** and **implementing**:
    - Support for **primary key constraint** and **batch insert/update while ensuring consistent reads in Hive** using **type-2 tables** and **multi-version concurrency control** concepts.
    - Support for **ACID properties in Hive**.
    - Support for non-blocking **compaction** (carefully discarding old data) to keep read operations efficient.
  - **Technologies**: Java | Hive | DynamoDB

- **Member Technical @ Arcesium India Pvt. Ltd., India** **Jul 2015 – Jul 2016**
  - Worked as a **primary Java developer** for the Straight Through Processing team. Important responsibilities:
    - **Migration of** scripts from **legacy** to **Java-based** infrastructure while **ensuring reusability** and **scalability**.
    - Adding support for **self-sanitization**, **self-recovery** and **fault tolerance** in the new infrastructure.
    - Adding a **self-aware triggering mechanism** for Blotters, greatly **minimizing** data completeness **issues**.
  - **Technologies**: Java | Spring Framework | MyBatis | Microsoft SQL Server

RESEARCH EXPERIENCE

- **Student Researcher under Prof. Julian McAuley @ UCSD** **Apr 2017 – Present**
  - Working towards **understanding structure of reviews'**, which include many signals like users' purchase motivation apart from product feedback, using **NLP techniques** and **designing models for recommendation**.
- **Research Intern @ Indian Institute of Technology, Madras** **Dec 2014 – May 2015**
  - **Scalable Bayesian Matrix Factorization algorithm:** reduces the cubic time complexity of existing Bayesian matrix factorization algorithm to linear. (C++ | Python | Matlab) (**Published**; **Link**: goo.gl/ou2B7f)
  - **Scalable Variational Bayesian Framework for Factorization Machines:** Supplements the existing framework with a scalable alternative that gives state-of-the-art performance. (C++ | Python | Matlab)
- **Summer Intern @ Indian Institute of Technology, Madras** **May 2014 – Jul 2014**
  - **Collaborative Tweet Recommendation:** Used Parametric Matrix Factorization to efficiently recommend relevant tweets to users. (C++ | Python)

TEACHING EXPERIENCE

- **Teaching Assistant @ UCSD**
  - **Recommender Systems and Web Mining** (CSE 258) | Prof. Julian McAuley **Sep 2017 – Dec 2017**
  - **Software Engineering** (CSE 110) | Prof. William Griswold **Mar 2017 – Jun 2017**
  - **Software Engineering** (CSE 110) | Prof. Gregory Kesden **Sep 2016 – Dec 2016**

PROJECTS

- **Review Ranking and Recommendation on Ciao Product Dataset** **Java**
  - Reviews ranked and recommended by **optimizing** the **Bayesian Personalized Ranking** measure on **biased Matrix Factorization** and **biased Tensor Factorization** models.
- **Music Generation using Character-level Recurrent Neural Networks** **Python | Keras framework**
  - **Trained** an **RNN** to **learn** the structure of **music** files in ABC format and then **generated music** from the trained network.
- **An Ensemble of CNNs for Traffic Lights Recognition** **Python | Keras framework**
  - **Ensemble** of custom built **CNNs** trained on **Nexar** traffic lights challenge **dataset** while **ensuring small model size** which allows for a **quick training** even with **scarce computational resources**.
- **Hotel Recommendation System** **Python | Scikit-learn | Pandas**
  - **Recommender system** trained on Expedia Hotel Recommendation Dataset to **recommend top 5 hotel clusters** to users, built using **ensemble** of **Random Forest**, **Naïve Bayes**, **SGD classifier** and **XGBoost models**.

GRADUATE COURSE WORK

- **Recommender Systems and Web Mining**, **Neural Network and Pattern Recognition**, **Big Data Analysis using Spark**, **Probabilistic Reasoning and Learning**, **Statistical Inference and Data Analysis**