Rishabh Misra

✉ rishabhmisra1994@gmail.com   ☎ +1 (650)-686-9605   💼 linkedin.com/in/misrarishabh   🖥 rishabhmisra.github.io

EDUCATION

University of California San Diego – Master's in Computer Science	**GPA**: 3.93/4.00	**[Sep 2016 – Jun 2018]**
Thapar University, India – Bachelor's in Computer Science	**GPA**: 9.88/10.00	**[Jul 2011 – Jul 2015]**

INDUSTRY EXPERIENCE

➢ **Software Development Engineer @ Amazon.com, Seattle, Washington**   **[Jul 2018 – Present]**
- Working for Amazon Global which enables customers to buy products internationally based on eligibility.

➢ **SDE Intern @ Amazon.com, Seattle, Washington**   **[Jun 2017 – Sep 2017]**
- Interned at Financial Intelligence Systems team which provides a platform for running big data operational workloads consistently within service level agreement. Worked towards designing and implementing:
    - Support for primary key constraint, batch insert/update and ACID properties while ensuring consistent reads 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 for the Straight Through Processing team. Important responsibilities:
    - Added support for self-sanitization, self-recovery and fault tolerance in the new infrastructure.
    - Added self-aware trigger mechanism for EOD trade files, minimizing data completeness issues by 30%.
- **Technologies**: Java | Spring Framework | MyBatis | Microsoft SQL Server

RESEARCH EXPERIENCE AND PROJECTS

➢ **Student Researcher with Prof. Julian McAuley @ UCSD**   **[Apr 2017 – Jun 2018]**
- **Decomposing Fit Semantics for Product Size Recommendation in Metric Spaces (Accepted at RecSys 2018)**
    - Proposed a framework that uses latent factor model and metric learning to predict fit of clothing products for recommending appropriate catalog sizes to customers. Contributed two public datasets as well.
    - Observed an improvement of around **18%** over an algorithm developed by Amazon.
- **Spoiler Detection using Probabilistic Graphical Model**
    - Working towards developing probabilistic graphical model with integrated language model to detect spoilers in book/movie reviews and contributing some interesting linguistic insights on what constitute spoilers.
    - Data obtained from two popular book/movie review websites, then cleaned, pre-processed and analyzed.

➢ **Research Intern @ Indian Institute of Technology, Madras**   **[May – Jul 2014, Dec 2014 – May 2015]**
- **Scalable Bayesian Matrix Factorization algorithm:** reduces the cubic time complexity of existing Bayesian matrix factorization algorithm to linear. (**Published at MUSE 2015**)
- **Scalable Variational Bayesian Factorization Machines:** Supplements the existing framework with scalable alternative that gives state-of-the-art performance.

OTHER PROJECTS

➢ **Sarcasm Detection using Hybrid Neural Network**   **Python | PyTorch**
- Collected a news headlines-based dataset that addresses gaps in existing literature.
- Developed interpretable hybrid neural network architecture with attention mechanism which improves baseline by **5%**.

➢ **Jointly Modeling Aspects, Ratings and Sentiments with Temporal Dynamics**   **Python**
- We jointly model aspects of the products, user sentiment on products, associated ratings and temporal information in a probabilistic graphical model to predict the review ratings. For interpretability, model produces insights on the various aspects of products and user sentiment on them which explains the rating.
- Improves upon the baseline by **1%** and additionally provides insight into how users' preference change over time.

➢ **Hierarchical Attention Network for Rating Prediction**   **Python | Keras**
- Implemented a hierarchal RNN with attention mechanism for rating prediction using product reviews. Attention mechanism allows the RNN to focus on sentences which best explain the rating given to an item.

➢ **An Ensemble of CNNs for Traffic Lights Recognition**   **Python | Keras**
- We construct an ensemble of CNNs with varied architectures and train it on Nexar traffic lights challenge dataset. Proposed architecture ensures small model size that allows for a quick training even with scarce computational resources.

➢ **Music Generation using Character-level RNN**   **Python | Keras**
- Proposed a character-level RNN that first learns the structure of music files in ABC format. Following the training we generate music from the trained network in ABC format which can then be converted into audio format.

➢ **Hotel Recommendation System**   **Python | Scikit-learn**
- We build a recommender system using an ensemble of Random Forest, Naïve Bayes, SGD classifier and XGBoost models and train it on Expedia Hotel Recommendation Dataset to recommend top 5 hotel clusters to users.