RISHABH MISRA

+1-650-686-9605 | rishabhmisra33@gmail.com | rishabhmisra.github.io | rishabhmisra | misrarishabh

R&D leader in Recommender Systems, NLP & Deep Learning domains. Specialized in architecting high-performance, low-latency, & massive-scale AI models for production usecases. Published & refereed publications at top AI conferences like ACL, RecSys, ICML, & KDD (w/ 900+ citations), and co-author of the book Sculpting Data for ML.

Experience

Attentive Mobile — *San Francisco, CA*

Staff Machine Learning Engineer / Tech Lead — *June 2023 - Present*

Founded the Personalization ML team, developing and scaling deep learning-powered personalization models for Conversational Commerce, impacting hundreds of millions of users and driving revenue for 8k+ leading brands.

- Architected Two-Tower & Multi-Task Learning Deep Learning pipelines from scratch, optimizing for relevance and business objectives at scale.
  - Leading model improvement through loss function customization, multi-objective optimization, advancing neural architectures (MLP stacks, Cross modules, Attention, MMoE, PLE, etc), data sampling & representation, and LLM-powered feature engineering.
  - Leading model scalability through distributed model training, model and data sharding, mixed precision training, heterogeneous GPU clusters, and ANN-based serving for hundreds of millions of users.
  - Optimizing the pipeline for relevance while balancing brand-specific requirements around privacy, freshness, coverage, and seasonality.
- Driving generative applications for personalization by leveraging prompt engineering, RAG architectures, and PEFT-based LLM fine-tuning, grounded in rich user and item representations derived from predictive deep learning models, to enable deep hyper-personalization.
- Technical leadership in defining a long-term personalization strategy across the company while overseeing execution across 5+ MLEs.
  - Partnering with platform teams to build feature stores, distributed training/inference, A/B experimentation, and ANN-powered serving for real-time personalization.
  - Facilitating model application for personalized product recommendations, audience targeting, Automated Campaigns, and so on.
- Technologies: Python, PyTorch, GPT, LLaMa, Ray/Anyscale, Metaflow, Argo, Kubernetes, StatSig, AWS Technologies, Snowflake, Tecton

Twitter — *San Francisco, CA*

Senior Machine Learning Engineer / Tech Lead — *July 2019 - April 2023*

One of the founding engineers and technical lead for the Reply Feed Ranking team that applied ML and NLP to drive meaningful conversations across Twitter. Also shared time working on Twitter's Home Feed Ranking to personalize user experience and recommend relevant tweets.

- Built Deep Learning models to predict various engagements to promote more engaging content and personalize user experience. Drove the progression from shallow models to deep models to multi-task learning architecture.
- Shipped graph embeddings (based on Heterogeneous Information Network) and text embeddings (based on fine-tuned BERT), among other sophisticated features to encode users' preferences and author's affinity. Drove >30% gains in key engagement metrics over tenure.
- Architected a low-latency Early Ranking system using ML signals to scale the ranking service to systematically handle the ranking of tens of millions of candidates per second. This led to a 20% gain in p9999 latency while improving product health metrics by 5% overall.
- Drove modernization of training and deployment stack by architecting KubeFlow pipelines to train next-generation native Tensorflow models.
- Built roadmaps guided by data-backed analyses & shipped measurement frameworks to quantify key product metrics improvements.
- Technologies: Python, Java, Scala, Tensorflow, Airflow, Kubeflow, BigQuery, DataFlow, Google Cloud Platform (GCP), Scalding.

Amazon — *Seattle, WA*

Software Development Engineer — *July 2018 - July 2019*

Worked in Amazon Global that enables customers to buy products internationally based on export eligibility.

- Improved the infrastructure scalability by designing solutions using Native AWS technologies.
- Conducted experiments to improve the eligibility prediction of products using Machine Learning models.
- Assisted courses taught in Amazon's Machine Learning University, aimed at upskilling Amazon engineers in ML.
- Technologies: AWS Technologies, Java, Python, Jupyter Notebook.

Amazon — *Seattle, WA*

Software Development Engineering Intern — *June 2017 - Sep. 2017*

Worked in the DataForge team that provides a platform for smartly scheduling Big Data operational workloads within SLAs

- Designed support for primary key constraint and batch inserts/updates, using append-only table and multi-version concurrency control concepts, while ensuring consistent reads in Hive.
- Support for transactional operations and compaction (carefully discarding old data) without blocking other operations in Hive.
- Technologies: Java, Hive, DynamoDB.

Arcesium (a D.E. Shaw Company) — *Hyderabad, India*

Software Engineer — *July 2015 - July 2016*

I worked in the Arcesium/Tech division's "Straight Through Processing" team. Some of my important responsibilities include:

- Adding support for self-sanitization, self-recovery and fault tolerance in the new infrastructure built using JAVA.
- Profiling and optimizing (around 40%) code (using concurrency) and database (using index and partitions).
- Technologies: Java, Spring, MyBatis, SQL Server.

Education

University of California, San Diego *La Jolla, CA*

Master of Science - Computer Science (Artificial Intelligence specialization), GPA: 3.93/4.0

Thapar University *Patiala, India*

Bachelor of Technology - Computer Engineering, GPA: 9.88/10.0, Rank 1, Gold Medalist.

Research Publications

900+ citations on Google Scholar.

- **Keeping it Low-Key: Modern-day approaches to Privacy-Preserving Machine Learning**: Book Chapter by Jigyasa Grover* and Rishabh Misra*, published in Springer book "Data Protection in a Post-Pandemic Society - Law Regulations, Best Practices and Recent Solutions", ISBN 978-3-031-34006-2. Jul. 2023. *equal contribution
- **Sarcasm Detection using News Headlines Dataset**: Rishabh Misra, Prahal Arora, in Proceedings of journal AI Open, Volume 4, 2023, ISSN 2666-6510, https://doi.org/10.1016/j.aiopen.2023.01.001, Feb. 2023.
- **Do not fake it till you make it! - Synopsis of trending fake news detection methodologies**: Book Chapter by Rishabh Misra and Jigyasa Grover, published in book "Deep Learning for Social Media Data Analytics" of Springer book series "Studies in Big Data", ISBN: 978-3-031-10868-6 Sep. 2022.
- **Sculpting Data for ML: The first act of Machine Learning**: Book by Jigyasa Grover* and Rishabh Misra*, Jan. 2021. Independently published. ISBN-13: 979-8585463570. *equal contribution
- **Addressing Marketing Bias in Product Recommendations**: Mengting Wan, Jianmo Ni, Rishabh Misra, Julian McAuley, in Proceedings of 2020 ACM Conference on Web Search and Data Mining (WSDM'20), Houston, TX, USA, Feb. 2020.
- **Fine-Grained Spoiler Detection from Large-Scale Review Corpora**: Mengting Wan, Rishabh Misra, Ndapa Nakashole, Julian McAuley, in Proceedings of 57th Association for Computational Linguistics 2019 (ACL'19), Florence, Italy, Jul. 2019.
- **Hotel recommendation system**: Aditi A Mavalankar*, Ajitesh Gupta*, Chetan Gandotra*, Rishabh Misra*, arXiv preprint arXiv:1908.07498 (2019). *equal contribution
- **Decomposing Fit Semantics for Product Size Recommendation in Metric Spaces**: Rishabh Misra, Mengting Wan, Julian McAuley, in Proceedings of 2018 ACM Conference on Recommender Systems (RecSys'18), Vancouver, Canada, Oct. 2018.
- **Scalable Variational Bayesian Factorization Machine**: Avijit Saha, Rishabh Misra, Ayan Acharya, and Balaraman Ravindran, preprint 2017.
- **Scalable Bayesian Matrix Factorization**: Avijit Saha*, Rishabh Misra*, Balaraman Ravindran, In Proceedings of 6th International Workshop on Mining Ubiquitous and Social Environments (MUSE), co-located with the ECML PKDD 2015. *equal contribution

Dataset Publications

150k+ downloads on Kaggle; Used in DeepLearning.AI's "Natural Language Processing in TensorFlow" course on Coursera & Youtube (taken by 300k+ people to date). Featured in best selling book AI and Machine Learning for Coders.

- **Politifact Fact Check Dataset**: Misra, Rishabh, DOI: 10.13140/RG.2.2.29923.22566 (2022).
- **IMDB Spoiler Dataset**: Misra, Rishabh, DOI: 10.13140/RG.2.2.11584.15362 (2019).
- **News Headlines Dataset**: Misra, Rishabh, DOI: 10.13140/RG.2.2.16182.40004 (2018).
- **News Category Dataset**: Misra, Rishabh, DOI: 10.13140/RG.2.2.20331.18729 (2018).
- **Clothing Fit Dataset for Size Recommendation**: Misra, Rishabh, DOI: 10.1145/3240323.3240398 (2018).

Research Committees

Invited Program Committee Member at some of the leading research conferences

- TheWebConf 2023, 2024, RecSys {2025, 2024, 2022}, ICML 2022, SIGKDD 2022, 2023, SIGIR {2025, 2024, 2023, 2022}, CIKM 2023, 2024, ICDM 2022, ECML-PKDD 2022, ICWSM {2024, 2023, 2022}, TORS 2022.

Other Notable Achievements

Blogs *Machine Learning blogs on Towards Data Science online publication have been seen by 135k+ people. Additionally, content on personal website has been visited 85k+ times by people from 169 countries.*

Media Coverage *Spoiler Detection research featured in TechCrunch, Gizmodo, Times of India, NBC, Geek.com, TechXplore, and UC News. which have up to ∼ 148M monthly readership.*

Expert Commentary *Featured in various media articles about thoughts on the impact of AI on various domains: Rise of AI bots, Future of AI in acting, AI tools for art generation, Workings of DALL-E, AI changing Art Industry, and so on.*

Awards *Recognized as an Outstanding Researcher by US Government for ML research contributions through EB1-B Green Card, Undergraduate University Medal, Research Fellowship from the Indian Institute of Technology, Madras, India*

Conference Talks *ACM Conference on Recommender Systems, Coalesce, PyCon US, RVA Tech Data Summit, Re-Work Enterprise AI Summit, All Things Open (+ book signing), LeadDev Live, ML Conference, This Week in ML & AI, and many more*