1131 Prestige Ferns Residency
Harlur Road, HSR Layout
Bangalore - 560102

Varun Tulsian

(+91) 9611185547
[LinkedIn](#)
varuntulsian2001@gmail.com
[My Blog, wity'AI](#)
[Github](#)

Principal-level AI leader & hands-on IC with 12 + years building and scaling retrieval-augmented generation (RAG), LLM agents and ranking systems that move revenue-critical metrics in finance, commerce and search. Equally comfortable shaping strategy for execs and writing the production code/infra that ships—patent-pending long-doc QA at Needl, 10 TB GPU-efficient RAG pipelines, and finance-grade agent orchestration. Proven record of hiring & mentoring small, high-leverage teams while remaining a top individual contributor driving core algorithmic breakthroughs.

Work Experience:

[Needl.ai](#), Bangalore
Apr 2023-Present

VP, Head of AI (Promoted from Director)

Leading AI strategy, team building, and cross-functional collaboration to drive Needl.ai's vision of creating AI-powered workflows that help customers stay informed and improve decision-making.

- **Team Leadership & Growth:** Expanded the AI team from 2 to 7 high-caliber engineers and researchers. Oversaw the end-to-end development of cost-effective AI solutions for search, large-scale QA (over 10TB of data), and agentic (automated) workflow automation.
- **Strategic AI Roadmapping:** Defined and executed the AI roadmap in partnership with executive leadership, product, and sales teams. Enabled strategic partnerships with Microsoft/PwC/BCG and others, aligning technical initiatives with long-term business goals.
- **Client-Centric Innovation:** Developed modular, customizable AI frameworks ("Lego blocks") frameworks for diverse customer use cases like question answering, alerting, and advanced document synthesis. Delivered agentic workflows to generate actionable insights from complex, unstructured documents.
- **Flagship Project – AskNeedl (Patent Pending):** Led creation of a QA system achieving 85% precision, surpassing OpenAI's performance on document question answering over complex financial documents (55 vs. 31 wins for AskNeedl solution). Document QA can handle 2500+ pages large PDF, Currently joint #1 on long doc QA [Kensho S&B benchmark](#).
- **Industry-Specific Solutions:**
  - **Energy-trading alerts:** Deployed generative-AI pipelines that sift real-time market data and flag actionable events, boosting trader confidence and response speed.
  - **Bank-underwriting assistant:** Built a proof-of-concept workflow that reads bank statements, financials and tax filings, then drafts credit-analysis reports—cutting manual review effort for analysts.
  - **Relationship-manager (RM) assistant:** Created an agent-orchestrated system that scans hundreds of funds, bonds and internal research notes to recommend investment options for private-bank clients.
- **Technical Depth:** Personally developed information retrieval solutions for SOTA [Long Document Question-Answering](#), fine tuning small LMs/ Bi-encoders/ Cross Encoders, parallel LLM inference with shared GPU over torch-serve/ vLLM, complex PDF parsing and retrieval, Agentic GenAI solutions for complex questions answering over Large Corpus, customized agentic solutions for Agentic Process Automation, Research Report Generation, and alerting. Personally optimizing llama models using structured generation, test time scaling and fine tuning to enhance query understanding capabilities.

Professional development, Break
Oct 2022-Mar 2023

Took a career break, utilized to deepen expertise in deep learning and generative AI and working on personal projects related to diffusion models. Built tutorials to demonstrate key ideas of diffusion models with 2D datasets and on EMINST dataset for Masters and PhD students. [View my work here](#). **60+ stars** on [Github](#). Starrers include PhD students working in this area.

Walmart Labs, Bangalore
Apr 2020-Sep 2022

Varun Tulsian

Senior Manager, Data Science, Workstream Lead, Walmart Search Marketplace

Managed a 5-member AI team delivering data-driven insights for 100K sellers and 60M products on Walmart's marketplace.

- **AI-Powered Recommendations:** Implemented content recommenders (NLP & transformers) adopted by 7K sellers across 400K products, significantly improving product discoverability.
- **Assortment Gap Analysis:** Leveraged semantic search embeddings to identify product gaps, enabling $20M incremental sales.
- **Search Conversion Optimization:** Drove multi-objective ranking improvements, increasing search conversion by 61 basis points.
- **Technical Depth:** Introduced semantic search ranking (Bi-Encoders), graph-based embeddings, spam identification, and taxonomy matching with transformer-based methods.

Sensara Technologies, Bangalore   Jun 2019-Apr 2020

Staff Data Scientist, Lead Content Personalization

Created a real-time personalization platform for consumer entertainment experiences.

- **Recommendation Frameworks:** Deployed user/content embeddings and knowledge graph models to power Sensara's "+You Page" for Xiaomi and Airtel.
- **Scalable Infrastructure:** Handled 1M events/hour, serving 300K+ users with personalized content recommendations.

Walmart Labs, Bangalore   Aug 2013-May 2019

Staff Data Scientist, Walmart Search

Lead and key contributor on Walmart Grocery, Sam's Club, and Walmart.com Search ranking teams.

- **Grocery Search:** Improved product relevance using CNN models for intent and attribute understanding, boosting grocery conversions.
- **SamsClub Search Personalization:** Developed Bayesian models for reorder-based personalization, increasing conversion rates by 1%.
- Led a team to optimize search ranking via clickstream data, resulting in a +1.9% increase in user orders.

Other Work: XGBoost Re-Ranking for Grocery, Related Searches via Markov models, Spell Correction, Typeahead Systems, Merchandising Modules for Walmart.com.

Microsoft Research, Bangalore   Jan 2013-March 2013

Intern

Worked on portfolio algorithmic selection with the Software Verification team, contributing to research that advanced model checker selection methodologies.

EDUCATION & ACHIEVEMENTS:

Masters in Engineering (ME), Computer Science, Indian Institute of Science   Aug 2011-Jul 2013

- CGPA of **6.8/8**. Top 10 percentile at IISc.
- **AIR 8** (100 percentile) among 136K in **GATE 2011 Computer Science Exam.**

1131 Prestige Ferns Residency
Harlur Road, HSR Layout
Bangalore - 560102

Varun Tulsian

(+91) 9611185547
LinkedIn
varuntulsian2001@gmail.com
My Blog, wity'AI
Github

Selected Publications & Patents:

1. Multi-Objective Learning: Explored label aggregation for e-commerce learning to rank, contributing to improved relevance and ranking models.
2. MUX: Co-authored a paper on algorithm selection for software model checkers (**40 citations**).
3. MintHint: Developed automated repair hints for software engineering (**120 citations**).
4. Diffusion Models: Created a tutorial on diffusion models with over **60+ stars** on GitHub.
5. Patent Pending: Enhanced Question Answering and Summarization over Long doc QA Kensho S&B benchmark.