HEMANTH KONGARA

Senior Data Scientist - Apna | M.Tech - IISc | B.Tech - NIT, Trichy

@ hemanthkongara35@gmail.com @ hemanthk@alum.iisc.ac.in +918919692372 in HemanthK Google Scholar

EXPERIENCE

Senior Data Scientist - Apna

Search and Recommendation, TNS (Trust and Safety)

Sep 2024 – Present Bangalore

- **Profile Unification**: Working on an agent-driven Profile Unification system that merges user-entered data with resume-parsed data to auto-update title, role, and skills of user profiles
- **Content Discovery and Personalization**: Independently Led efforts across Search and Recommendation systems, collaborating cross-functionally to drive content discovery and personalization improvements.
  - Transitioned from entity-based to vector retrieval with a taxonomy-aware two-tower model.
    - Resulted in **11%** SL uplift per user, **8%** TU increase, **11%** CTA/impression boost and **5%** rise in matched screen leads
    - JobFeed PNs: Delivered **6%** CTR uplift and **20%** SL gain
    - RHR: Achieved **38%** CTR growth and extended coverage by 3M users
    - SHR: Achieved **20%** SL uplift and **25%** increase in matched screen leads
  - Trained a self-attention fusion layer (with trainable Q,K), improving SL/User by **7%**
  - Identified suboptimal inference latency in engineering pipeline; replaced with ONNX-based models, achieving **50%** latency reduction (2× faster).
  - Designed and deployed an L4 cross-encoder with advanced hard negative mining, achieving a **5.5%** improvement in SL/U for experienced candidates; presently optimizing the latency–performance tradeoff for production efficiency.
- **Xgboost Ranker**: Trained an XGBoost Ranker achieving a **10%** uplift in offline NDCG@10, rolled out to 5% traffic in production, with further scaling planned based on online metrics.
- **Taxonomy Model (Custom)**:Fine-tuned a taxonomy model for internal ontology using UAE Large V1 with LoRA; leveraged anchor-positive pairs and in-batch negative sampling to improve entity representation. Increased average NDCG@10 for the top 3K entities by **20%**, from 0.75 to 0.90
- **Fraud Risk Prediction Model (TNS)**: Developed and deployed a fraud classification model using classical machine learning techniques; leading to an **8%** increase in recall.
- **Company-Industry Mapping**: Designed a robust company-industry mapping system using RAG with grounding and few-shot Chain-of-Thought (CoT) reasoning to ensure accurate attribution for each company; established a daily Airflow DAG to automate mapping for newly added companies.

Senior Data Scientist - Monster.com (Foundit.ai)

Search and Recommendation

Aug 2022 – Sep 2024 (Promoted in Apr 2024) Bangalore

- **Search and Recommendation – Learn-to-Rank (LTR)** : Developed a proof-of-concept using open datasets to train custom Learning-to-Rank (LTR) models, enhancing personalized reranking for both search experience for candidates and recruiters
- **Cohort Modeling for Enhanced Personalized Search**: Engineered GNN-based vector representations to form hiring and industry similarity cohorts, boosting NDCG@5 by **13%**and NDCG@10 by **17%**.Deployed API to generate cohorts from hiring data and industry embeddings.
- **Personalized Job Recommendation**: Improved apply rate by **50%** by deploying BiVAE for collaborative filtering (indexed profiles) and content-based recommendation (non-indexed profiles).
- **Core-Koyo Integration (NLP)** :Developed a custom SBERT using Siamese networks for profile embeddings and semantic matching with USE + Elasticsearch. Deployed via FastAPI and Docker, achieving **6%** higher coverage and **15%** fewer false positives.
- **Salary and Notice Period Prediction**:Implemented semantic matching with all-MiniLM-L6-v2 using Elastic vector search. Improved coverage of the notice period by **45%** and salary by **20%** using diverse datasets.

Associate - Irunway

Intellectual Property

Jul 2018 – Jun 2019 Bangalore

- Worked on patent landscapes, prior-art searches, evidence-of-use claim charts, and 5G patent analysis.

HEMANTH KONGARA

Senior Data Scientist - Apna | M.Tech - IISc | B.Tech - NIT, Trichy

@ hemanthkongara35@gmail.com @ hemanthk@alum.iisc.ac.in +918919692372 in HemanthK Google Scholar

RESEARCH PROJECTS

Research Scholar - Spectrum Lab in association with Carl Zeiss

Visual Explanation for abnormality prediction - OCT Images (CNN - XAI)

Aug 2021 – Jul 2022 Spectrum Lab, IISc, Bangalore

- Explored various pre-processing methods including Noise modelling for denoising, denoising using k-SVD, Optimal Bilateral filter using PURE, Guided filter and Rolling Guided Filter
- **Grad-CAM, Grad-CAM++, Score-CAM, Ablation-CAM** are the Visual Explanation methods of CNN's explored CNN in this project.
- **"Denoising Enhances Visualization of Optical Coherence Tomography Images"** - presented in Medical Imaging Workshop at **NeurIPS, 2022**
- **"Deep Learning based Visualization and Volumetric Analysis of Fluid Regions in OCT Scans"** - published in the prestigious **MDPI** Journal under advances in the diagnosis of Retinal Diseases.
- **"ENSEMBLE-CAM: Robust Visualization for Optical Coherence Tomography Image Classification"** - submitted for **IEEE International Symposium on Biomedical Imaging (ISBI), 2023**

SKILLS

- **Tools/ Frameworks**: Tensor Flow, Pytorch, JAX, ONNX, MLflow, LangChain, LangGraph, Amazon SageMaker, Azure, Databricks, Spark, Airflow, Docker, Git, FastAPI
- **Languages**: Python, SQL, R, Scala, MATLAB, C
- **Databases / Vector Stores**: SQL, MongoDB, Elasticsearch, Apache Solr, Chroma, Qdrant
- **Machine Learning**: Retrieval (Candidate Generation), Ranking, Fraud Detection, Classification, Clustering, SVMs, Time-Series Models, Bagging and Boosting
- **Deep Learning**: Bi-Encoder, Cross-Encoders, Transformers, Attention Mechanisms, Graph Neural Networks (GNN), CNNs

EXTRA-CURRICULAR ACTIVITIES

- Organised an event, **IDEAX** sponsored by **Accenture** to promote green computing in **Raphsody'22'**- an annual fest of IISc, Bengaluru
- **Head member** of 'Public Relations and Hospitality Team' of **Currents'18'**- an annual Technical Symposium of Department Electrical and Electronics Engineering of NIT Trichy
- Coordinator of 'Public Relations and Hospitality Team' of **Festember'16'**- National Level annual cultural festival of NIT Trichy

MINI PROJECTS

- A Graph Multi-Attention Network (GMAN) for Traffic Prediction
- DiffNet++: A Neural Influence and Interest Diffusion network for Social Recommendation
- Super Resolution using Generative Adversarial Networks
- Panoramic Image Mosaic
- Classification and Segmentation using YOLOv5

COURSES/WORKSHOPS

- Machine Learning with Graphs (**Stanford Online**)
- Advanced Deep Learning
- Generative Deep Learning (Coursera)
- Data Analytics
- Computer Vision
- Pattern recognition and Neural Networks
- Machine Learning (Coursera)
- Building Transformer-Based Natural Language Processing Applications (Nvidia Deep Learning Institute)
- Fundamentals of Deep Learning (Nvidia Deep Learning Institute)
- Applied Deep Learning for Medical Data Analysis (EDUXLABS)

ACHIEVEMENTS

- **Reliance Foundation Scholar** (1 in 40 post graduate students selected in India)
- **All India Rank 92** in GATE (EE) 2020

EDUCATION

M.Tech in Artificial Intelligence

Indian Institute of Science

2020-2022 Bangalore

CGPA: 8.30/10

B.Tech in Electrical & Electronics Engineering

National Institute of Technology, Tiruchirappalli

2014-2018 Tiruchirappalli

CGPA: 8.60/10