# Mani Kumar R

ksrmanikumar@gmail.com | manikumar-ksr | Manikumarksr

+91-9480225156

## PROFILE

Data scientist at Meesho and post-graduate from the Indian Institute of Science with expertise in Computational Biology and AI solutions. Skilled in deep learning, and natural language processing with a proven track record of delivering impactful industrial projects and academic publications.

## TECHNICAL SKILLS

| | |
|---|---|
| **Programming Languages:** | Python \| Bash(Linux) \| C++ \| R |
| **Data Science & Machine Learning:** | Supervised and unsupervised algorithms \| Deep learning (CNN, RNN, LSTM, GANs, VAEs) \| Regularisation methods \| Fine-tuning \| Explainable AI (SHAP, LIME) \| Natural Language Processing (NLP) \| Generative AI |
| **Web Technologies and DBMS:** | HTML\| CSS \| FLASK \| React \| Bootstrap \| CytoscapeJS \| mongoDB \| SQL |
| **Libraries & Frameworks:** | Databricks \| Google cloud \| PyTorch \| Scikit-learn \| TensorFlow \| OpenCV \| Transformers \| Langchain \| Spacy \| Pandas\| Pyspark \| NumPy \| Matplotlib \| Seaborn \| Networkx \| Plotly \| Streamlit |
| **Cloud Technologies:** | Microsoft Azure \| Google cloud \| Docker \| Kubernetes |

## EDUCATION

- **Indian Institute of Science** *2020 - Present*
  *BS-MS(Research) CGPA - 7.7/10* Bangalore, India
  - **Key Courses:** Probability and Statistics | Machine learning for Data science | Linear algebra | Deep Learning | Generative AI | Computational Epidemiology
- **Blooms Pre-University College** *2018 - 2020*
  *Pre-University Education 96.7%* Bagepalli, India
  - **Key Subjects:** Physics | Chemistry | Mathematics | Biology | English

## EXPERIENCE

- **Meesho** *May 2025 - Present*
  *Data Scientist (Ranking Team)* Bengaluru, India
  - Developed and optimized a deep learning **prepaid ranke**r using the **FT-Transformer** architecture, specifically designed to increase the share of prepaid orders on the platform.
  - Engineered complex features and processed large-scale datasets for ranker training, contributing to improved model accuracy and relevance.
  - Achieved a significant impact on platform metrics with an **Obyv (Orders by Views) increase of +0.6%.**
- **Fujitsu Research of India Private Limited (FRIPL) [🌐]** *May 2024 - Aug 2024*
  *AI/ML research intern* Bengaluru, India
  - Conducted survey on the Explainability of time series forecasting models.
  - Developed explainable AI models for time series forecasting, reducing MSE to **0.25**.
  - Built and deployed scalable, real-time dashboards to monitor model insights using Streamlit.
- **Nanyang Technological University [🌐]** *Aug 2023 - Dec 2023*
  *Research intern at Mutwil lab, NTU* Singapore
  - Spearheaded a project on GPT-based gene function retrieval from 85,000 scientific articles, creating a yeast knowledge graph available at yeast.connectome.tools.
  - Leveraged OpenAI APIs and natural language processing to automate large-scale biological data integration.

## PROJECTS

- **Vein Visualizer: Segmenting Hand Vein Images Using Deep Learning [Master's Thesis]** *Oct 2024 - April 2025*
  *Indian Institute of Science | Supervisor: Prof. Hardik J Pandya*
  - Objective: Developed a vein visualization system leveraging image processing and deep learning techniques to segment hand vein patterns captured by infrared cameras.
  - *Approach:*
    – Initially explored OpenCV's morphological operations and filtering techniques for preprocessing and segmentation, but results were inconsistent due to image heterogeneity.
    – Trained a DeepLabV3 model with an attention mechanism and a MobileNet backbone, achieving an **IoU score of 0.78** for vein segmentation.
  - Result: This work successfully improved the accessibility of non-invasive vein mapping for medical diagnostics.

- **Skills:** Python, OpenCV, Deep Learning, PyTorch, Image Segmentation, Computer Vision, MobileNet, DeepLabV3, Attention Mechanisms

- **Visual Taxonomy: Predicting Product Attributes (Kaggle Competition)** *2024*
  *Meesho Data Science Challenge.* [🌐]
  - *Objective:*Developed a machine learning model to predict product attributes like color and pattern from images for e-commerce cataloging.
  - *Approach:*
    –Utilized **EfficientNet** with a custom linear classifier, applied image augmentations, and one-hot encoding for labels.
    –Employed one-hot encoding for attribute labels and applied image augmentations to enhance model generalization.
  - *Result:* Achieved a competitive f1 score of **0.73** in the evaluation metric, demonstrating model robustness in attribute prediction. Thereby reducing manual cataloging time, improving catalog accuracy, and enhancing operational efficiency and user experience in e-commerce.
  - **Skills:** Python, Deep Learning, Computer Vision, PyTorch, EfficientNet, Image Classification, Image Augmentation, One-Hot Encoding, Model Optimization

- **The Yeast connectome: [Bachelor's Thesis]** *Aug 2023 - Dec 2023*
  *School of Biological Sciences, NTU, Singapore | Supervisor: Prof. Marek Mutwil.* [🌐]
  - Spearheaded a dynamic project titled: "GPT-based Gene Function Information Retrieval from Scientific Literature" to build a yeast knowledge graph from 85000 research articles.
  - Employed Python programming, Natural language processing, and prompt engineering in conjunction with the OpenAI API to harness the capabilities of GPT-based methods.
  - **Skills:** Python, Natural Language Processing (NLP), OpenAI API, Prompt Engineering, Knowledge Graphs, Text Mining, Machine Learning

- **Computational Genomics Project | CBRAIN Internship** *May 2023 - July 2023*
  *Centre for Brain Research, IIsc | Supervisor: Prof. Shweta Ramdas*
  - Conducted an intensive research internship at the Centre for Brain Research, contributing to the project "Discovery of novel eQTLs using gene expression data in smokers and non-smokers."
  - Leveraged a combination of **Linux**, **Python**, and specialized bioinformatics tools such as **qtltools** and **bcftools** to analyze complex genetic datasets and identify significant expression quantitative trait loci (eQTLs).
  - **Skills:** Python, Linux, Bioinformatics, Genomic Data Analysis, QTL Analysis, QTLtools, BCFtools, Computational Biology, Statistical Genetics

- **Object detector for blinds** *2022*
  *Tools: [Arduino, ESP32 Cam, OpenCv, Kivy]* [🐙]
  - I prepared a prototype of an object detector that will assist blind people while walking using Arduino, an Ultrasound sensor, ESP32 CAM, and OpenCV-python. This will detect the objects and measure the distance of an object, and voice output is obtained from the application running on the phone. A sample video of this project can be found **here**.
  - **Skills:** Computer Vision, OpenCV, Arduino, ESP32-CAM, Kivy, Python, IoT, Object Detection, Ultrasonic Sensors, Assistive Technology

## Publications

**[1]** **Mani R Kumar**, Karthick Raja Arulprakasam, An-Nikol Kutevska, Marek Mutwil, Guillaume Thibault. Yeast Knowledge Graphs Database for Exploring Saccharomyces cerevisiae and Schizosaccharomyces pombe. DOI: **10.1016/j.jmb.2025.169072.**.

**[2]** Karthick Raja Arulprakasam, Janelle Wing Shan Toh, Herman Foo, **Mani R Kumar**, An-Nikol Kutevska, Emilia Emmanuelle Davey, Marek Mutwil, Guillaume Thibault. Harnessing full-text publications for deep insights into C. elegans and Drosophila biomaps. BMC Genomics 25, 1080 (2024). DOI: **10.1186/s12864-024-10997-6**.

## Scholarships and Awards

- **Kishore Vaigyanik Protsahan Yojana (KVPY) Fellow.** **2020–present**
- **NTU-India connect Fellow.** **Aug 2023–Dec 2023**
- **Winner: TATA Building India Essay Competition.** **2017**
- **District-level Winner: Scientific Model on Natural Resource Management.** **2016**

## Certifications

- **Introduction to Retrieval Augmented Generation (RAG)** *Duke University*
- **Introduction to Generative AI** *Duke University*
- **Generative Pre-trained Transformers (GPT)** *University of Glasgow*