Santhosh Manohar Gouda Patil

| +1-812-778-4649 | sanpati@iu.edu | LinkedIn-Santhosh | GitHub | Portfolio | Medium Articles |

EDUCATION

Indiana University, Bloomington AUG 2022 – MAY 2024

Master of Science in Data Science; GPA: 3.93/4.0

Coursework: Machine Learning, Statistics, Advanced Database Concepts, Computer Vision, Approximate Inference in Graphical Models, Advanced NLP

R V College of Engineering, Bengaluru, India AUG 2016 – JUN 2020

Bachelor of Engineering in Computer Science and Engineering; CGPA: 8.04/10

SKILLS

Languages and Frameworks: Python, R, C++, SQL, PyTorch, Keras, Streamlit, FastAPI
Libraries: TensorRT, NumPy, Scikit-learn, Pandas, Statsmodels, OpenCV, Matplotlib, NLTK, PIL, Multiprocessing, DDP, PySpark, SciPy
Developer Tools: AWS, Docker, Kubernetes, PostgreSQL, MySQL, MongoDB, Weights & Biases, Git
Hard Skills: Bayesian Inference, Variational Bayesian Methods, Time Series Analysis, Graphical Models, Generative Models, A/B Testing

WORK EXPERIENCE

Graduate Research Assistant: Indiana University, Bloomington - Prof. Roni Khardon MAY 2023 – present

- Integrated **Convolutional LSTM** and **Transformer** models for improved tropical cyclone genesis forecasting in the Northern Pacific.
- Enhanced Earthformer model with **hierarchical encoder-decoder** structure and **cuboid attention** for accurate atmospheric predictions, reducing RMSE by 15%.
- Developed a LSTM model for locating tropical cyclones, boosting Nowcasting with a 0.82 F1 score through advanced convolutional transitions.
- Crafted an efficient ensemble modeling with **multiprocessing** and **multi-GPU distributed data parallelism**, achieving a 0.75 score for 42-hour forecasts.
- Implemented a **Variational Bayes** extension for **probabilistic forecasting**, enhancing accuracy by 10% through better uncertainty estimation.

Associate Software Developer: IBM Private LTD, India JAN 2021 – JUN 2022

- Developed and deployed an **XGBoost**-based **churn prediction** model using **AWS SageMaker**. Utilized **AWS Glue** for **ETL processes** and data integration from Salesforce, enhancing **predictive analytics by 13%**.
- Conducted comparative experiments on XGBoost and **Random Forest** algorithms to assess model accuracy and robustness, selecting XGBoost for deployment due to its higher efficiency and scalability in churn prediction.
- Established a secure data retention policy to protect production data, **leveraging SQL** for effective data handling and compliance, minimizing unauthorized access risks.

Research Assistant: Indian Institute of Science, Bengaluru, India - Prof. Raghu Krishnapuram OCT 2019 - JUN 2020

- Crafted a sophisticated **U-Net**-based deep learning model for monocular **depth estimation**, incorporating innovative loss functions, which led to a **5% decrease in RMSE** and enhanced depth map quality on the KITTI benchmark, outperforming prior models.
- Led the exploration of **Visual Odometry** using **Deep Recurrent CNNs**, enhancing image sequence modeling for autonomous navigation.
- Optimized the **Mask R-CNN semantic segmentation** model, enhancing accuracy in object detection and classification within intricate visual settings, achieving a **20% improvement** in model precision.

Research Intern: Indian Institute of Science, Bengaluru JUN 2019 - JUL 2019

- Expanded the capabilities of PolygonRNN Auto - Annotation model by incorporating the **YOLO Algorithm** for precise object recognition and classification, accelerating annotation **speed by 40%** and elevating overall annotation quality.

PROJECTS

NLP-Driven Market Sentiment & Valuation Analysis

Established a sophisticated NLP pipeline to enhance market sentiment and company valuation insights from SEC filings and financial news, integrating **Amazon SageMaker**, **Hugging Face**, Pegasus, and **FinBERT** with **CI/CD** in SageMaker Pipelines.

Embedded Topic Modeling (ETM)

Improved **topic modeling** accuracy and **posterior approximation** by combining word and topic embeddings with a **Variational Autoencoder** and analyzed LDA vs. ETM models using Topic Perplexity and Diversity for superior topic extraction.

Bayesian Approach to Movie Recommendation Model

Refined and applied a **Variational Probabilistic Matrix Factorization** model for movie rating prediction, enhancing **performance by 10%** through strategic initialization and training of an **EM Algorithm**, with detailed assessment of RMSE improvements.

CONFERENCES

Yadi Wei, **Santhosh Patil**, Roni Khardon, Chanh Kieu, "Predictability of Tropical Cyclone Formation with Large-Scale Memory Using Deep Learning Transformer" is being presented at the American Metrological Society, Conference on Hurricanes and Tropical Meteorology, May, 2024: Abstract