MANJULATA GARIMELLA

Bargersville, IN 46106 | manju.garimella@lilly.com | 812-606-2781 | LinkedIn: manjulata-chivukula | Github: mchivuku | website

DATA SCIENTIST | MACHINE LEARNING

Over ten years' experience delivering data science solutions to clients leveraging solid computer science background and strong analytical and problem-solving skills. Focuses on data engineering, data science, and machine learning as a data scientist.

EDUCATION

Doctor of Philosophy (PhD), Informatics (expected 2026) INDIANA UNIVERSITY, Bloomington, IN, *Aug 2022 - current*

Master of Science (M.S.), Data Science, INDIANA UNIVERSITY, Bloomington, IN, December 2019. GPA: 3.93/4.0

Relevant Coursework: Advanced Machine Learning, Machine Learning with Signal Processing, Image Processing, Computer Vision, Deep Learning, Learning Theory and Probabilistic Graphical Models, Advanced Database Management Systems, Elements of Artificial Intelligence, Data Mining, Introduction to Statistics, Data Visualization, Algorithm Design and Analysis, High Performance Big data systems.

Master of Science (M.S.), Computer Science, INDIANA UNIVERSITY, Bloomington, IN December 2009; GPA: 3.80/4.0

Bachelor of Technology, Electronics/Communication Engineering, JAWAHARLAL NEHRU TECHNOLOGICAL UNIV., India, 2005

SKILLS

Languages: Python, R, C, SQL, Java, C#, C++, PHP, Javascript, CSS, XML, XSLT, Apex, Visualforce

Technologies: Amazon Web Services (AWS), Azure, Tableau, Alteryx, GCP, AWS, GitHub, Apache Spark, Apache Hadoop, Docker, Microsoft Team Foundation Services (TFS)

Libraries: PyTorch, Scikit-Learn, Numpy, Scipy, Pandas, Jupyter, OpenCV, PIL, ASP.NET framework, .NET framework

Methodologies: Certified Scrum Developer (Agile Engineering)

PROFESSIONAL EXPERIENCE

ELI LILLY AND COMPANY, Indianapolis, IN *March 2020 - current*

Manager, Lilly Research Laboratories Information and Digital Services (LRL IDS)

- Served as technical lead on several ML based projects and PoCs by contributing towards the architecture design of various application components as well as acting as a technical liaison on the project by providing technical expertise, help to developers (for code reviews, brainstorming etc.,), and other key stake-holders in the project. This helped build credibility with the team, the business stakeholders, and personal self-confidence to deliver projects inside Lilly.
- Collaborated with the AADS (Research) team to build an innovative search solution that enabled searching across multiple search engines (e.g., elastic search, semantic search) to return the most relevant results to the user.
- Developed and delivered the data-ingestion pipeline in AWS to ingest data into the application. The code and design for the data ingestion task is a reusable component that can be used for building future applications.
- Developed and delivered successful end-of-the sprint demonstrations to the business, ensuring business requirements for the project were met.
- Setup ML Ops pipeline in AWS to build sagemaker endpoints for the machine learning models.
- Received the NLG innovation award for the PoC project NLG regulatory text summarization which used NLG techniques to glen usable insights from regulatory documents and submissions.
- Implemented and demonstrated the PoC using python's library called Streamlit using various visualizations and results from running machine learning models to describe correlation analysis between the parameters in the dataset that most describe the slippage on the clinical milestone projec. Created a detailed report for describing the dataset/approach/methodology and conclusion/discussion and future work for the implementation of correlations to understand the parameters that determine the project slippage. The document helped provide a baseline understanding of the PoC work and means to determine business value in progressing the work beyond PoC.
- Developed R-Shiny application to cluster Trackwise observation data (with Manufacturing Team MQ) to identify root cause segments inside deviations. This will help automate the root-causes in the deviation data in a more effective and efficient manner.
- Mentor, coach and direct developers to ensure best practices of development are followed in projects.
- Coached students from LAIDEL (Luddy School of Artificial Intelligence Program) program in IU to deliver Computer Vision and NLG POC as a part of summer session.
- Coached Summer Intern from a cross-functional team in MD to help provide predictive analytics for the IIR project. Helped the intern to scope the project and directed him to the deliver successful intern project that befitted Lilly and also the intern to learn and apply the knowledge acquired in school to real-world application in Lilly.
- Currently implementing machine learning models that helps in identifying the clinical sites that are at-risk for enrollment that helps trail capabilities personnel at Lilly to be proactive and perform necessary measures.
- Currently, implementing automatic content generation project that will enable trail capabilities team personnel build site contract draft efficiently providing efficiency to the business processes.

INDIANA UNIVERSITY STUDIOS (IU Studios), Indianapolis, IN **(Github Link)** *May 2019 - Dec 2019*

Data Engineer (part time)

- Developed mechanisms to extract, transform, and load data related to Indiana University from various social media channels (e.g., Twitter, LinkedIn etc.,) into a staging area for analysis using Python APIs and technologies like ETL (Alteryx)
- Implemented a Naive Bayes algorithm for sentiment analysis on the data feed from Twitter to model the impact of tweets from the Indiana University official account
- Delivered social media dashboards using Tableau to the Indiana University Marketing Team (IU Studios) with the goal of providing insights into the social media traffic for IU. Developed reports that are utilized both by designers and decision makers in setting up a successful social media strategy for IU which maximizes brand awareness and positioning
- Developed a comprehensive research report on the qualitative and quantitative data available at Indiana University - Purdue University Indianapolis (IUPUI) for drawing insights on student enrollment and provided recommendations on improvements that could be made to the data collection and data inventory systems to improve predictions of student attractiveness

PACERS SPORTS AND ENTERTAINMENT (NBA), Indianapolis, IN **(Github Link)** *Aug 2018 - May 2019*

Data Analyst (part time)

- **WNBA (Indiana Fever basketball) Visualization**: Developed visualizations for the WNBA game survey data to help understand the impact of in-game entertainment, food and beverage services, and other in-game services on ticket sales; analyzed the data to help with identifying potential season ticket purchasers and the promotions that need to be provided for increasing the renewal rate;
- **WNBA Customer Renewal Profile**: Developed a "renewal profile" for the users by performing cluster analysis of the user data coming from various sources likes Salesforce (for demographic data), internal databases (for ticket transaction data), and user survey data. This was to help the Pacers marketing staff in understanding the various segments of users and customizing offerings to individual segments to help improve ticket sales.

INDIANA UNIVERSITY, Bloomington, IN *Jan 2010 - Aug 2018*

Senior Systems Analyst

- Developed and delivered a number of responsive web applications and websites using state of art web design frameworks like Bootstrap and Zurb to multiple clients across I.U., e.g., Kelley School of Business, Marketing-lockup, IU Health Center, etc.
- Developed applications using a variety of tools and technologies like Salesforce in Apex, Visual force code, ASP. NET, PHP etc.,
- Worked with clients to understand their needs and translate them into requirements for the various applications.

INDIANA UNIVERSITY, University Information Technology Services (UITS) Bloomington, IN *May 2009 - Aug 2009*

Gerald L. Bepko Student Intern, Biomedical Applications, Research Technologies

- Worked on National Gene Vector Laboratories, a Java J2EE-based web application.

BROADCOMM INC., Hyderabad, India *May 2005 - Jan 2007*

Software Engineer

DATA SCIENCE PROJECTS

Learning Disentangled Latent Spaces using Variational Autoencoder (Github Link): A probabilistic graphical model for disentangling latent spaces in images

- Studied and implemented various models using variational autoencoder (VAE) like basic VAE, β-VAE, VAE+GAN, Adversarial symmetric VAE using MNIST dataset to understand the role of VAE in disentangling latent spaces in the image and analyzed the performance of the learnt latent code in image classification tasks

Variational Bayesian Approach to Movie Rating Prediction (Github Link):

- Implemented Variational EM algorithm proposed in the literature using probabilistic matrix factorization to perform movie rating prediction and computed RMSE values using various datasets and rank values

Contributions to Open Source Software (Dipy) (Github Link):

- Contributed to Dipy open source software by implementing SSD (sum of square differences metric) for affine registration of 3D MRI images.

Deep learning models for Computer Vision tasks:

- Used Pytorch framework to build Resnet image classification model and analyzed accuracy of the model in semi-supervised setting using synthetic data
- Implemented deep learning models for image-to-image translation tasks like Multimodal Unsupervised Imageto-Image translation (MUNIT), Unsupervised Image-to-Image translation (UNIT) on amazon and office datasets
- Implemented various GAN models like BicyleGAN, CycleGAN etc. using amazon and office datasets to perform data augmentation using GANs network

Deep learning for Audio Processing:

- Implemented deep learning models such as Siamese network to perform speaker verification task
- Implemented RNN model using LSTM for Speech denoising
- Privacy Preserving for Audio Synthesis - Implemented GAN and Cycle GAN (generative adversarial network) models in order to perform audio synthesis on the input audio signal data

Animal Care and Control of Adopted Animals - Bloomington Animal Shelter:

- Developed visualizations using the animal shelter data in order to draw insights on the outcomes of the animals based on the intake reasons, and incoming months

ACADEMIC EXPERIENCE

INDIANA UNIVERSITY, Bloomington, IN *Aug 2018 – Aug 2020*

Research Assistant *(Jan 2019 – Aug 2020)*

- Developed deep learning based convolution neural network models to improve the prediction time of the failure of electronic circuits based on electromigration analysis of CCD and thermal images
- Developed a deep learning classification model using PyTorch for facial recognition of Asian elephants to help with animal conservation efforts (**Github Link**)

Teaching Assistant for Graduate Course in Elements of Artificial Intelligence *(Aug 2018 – Dec 2018)*

Research Assistant, Computer Science Department *(Jan 2009 – Aug 2009)*

PUBLICATIONS

Garimella, Manju, et al. "A Newborn Embodied Turing Test for Comparing Object Segmentation Across Animals and Machines." The Twelfth International Conference on Learning Representations. 2023.

M. A. Reza, **M. Garimella**, et al., "Microelectronics Failure Prediction and Localization from Optical and Thermal Imagery," 2021 IEEE Physical Assurance and Inspection of Electronics (PAINE), 2021, pp. 1-8, doi: 10.1109/PAINE54418.2021.9707706.

ACHIEVEMENTS

- Winner of Grace Hopper Scholarship for GHC#19
- Winner of state merit scholarship from Govt. of Andhra Pradesh, India.
- Awarded certificate of merit at the end of senior year of undergraduate education for standing second in the school.