Kenny Lawrence Swamy

+1 (812) 778-4957 | kenswamy@iu.edu | linkedin.com/in/kenny-swamy/ | https://kennyswamy.github.io/

EDUCATION

Indiana University Bloomington — **Aug 2022-May 2024**
Master of Science in Data Science; CGPA: 3.67/4 — *Bloomington, Indiana, USA*

JSPM's Rajarshi Shahu College of Engineering, Pune University — **Aug 2016-May 2020**
Bachelor of Engineering in Information Technology; CGPA: 8.47/10 — *Pune, Maharashtra, India*

Work Experience

Tata Consultancy Services — **Aug 2020 – Jul 2022**
System Engineer — *Pune, Maharashtra, India*

- Led the creation of a high-throughput ETL data pipeline using Apache Spark and Airflow, with Python for efficient processing, achieving a **20% reduction** in data processing time.
- This enhancement boosted real-time analytics, supporting strategic decisions, and resulting in a **15% improvement** in analytical decision-making speed.
- Directed the smooth transition of on-premises data systems to AWS, optimizing costs for scalable cloud infrastructure. Automated migration processes with Python scripts led to a **10% cost saving** in cloud operations.
- Enhanced data security with robust encryption and AWS security features, reducing security incidents by 25% and ensuring **100% compliance** with industry standards.
- Established CI/CD pipelines for data applications using Python, reducing data inconsistencies by **30%.** Created custom Tableau dashboards integrated with AWS, enhancing data-driven decisions and contributing to a 20% increase in stakeholder satisfaction.
- Proficiently managed IT infrastructure at Maersk, leveraging SCCM for application deployment, which improved system efficiency by 20% and reduced **downtime by 30%**. Achieved a system reliability score of **99.5%**, ensuring high performance and system reliability.
- Developed resilient network infrastructures for projects like Hamburg Süd and Win365, improving network efficiency by **25%** with the engineering of virtual desktop environments. Earned the "Dynamic Employee of the Team" award within three months, showcasing notable enhancements in team training effectiveness and project delivery speed.

Projects

Virtual Tissue Simulation| *Bloomington, Indiana, USA* — ***Aug 2023 - Dec 2023***

- Implemented a Retrieval-Augmented Generation (RAG) framework, with Large Language Models (LLMs) for information retrieval, **acheiving 20% accuracy** and efficiency boost.
- Developed a dynamic text summarization with a rolling context window, enhancing prompt optimization by **25%** for superior results. Created a Python script for t-SNE-based visualization of LLM embeddings on academic paper fragments, achieving a **30% increase** in clustering precision.

Social Media Community Identification and Analysis | *Bloomington, Indiana, USA* — ***Aug 2023 - Dec 2023***

- Pioneered analysis with Python's Praw library for Reddit data extraction**,** increasing filtering **accuracy by 30%.**
- Employed K-Means clustering and Random Forest within a machine learning framework for efficient community classification, achieving a **35%** accuracy boost.
- Enhanced social media analysis by integrating NLP-based sentiment analysis and trend identification for Reddit. This led to a **25% improvement** in understanding user behavior and a **15% increase** in trend predication accuracy.

Evolution Of Wildfire| *Bloomington, Indiana, USA* — ***Aug 2022 - Dec 2022***

- Developed a Virtual Reality (VR) model with Python and R, processing an expanded 85 GB dataset, which increased data processing **efficiency by 20**%.
- Advanced EDA techniques improved environmental pattern recognition by 25%, with a 30% faster preprocessing time.
- The model accurately visualized wildfire evolution across terrains with an **88% accuracy**, enhancing its academic and research utility by 35%.

Job Recommendation System| *Bloomington, Indiana, USA* — ***Aug 2022 - Dec 2023***

- Developed a model to predict the best-suited IT role based on skills extracted from resumes, using ML algorithms like SVM, Decision Trees, Random Forest, NN, Naive Bayes (Gaussian and Multinomial), collected via web scraping from LinkedIn/Indeed.
- Achieved **92% accuracy with the Random Forest** model, leading to a **15% increase** in user satisfaction and engagement. Visualized outcomes through a network graph, showcasing a profound understanding of community dynamics and intricate topic clustering patterns.

SKILLS & CERTIFICATIONS

Programming Languages: Python, R, Java, SQL, C, C++, HTML, CSS, JSON, XML, PHP, D3, JavaScript

Databases: Oracle, MySQL, NoSQL, MongoDB, PostGre, Neo4j

Libraries: Matplotlib, Seaborn, BeautifulSoup, Scrapy, Scikit-learn, NumPy, PyPDF2, Pytorch, Keras, Pandas, Streamlit, Theano, Tensorflow

Others: Microsoft Azure, AWS, Git, Tableau, Microsoft Power BI, Paraview, Unity, Advanced Excel

Certifications: AWS Certified Machine Learning, IBM Python for Data Science, Business Analytics from LinkedIn Learning, Tableau Desktop Advanced.