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*

Coursework: Advanced Database Technologies, Statistics, Elements of AI, Information Visualization, Exploratory Data Analysis

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*

- Designed and led the implementation of a high-throughput ETL data pipeline using Apache Spark and Airflow, which decreased data processing time by **10%** and enabled **15%** faster analytical decision-making, significantly enhancing support for real-time business strategies.
- 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.
- Enhanced network infrastructures for Hamburg Süd and Win365, increasing efficiency by 25% with virtual desktops and earned the 'Dynamic Employee of the Team' award for rapid improvements in project delivery.
- Optimized multiple SQL queries, significantly reducing execution times by **30 minutes**, which enhanced database performance and efficiency.
- Established robust CI/CD pipelines for data applications using Jenkins, streamlining processes, and reducing data inconsistencies by **30%**, thereby enhancing overall data quality and reliability.
- Created custom **Tableau dashboards** to analyze user interactions with **Ultimatix** application, enhancing data-driven decisions and leading to a **35%** increase in stakeholder satisfaction.

Zensar Technologies   **May 2019 – May 2020**

Software Developer Intern   *Pune, Maharashtra, India*

- Specialized in designing and optimizing **MySQL** database schemas, implementing efficient data structures that improved query performance and system responsiveness.
- Maximized quantitative reporting with batch processing pipelines, cutting data processing time by **18%**, and enabling faster data transformations, visualizations, and statistical analysis for improved decision-making.
- Advised on refining the supplier data modeling system, which boosted supplier competitiveness by **15%** and raised conversion rates by **10%**, enhancing overall business performance.

Projects

Virtual Tissue Simulation| *Bloomington, Indiana, USA*   *Aug 2023 - May 2024*

- Implemented a Retrieval-Augmented Generation (RAG) framework, with Large Language Models (LLM) for information retrieval, **achieving 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.

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

- Developed a Virtual Reality (VR) model with Python and R, processing an expanded **135 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 **88% accuracy**, enhancing its academic and research utility by 35%.

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

- 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 and visualized outcomes through a network graph, showcasing a profound understanding of community dynamics and intricate topic clustering patterns.

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 **precision 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 leading to a **25% improvement** in understanding user behavior and a **15% increase** in trend predication accuracy.

SKILLS & CERTIFICATIONS

Programming Languages: Python, R, Java, SQL, C, C++, HTML, CSS, JSON, XML, PHP, D3, JavaScript
Databases: Oracle, MySQL, RDBMS, SQL Server, SSIS, SSRS, NoSQL, MongoDB, PostgreSQL, Neo4j
Libraries: Matplotlib, Seaborn, BeautifulSoup, Scrapy, Scikit-learn, NumPy, PyPDF2, Pytorch, Keras, Pandas, Streamlit, Theano, Tensorflow
Others: Microsoft Azure, AWS, Git, Tableau, LucidChart, Microsoft Power BI, Paraview, Unity, Advanced Microsoft Excel, Powerpoint, Microsoft Office, Visio
Certifications: AWS Certified Machine Learning, IBM Python for Data Science, Business Analytics from LinkedIn Learning, Tableau Desktop Advanced.