Hasaranga Jayathilake

Website Graduate Highlight hdj1414@gmail.com LinkedIn GitHub Google Scholar (317)-334-5713

Education

MS Applied Data Science - Indiana University, Indianapolis **(GPA: 4.0)** **Dec 2024**
Relevant Coursework: Data Analytics, Natural Language & Deep Learning, Cloud Computing, Data Visualization, Database, Statistics

MSc. Logistics and Supply Chain Management (Hons) - Sheffield Hallam University, Sheffield UK **(GPA: 3.55 - Merit Pass)** **Jan 2017**
Relevant Coursework: Enterprise Information System, Lean Operations & Six Sigma, Manufacturing Strategy, Modelling & Simulation

BSc. Logistics Management (Supply Chain Mgt.) (Hons) - Kotelawala Defence University, Sri Lanka **(GPA: 3.89 - First Class)** **Oct 2015**
Relevant Coursework: Remote Sensing & Geographical Information System, Computer Application on Supply, E-Commerce, Mathematics

Experience

Lead Data Analyst, COMET Lab, Indiana University Indianapolis **Aug 2023 - Present**
- Lead a team of 40, including 30 data analysts, 5 UX researchers, & 5 UI designers, across 19 projects: 1 visualization software development, 15 data analysis projects, 1 dashboard development, and 2 UI developments.
- Applied text mining techniques with python (NLTK, spaCy, and scikit-learn) to identify user requirements from user testing researcher transcripts to analyze open-ended survey responses, extracting actionable insights to develop dashboard on dementia patient tracker for FSSA Indiana to
- Automated ETL processes for CMS data using Python Selenium, ensuring accurate and up-to-date data for analysis.
- Performed bigdata and geospatial analysis using Python, R, and SQL, deriving insights from U.S. Census, CDC, and NIH data.
- Created interactive dashboards with Tableau, PowerBI, D3, Dash, and Plotly, and prototyped designs using Figma for client approvals.
- Collaborated with clients like FSSA Indiana and local universities, aligning projects with requirements and implementing feedback.
- Developed a centralized GitHub repository system for team collaboration, including Standard Operating Procedures (SOPs) for project workflows, enhancing on time delivery by 98%, & ensuring 80% improvement in code efficiency to meet customer requirements.

Graduate Adjunct Instructor, Indiana University Indianapolis **Jan 2023 - Present**
- Instruct data visualization on big data and geospatial data using Python, Plotly, Dash, D3.js, and Tableau for INFO-H517: Visualization Design, Analysis, & Evaluation Course.
- Guide students in creating impactful data visualizations, integrating real-world datasets to solve practical problems.
- Teach database design with SQL on MySQL (Maria DB), Oracle and SQL Server for INFO-I308: Information Representation Course.
- Mentor students in developing normalized database structures and efficient SQL queries for real-world applications.

Visiting Lecturer, Human Resource Management Institute (HRMI) for University of Northampton, UK **Jun 2022 - Present**
- Developed & delivered curriculums for supply chain, project management, business analytics, & research methodologies for master's & undergraduate programs offered by the University of Northampton, UK, leading to a 70% increase in student course completion rate.

Lecturer, Human Resource Management Institute (HRMI) **Aug 2021 – Jun 2022**
- Developed & delivered curriculums on procurement, human resources, logistics, HR analytics, operations analytics, & information management for HND, undergraduate, & master's level students.
- Designed & implemented a Moodle-based self-hosted LMS system, reducing academic support staff workload by 50% & increasing overall efficiency by 35%.
- Facilitated workshops & training sessions on business analytics & project management tools, improving students' practical knowledge & readiness for industry roles, resulting in a 75% increase in job placement rates within the first 6 months from the graduation.

Business Analyst, MSS Sri Sai Enterprise **May 2017 - Aug 2021**
- Developed an SQL-based transaction system using MySQL & Microsoft Access, improved procurement accuracy & reduced overvaluation by 70%.
- Applied text mining with Python, NLTK, & Pandas on procurement & feedback data, standardizing tender evaluations & boosting success by 80%.
- Built an in-house timesheet program using Python & Flask, improving employee management & satisfaction by 80% & reducing turnover by 40%.
- Created an in-house Warehouse Management System using JavaScript, Python, Flask, and ReactJS, reducing inventory turnover time by 67%.
- Designed a dashboard with Excel & Python to optimize inventory visibility, reduce warehouse stock holding costs & stock turnover days by 50%.

Demand Planner, Industrial Clothing Private Limited **Jun 2016 - Apr 2017**
- Contributed to a 65% increase in on-time shipments by implementing an SAP ERP system into the shipment and order planning process from East Asian suppliers to North American customers.
- Extracted ERP data using Python & SQL, structured reporting with Power BI & SQL Server, improving production efficiency by 70%.

Operation Development Management Trainee, DP Logistics Private Limited **Jul 2015 - May 2016**
- Designed operational workflows for an in-house ERP solution using Visio, reducing operational costs by 65%.
- Established standardized operational practices (SOP) with Lean principles to optimize logistics & supply chain operations.
- Developed inventory tracking module using SQL & Excel, increasing inventory accuracy by 30% & reducing stock discrepancies.

Volunteer Experience

Tutor Volunteer, Brightlane Learning, Indianapolis, IN   **August 2022 - Present**
- Tutoring Mathematics and English for homeless elementary pupils at Wendell-Phillips 63 and IPS 31 in Indianapolis.

Board Member, Biomedical Informatics Club, Indiana University Indianapolis   **Dec 2022 - Dec 2023**
- Organized events and social media initiatives that drove a 110% increase in membership year-over-year.

Journal Article Publications

- Employee Development and retention of Generation-Z employees in the post-COVID-19 workplace: a conceptual framework - **Benchmarking: An International Journal (2021)**
- Exploring HRM Practices to Retain Generation-Z in Information Technology Sector: A Systematic Literature Review - **International Journal of Business and Technopreneurship (2020)**
- The Impact of Unemployed Youth on National Development in Sri Lanka - **Journal of Emerging Issues in Economics, Finance and Banking (2020)**
- Retention of Generation-Z in Information Communication Technology Sector of Sri Lanka: A Conceptual Paper - **Kelaniya Journal of Human Resource Management (2019)**

Conference

- Gamified Feedback and Job Satisfaction of Generation Z Software Professional in Post-COVID-19 Environment - **International Conference in Data Science, Colombo Sri Lanka (2021)**
- Retention of Generation-Z in ICT Sector in Sri Lanka: A Conceptual Paper - **International Conference on Business and Information Kelaniya, Sri Lanka (2019)**
- Importance of Leadership to Success of Digital Workplace - **International Conference on Business Research, Moratuwa, Sri Lanka (2019)**

Poster Publications

- National Health and Nutrition Examination Survey Data Analysis – **Annual Meeting of The Polis Center, Indianapolis, IN, (2024)**
- Indiana State Life Expectancy Calculation – **Annual Meeting of The Polis Center, Indianapolis, IN, (2024)**
- Centers for Disease Control and Prevention Places Data Analysis – **Annual Meeting of The Polis Center, Indianapolis, IN, (2024)**
- savi.org Outlier Analysis Using R - **Annual Meeting of The Polis Center, Indianapolis, IN, (2024)**

Projects

Enhanced Sentiment Analysis Using BERT Models for Multilingual Applications
- Utilized BERT, PhoBERT for Vietnamese, and mBERT for multilingual tasks, integrating LSTM to enhance sentiment classification.
- Improved sentiment analysis for English and Vietnamese datasets by fine-tuning pre-trained models with domain-specific knowledge.
- Achieved 96.7% accuracy and 93.6% F1 score with the PhoBERT + LSTM model for Vietnamese, significantly outperforming baseline models. Set new benchmarks in multilingual sentiment analysis, achieving an F1 score of 83.2% for English with PhoBERT + LSTM.

Centers for Medicaid and Medicare Services (CMS) Data Download Automation and Analysis
- Developed automation scripts using Python and APIs to efficiently download and process CMS data.
- Streamlined data retrieval from CMS systems, enabling faster and more accurate analysis of healthcare datasets.
- Reduced data download time by 40% and improved data accessibility, facilitating quicker insights for healthcare decision-making.

Text Classification model to identify AI vs Human-Generated content.
- Developed a model combining statistical linguistic features and a neural network to classify AI vs. human-generated text
- Achieved 90.2% accuracy in distinguishing AI-generated content, using GloVe and BERT for feature extraction.
- Awarded 'Best Deep Learning Project' at the IUPUI Datathon 2024.

Text Classification Model with FastText and GloVe Embeddings
- Developed a text classification model using FastText and GloVe embeddings, achieving 92.5% accuracy on sentiment datasets.
- Leveraged GloVe embeddings for word representation, improving feature extraction and model performance.

Asset Limited, Income Constrained, Employed (ALICE) Budget Data Scraping and Analysis
- Utilized Python, BeautifulSoup, and Selenium for web scraping and data extraction from ALICE Budget datasets.
- Automated the collection of budget data to support financial analysis and insights for policy-making decisions.
- Successfully extracted and structured complex budget data, enabling 25% faster analysis and enhanced decision-making capabilities.

Credit Card User Clustering for Market Segmentation using R
- Utilized R to perform clustering analysis on credit card user data, identifying distinct consumer segments for targeted marketing.
- Implemented model-based clustering and PCA, achieving effective segmentation of users based on spending behavior.

Mentoring System Development Database
- Developed a mentoring system to track program progress, interaction levels, and feedback, improving employee engagement by 30%.
- Integrated SQL and Python to create a database for tracking mentor-mentee interactions, ensuring detailed progress monitoring.
- Implemented rewards for high-performing mentors, increasing program effectiveness and motivation by 40%

Alaska Fish Catchment Area Optimization Using Pyspark on Could Computing

- Applied regression analysis on bigdata using PySpark to predict catch per unit effort (CPUE) for different fish types in Alaskan waters.
- Identified high productivity fishing areas, optimizing fishing operations and increasing catch efficiency by 25%
- Built models with PySpark, comparing performance metrics to determine the best predictive approach for CPUE analysis

Indiana Life Expectancy Analysis

- Conducted comprehensive analysis of life expectancy in Indiana using demographic, socioeconomic, and health data using R
- Built predictive models identifying key determinants of life expectancy, resulting in targeted intervention areas.
- Utilized data visualization with Tableau to highlight regional disparities, aiding policymakers in formulating effective health strategies.

Savi user data monitoring Live Data Dashboard Creation

- Combined Google Analytics with Power BI, Looker Studio, and Tableau to create dynamic dashboards, improving user behavior insights
- Built user-friendly, interactive dashboards to visualize key metrics like page views and session durations, enhancing stakeholder decision-making.
- Identified high engagement areas and optimized content strategies, leading to a 20% increase in user retention on the SAVI platform

FSSA Indiana State Dementia Data Dashboard Creation

- Developed dashboards using Tableau to visualize dementia prevalence, Medicaid spending, and caregiver support across Indiana
- Enabled policymakers to assess service gaps and allocate resources effectively, improving community support programs by 30%.
- Created interactive dashboards to track Medicare spending and demographic risk factors, aiding data-driven decision-making.

National Health and Nutrition Examination Survey (NHANES) Data Analysis

- Conducted analysis on NHANES data to understand public health trends and identify risk factors across different population groups.
- Created visual dashboards using R & Python to provide insights on health metrics, improving stakeholders' understanding of community health.
- Supported health initiatives by identifying actionable patterns, leading to a 15% improvement in targeted health interventions.

Social Vulnerability and Disparity Index Analysis

- Employed Python, Tableau, and statistical libraries to analyze and visualize Social Vulnerability Index (SVI), Area Deprivation Index (ADI), and Social Disparity Index (SDI) data.
- Identified and mapped social disparities to support data-driven decisions in resource allocation and community development.
- Delivered actionable insights through comprehensive analysis & visualizations, enhancing understanding of social vulnerability trends in regions.

Skills

Languages:	Python, SQL, R, JavaScript, HTML, CSS, SASS
Platforms:	MS Windows, Linux, Google Colab
Data Processing & Cloud:	ETL Processes, Data Wrangling, Data Pipelines, APIs, CUDA. Selenium, BeautifulSoup, Hadoop, Spark, Azure, AWS,
Data Management Tools:	SQL Server, Oracle, MySQL, MongoDB, MS-Office, XML, JSON, Jira, SPSS, MATLAB, Snowflake, Git, PHP
Data Visualization:	PowerBI, Tableau, D3.js, Google Looker Studio, Plotly, Matplotlib, Render, Dash, Excel, PowerPoint, Visio, ArcGIS, Figma
AI/ML/NLP Frameworks:	Pyspark, PyTorch, NLTK, spaCy, NumPy, Pandas, Scikit-Learn, TensorFlow, Predictive Modeling, Text Mining
Analytics:	Regression Analysis, Sentiment Analysis, Clustering, Time Series, A/B Testing, Hypothesis Testing, Geospatial Data Analysis
SCM & Logistics Tools:	SAP (S/4HANA, SCM, EWM, Ariba), Oracle ERP, MS Dynamics 365, Infor Nexus, Manhattan TMS, Coupa, Blue Yonder, O9
Soft Skills:	Problem-Solving, Project Management, Team Collaboration, Scrum, Agile, Communication, Leadership, Documentation

Certifications

- **Data Analytics Professional Certificate** – Google
- **IT Support Professional Certificate** – Google
- **Google Workspace Certificate** – Google
- **Text Mining and Analytics** – University of Illinois at Urbana-Champaign
- **Programming Foundations with JavaScript, HTML and CSS** – Duke University
- **SASS Certification** Course – Google Programming Hub
- **Managing Big Data with MySQL** – Duke University
- **Mathematics for Machine Learning: Linear Algebra** – Imperial College London
- **Customer-Centric IT Strategy** – University of Virginia
- **IT Project Management** – Indian School of Business, Hyderabad, India
- **Excel Skills for Business Specialization Certificate** – Macquarie University, Sydney, Australia
- **Agile Development Specialization and Digital Product Management Specialization Certificates** – University of Virginia
- **Six Sigma Yellow and Green Belt Specialization Certificates** – Kennesaw State University
- **Supply Chain Analytics** – Rutgers University, New Jersey
- **Marketing Analytics** – University of Virginia
- **Human Resources Analytics** - –University of California, Irvine
- **Business Analytics for Decision Making** – University of Colorado Boulder
- **Communicating Business Analytics Results** – University of Colorado Boulder