KRITI SHUKLA

kas140472.github.io/myWeb | linkedin.com/in/kriti-shukla-258396182/

Education

University of Southern California, Los Angeles, CA Dec 2024
Master of Science in Computer Science *GPA: 3.82/4.0*
Relevant Coursework: Web Technologies, Database Systems, Information Retrieval and Web Search Engines, Scientific Computing, Operating Systems, Computer Networks, Algorithms Analysis, Machine Learning
***Awards**: Society of Women Engineers (SWE) Los Angeles Graduate Scholarship Winner*

National Institute of Technology Karnataka, India May 2022
Bachelor of Engineering, Mechanical Engineering (Minor in Computer Science) *GPA: 9.65/10.0*
***Honors**: Institute Gold Medals (outstanding academic performance), Best Outgoing Student Award*

Experience

Data Science Intern | ***Ford Motor Company, Dearborn, MI*** May 2024 - Aug 2024
- Built and optimized data pipelines for handling big data on Google Cloud Platform (GCP) using BigQuery SQL and Dataform, reducing query runtime by **50%**, while contributing to Ford's Electric Vehicle (EV) Charging Taskforce.
- Expanded analytics capabilities to a global scale by analyzing Germany's EV charging data, extracting critical insights to enhance the public charging experience.
- Conducted topic-wise sentiment analysis on charging station reviews, explored advanced Aspect-Based Sentiment Analysis (ABSA) techniques, and improved the user sentiment dashboard.
- Led the annotation of user reviews for model training data, using Python to calculate agreement scores; facilitated consensus-building meetings, improving consistency by **12%**.
- Improved data accuracy and integrity by resolving discrepancies in data pipeline, ensuring high-quality, real-time insights for executive decision-making. Applied data transformation techniques to automate data collection and processing.
- **Tech Stack**: Python, SQL, Google Cloud Platform (GCP) BigQuery, Machine Learning/NLP, Agile Methodologies

Software Engineer Intern | ***Information Sciences Institute (ISI), Los Angeles, CA*** Jan 2024 - Apr 2024
- Built front-end components using **React** for FaceBase Dataset platform (data source for craniofacial research)

Data Science Intern | ***LISUS Energy, Los Angeles, CA*** Jun 2023 - Jul 2023
- Analyzed remote-sensing satellite data to develop data-driven tool for sustainable exploration of critical minerals.
- **Tech Stack**: Python, Pandas, NumPy, Matplotlib

Software Engineer Intern | ***Caterpillar Inc., India*** May 2021 - Jul 2021
- Worked in a team of 4 to develop an Ignition Delay Prediction Model for compression ignition engines, improving engine failure analysis and diagnostics.
- Developed data management solutions to efficiently handle and analyze large volumes of engine test data using Pandas, NumPy, and Matplotlib, improving data processing accuracy and enabling data-driven decision-making.
- Identified and resolved discrepancies in the Ignition Delay Prediction within the company's proprietary software, enhancing the reliability of engine performance insights.
- Streamlined code for over **50%** reduction in runtime, optimizing the performance of the engine analysis software and improving overall program efficiency.
- **Tech Stack**: Python, SQL, Cantera

Software Engineer Intern (Team Lead) | ***Virtual Labs, NIT Karnataka, India*** Nov 2020 - Jan 2021
- Led team of three in the design and development of 12 interactive simulations for Thermodynamics.
- Applied Agile methodologies and software best practices to deliver high-quality simulations on time.
- Developed online tools that facilitated remote learning and improved accessibility to engineering concepts, fostering a hands-on approach to learning through visual animations and dynamic content.
- **Tech Stack**: HTML, CSS, JavaScript

AP Computer Science Teaching Assistant for high schools | ***Microsoft TEALS, CA*** Sep 2023 - Present

Projects

Data Science/ML/AI/Analytics

Recommender Systems
- Built a content-based movie and fanfiction recommender system using NLP techniques like CountVectorizer and cosine similarity for text vectorization and similarity computation.
- Preprocessed data by cleaning, handling missing values, removing duplicates, and engineering features.
- Developed an interactive web app with Streamlit.
- **Technologies**: Natural Language Processing (NLP), Streamlit, Scikit-learn, NLTK, Data Engineering

Single Track Deposit Geometry Prediction Model

- Led a team of four in developing machine learning models with 98% accuracy to predict single-track deposit geometry for Laser Directed Energy Deposition, enhancing additive manufacturing quality.
- Applied data engineering techniques including data pre-processing (randomization and normalization), hyper-parameter optimization, and feature importance analysis to improve model performance.
- **Technologies**: Machine Learning, Data Preprocessing, Feature Engineering, Hyperparameter Optimization, Neural Networks

Multi-Class and Multi-Label Classification & Clustering Analysis

- Implemented multi-class and multi-label classification on the Anuran Calls dataset using SVMs with Gaussian and L1-penalized kernels, applying SMOTE for class imbalance and evaluating performance with precision, recall, AUC.
- Conducted K-means clustering to identify clusters and evaluated clustering accuracy using Hamming distances and scores between true labels and cluster assignments.
- **Technologies**: SVM, SMOTE, K-Means clustering

Class Imbalance and Model Evaluation

- Applied data imputation and coefficient of variation (CV) for feature selection on the APS Failure dataset, trained and evaluated Random Forest and XGBoost models with and without SMOTE for class imbalance, and analyzed performance metrics including confusion matrices, ROC, AUC, and cross-validation.
- **Technologies**: Random Forest, XGBoost, Data Imputation, Co-efficient of variation

Decision Trees and Regression Models Analysis

- Built and pruned decision trees for the Acute Inflammations dataset, converted decision rules into IF-THEN rules, and performed data imputation, correlation analysis, and feature selection; fitted and evaluated linear, ridge, LASSO, PCR, and boosting regression models with cross-validation.
- **Technologies**: Decision trees, Regression models, Correlation analysis, Feature selection

Full-stack/App/Web Development

Stock Trading Web and Android App - Virtual Trading & Market Insights

- Responsive cloud-hosted stock trading web application and Android app featuring real-time stock information through Finnhub APIs, virtual trading, portfolio management, and watchlist functionality.
- Enabled users to search stock symbols, view detailed stock information (company news, historical and hourly charts, recommendation trends), and manage investments with interactive features across web and mobile platforms.
- **Technologies**: Angular, AJAX, Bootstrap, Android Studio, XML, Java, Node.js, Google App Engine, MongoDB

Game-based Learning Web App

- Anime-themed interactive game-based learning web app; integrates puzzles designed to boost coding skills and problem-solving, with leaderboard and users earning rewards to progress through levels.
- **Technologies**: HTML, CSS, JavaScript, Python, Flask

Networks & Operating System

Optimized Meeting Scheduling Application using C/C++ and Socket Programming

- Meeting scheduling application on Unix-based system, employing TCP and UDP protocols to establish efficient communication between servers. Identifies optimal meeting times based on participant availability.
- **Technologies**: C/C++, Linux, Internet Protocol Suite (TCP/IP)

Token bucket emulation in C using Multi-threading

- Emulated traffic shaper: transmits packets controlled by a token filter using multi-threading within a single process.

Skills & Technologies

Programming Languages: Python, Java, JavaScript, TypeScript, R, MATLAB, SQL, C, C++, Bash, PowerShell;
Data Science and Analytics: Pandas, NumPy, Scikit-learn, Tableau, Power BI, Matplotlib, Seaborn, BigQuery, Advanced Data Analysis, Statistical Modeling, Predictive Modeling, Advanced Excel (Pivot Tables), Databricks;
Machine Learning/NLP: TensorFlow, Keras, Natural Language Processing (NLP), NLTK, Azure ML, PyTorch;
Data Management: Big Data, ETL (Extract, Transform, Load), Data Pipelines, Data Engineering;
Databases: SQL, RDBMS, MongoDB, NoSQL; **Version Control**: Git, GitHub, Bitbucket;
Cloud Platforms & Containerization: Azure, AWS, Google Cloud Platform, Google App Engine, Docker;
Front-End Development: HTML, CSS, XML, React, Angular, Bootstrap, jQuery;
Back-End Development: Node.js, Express.js, Flask, RESTful APIs, AJAX, JSON;
Mobile Development: Android Studio, Kotlin, iOS;

Certificates: Google Cloud Big Data and Machine Learning Fundamentals (Jun 2024), IBM Data Science Professional Certificate (In Progress), Microsoft Azure AI Fundamentals (AI-900) | Dec 2024

Leadership and volunteer experience

- **USC Viterbi Graduate Orientation Leader** | Jun 2023 - Dec 2024 | Led orientation for new students, delivered presentations, organized campus-wide events, provided training on inclusive practices.
- **Marketing Director & Webmaster, USC Women in Engineering** | Mar 2023 - Dec 2024 | Led event promotion and marketing efforts, created content for social media platforms, developed and managed the website. Liaised between WIE and the Graduate Student Association, organizing joint events.