SHRISTI SHRESTHA

Email: shristi.stha33@gmail.com | Linkedin: https://www.linkedin.com/in/shristi-stha/

EDUCATION

University of Memphis (UofM), Memphis, TN, USA   Expected Grad. Dec 2025
Master of Science in Data Science - 4.0 GPA
Relevant Courses: Advanced Statistical Learning I, Advanced Statistical Learning II, Fundamentals of Data Science, Machine Learning, Advanced Database Systems, Biostatistical Methods I, Biostatistical Methods II, Image Processing

Louisiana Tech University (LaTech), Ruston, LA, USA   Mar 2018 - Mar 2020
Master of Science in Computer Science - 4.0 GPA
Relevant Courses: Database Management Systems, Data Mining and Knowledge Discovery, Distributed and Cloud Computing

PROFESSIONAL EXPERIENCE

IBM, Chicago, IL, USA (Remote)   Oct 2021 - Apr 2023
Software Quality Assurance Engineer
- Worked in IBM Cloud Object Storage (COS) System Integration team to develop Python Automation test scripts for Cloud Storage solutions
- Majorly involved in COS replication to develop automation scripts for Object Replication using IBM COS Python SDK
- Performed Stress & Stability testing using IBM Object Generator (OG) scripts for the developed feature and analyzed the system behavior in Kibana and Grafana dashboards
- Set up systems for Upgrade/Downgrade testing for Object Replication and Context-Based Restriction (CBR) requests
- Collaborated with cross-functional teams to investigate software defects across the IBM COS application
- Proactively adapted to shifting priorities and effectively managed multiple competing projects, consistently delivering high-quality results on time
- Involved in peer code review of Github PRs and Confluence Wiki documentation of application features
- Actively engaged in Agile practices such as Stand-ups, Sprint Planning, Backlog Grooming, and Sprint retrospectives to ensure quality and timely Sprint deliverables

Yes Energy, Boulder, CO, USA (Remote)   Nov 2020 - Oct 2021
Software Quality Assurance Engineer
- Worked in an evolving power market domain that involved highly complex and intricate business logic to deliver robust and high-quality energy data to customers in real time
- As a member of the Big Data team, designed and executed PL/SQL Oracle and Snowflake scripts to verify complex backend business logic, and ensure data accuracy in the Data Cloud and AWS Data Lake respectively
- Designed & developed Python automation scripts using a multiprocessing library to simulate multiple concurrent requests for testing system performance and subscription-tier throttling limits
- Created Tableau dashboards to report Power Grid data for generating insights into Power Trading
- Collaborated in full SDLC from collecting business requirements, analyzing user stories, designing and documenting test cases for use-case scenarios along with test results
- Onboarded and provided technical guidance to new QA hires along with constructive feedback to the team members
- Praised for excellent work attitude and attention to detail by fellow team members in a numerous Sprint retrospectives

Gerer InfoTech Pvt. Ltd, Nepal   Jan 2015 - July 2017
Software Engineer
- Analyzed business requirements and user stories to design and execute manual and automated test cases, ensuring the reliability and accuracy of enterprise-level data-driven purchase and pay financial web application
- Developed SQL queries for data verification for back-end testing using Java Database Connectivity (JDBC)
- Worked together with the DevOps team to schedule cron jobs that automated test suite execution using CI/CD Jenkins

Deerwalk Inc., Nepal   Mar 2014 - May 2014
Associate Software Engineer
- Worked in a Big Data healthcare analytics company to build Java REST API services to query Elasticsearch indexes, enabling efficient retrieval and analysis of large healthcare datasets
- Got familiarized with the US healthcare domain, including HIPAA compliance, to handle sensitive personal health data

ACADEMIC EXPERIENCE

Louisiana Tech University, Ruston, LA, USA   Jan 2020 - Mar 2020
Graduate Teaching Assistant
- Assisted in grading graduate & undergraduate-level Computer Science courses under the supervision of Dr. Ben Choi
- Developed Python test cases to automatically evaluate the student's Python scripts and grade the test scores

Louisiana Tech University, Ruston, LA, USA   Mar 2018 - Sep 2019
Graduate Research Assistant
- Designed and developed a prototype for Smart Cities framework using IoT devices and Cloud computing that encompasses all the layers of IoT sensor nodes, Fog Server, Cloud, and Web dashboard
- Built a data analytics pipeline on AWS to process, analyze, and visualize IoT-collected data, enabling efficient data processing and management and ensuring actionable insights for decision-making from IoT data
- Technologies & Tools used: AWS (IoT Core, IAM, DynamoDB, EC2, Lambda, EMR, CloudWatch, S3, SNS), MySQL, HTML, CSS, JSON, Rest APIs, Leaflet, Linux, C++, MQTT, Waspmote IDE

RELEVANT PROJECTS
House Price Prediction on Real Estate Valuation Dataset using Linear Regression Models, UofM   Sep 2024 - Dec 2024
- Built predictive models using linear regression models along with regularization techniques for predicting house prices
- Performed data preprocessing and exploratory data analysis (EDA), including summary statistics, scatter plots and correlation analysis to understand variable relationships and identify potential outliers
- Implemented linear regression models including full, reduced, and subset selection models, polynomial regression with interaction terms, and regularization techniques including Ridge, Lasso, and Elastic Net
- Technologies & Tools used: R, R Studio

Predictive Policing in Memphis using MPD Public Safety Incidents Data, UofM   Sep 2024 - Dec 2024
- Applied Data Science approaches to derive insights from Memphis Police Department (MPD) Public Safety Dataset for strategic deployment of law enforcement resources
- Executed data science lifecycle, encompassing data mining, data cleaning, EDA, feature engineering, model development, performance evaluation, and reporting
- Cleaned and preprocessed a large-scale dataset of over 640,000 crime incident records, addressing missing values, inconsistencies, data entry errors, and duplicates to ensure data quality and reliability for analysis and modeling
- Developed Decision Tree and Random Forest models to predict crime severity
- Technologies & Tools used: Python, Pandas, sklearn, NumPy, Matplotlib, Jupyter Notebook

Exploring ML and DL Techniques for Named Entity Recognition (NER), UofM   May 2024 - Aug 2024
- Evaluated and compared the performance of different models, including Conditional Random Fields (CRF), Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), and a hybrid BiLSTM-CRF, in identifying and classifying named entities from unstructured text data using Groningen Meaning Bank corpus
- Technologies & Tools used: Python, Pandas, sklearn, Tensorflow, Keras, K-fold cross validation, Jupyter Notebook

Investigating Wildfire Severity using Data Mining and Classification Models, LaTech   Sep 2019 - Nov 2019
- Handled large and complex datasets, collected from multiple data sources, to investigate the independent variables causing the spread of wildfire and classify their severity in terms of wildfire size
- Constructed Machine Learning Classification Models such as Decision Tree, Random Forest and SVM
- Technologies & Tools used: Python, Pandas, sklearn, imblearn, SMOTE, Jupyter Notebook, RapidMiner

Nursing Home Database Management, SIUE   Aug 2017 – Dec 2017
- Developed a Spring Boot project using MVC framework for performing CRUD operation on a MySQL Database
- Formulated complex MySQL queries including joins, subqueries, aggregates, and control flow functions
- Technologies & Tools used: Spring Boot, JSP, Java, JDBC, MySQL, UML, Maven

RELEVANT SKILLS
Languages:	Python, R, SQL, Bash, Java, SAS
Databases/Warehouse:	MSSQL, MySQL, Oracle, Snowflake
Methodologies/Concepts:	Statistical Analysis, Hypothesis Testing, Regression Analysis, Machine and Deep Learning
Build Integration Tools:	Git, Bitbucket, Maven, Gradle, Jenkins
Others:	RStudio, VS Code, PyCharm, Jupyter, Toad, Kibana, Grafana, Tableau, Excel, Matplotlib, JIRA

CERTIFICATIONS
IBM Developer Jumpstart - Practitioner | IBM Enterprise Design Thinking Practitioner Course | IBM Agile Explorer | IBM Certified Advocate Cloud v2

HONORS & ACTIVITIES
TL James Endowed Scholarship, LaTech (2018-19) | Graduate Scholar Award, SIUE (2017)

Graduate Mentor & Presenter at Google CREWE workshop on topic on Smart Cities Framework, Bossier City, LA (2019)