SHREY SHAH

(447) 902-1978 | sshah023@illinois.edu | linkedin.com/in/shreysshah | github.com/shreyss99

EDUCATION

University of Illinois at Urbana-Champaign, USA Aug 2022 – May 2024
Masters of Science in Information Management (Data Science and Analytics) **GPA – 4.0/4.0**

Sardar Patel Institute of Technology, India Jul 2015 – May 2019
Bachelor of Engineering in Computer Engineering **GPA – 3.9/4.0**

TECHNICAL SKILLS

Languages/Data Formats	Python, Java, R, SQL, HTML/CSS, Ruby, Shell Scripting, C++
ML/Analytics	PyTorch, Keras, Tableau, PowerBI, Jupyter, Power Query, Talend
Databases/Frameworks	MySQL, PostgreSQL, Oracle, MongoDB, Neo4j, Spring Boot, Django
Tools/Platforms	IDE, Postman, Git, Chef, AutoSys, Kibana, AppDynamics
Others	Jenkins, OpenShift, Kafka, JIRA, Confluence, AWS, Jinja, Agile/Scrum

WORK EXPERIENCE

University of Illinois at Urbana-Champaign Mar 2023 – Present
Center of Innovation in Teaching and Learning – Data Analyst Champaign, USA
- Modeled a solution to redact PII and flag inappropriate responses with 20% accuracy using Python NLTK
- Created accessible HTML reports to depict college survey findings via R markdown automation
- Constructed dashboards to illustrate climate survey results and demographic summaries using Tableau
- Formulating custom weighting function for survey response statistics using R and Power BI M functions

Yahoo May 2023 – Aug 2023
Software Engineering Intern Champaign, USA
- Built a CLI to automate Amazon EMR cluster deployment, decreasing manual effort by **60%** using Python
- Boosted operational efficiency and scalability by **35%** by leveraging AWS services and boto3 SDK
- Designed a template immersing users in editor mode for EMR cluster configuration to their choices via Jinja
- Incorporated automated Git repository, branch and pull request creation cutting setup time by **40%** using CI/CD
- Streamlined verification process with lucid feedback to troubleshoot and replace 10+ minutes to **30 seconds**
- Engineered a recipe for dynamic installation of automation agent on **50000+** Grid PE hosts using Chef

Barclays Global Service Center Jul 2019 – Jul 2022
Software Engineer Pune, India
- Worked on OpenShift PaaS v3 to v4 migration for payment search microservice to reduce pod startup time by **30%**, add security vulnerability scanning, and create v4 compatible jobs
- Acted as SME for payment search REST APIs and increased code coverage to **80%** benchmark using Spring Boot
- Migrated from monolithic to event-driven microservices architecture with a **10x** performance increase using Kafka
- Enhanced SQL queries for non-functional test performance to reduce fetch time by **25%** using Oracle Explain Plan
- Virtualized physical workloads to improve resiliency and horizontal scaling with configuration management via Chef
- Automated CI/CD pipelines for **1-click** deployment in multiple environments using parameterized Jenkins jobs
- Collaborated with Level 2 team for successful deployment to production environment with live issue remediation

Software Engineering Intern May 2018 – Jul 2018
- Developed test scenarios for trade booking models using Cucumber, JUnit, and Behaviour Driven Development
- Mapped 9 message types, including fields and interconversion formats, for COBOL based TradeRT utility using Java

PROJECTS and PUBLICATIONS

Fashion Apparel Classification System Jul 2023 – Present
- Examined existing datasets and selected one consisting of over **3000+** coloured images featuring fashion apparel
- Performed image augmentation and preprocessing to enhance data quality, quantity and generalizability using PyTorch
- Obtained a current accuracy of **68%** using the ResNet CNN classification model with computational efficiency

Forbes & NYSE Company Data Analysis, Exploration, Visualization Apr 2023 – May 2023
- Harmonised 4 datasets – NYSE, Forbes companies and prices with Python and Talend ETL into **1M+** rows dataset
- Built intuitive dashboards – YoY Profit change, stock trends, share volatility, geographical spread via Tableau

Online News Popularity Analysis and Prediction Oct 2022 – Dec 2022
- Analyzed factors like domain, sentiment, shares that affect popularity of online news using statistical techniques
- Reduced feature set from **61** to **5** with data explanation of **70%** by applying PCA algorithm in R
- Classified popular and unpopular articles with a **72%** accuracy using the XGBoost model

Wafer Sensor Fault Detection Aug 2021 – Nov 2021
- Validated training data against schema and used KMeans clustering for customised ML approach using sklearn
- Trained XGBoost and RandomForest models using GridSearchCV tuning and chose model based on AUC score
- Achieved an overall accuracy of **82%** on the prediction dataset and deployed the application on Heroku

Wireless IoT based Solution for Women Safety in Rural Areas IEEE 2019
- *ICCES*, Coimbatore, India, 2019, pp. 232-237, doi: 10.1109/ICCES45898.2019.9002392