SHREY SHAH

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

EDUCATION

University of Illinois at Urbana-Champaign   Aug 2022 – May 2024
Master of Science in Information Management (Data Science and Analytics)   **GPA – 4.0/4.0**
Courses: Statistical Models, Data Visualization, Data Warehousing & Business Intelligence, Database Administration, Data Mining

Sardar Patel Institute of Technology, University of Mumbai   Jul 2015 – May 2019
Bachelor of Engineering in Computer Engineering   **GPA – 3.9/4.0**
Courses: Object Oriented Programming, Data Structures & Algorithms, Database Systems, Operating Systems, Big Data Analytics

SKILLS

Programming Languages/Web	Python, Java, R, SQL , HTML/CSS, Ruby, Shell Scripting, JSON, XML, REST
ML/Data Analytics	NLTK, TensorFlow, Keras, OpenCV, Excel, Tableau, Power BI, Jupyter, Talend
Databases/Frameworks	MySQL, PostgreSQL, Oracle, MongoDB, Neo4j, Spring Boot, Django
Tools/Platforms	IDE, Postman, Git, Chef, AutoSys, ELK, AppDynamics, Agile/Scrum
DevOps/Cloud/Other	Jenkins, OpenShift, Linux, Docker, Kafka, JIRA, Confluence, AWS, Jinja, Prisma

WORK EXPERIENCE

Yahoo | *Software Engineering Intern* | Champaign, USA   May 2023 – Aug 2023
- Built a CLI to automate **AWS** EMR cluster deployment, decreasing manual effort by **70%** using **Python**
- Enhanced operational efficiency and scalability by **35%** using **AWS** services, **boto3** SDK, and Infrastructure as Code
- Incorporated automated GitHub repository, branch and pull request creation, cutting setup time by **40%** using CI/CD
- Refined verification flow with intuitive feedback replacing 10+ manual minute process to **30 seconds**
- Engineered a recipe for dynamic installation of automation agent on **50000+** Grid PE hosts using **Chef** and **Ruby**

UIUC | *Data Analyst* | Champaign, USA   Mar 2023 – Nov 2023
- Modeled a first-pass solution to redact PII and flag sensitive responses with 50% accuracy using **Python NLTK**
- Created accessible HTML reports to depict college survey findings via **R** markdown and IBM SPSS
- Constructed dashboards about climate survey summaries to increase customer engagement by **35%** using **Power BI**

Barclays Global Service Center | *Software Engineer* | Pune, India   Jul 2019 – Jul 2022
- Upgraded payment search service to **OpenShift** PaaS v4, reducing pod startup time by **30%** with **Prisma** vulnerability scans
- Acted as SME for payment search REST APIs and increased code coverage to **80%** benchmark using **Java**, **Spring Boot**
- Developed **5–6** user stories for payment validation with **AutoSys** job scheduling and alert monitoring via **ELK** stack
- Migrated from monolithic to event-driven microservices architecture with a **10x** performance increase using **Kafka**
- Optimized **SQL** queries to improve non-functional test performance by **25%** using **Oracle** Explain Plan and **AppDynamics**
- Virtualized physical workloads to improve resiliency and horizontal scaling with configuration management via **Chef**
- Automated CI/CD pipelines to streamline software builds and **1–click** deployment using **Jenkins** and **bash** scripts
- Collaborated with Level 2 team for successful release management to production environment with live issue remediation
- Contributed to team growth by conducting knowledge transfer sessions with presentations for Chef technology

Barclays Global Service Center | *Software Engineering Intern* | Pune, India   May 2018 – Jul 2018
- Developed test scenarios for trade booking models using Cucumber, JUnit framework and **Test Driven Development**
- Mapped 9 message types, including fields and interconversion formats, for COBOL based TradeRT utility using **Java**
- Demonstrated proficiency in financial terminology, post-trade procedures, bank settlement and payment processes

PROJECTS and PUBLICATIONS

Leetcode Problems Search Application
- Created a web application for retrieval of Leetcode problems based on user choices via **Django** with a **200ms** response time
- Implemented secure CRUD operations with access control mechanism via database roles in **AWS Aurora** database
- Transformed database management for **99.9%** uptime, rapid recovery, consistent backups, and reduced annual cost of **$10,000**

Amazon Reviews Product Categorization
- Conducted text preprocessing using **Python NLTK** for multilingual product category classification on customer reviews
- Employed Count Vectorizer with chi-square thresholding and word2vec embeddings to engineer the top **5000** features
- Achieved mean F1 scores of **0.56** and **0.78** for **31** classes in English and German with traditional and neural models, respectively

Fashion Apparel Classification
- Performed image augmentation and data preprocessing over **44000+** coloured fashion apparel images using **OpenCV**
- Obtained an F1 score of **0.88** using the pre-trained ResNet50 classification model with computational efficiency in **TensorFlow**

Forbes & NYSE Companies Analysis
- Harmonised 4 datasets with 44 features and 150 companies using **Python** and **Talend ETL** into **1M+** instances
- Built **a** dashboard with **8** interactive graphics – YoY profit change, stock trends, share volatility, geographical spread using **Tableau**

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