Keshav Shah

Houston, TX | keshav@rice.edu | https://www.linkedin.com/in/keshavkshah/ | 281-644-9320

EDUCATION

Rice University — May 2026
Bachelor of Science in Computer Science, Minor in Data Science and Statistics — *GPA: 3.84*
Relevant Courses: Low Level Systems, Advanced Data Structures and Algorithms, Object Oriented Program Design, Concurrent Program Design, Discrete Math, Database Systems, Distributed Computing, Reinforcement Learning, Multivariable Calculus, Linear Algebra, Linear Regression, Optimization

TECHNICAL SKILLS

Languages: Python, Java, Scala, MySQL, JavaScript, Typescript, HTML, CSS, Go, C, C++, R
Frameworks: React, Node.js, Next.js, Material-UI, Flask, JUnit, FastAPI, REST API, Tensorflow, Maven, NumPy
Developer Tools: Git, AWS, Docker, VS Code, Visual Studio, PyCharm, IntelliJ, Eclipse
Big Data: Apache Spark, Apache Hadoop, Apache Hive, Apache Kafka

EXPERIENCE

Data Engineering Intern — May 2024 – Aug. 2024
Visa — *Austin, TX*
- Implemented Spark ETL pipeline in Scala to process 300 million daily transactions from Kafka, storing data in Hadoop and aggregating to create a report comparing offline and real-time models that detect enumeration attacks
- Automated overlap report generation, reducing overall process time by 88% and ensuring accurate data loading into MySQL and Hive in the UAT environment, with successful pen and integration testing and production deployment
- Designed attack details table in Hive and MySQL to track over 536,000 merchants and 2.98 million daily transactions flagged for potential fraud by the real-time model, using Spark coalesce to enhance query performance

Software Developer — June 2023 – May 2024
RiceApps — *Houston, TX*
- Developed SMS capabilities, interactive map, and vendor directory for 117,000+ Houston Ballet Market attendees
- Reduced time to locate vendors by 75% by utilizing Next.js to create a landing page for attendees to submit contact information, Twilio API to send text notifications, and MappedIn SDK to create deep-linked map of the market
- Increased customer satisfaction ratings by 20% and received positive feedback from Houston Ballet supervisors

Machine Learning Intern — May 2023 – Sep. 2023
University of Houston HULA Lab — *Houston, TX*
- Reduced anxiety symptoms by 20% in patients after 10 weeks by creating a personalized mental health chatbot
- Improved response accuracy by 15% through leveraging GPT-3.5 LLM, distributed computing, developing React UI widgets for user navigation, and REST API endpoints connected to a MySQL database for context storage
- Achieved 92% accuracy in chatbot performance by spearheading a regression test phase through 50 case studies

Software Engineering Fellow — Jan. 2023 – Aug. 2023
Uber Career Prep — *San Francisco, CA*
- Selected among top 4% of candidates to improve technical interview performance and algorithmic thinking through bi-monthly mock interviews and 1x1 mentorship with an Uber Software Engineer

Software Engineering Intern — June 2022 – Aug. 2022
JP Morgan Chase and Co — *Houston, TX*
- Enhanced data retrieval efficiency on 1 million bank transactions by 25% by extracting cluster metrics from a Cassandra NoSQL database using the CQL to derive cluster size, partitioning data, and projected growth rates
- Boosted user interaction efficiency with transaction data by 30% by developing 10+ UI components with React
- Achieved a 98% accuracy rating in data metric aggregation after presenting UI to 15+ JPMC project managers

PROJECTS

MegaDocs | *Python, Apache Spark, AWS, NumPy* — May 2024 – July 2024
- Designed a scalable ETL pipeline using Apache Spark on AWS EC2 to process and classify over 19,000 documents
- Constructed a dictionary of the 20,000 most frequent words using Spark RDDs for feature extraction and converted documents into TF-IDF vectors to enhance model input quality
- Developed a kNN classifier using Python and Spark, applying L2 normalization for precise document classification
- Increased classification efficiency by through feature engineering and distributed computing by leveraging Spark