

Data Engineer | Software Engineer | Automation Engineer | Product Owner

Zeren Yan

"I am the one to make the impossible possible"

(669)278-9865

zyan30@asu.edu

<https://linkedin.com/in/zerenyan>

<https://github.com/yaninsanity>

Self-driven technical professional with experience in developing predictive models on large-scale datasets to address strategic business problems through leveraging advanced statistical modeling, machine learning, and data mining techniques with automatic self-healing solutions.

PROFESSIONAL EXPERIENCE

Software Developer Engineer (Data Engineer)

LexisNexis | Chandler, Arizona | February 2022 – Now

- Spearheaded the design, development, and optimization of a cloud-based, scalable CI/CD event driven automated data pipeline using AWS services (ECS/EMR/Step-Function/Redshift/ECR/DynamoDB/Lambda Function/CloudWatch/Glue/CloudFormation), leading to streamlined data ingestion and enhanced performance.
- Upheld data quality and integrity through the implementation of rigorous validation checks and model validation mechanisms, leading to a reduction in data errors and discrepancies.
- Collaborated with cross-functional teams to develop machine learning models, automated ETL processes, and data integration solutions, supporting the organization's business intelligence and analytics needs.

Full Stack Software Engineer

Find your influence | Scottsdale, Arizona | January 2020 – January 2022

- Cooperated with UI/UX designers to implement business logic from design to production using Angular, ASP.NET, and T-SQL.
- Leveraged Apache Airflow to collect social data feeds from platform Key Opinion Leaders (KOLs), scheduling scripts and persisting data in the database.
- Designed a weekly KOL token status alert system via AWS Simple Email Service.
- Managed Zendesk support tickets, improved ad-hoc queries, and unit testing via Postman and CURL and designed a linear programming model to maximize KOL influence within clients' budgets.
- Created a web crawler using web request technology and built an hourly KOL influence time-series tracking system, guiding digital marketing strategies and increasing revenue.

Research Assistant

Arizona State University | Tempe, Arizona | August 2019 – May 2020

- How to Create an Effective Crowdfunding Campaign: developed text processing pipelines and topic model to detect themes in Kickstarter project and campaigns examined their impact on project success.
- How Informative YouTube Healthcare Related Videos are: devised a deep learning based medical entity recognition model with Keras and TensorFlow to extract medical terminologies from YouTube videos.
- Course Evaluation Analysis: evaluated and re-architected 10 years of ASU course evaluation data and designed a MySQL DB to reduce redundancy.

EDUCATION

Engineering Management (MSEM)

Trine University | GPA 3.8 | May 2022

Business Analytics (MSBA)

Arizona State University | GPA 3.50 | May 2020

Computer Information System (CIS)

Arizona State University | GPA 3.65 | May 2019

TECHNICAL SKILLS

Programming Languages: Python | Java | C# | R | C++ | SQL | HTML | CSS |

Tools: MySQL | SQLite | Hadoop | Jupyter Notebooks | Tableau | Excel Solver | @Risk Analytics | SAS Enterprise | Palisades Decision Suite | SPSS | FileZilla | GIT

Library: Pandas | Sklearn | Seaborn | NumPy | Surprise | Guided-LDA | Keras | Tensorflow | bs4 | Matplotlib | itchat | Pillow | NLTK | SpaCy | TextBlob | StanfordNLP | LightGBM

Analysis Model: Predictive analysis | Cost Analysis | Decision analysis | Financial Analysis | Time-series Analysis |

Project Driving Capabilities: Waterfall | Scrum | Automation | Six Sigma |

PRODUCT INITIATIVES

<https://www.jl-blog.com> | Product Owner

- Built a personal website using WordPress, emphasizing user experience and performance optimization.
- Containerized server infrastructure with Docker to streamline deployment and automate tasks through CI/CD pipelines.
- Implemented 1Panel as a DevOps monitoring dashboard to bolster cybersecurity defenses and system visibility and Affine for DMS
- Shared knowledge and best practices via open-source prototype projects, promoting community engagement and collaboration.

OPEN-SOURCE PROJECTS

<https://pypi.org/project/django-user-starter>

<https://github.com/yaninsanity/TubeWhale>

https://github.com/yaninsanity/Sutra_QAS