

KAJOL TANESH SHAH

Email: kajolshahsm@gmail.com | Ph: +1 872-281-2046 | Chicago, IL - 60608 | LinkedIn: <https://www.linkedin.com/in/kajol-t-shah>
GitHub: - <https://github.com/kajolshah310> | Medium: - <https://medium.com/@kajolshah>

Profile Summary

AI/ML Engineer with **4+ years** of work experience and **Masters in Artificial Intelligence** with **3.7 GPA**. Successfully completed **10+ ML, CV, and NLP** projects within a span of 2 years, displaying my proficiency in **Python, R, SQL, Tableau**, and leveraging ML/DL frameworks like **Keras, TensorFlow** and **Scikit-learn**. Led development to deployment of a critical project for **Atos Syntel** with 0% defect leakage in Production environment and awarded with **AGILE Entrepreneurial award** for the same. **AWS Certified Cloud Practitioner** and **Microsoft Certified Azure Fundamentals**.

Skills

- Python, R, Bash, Perl, C++
- Machine Learning, Deep Learning, Computer Vision, MLOps
- Natural Language Processing, LLMs, Data Analysis, Data Mining, ETL
- **Data Visualization**: - Tableau, PowerBI
- **Databases**: - MySQL, Db2, PostgreSQL, MongoDB (NoSQL)
- **Cloud**: - Azure, AWS - S3, EC2, EMR, AppSync, RDS, Athena, Kinesis, Glue, SageMaker, Lambda
- **Big Data**: - Hadoop, Spark
- Jira, Docker, Git, Linux, Postman, SDLC, OOP, CI/CD
- Jupyter Notebook, Anaconda, VS Code, RStudio, MS Excel, MS Office

Work Experience

SOFTWARE & AI/ML ENGINEER - TRANSPERFECT, USA

JUN '22 - PRESENT

- Spearheaded the development of keyword specific extractive text summarizer using **NLP** and unsupervised **deep learning** model in **Python** for clinical documents, reducing perusing time by **70%**
- Established a robust pipeline for an **ETL** service to extract data from database on **AWS Glue**, push data into **AWS Athena** and execute queries
- **Automated** data conversion process by developing Python codes to transform data from Excel sheets into JSON script, saving project time by **25%**
- Designed and implemented **AWS Lambda** functions in Python to **scrape** data from multiple **APIs, SQL** DBs and remove access of terminated users
- Conceptualized, built, and deployed a serverless GraphQL API using **Docker**, Terraform, AWS AppSync and RDS

AI/ML ENGINEER INTERN – DIYVERSITY, USA

MAY '22 – JUN '22

- Conducted extensive **data mining** on music dataset and developed python scripts to extract features such as chord progressions using **NLP**
- Actively contributed to the development of **RNN** model to predict patterns within songs to enhance music learning experience for users

SOFTWARE ENGINEER & ANALYST - ATOS SYNTEL, INDIA

JUN '18 – JUL '21

- Conducted extensive exploratory data analysis on healthcare data and leveraging **Tableau** to create insightful dashboards that enabled our team to achieve targets efficiently, resulting in a **20%** reduction in project costs
- Performed data wrangling and developed a highly accurate supervised ML model using **Logistic Regression** to predict severity of mental illnesses using data collected from mental health assessments (accuracy of **>90%**)
- Designed and implemented a prediction model using **KNN** to predict healthcare plans for new individuals based on their PII and medical history
- Led the **database migration** process from IDMS to Db2 **SQL** and successfully designed and implemented the migration strategy
- Developed, deployed, and maintained high-quality codes in the Production environment and **optimized** their performance

Projects

- **Multi-Collection Style Transfer with GANs** (*CNN, Keras, TensorFlow, OpenCV, NumPy, Pandas, Python*)
Devised and developed a **deep learning** model to transfer 4 artistic styles on real images using **Generative AI** inspired by an **IEEE** paper.
- **Fake News Classification** (*Scikit-Learn, NLTK, TextBlob, Naïve Bayes, SVM*)
Applied NLP techniques and implemented **LSTM** and ML models like to detect fake news in social media (accuracy **> 85%**).
- **Prediction of MSRP of cars using PySpark and MLLib** (*Seaborn, Matplotlib, Linear Regression, Decision Tree, Gradient Boosted*)
Performed data cleaning, **visualization** to get meaningful insights and implemented ML models for forecasting MSRP of cars.
- **Communities Crime Regressor** – Implemented fully connected neural network to predict community crime rates, achieving accuracy **>90%**.
- **Covid-19 Tweets Classification** – Applied NLP techniques and implemented ML models, achieving accuracy **> 79%**
- **IMDB Sentiment Classification** - Developed a sentiment classification model of IMDB movie reviews using NLP and **RNN** with accuracy **>80%**.
- **CIFAR10 Image Classification** - Achieved **>93%** accuracy in image classification on the CIFAR10 dataset using fully connected neural networks.
- **Hazardous Asteroids Prediction** - Developed an ML model using Random Forest in R with **>80%** accuracy to classify hazardous asteroids.

Awards and Certifications

- Data Science with Python - Udemy
- AGILE Entrepreneurial Award - Atos Syntel
- Quality Bronze Award - Atos Syntel
- AWS Certified Cloud Practitioner
- Microsoft Certified Azure Fundamentals
- Microsoft Certified Azure AI Fundamentals

Education

Illinois Institute of Technology - Chicago, US
Masters in Artificial Intelligence - GPA – 3.7

Aug '21 - May '23