

Lokesh Kank

San Diego, CA, USA | +1 (858)-744-2548 | [LinkedIn](#) | kank.lokesh75@gmail.com | [GitHub](#) | [Portfolio](#)

EDUCATION

San Diego State University (SDSU)

Aug 2022 - May 2024

Master of Science in Big Data Analytics (BDA)

San Diego, CA, USA

- *Big Data Science and Analytics Platforms.* - *GIS Programming with Python.* - *Business Analytics.*
- *Database Theory and Implementation.* - *Machine learning Engineering.* - *Data Management for GIS.*

Savitribai Phule Pune University

Aug 2016 - May 2019

Bachelor in Mechanical Engineering – GPA 3.4/4

Pune, MH, India

SKILLS

Programming & Technology: Python, R, Matlab, SQL, Postgres, HTML, CSS.

Machine Learning: Pandas, Numpy, Plotly, Pyspark, Matplotlib, Pytorch, Tensorflow, Scikit-learn Algorithms, ETL, Data Collection, Supervised & Unsupervised learning, Statistics, Hadoop Ecosystem, Sparks.

Toolkit: AWS, Power bi, Git, Docker, Databricks, Relational Database, MongoDB (NoSQL), Github, Pycharm, Excel.

Soft Skills: Verbal and Written Communication, Attention to detail, Analytical and Problem-solving, Creative.

EXPERIENCE

Data Science Intern

Feb 2023 – Present

Data Glacier

San Diego, CA, USA

- Developing a predictive model to determine the likelihood of customers buying the Bank's term deposit product, using Agile methodologies (Scrum and Kanban).
- Implementing data cleaning, and exploratory data analysis tasks to achieve an accuracy of **92%** or higher.
- Maintaining code quality by using Bitbucket. Developing a Flask web application, deploying it to the cloud, and integrating with APIs for seamless data retrieval and processing.

Graduate Research Assistant

Jan 2023 – Present

Department of Computer Science | SDSU

San Diego, CA, USA

- Working on Anomaly Detection of microvascular RBC flow and velocity analysis with pupil tracking or timing.
- Researched the use of CNN models using data in .avi video format. Optimized data pre-processing by utilizing Python, subprocess, & FFmpeg modules and converted “.avi” to “mp4” format.
- Leveraged OpenCV2 to extract metadata, and images from videos and created datasets. In collaboration, designing a future approach for the project using the [ideal](#) mechanism.

Software Engineer

Feb 2020 – Jul 2022

Accenture, Advance Technology Centre

Pune, MH, India

- Exploited SQL, and Python for failure analysis, and root cause analysis following best practices for fetching business insights for reporting. Saved **80** hrs of work by collaborating to develop automation in python.
- Experienced in migrating applications on Linux machines & work with cross-functional teams with a vision.
- Implemented app releases in the prod environment, contributing **3.2 % (\$ 7.2 Million)** of quarterly profit.

PROJECTS

Prediction of energy star scores of new buildings

- Conducted data pre-processing, EDA, and feature engineering (one hot encoding, removed collinear features with collinear coefficient > **0.6**). Noted that few predictors are negatively correlated hence selected essential predictors.
- Established baseline, and compare ML models on the performance metric. Performed hyperparameter tuning on the best model. Evaluated the best model on the testing set. Used Random Forest with MAE **9.044**.

Classification of pharmaceutical drugs for prevention of medication error ([website](#))

- Implemented Image Pre-processing to raise the image's quality and reduce medication errors by **56%**.
- Analysed the problem of Medication errors on various drugs and developed Deep Learning models – VGG 16, Xception, and Inception with **95 to 99%** accuracy to classify drugs efficiently.
- Integrated Tableau interactive dashboard for business intelligence about errors in the medical field.

Data science salary recommendations

- Scrapped about **1000** job listings from Glassdoor using Selenium to complete an end-to-end project.
- Quantified the value that businesses place on Python, R, Excel, AWS, and Spark using Exploratory Data Analysis and created features that were engineered from the content of each job description.
- Applied Grid search-CV to optimize Linear, Lasso, and random forest regressors and found the ideal model

CERTIFICATION

AWS - Certified Solution Architect

May 2022 – Jul 2025