

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 (STEM)

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

EXPERIENCE

Data Science Intern Skills: Pandas, Numpy, Data Cleaning, Scrum and Kanban, Flask.

Feb 2023 – May 2023

Data Glacier

San Diego, CA, USA

- Developing a predictive model for determining the likelihood of customers buying the Bank's term deposit product. Increased productivity by **16%** as measured by burn-down chart by implementing Agile methodologies.
- Attained a **7%** increase in model accuracy by handling missing values, outliers, and performing normalization to reduce predictor biases. Utilized correlation matrix for getting insights, resulting in a model accuracy of over **92%**.
- Achieved a **50%** increase in stakeholder trust & reliability measures by effectively communicating project progress and statistical insights, which allowed for a reduction in meeting frequency from twice a week to once a week.

Software Engineer Skills: Data Analysis, SQL, Databases, Automation, RHEL Linux, AWS.

Feb 2020 – Jul 2022

Accenture

Pune, MH, India

- Obtained a **35%** increase in data retrieval speed, by streamlining SQL queries and using indexing strategies.
- Exploited SQL, and Python for troubleshooting, and created business reports for client presentations. This resulted in a **26%** increase in client satisfaction as calculated by NPS (Net Promoter Score).
- Saved **60hrs** of monthly work by developing automation in python resulting in smooth processing of cross-functional.
- Concluded **18%** improvement in the application performance as gauged by the Smile tool after cloud migration.

PROJECTS

Prediction of energy star scores of new buildings.

Skills: Statistics, Exploratory Data Analysis, Model Optimization.

- Accomplished a **12%** increase in model accuracy examined by error reduction and by conducting data processing, EDA, and feature engineering, including one hot encoding and removal of collinear features with a coefficient **> 0.6**.
- Established baseline error of **25%** by taking a median of the training label set which helped compare ML models.
- Performed hyperparameter tuning on the best model causing an increase in model accuracy by **8%**. Evaluated the best model on the testing set and went with Random Forest with MAE **9.044**.

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

Skills: Tableau, CNN, TensorFlow, Keras, Image Processing, Google sites.

- Reduced medication errors by **56%** after implementing Image Pre-processing in detail to raise the image's quality.
- Analysed the problem of Medication errors on various drugs and developed Deep Learning models (Neural Networks) – VGG 16, Xception, and Inception with **95 to 99%** accuracy to classify drugs efficiently.

Data science salary recommendation.

Skills: Data Collecting, Scripting, Regression, Model Evaluating.

- Scrapped about **1000** job listings from Glassdoor using Selenium for datasets by understanding data architecture.
- Cleaned **80%** of scraped data and turned categorical data into numerical labels using label encoding for model building
- Achieved MAE of **11.22** by using the Random Forest model. Used linear regression as a baseline model.

Automation of sales dashboard using python.

Skills: Python, Excel, Quantitative Analysis, ETL.

- Effected a **50%** reduction in manual business process as determined by time to build a dashboard, by automating data collection and statistical analysis using Python.
- Implemented advanced analytics and visualization capabilities in the dashboard using Python libraries such as matplotlib and seaborn, resulting in a **30%** increase in decision-making efficiency

SKILLS

Programming and Technology: Python, R, SQL, PostgreSQL, Scala, HTML5, CSS

Machine Learning: Supervised and unsupervised learning, Pyspark, PyTorch, Scikit-learn algorithms, Data Visualization, Apache Spark, Clustering, Statistics, Natural Language Processing, A/B testing, Hadoop Ecosystem.

Toolkit: Pycharm, Tableau, Databricks, Git, Github, Hive, Excel, PowerPoint

CERTIFICATION

AWS - Certified Solution Architect

May 2022 – Jul 2025