Lokesh Kank

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

EDUCATION

San Diego State University Aug 2022 - Aug 2024

Master of Science in Big Data Analytics (BDA) San Diego, CA, USA

Coursework: *Big Data Science and Analytics Platforms, GIS (Geospatial Information Systems) Programming with Python, Business Analytics, Advance Machine learning, GIS Database management.*

Savitribai Phule University Aug 2016 - May 2019

Bachelor in Mechanical Engineering – **GPA 3.3/4** Pune, MH, India

SKILLS AND KNOWLEDGE

Languages and technology: Python, R, HTML, Scala.

Big Data & Machine Learning: Pandas, NumPy, Scipy, MongoDB, Plotly, Matplotlib, Tensorflow, Keras, PyTorch, Scikit learn algorithms, Supervised & Unsupervised learning, Ensemble learning (Random Forest, Bagging, Boosting) Resampling techniques, Hadoop, Hive, MapReduce, HDFS, Spark.

Data Science Technologies: A/B testing, ETL, data wrangling, statistical modeling, data mining, data visualization, Geospatial techniques (GDAL, OGR), Time Series Analysis, and Hypothesis Testing.

Toolkit: Jupiter Notebook, Tableau, Google Analytics, PyCharm, Power BI, GitHub, Git, AWS Sagemaker, Confluence, nBHB (Official Document management), WinSCP, Putty, HP-ALM, JIRA.

Soft skills: Strong verbal and written Communication, learning quickly, self-motivated.

PROJECTS

Prevention of medication error using Deep Learning (https://sites.google.com/sdsu.edu/**)**

- Analysed the problem of Medication errors on various drugs and developed Deep Learning models (CNN) - **VGG-16, Xception, and Inception** with **95–99%** accuracy to **classify** pharmaceutical drugs efficiently.
- Implemented **Image Preprocessing** (Edge detection, Masking, standardization, and normalization) to raise the image's quality and **reduce** medication **errors** by **56%**.
- Integrated **Tableau** interactive dashboard to display information about medication errors in the medical field.

Infer energy star scores of new buildings

- Conducted **data exploration**, **Data pre-processing**, and Built **regression/classification models** that can estimate a building's Energy Star Score based on selected features after **feature engineering** (one-hot encoding, Removed collinear features with collinear coefficient > **0.6**).
- Established baseline error scores, and compare different models. Used **Random Forest** with **MEA 9.044**.

Time series forecasting of energy generation data

- Reduced the loss due to delays in the production system by **68%** by forecasting the undesirable breakdowns due to various factors (grid failure, scheduled maintenance, unscheduled maintenance, gearbox failure, etc.)
- Increased energy production by **47%** by analyzing the past energy generation data of the wind turbine generator.
- Fitted, evaluated, and made predictions with the - **Random Forest** regression model for **time series forecasting**;
- Created a **Power BI** Dashboard to display the forecasted energy and breakdown data. Combined and converted multiple DGR with different data formats to a single unified **SQL database** for ease of **analyzing** and **forecasting**.

Data science salary estimator

- Completed an **end-to-end project** to calculate data scientist salaries and negotiate their compensation package.
- **Scrapped** about 1000 job listings from Glassdoor using Python and Selenium.
- Quantified the value that businesses place on Python, R, Excel, AWS, and Spark using **Exploratory Data Analysis** and features that were engineered from the content of each job description.
- Applied Grid search-CV to optimize **Linear, Lasso,** and **Random Forest Regressors** and find the ideal model.

PROFESSIONAL EXPERIENCE

Accenture, Advance Technology Center Feb 2020 – Jul 2022

(Application Development Analyst**)** Pune, India

- Exploited **SQL**, and **Python** for **Data Analysis**, fetched business insights, and **communicated** with clients. Saved **80 hrs** of additional work by **collaborating** with the team to develop automation in **python**.
- Solved high-impact incidents using **SQL** within breach time. Experienced in migrating applications on **Linux machines**, and configuring server parameters on Linux and Windows machines. Experienced cross-functions.
- Establishing connections with its interfaces, and deploying **hotfixes**, and releases in the PROD environment. The release has contributed **3.2 % ($ 7.2 Million)** of quarterly profit made by clients for the Europe market.
- Tools used: WinSCP, Putty, SC3, HP-ALM, JIRA, RDBMS (MS-SQL, Oracle), Microsoft Excel, Confluence, nBHB (Official Document management).

Certification

AWS - Certified Solution Architect Issue on – May 2022 – Expire on – July 2025