Vatsal Jigar Kadakia

Hoboken, NJ (813) 843-9863 vkadakia@stevens.edu linkedin.com/in/vatsal-kadakia-760a3b1ab/

Education

Stevens Institute of Technology, *New Jersey* **Aug 2022 – May 2024**
Masters of Science in Data Science *GPA: 3.8/4.0*
Coursework: Time Series Analysis, Deep Learning, Applied Machine Learning, Statistical Methods, Optimization Methods

University of Mumbai, *India* **Aug 2018 – May 2022**
Bachelor of Engineering in Computer Engineering *GPA: 3.82/4.0*
Coursework: DBMS, Data Structures, Data Analytics, Software Engineering, Object-Oriented Programming

Technical Skills

Languages: Python, R, SQL, C++, JavaScript, XSL, XML, JSON, Dart, HTML, CSS
Database: MySQL, NoSQL, Firebase, MongoDB, Hadoop, Apache, Airflow, Spark
Developer Tools: Git, GitHub, AWS, Jira, Confluence, Microsoft Power BI, Tableau, Oracle Analytics Cloud
Technologies/Frameworks: Docker, React, Flutter, Flask, Excel(VBA, Macros, Pivot tables)

Experience

Minerals Technologies Inc, Bethlehem, PA **Jun 2023 – Dec 2023**
*IT Intern - **MTI Oracle***

- Collaborated with the **Oracle Analytics** team of the USA, Europe and India to migrate client data from **Oracle E-Business suite** to **Oracle Analytics Cloud(OAC)**.
- Developed **composite visualizations** of financial portfolio on OAC dashboard **(Sales Data Mart and ADW Sales Voice - Datasets)** transacting CY expense of $893 Millions, PY expense of $629 Millions and 42% Change expense.
- Revamped data project with OAC's **Machine Learning module**, reducing duplicate data by 40% and facilitating timely report and dashboard creation.
- Automated incident and request handling of company's **ServiceNow Chatbot**, reducing employee manual work by 35% and improving ServiceNow accessibility by 50%.
- Mastered and deployed **Java Scripts** using in-built **Natural Language Processing(NLP)** techniques to improvise plugins(ITSM, Conversational Appointment booking).

Projects

Demand Forecasting - ARIMA, SARIMA, ARIMAX | *Time Series Analysis, Python* **Academic Project**

- Conducted **EDA** on Store Item dataset to check if series is stationary or not by plotting **ACF** and **PACF** graphs and took their first differencing. Obtained p-value < 0.05.
- Implemented ARIMA, SARIMA & ARIMAX models to compare **AIC** and **MSE** scores. **ARIMAX(6,1,7)** was the best model with AIC score 10110.465 and MSE 22.95. Plotted diagnostics to predict forecast of 50 items in 10 stores.

Pyspark weather API data streaming and analysis | *Python, Machine Learning, SQL* **Academic Project**

- Engineered a **PySpark** streaming application integrating **SQL** and machine learning algorithms to dynamically fetch weather data from custom-selected locations and dates via a **weather API**(Processed over 25 fields).
- Enhanced functionality with **PyML operations** to compare weather data from two locations simultaneously, aiding event scheduling teams and skincare brands in strategic decision-making.

Grocery Store Optimization | *R,Python, Machine Learning* **Academic Project**

- Conducted **Z-test**,**Chi-Square** test to analyze significant difference between sign up rate for cheaper expensive mailer of delivery club. Performed **ANOVA** to find p-values of all columns(response variable as Sales column).
- Obtained 96% loyalty score accuracy, 93% behavior classification with **Regression** algorithms, significantly transforming store layout and marketing.

Publications and Courses

Data Science Infinity - Andrew Jones (Course)

Dr.Megharani Patil, Vatsal Kadakia, Yash Gupta, Harsh Doshi(2022); Automated Healthcare System using Gamification and AI based Chatbot; Computer Vision (Applications of Visual AI and Image Processing),vol 15,De Gruyter

Mendon, A., Patil, M., Gupta, Y., Kadakia, V., Doshi, H. (2023). Automated Healthcare System Using AI Based Chatbot. In: Balas, V.E., Semwal, V.B., Khandare, A. (eds) Intelligent Computing and Networking. Lecture Notes in Networks and Systems, vol 632. Springer, Singapore