

Abhijit Gokhale

710 South Beech Street, Syracuse, NY, 13210. | +13152784351 | agokhale@syr.edu | abhijit5october@gmail.com
[LinkedIn](#) | [Github](#) | [Handshake](#)

Objective

A skilled and dedicated IT professional with diverse experience in handling structured and unstructured data by implementing complete Software Development Life Cycle (SDLC) with the help of data preprocessing, predictive analytics and creating business solutions. Currently pursuing MS in Applied Data Science from Syracuse University, New York. Looking for open positions to work in the field of data science and data analytics to provide low-cost business solutions.

Technical Skills

- 1) Supervised Learning – Null Hypothesis Testing, Student-t Test, Bayesian analysis, ANOVA, MCMC Analysis, Chi-Square test, Linear Regression, Logistic Regression, Support Vector Machine learning.
- 2) Unsupervised learning – PCA, Association Rules, Clustering,
- 3) Structured and Unstructured data analysis, Deep Learning, Time Series Analysis.
- 4) ETL Development, Testing, Support, and database management.
- 5) SQL, Python and R language code building and debugging.
- 6) Tools - R-Studio, Jupyter Notebook, Microsoft SQL server, Microsoft PowerApps, Azure Data Studio, Informatica PowerCenter, Informatica BDM, Netezza SQL, Ambari HiveQL, MS Word, MS Excel, MS PowerPoint, Draw.io, Jeera.

Work Experience

Accenture Solutions Pvt. Ltd. (Dec 2016 – Aug 2020)

Role – Application Development Analyst

1) Job Requirements - Performing Complex Informatica Powercenter ETL applications using Microsoft SQL server as DBMS to store the 'US based Energy Production Client' data into the DW for reporting and data tracking purposes.

Job Deliverables - a) Developed, Enhanced and production support to the daily 50+ Informatica ETL jobs of HR and Finance Modules.
b) Resolved and Deployed 30 Jeera issues linked to client requests(update/insert/delete) and new enhancements as well as accounting to 0 SLA for critical and high issues.
c) Handled the ETL backups (Technical Checkouts) in Dev, Test, QA, QASUP, TestSUP, Prod environments before respective server refreshes.

2) Job Requirements - Performing Complex Informatica BDM ETL applications using DBMS such as HiveQL DB with Ambari , Netezza SQL to store the US/EU/Asia based Pharmaceutical client's data into the DW for reporting and data tracking purposes.

Job Deliverables - a) Developed 50 ETL jobs corroborating the approved and successful test case scenarios.
b) Resolved and handled 50 ETL Jeera issues occurred during testing.
c) Conducted Knowledge transfer sessions and appreciated as a crucial member while conducting SIT.
d) Developed R language script to handle the columns with decimal numbers and 'date-month-year' columns.
e) Created functional design and test cases documentation (avg 15 per file) for reporting purposes, capturing, and correcting data error.

Academic Projects

1) Objective - Statistically analyzing ongoing trends in the given dataset of vaccination rates in the California School Districts and recommending the government about steps to be taken to increase the total vaccinations.

Contributions - a) Exploratory analyses of dataset of 1000 observations (sample of original dataset) with histograms, bar-plots, and boxplots/violin plots to understand data distribution and then performed data cleaning. Identified the past trends in the vaccine rate variations.
b) Implemented the Descriptive and Inferential statistical results to report the analyses with 95% accuracy.
c) Incorporated unbiased data along with data transformation to understand the underlying correlation between dependent (vaccination rates) and independent variables.
d) Consolidated and presented the statistical analyses report confirming the results obtained with numerous tests and recommended important factors to be focused on to increase the total vaccinations in California Districts which are lacking.

2) Objective - Analyzing given dataset of European Hotel Business and recommending the most important factors affecting the Average Daily Rate(ADR) of the hotels at Resort location and City location along with recommendation for the better revenue generation at both locations.

Contributions - a) Executed exploratory, descriptive and inferential statistical analyses, on City hotels data of 80000 observations and resort hotels data of 40000 observations, to understand data distribution, correlation between Dependent (ADR) and several Independent variables followed by data Cleaning.

b) Showcased the test results with 95% accuracy in the statistical analyses report confirming the possible traits in data and recommended important factors to be focused on to increase the average daily rate (ADR) based on current business model of hotels at City and Resort locations.

3) Objective - Building user friendly informative application in Microsoft PowerApps for the Soccer dataset, for companies or managers to bid on soccer players.

Contributions - a) Utilized Microsoft Azure to build Microsoft SQL server followed by DB creation, tables creation and , views creation in Azure Data Studio as DBMS.

b) Incorporated business functional design with E-R requirements and diagram to build conceptual and logical data model.

c) Performed data acquisition of 20 to 30 observations on avg in 11 tables and 16 views.

d) Presented the final working business model solution by publishing an application in Microsoft PowerApps to bid on soccer players of choice.

Education

1) Masters in Applied Data Science, Syracuse University (*Jan 2021 - Present*)

GPA – 4.0

2) Bachelors in Electronics Engineering, Mumbai University (*July 2012 – May 2016*)

GPA – 3.3

Personal Data Science Projects

1) Recommending important factors, in the given dataset of HR department, affecting the employee attrition in a company.

2) Performing EDA, Statistical Analyses, Hypothesis Testing, and training Linear Regression model for predicting US insurance premium charges.

3) Performing EDA, Statistical Analyses, Hypothesis Testing, and training data using appropriate Regression model for predicting NYC Airbnb hotels prices.

4) Building a model of Deep Neural Network for MNIST data Classification.

5) Building Amazon Price tracker to intimate the user once the price of a certain product reaches below threshold value.

6) Showcasing the gist of any news article by building the clean word-cloud which focusses on most occurring although important words.

Hobbies

1) Play and perform with Indian Drums (Tabla) as musical instrument,

2) Play Table Tennis, Carrom.

3) Cooking and Baking.

References

[Available Upon Request.]