GOKUL KRISHNASWAMY

gokul.krishnaswamy.1996@gmail.com | 720-609-1247 | linkedin.com/in/gokul-krishnaswamy | Edison, NJ

PROFESSIONAL SUMMARY

Data Scientist with 3 years of work experience. Graduate student in Data Science with strong analytical skills, demonstrated leadership experience and proven ability to efficiently solve business critical problems. Area of specialization includes data structures and algorithms, data analysis and building predictive models.

SKILLS

Languages : Python, SQL, R, Java, Hive, C++, XML, HTML
Database : Oracle, MySQL, PostgreSQL, MongoDB
Tools : Git, Airflow, Spark, Tableau, Oracle SQL Developer, Databricks, AWS, JIRA
Packages : Scikit-learn, NumPy, Pandas, SciPy, Keras, Matplotlib, Seaborn, PySpark
Certifications : AWS Machine learning speciality, AWS Cloud practitioner

EXPERIENCE

Data Scientist | AT&T | Plano, TX, USA | Nov 2020 – Present
Technologies used: Python, SQL, Airflow, SparkSQL, Azure, ETL, Git, Linux, Active Directory, LDAP, JIRA

- Setup and maintained a robust Airflow application in production, which served as backend for ATT internal web forms.
- Automated ETL workflows using airflow DAGs, which saves almost 20 work hours of a user per week.
- Automated active directory resource management to meet security compliance by creating data pipelines in python.
- Mentored new team members and planned application migration to Azure cloud services.

Data Acquisition and Developer intern | Hitachi ABB Powergrids | Broomfield, CO, USA | May 2020 – August 2020
Technologies used: Python, SQL, Machine learning, ETL, PySpark, Linux, Git, Azure DevOps

- Developed efficient python scripts to collect unstructured data, which powered an analytical software generating $8M.
- Rewrote a data integration algorithm and reduced its runtime by factor of 3, while using git.
- Improved machine learning classification model's performance (F1 score) by using efficient data cleaning process
- Collaborated with data analysts for writing custom ETL jobs in SQL, and so ensuring normal business operations.
- Designed architecture for a new tool to collect Japan's energy market data.

Data Collection and Research Coordinator | University of Denver | Denver, CO, USA | July 2019 – May 2020
Technologies used: Python, Statistics, Leadership, AWS

- Head of a diverse team of 15 members. Redesigned the architecture of the project and reduced its complexity.
- Generated qualitative data from more than 300 research videos and derived statistical measures from it.
- Implemented efficient solutions and eliminated manual work in tasks which otherwise needed 150 work hours.

Data Analyst | Onceptual Business Solutions **|** Coimbatore, TN, India | July 2018 – June 2019.
Technologies used: Python, SQL, Tableau, Machine learning, ETL, Git

- Designed and created over 15 normalized tables in Oracle database and migrated data through ETL pipelines in python.
- Conducted data analysis using Tableau. Created over 100 ad hoc SQL queries and validated the results of dashboards.
- Improved customer activity by 30%, by recommending business changes based on results of data analytics.
- Prepared and presented data reports to clients and communicated insights.

Software Developer intern | Oracle | Bangalore, KA, India | January 2018 - June 2018
Technologies used: Python, Selenium, HTML, XML, Linux

- Automated tests of an analytical software by creating over 60 XML scripts and testing them in linux server.

EDUCATION

MS, Data Science | University of Denver | Denver, CO, USA
GPA: 3.85 out of 4 | August 2018 – August 2020

BTech, Computer Science and Engineering | Amrita Vishwa Vidyapeetham University | Coimbatore, TN, India
GPA: 8.57 out of 10 | August 2014 - August 2018

RELEVANT COURSES

Python, Database (SQL), Machine learning, Data Structures and Algorithms, Big Data, Data Visualization, Statistics and Probability, Data mining, Parallel and distributed computing.

RELEVANT ACADEMIC PROJECTS

Insurance cost prediction: Performed exploratory data analysis, feature engineering and data manipulation to create LASSO regression model to predict insurance cost based on patient medical records. Achieved a record low RMSE value on test data.

Churn modelling: Created a customer churn model in python using artificial neural network. Used ROC curve to find ideal trade-off value. Optimized the model through hyper-parameter tuning. Achieved accuracy of 85% and recall score of 94%.

Simple image classification: Performed dimensionality reduction and developed a logistic regression model using TensorFlow library that classifies the handwritten digit to a numerical value. Achieved accuracy of 91% on test data.

Bacteria cases prediction: Forecasted the number of bacteria cases in Melbourne city by developing a seasonal ARIMA model in python. After optimization, achieved a record low RMSE value.

RESEARCH PAPERS PUBLISHED

K. Gokul, R. Pooja, K. Gowtham and G. Jeyakumar, "A self-switching base vector selection mechanism for differential mutation of differential evolution algorithm," *2017 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, 2017, pp. 1545-1549, doi: 10.1109/ICCSP.2017.8286647.

- Developed a strategy (RandtoBest) to optimize Differential Evolution Algorithm for uni-modal functions. Presented the paper at an IEEE conference.

K. Gokul, R. Pooja and G. Jeyakumar, "Empirical Evidences to Validate the Performance of Self-Switching Base Vector Based Mutation of Differential Evolution Algorithm," *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, 2018, pp. 2213-2218, doi: 10.1109/ICACCI.2018.8554928.

- Studied the critical parameters and provided mathematical evidence to validate the performance of RandtoBest strategy.