

Vidya Venugopal

vidyav.vgopal@gmail.com • <https://github.com/vidyavcodes>

EXPERIENCE

Data Scientist

Jan 2022 - Present

Capgemini

- Developed end-to-end pipelines in Azure Synapse Analytics to extract data from various sources and wrote SQL scripts for data transformation which reduced pipeline runtime by 50%
- Identified key consumer trends in the market for a client's product line with the use of clustering algorithms
- Re-designed and maintained PowerBI dashboards that were instrumental in supporting a USD 1 million deal for a client's new product
- Created a proof of concept for a project which featured deep learning as a solution and presented it to senior leadership and other key stakeholders

Business Data Analyst

March 2019 - Dec 2020

Varian, A Siemens Healthineers Company

- Created PowerBI dashboards that visualised key metrics for product orders and instalment status, leading to a 9% increase in gross orders
- Built appropriate data models and wrangled data using R and DAX in PowerBI which led to the report loading time decreasing by 70%
- Developed a system to identify anomalies in customer data that was entered manually and sent out alerts to the sales team which resulted in a 90% decrease in inaccuracies

Business Research Intern

May 2018 - Aug 2018

SAP

- Built dashboards to track KPI and found potential growth opportunities for 100 existing customers
- Collected data on customers' opinions of different products and presented findings to key stakeholders and different teams

SKILLS

Programming Languages: SQL, Python (Sci-kit learn, Tensorflow, Pandas, Numpy, Plotly), PySpark, R

Cloud Platforms: Amazon Web Services, Microsoft Azure and Dataiku

Data Science and Miscellaneous: Machine Learning, Data Pipelines, Apache Airflow, Spark, PowerBI, Statistics, A/B testing, Agile Project Management, Git, Jira

Certifications: AWS Certified Machine Learning - Specialty, Dataiku ML Practitioner, Dataiku MLOps Practitioner, Power BI Data Analyst Associate, Professional Scrum Master™

EDUCATION

Erasmus University Rotterdam | Master of Science in Data Science & Marketing Analytics

Sept 2021

National University of Singapore | Bachelor of Science (Honours) in Economics

Dec 2018

PROJECTS

Predicting Customer Behaviour Online

- Used the Random Forests algorithm in R to classify a customer's online purchasing intention and determined that the number of pages a customer visits can be a significant indicator of their likelihood to purchase an item
- Accuracy of the model was 3% higher than the average accuracy of other algorithms used

Impact of the Emerging Middle Class on Consumption

- Examined the impact of the emerging middle class on conspicuous consumption patterns in Southeast Asia by using k-means clustering and multiple regression analysis
- Identified 2-3 key market segments and suggested strategies that companies can use to capture the different segments