SARTHAK VAJPAYEE

Dallas, TX 75248 | 469-347-9198 | sxv220020@utdallas.edu | WWW: Online Digital Resume

Website, Portfolio, Profiles

- https://sarthakv7.github.io/my_folio/
- https://github.com/SarthakV7
- https://www.linkedin.com/in/sarthak-vajpayee/

Professional Summary

Well-qualified Data Scientist experienced working with vast data sets to break down information, gather relevant points and solve advanced business problems. Skilled in predictive modeling, data mining and hypothetical testing. Offering 4 years of experience in improving business operations.

Skills

- Data Analysis & Data Mining: Scikit-Learn
- Probability & Statistics, Multivariate Calculus & Linear Algebra
- Machine Learning & Deep Learning
- Tensorflow, Keras, PyTorch
- Python, MySQL, MongoDB, PySpark
- Computer Vision, Natural Language Processing (NLP)
- Data Visualization
- Git, Docker, Spacy, NLTK, Plotly, OpenCV
- Data Wrangling
- Cloud Computing (AWS & GCP)
- Demand Forecasting
- Good Communication Skills

Work History

Senior Data Scientist 12/2021 to 07/2022

Ernst & Young Global Limited – Bengaluru, India

- Designed **predictive modeling** framework for world's 3rd largest FMCG, which achieved an **accuracy** of over **90%** and **bias** under **15%**, **outperforming** previous demand planning techniques across all categories in the Nordics market.
- Delivered Gen-2 version **(70% Volume and SKU coverage)** for the Norwegian market within 15 days, **benchmarking** among European, South-East Asian, and South American markets.
- Built automated **data discovery** streams and **data pipelines** and implemented code **modularization** along with **E2E testing** for production on **Azure Data-Lake** utilizing **Apache Spark** and **Python**.
- Leveraged highly **intermittent time-series** data sources to extract deep insights on demand, such as previous sales, interior trends, external trends, events, and promotions.
- Performed **regression and variation analysis** across different FMCG categories and conducted **exploratory data analysis** on variables, such as **product volume** and **market cannibalization.**

Machine Learning Engineer 09/2019 to 11/2021

Applied Roots – Hyderabad, India

- **Designed, released,** and **maintained** six **python** based **Machine Learning** and **Deep Learning** projects in Applied Roots (an e-learning platform), along with mentoring and evaluating assignments of over **8000 students.**
- Contributed to **100% year-over-year** revenue growth by creating **operational strategie**s and technical development of the platform.
- Designed **plagiarism checker** for student coursework by leveraging various **NLP and Deep Learning**

techniques such as **Word-to-Vector, Code-to-Vector, Pre-trained Language Models, and LSTM.**

Research Project Assistant 03/2021 to 11/2021
Stellenbosch University – Stellenbosch, South Africa

- **Designed, trained, fine-tuned,** and **validated** several **deep-learning** and **transfer-learning** based models such as **MLP, RNN, CNN, VGG-16,** and **RESNET-50** for predicting **COVID-19** using patient's **cough audio**.
- Achieved **specificity, sensitivity, AUC,** and **accuracy** scores of **0.9, 0.93, 0.96,** and **92.78%** resp. Using **RESNET-50** model, making COVID-19 diagnostic through this tool **on par with the RT-PCR test.**
- Developed **web-application** using **Python** and **Javascript** for users to get diagnosed with COVID-19 **within 5 seconds** through their cough audio.

Freelance Data Scientist 10/2020 to 01/2021
ViVi Green Innovators (Viscon Group) – Amsterdam, Netherlands

- Utilized **python** and **image processing** for **object skeletonization**, **localization,** and **unsupervised machine learning (DBSCAN clustering)** to develop and test a **machine vision and robotic process automation-based** product for **extraction** of **plant cuttings**, used in **horticulture** for **vegetative (asexual) propagation.**
- Developed and delivered a prototype within **70%** of deadline duration with **5/5 client satisfaction.**

Education

Master of Science: Business Analytics (Data Science Specialization) Expected in 05/2024
University of Texas At Dallas - Dallas, TX

Bachelor of Technology: Electronics And Communications Engineering 06/2018
JSS Academy of Technical Education - Noida, India

Certifications

- **Google Cloud:** Advanced machine learning on Google Cloud specialization.
- **Coursera:** Deep-learning specialization.
- **Applied roots:** Certified Data Scientist.
- **Udacity:** AI for Trading Nanodegree.
- **Stanford University:** Machine learning.
- **Amazon:** Machine learning on AWS.