Sumit Vaise

✉ sumitvaise29@gmail.com

📞 +15147016472

in linkedin.com/in/sumit-vaise-6644514b

🔗 https://sumit1673.github.io/

Summary

Sr. Data Scientist and Masters in Electrical and Computer Engineering with fascinating knowledge in building data-intensive applications, overcoming complex architectural, and scalability issues in diverse industries.
What I like to do is predictive modeling, data processing, data visualization, Computer Vision application development, and Applied Machine Learning.

Experience

Machine Learning Engineer

Quantiphi

Jul 2021 - Present (1 year 2 months +)

- Major tasks involved Machine learning application development on GCP, creating solution architecture for different problem statements based on computer vision or data science, and weekly basis client interaction.
- Lead a team of three ML engineers in building an inventory allocation model for a luxury goods conglomerate to prioritize shipment locations across the US while maximizing sales volume and marginal income. Architecture implements a GCP pipeline starting from Cloud Storage for user input, Cloud Function for triggering multiple scripts, and BigQuery to store the results.
- Designed GCP-based solution architecture for object detection application for detecting small objects inside a restaurant's kitchen.
- Utilize Kubernetes microservices to run different services related to the project.
-Performed EDA in BigQuery and Python and created visualizations to obtain insights on datasets
Projects:
- Automatic Document Classification System for Public Sector using ML, GCP, Kubernetes.
- Inventory Allocation for Maximizing Sales.

Senior Data Scientist

SalonEverywhere

Jan 2021 - Jul 2021 (7 months)

- Leading a team of data scientists and developing a framework to create a Professional Website for Salon Professionals using Machine Learning, AWS, and Python.
- Text classification to obtain the words related to salons, image classification to identify the hairstyles or arts a salon professional is experienced in, and image segmentation to find out the color they use most in their salon.
- Utilizing Tensorflow and Tensorflow lite for developing the predictive models and converting them to mobile optimized and quantized models.

ML: regex and NLTK stop words. LSTM-based model training, CNN, semantic segmentation, MaskRCNN, and VGG.

CLOUD: AWS Lambda, S3 buckets, Step Functions, and EC2 instances.

Programming: Python

Research Assistant
Concordia University
Jul 2019 - Dec 2020 (1 year 6 months)
Developed a French pronunciation application for desktop based on Google Cloud Speech APIs
Speech-To-Text and Text-To-Speech using Python and PyQt5 for Windows platform.
Utilized Model View Controller design patterns for developing the app.
Used Docker to distribute the application.

Python Trainer
Montreal Networks Inc.
Feb 2019 - Apr 2019 (3 months)

Computer Vision Engineer
Bosch Global Software Technologies
Sep 2017 - Dec 2018 (1 year 4 months)
Designed and Developed an end-to-end Deep-Learning-based computer vision pipeline to remove
densely labeled datasets and convert them into sparsely labeled datasets. The Sparsely Labelled
dataset has reduced the training time and computational cost by 32% and is currently used in the
project.
Developed a 2D cubic shape on car images using OpenCV and Python which was further used as a
label in training a model.
Developed Qt and C++-based semantic segmentation labeling tool.
Collaborated with Bosch's Germany team on an Autonomous driving project for data visualization and
training of the VGG network.
Project management uses tools like JIRA to create, assign, and submit tasks, and manage task
descriptions and task timelines.

Use SVN for version control.
Projects:
- Image Segmentation labeling tool development using Qt and C++.
- Model training and testing for object detection (cars, sign boards, trees and pedestrians) using Python
and PyTorch

Computer Vision Software Engineer
L&T Technology Services Limited
Jan 2014 - Aug 2017 (3 years 8 months)
Projects:
- Developing a novel algorithm for stitching two radiographic images using an intensity-based image
registration method in C++.
- Developed application for human segmentation on thermal images using OpenCV and C++ on
RaspberryPi.
- Developed end-to-end deep learning-based multiple object detection application for the marine
environment. The model was trained and deployed on the NVIDIA Jetson-TK1 development board. The
whole application was automated using SHELL scripting on a Linux operating system.

Tech: Python, C++. PyTorch, Shell Scripting, OpenCV

- Developed MATLAB-based data labeling tool for cloud and mountain segmentation for France's ministry of defense. The dataset was further used as an input for the Optical flow algorithm.

Education

Concordia University
Master of Engineering - MEng, electrical and computer engineering
2019 - 2020

uttar pradesh techinal university
Bachelor of Technology (B.Tech.), Electrical, Electronics and Communications Engineering
2008 - 2012

st. mary's senior secondry school
12th, maths and science
2007 - 2008

Licenses & Certifications

TensorFlow in Practice - Coursera
FSQPH6ERELWY

Sequences, Time Series and Prediction - Coursera
M8CARPGDTF6F

Natural Language Processing in TensorFlow - Coursera
SU45EKPGB948

Convolutional Neural Networks in TensorFlo - Coursera
C7W29QLQW79G

Deep Learning Specialization - Coursera
4QAC8DQXA69E

Skills

Image Segmentation • CI/CD • Google Kubernetes Engine (GKE) • Image Processing • Computer Vision • Linux • Shell Scripting • Deep Learning • Data Science • Python (Programming Language)

Honors & Awards

Batch Topper

Jun 2014
Topper of the DEPD course in CDAC-HYD with a percentage of 80.17%.