Sarthak Singh

asarthaks@gmail.com

Get Latest version

asarthaks

</> Python

See dark theme

+91-9868862862

Experience

- **Gigaforce Private Limited** — Noida, India
  Senior Software Engineer - AI/ML — *October 2022 - Present*
  - Utilized expertise in US insurance claims data to **identify and maximize subrogation potential, resulting in a 30 percent increase in revenue** for clients
  - **Collaborated with subject matter experts to incorporate domain knowledge and tailor subrogation process** to meet specific client needs
  - **Applied natural language processing techniques to understand context** and accurately determine parties at fault in loss descriptions
  - Implemented **digitization of insurance claims documents** to streamline information extraction and improve overall claim process
  - Designed and **developed scalable, efficient ingestion pipelines** to integrate predictive model output into databases
  - Created and **implemented architecture for various products** to ensure proper monitoring and scalability
  - Developed and **applied advanced algorithms for subrogation opportunity identification**, leading to increased efficiency and revenue for clients
  - **Built and maintained strong relationships with clients and subject matter experts** to understand and meet their unique business needs
  - Conducted **extensive data analysis and visualization to identify patterns and trends in claims data**, informing the development of improved subrogation models
  - **Led the development of software solutions** that automates the subrogation process, resulting in increased speed and accuracy
  - **Worked with cross-functional teams** to develop and implement best practices for data management and quality control
  - **Utilized Machine Learning techniques to improve the accuracy** of subrogation identification and increase the revenue

  Technologies: Python, SQL, Tensorflow, Scikit-Learn, Pytorch, Docker, GCP, Spacy, Word2Vec, Large language models.
  Theory: Data Preprocessing, Statistical Modeling, Deep Learning, Linear Algebra, NLP.

- **Navia Life Care** — Gurgaon, India
  Machine Learning Engineer — *August 2019 - September 2022*
  - Led the development and **implementation of an ETL pipeline using PostgreSQL and Django**, including a reporting system and data visualizations using HTML, CSS, and JavaScript
  - Built an **end-to-end system for collecting and labeling doctor's handwritten prescriptions**, achieving a **90 percent accuracy rate in offline recognition**
  - Leveraged collected data to enhance **handwriting and character recognition on doctor's prescriptions** through the utilization of CNN, RNN, and CTC techniques
  - **Implemented personalized recommendation models for doctors** based on usage data
  - Contributed to the **development of a Clinical Decision Support System** for the EHR system using SNOMED CT Database and healthcare data
  - **Created data pipelines and automated the reporting process** and data management for thousands of doctors and millions of patients
  - **Developed APIs** for insights and and analytics dashboards for doctors
  - Led the design and implementation of architecture for various products to ensure scalability and proper data reconciliation

  Technologies: Python, SQL, Django, Tensorflow, Docker, GCP, AWS, Javascript, MongoDB.
  Theory: Data Preprocessing, Statistical Modeling, Deep Learning, Linear Algebra, NLP, Recommendation systems.

Education

- **Guru Gobind Singh Indraprastha University** — New Delhi, India
  B.Tech. Computer Science and Engineering *(7.37 CGPA)* — *July. 2015 – June. 2019*

Certifications

- **Applied AI Course**

Fundamentals of Programming	*Issued August 2021*
Data Science: Exploratory Data Analysis and Data Visualization	*Issued October 2021*
Foundations of Natural Language Processing and Machine Learning	*Issued January 2022*
Machine Learning - II (Supervised Learning Models	*Issued February 2022*
Feature Engineering, Productionization and deployment of ML Models	*Issued March 2022*
Machine Learning Real-World Case Studies	*Issued April 2022*
Data Mining(Unsupervised Learning) and Recommender Systems + Real-World Case	*Issued May 2022*
Neural Networks, Computer Vision and Deep	*Issued June 2022*
Deep Learning Real-World Case Studies	*Issued July 2022*

Projects

- **Building a Machine Learning Pipeline for Image Classification using Apache Kafka and Neural Networks**
  - Developed a machine pipeline using Apache Kafka for classifying fashion clothing images using deep neural networks
  - Utilized the MNIST Fashion dataset for training a convolutional neural network
  - Built a complete pipeline for the application using Kafka, including a producer to generate predictions and a consumer to consume the outputs
  - Created a retraining pipeline by setting a threshold for predictions and retraining the model
  - Utilized Hyperopt for hyperparameter tuning of the models

- **Development of a Captcha Solver using Computer Vision Techniques**
  - Developed a captcha solver using computer vision techniques, including image preprocessing, thresholding, and contour detection
  - Implemented a convolutional neural network to recognize individual characters within the captcha
  - Utilized the OpenCV library for image processing and contour detection
  - Improved captcha solving accuracy through the use of image preprocessing techniques and neural network training
  - Created a end-to-end pipeline for solving captchas using the developed model

- **Development of a Facial Detection and Recognition System using OpenCV and Machine Learning Algorithms**
  - Designed and implemented a facial detection and recognition system using Python and the OpenCV library
  - Utilized Haar cascade classifier for facial detection in images
  - Constructed a dataset of facial images by capturing them through a laptop webcam
  - Implemented multiple classification algorithms like EigenFaces, FisherFaces, and LBPHFaces for recognizing faces in the dataset
  - Developed a script to perform live facial detection and recognition through a laptop webcam

Achievements

- Kaggle Expert
- Achieved top 40 ranking out of 6000 students in the ZS Data Science Challenge 2018.

Other programming tools

- **Cloud services:** AWS(EC2, S3), GCP.
- **Deep learning frameworks:** Keras, Tensorflow, Pytorch.
- **Database:** PostgreSQL, MongoDB, ElasticSearch.
- **Python:** numpy, pandas, sci-kit, django.
- **Webserver:** Gunicorn, Apache, Docker, Kafka.