

Sahil Kumar

sahil210695@gmail.com
+918447526767

[LinkedIn](#)
[GitHub](#)

Data Scientist
Delhi, India

SUMMARY

- Data Scientist with experience in executing data-driven solutions to increase efficiency and solving challenging business problems
- Highly skilled in machine learning, deep learning, computer vision, python, and deploying data science solutions

SKILLS

- | | | | |
|-----------------|----------------------|-------------------------|-------------------|
| • Python | • Machine Learning | • Deep Learning | • Computer Vision |
| • Pandas | • Data Visualization | • TensorFlow | • MongoDB |
| • NumPy | • Scikit Learn | • TensorFlow Serving | • SQL |
| • Docker | • Seaborn | • OpenCV | • MLflow |
| • Git | • Geospatial Data | • Web Scraping | • Nvidia Digits |
| • Flask | | • Google Cloud Platform | • Streamlit |
| • Data Analysis | | | • JupyterLab |

WORK EXPERIENCE

Data Scientist

Asset Management, UBS

Sep 2019 – Current

- Developed an Integrated framework for the Systematic Investment team for climate aware fund that looks at investments by reducing portfolios' carbon footprint, offering the ability to invest in new technologies, and helping investors align their portfolios to their chosen climate glide path
 - Generate Quantitative Factors by applying statistical transformation like winsorization, nanmean, and nanmedian using scalable sklearn pipeline architecture
 - Portfolio Optimization with objectives and constraints
- Built OCR pipeline using tesseract and python to convert scanned documents into searchable and trackable text documents which reduce the time and efforts required for the manual task
- Developed a collaborative data science platform for scalable model development and tracking to be used by data science teams across UBS AM Team using JupyterLab with python and spark kernels, MLflow, and Git
- **AutoDQ:** In-house dashboard a tool to visualize data from multiple sources over time to see the drift of point in time data with respect to multiple factors compared to data's benchmark over time

- Dashboard to visualize earnings call transcripts sentiment data for multiple teams of portfolio managers and research analysts which will help them to track their companies of interest and see the trend in the company's data over time with respect to multiple key drivers

Data Scientist

iGnite Labs, Tata Consultancy Services

Jun 2015 – Jun 2019

- Created a financial spreading automation tool in which extracted data from a large collection of financial statements (PDF format) and built a text classifier using tf-idf and SVM in python to map extracted data for a large UK client. Savings of time and cost associated with the manual process
- Built a deep learning-based platform to detect and count the number of hamlets (small human settlement) in the satellite imagery using YOLOv2 in python. An intelligent way of data acquisition and verification of census data for the Indian Government
- Created an AI-based application that provides the nutritional content of the food images by classifying the images using transfer learning on Inception V3 with TensorFlow and deployed using TensorFlow-serving to serve trained model
- Created a generic web-based platform to build multilingual chatbots using AIML (Artificial Intelligence Markup Language) which also gives a plugin to incorporate the bot in other apps
- A data-visualization portal that visually represents demographic details, Socio-Economic, Education, Skills, Agriculture, Religions in India at the state, district, town, and village level. Data collection was done by scraping, parsing Google earth data, OSM, Census, and from various other Government sources
- Trained a team of 24 employees on Deep Learning, Machine Learning, and Python

EDUCATION

Master of Computer Applications

Sastra University

Bachelor of Science

University of Delhi

CERTIFICATIONS

- Deep Learning, a 5-course specialization by deeplearning.ai on Coursera (04/2019)
- Fundamentals of Deep Learning for Computer Vision on NVIDIA Deep Learning Institute (02/2019)
- Google Cloud Platform Big Data and Machine Learning Fundamentals by Google Cloud on Coursera (02/2018)
- Google Cloud Platform Fundamentals: Core Infrastructure by Google Cloud on Coursera (02/2018)
- Machine Learning by Stanford University on Coursera (09/2017)