

Siddharth Tandon

AI Architect/ Applied Scientist/ Data Scientist

SUMMARY

Results-driven AI/ML Specialist with over 14 years of experience in developing and deploying machine learning models. Expertise in data analysis, algorithm development, statistical models and project management, with a commitment to leverage technology for impactful solutions. Proven track record of solving complex challenges and driving innovation through advanced analytical skills and strategic problem-solving.

EMPLOYMENT

Strategic Cloud Engineer, Google

JUN 2022 - PRESENT

- Led the NI2SQL implementation for a customer using ADK framework.
- Led the program for entire team to enhance fine tuning expertise (using LoRA, QLoRA, Vertex AI) within team, and guided more than 10+ pre-sales for customers.
- Presented a poster in Google Research Conference 2023 for "Explicit content detection using large vision models with contrastive objective". This also got published as one of the case [studies](#) in public domain.
- Distributed training/ inferencing of large language models on TPU/ GPUs using [lingvo framework](#).
- Architected AI solutions for strategic customers.
- Lead and implemented multiple Gen AI solutions across various use cases like summarization, RAG, etc.
- Build and deployed multiple agentic systems leveraging frameworks like Langchain, Langgraph.
- Deep expertise on Data Valuation using SHAP methodology.

AI & ML Solution Architect, Microsoft

APR 2020 - JUN 2022

- Implemented yield forecasting to assist in supply chain by leveraging vegetation indices and land images from satellite data.
- Implement ML models/ MLOPS to help customers in their AI journey.
- Implemented deep learning models to detect defects in parts (using transfer learning), brain segmentation using FCNNs.
- Implemented predictive/ forecasting models for manufacturing customers.
- Leveraged cognitive APIs for search, text translation for various use cases.
- Certified Microsoft Enterprise Data Science practitioner/ proctor.
- Drive, define, and implement AI initiatives for customers on Azure.
- Build end to end solution architecture for large scale technology projects.
- Implement/ assist customers with big data solutions to digitize their data estate.
- Manage client stakeholders, and their expectations, with regular cadence of weekly meetings and status updates.

DETAILS

(+91) 905 269 8726

siddharth.tandon1@gmail.com

LINKS

[Google Scholar](#)

[GitHub](#)

[LinkedIn](#)

SKILLS

Generative AI

Agentic AI

RAG

Text Generation

Image Generation

Machine Learning

Probabilistic Models

Non-Probabilistic Models

Linear and Logistic Regression

Trees and Ensembling

SVM

Neural Networks

Clustering

Regularization

Dimensionality Reduction

Linear Algebra

Deep Learning

CNN

NLP

Image Classification

Image Segmentation

Large Language Models

Generative AI

Inference Optimization

Frameworks/ Languages

Python

SQL

Pytorch

ADK

Langchain

Langgraph

Reinforcement Learning

- Define business case for technological investments to be made for future needs and to remain competitive.

Consultant, Microsoft

NOV 2014 – MAR 2020

- Architected and implemented big data solutions using Delta, Lambda and Kappa architecture leveraging Spark framework and Azure technologies.
- Assisted data scientists to build ML models like detect fraudulent transactions, lead generation for customers, etc.
- Lead multiple Data & AI projects and moved them to production successfully.
- Delivered sessions and mentored peers on big data technology landscape.
- Assisted in the recruitment process for the team.

Sr. Software Engg., MAQ Software

JUL 2011 – OCT 2014

- Implemented data warehouses and OLAP cubes.
- Collaborate with stakeholders/ clients to gather and analyze business requirements.
- Actively involved in effort estimation, deployment plans, test plans and peer code reviews.
- Break the requirements into tasks and allocate them within the team.
- Responsible for delivering the project within a specified time frame.
- Analyze changes and incorporate them in sprint as per the bandwidth and complexity.

EDUCATION

IIT, Hyderabad MTech
2020 – 2023

AKGEC BTech
2007 – 2011

PAPERS

FW-SHAPLEY: REAL-TIME ESTIMATION OF WEIGHTED SHAPLEY VALUES (ICASSP 2024)

- Accepted in ICASSP 2024 under Explainable and Interpretable Machine Learning area.
- Introduced an amortized framework for weighted Shapley value computation for feature attribution and data valuation setting in real-time via a learned estimator without actual target variable.

Explicit content detection (Google Research 2023)

- Detect explicit content using large vision models with contrastive objective.

Dynamic Programming
Monte carlo methods
TD methods
Deep Q Networks
RLHF/ RLAIF
Policy gradient (GRPO, DPO)

Use cases

Predictive Modelling
Forecasting
Regression and classification
Defect Detection
Advanced Search

CERTIFICATES

Google Cloud Professional Machine Learning Engineer
Microsoft Certified: Enterprise Data Science Practitioner
Microsoft Certified: Azure Data Scientist Associate
Microsoft Certified: Azure Data Engineer Associate
Microsoft Certified: Azure AI Engineer Associate