SACHIN VERMA

+91 (886) 188-8711 ⋄ Gurugram, India

vermasachin723@yahoo.com ⋄ linkedin.com/in/vermasachinds/

PROFESSIONAL SUMMARY

AI/Data Science professional with 9 years of experience, including 2+ years in LLMs and Generative AI. Led multiple ML and LLM-based projects across domains like finance, healthcare, compliance, commodities, transport and maritime. Experienced in team leadership, cross-functional collaboration, and stakeholder communication. Skilled in prompt engineering, RAG pipelines, and building AI agents. Hands-on with NLP, OCR, deep learning, and statistical modeling. Strong expertise in Python, SQL, PySpark, AWS, and MLOps practices. Deployed scalable end-to-end solutions with measurable impact and reduced manual efforts. Continuously upskilling via Kaggle.

SKILLS

- **Gen AI:** Language Modeling, LLMs, Prompt Engineering & Tuning, Advanced RAG, AI Agents.
- **Data Science:** Deep Learning, NLP, OCR, Statistical Modeling, Machine Learning, MLOps, Streamlit
- **Data Engineering:** ETL, Cloud Engineering, Pipeline automation
- **Languages:** Python, R, SQL, PL/SQL, HiveQL
- **Frameworks:** LangChain, Llama Index, PySpark, PyTorch, Power BI, AWS, Azure DevOps, Docker, REST API, CI/CD
- **Soft Skills:** Leadership & Mentorship, Stakeholder Communication, Cross-functional Collaboration, Conflict Resolution, Decision-Making Under Pressure

HIGHEST EDUCATION

- **M.Sc, Statistics:** Indian Institute of Technology Kanpur(IITK), India — 2014 - 2016
- **B.Sc, Statistics, Mathematics:** University of Lucknow, India — 2011 - 2014

WORK EXPERIENCE

Sr. Data Science - Sr. Team Lead — April 2022 - Till Date
S&P Global Commodity Insights — *Bangalore, India*

- Led an LLM-based project to extract data from documents. Achieved 85% accuracy and 70% less manual effort.
- Leading a team to develop & maintain flagship market leading shipping products along with essential client facing APIs and DevOps pipeline.
- Cross functional leadership collaboration with data providers, engineering, research team and product managers to ensure seamless ingestion and timely delivery.
- Spearheaded development of the ETL data ingestion platform powering **25+ diverse data sources (PDF, web, Excel, APIs)**. Designed monitoring systems for extraction health and resolution workflows, ensuring **99%** data availability and reliability for the flagship shipping analytics product.
- Implemented ML model to predict commodity grades information from unstructured commodity descriptions coming from bill-of-lading dataset **100% less manual effort.**

Methods & tech stacks used: Statistics, Machine Learning, Deep Learning, Python, MLOps, AWS EC2, S3, cloudfront, lambda, Pyspark, R, Prompt Engineering, ETL pipelines

Associate Manager - FCC Data Science — March 2020 - March 2022
Standard Chartered GBS — *Bangalore, India*

- Designed and built data science ML solution in python for suspicious transaction screening, **60% time saved.**
- Lead the team of scenario tuning framework and trained junior employees.
- Managed scenario tuning for AML model for suspicious transactions.
- Worked on 43 AML scenarios and their migration from PL/SQL to HiveQL.

Methods & tech stacks used: Statistics, Machine Learning, Python, PL/SQL, HiveQL, SQL, MLOps

Data Scientist   August 2016 - March 2020

IBM India Pvt. Ltd.   *Bangalore, India*

- Engineered a tool to streamline project allocation for the idle employees using text classification and gained efficiency in process, **40% time saved in allocation.**
- Developed and deployed an NLP and deep learning-based framework using LSTM and fine-tuned BERT to automate tech support ticket analysis, achieving high accuracy, securing a $200K deal, and enhancing client satisfaction across industries.
- Worked in a product sales prediction modeling project for a big pharmaceutical client.
- Designed a solution for creating life cycle prediction of the electric pole on a subway grid pole.
- Designed and deployed the prediction model for the electricity load movement in north India region.
- Worked in a development project for a shipping client to develop a back end to extract data from images and populate on dashboard, **80% manual intervention reduced.**

Methods & tech stacks used: Deep learning for NLP, Machine Learning, BERT, Docker, REST API, R, Rshiny

CURRENT NOTICE PERIOD

- 2 months (official) - negotiable

CERTIFICATIONS

- Applied Generative AI - Johns Hopkins University ↗
- Complete Guide to Building, Deploying, and Optimizing Generative AI with Lang-chain and Huggingface ↗

EXTRA-CURRICULAR ACTIVITIES

- Actively participating in Kaggle competitions to constantly up skilling and keeping up to date with industry knowledge.↗