

WORK EXPERIENCE

Data Scientist, Grainger Canada, Vancouver 2022 – present

- Led the development and integration of end-to-end ML services within the existing SOLR-based infrastructure for Grainger Canada's website search, deploying production-ready models to address business-critical functions
- Deployed an ML API to evaluate parsed search queries using LLM parsed queries, leveraging **LLM Validator** with LangBear for prompt management, significantly reducing assessment time for incorrect query labels
- Refactored the product category prediction **fastText** model from an on-prem server to the cloud using **MLflow** and Databricks, reducing model training time by 37% and improving recall by 8%
- Launched Taskify, a **Streamlit** app leveraging Azure-OpenAI API and product information to create tasks and curate necessary products, driving product discovery
- Developed an API testing framework using GitHub **Copilot** to generate unit tests and provide in-line assistance
- Executed health check scripts and stress testing for REST APIs using Locust, implemented efficient logging with Splunk, and utilized Grafana dashboards to monitor model drift and performance

Data Analytics Manager, SBI Funds Management Pvt. Ltd., Mumbai 2018 – 21

- Developed and fine-tuned **FinBERT** for classifying sentiment of 20,000 user reviews, integrating the result in the internal CRM tool, thereby improving the support team's response time to negative feedback
- Implemented clustering and segmentation for over 10 million investors and 37,000 brokers using **K-Means** and **DBSCAN**, for sending targeted notifications on app, increasing app usage by 25%
- Implemented item-item collaborative filtering model to recommend products to customers in the broker platform for upselling and cross-selling schemes, leading to an 11% increase in digital sales
- Built Power BI visualization reports with data from SQL Server & Google Analytics within an ETL framework

EDUCATION

Master of Data Science, University of British Columbia, Vancouver, Canada 2021 – 22

M.Sc. Mathematics, Indian Institute of Technology, Bombay, India 2016 – 18

B.Sc. Mathematics, Hindu College, University of Delhi, India 2013 – 16

ACADEMIC PROJECTS

Head Impact Detection in Sports Videos, UBC Vancouver Jun 22

- Generated a reproducible machine learning pipeline in Google Colab using **computer vision** algorithms to detect the occurrence of head impact event using a combination of **Object-detection** and **Activity-recognition** models
- Using **YOLOv5** & **Resnet 3D**, the manual analysis time for videos decreased by 60%, achieving a recall score of 0.91

TECHNICAL SKILLS

Programming Languages & Databases – Python, R, SQL, PostgreSQL, MongoDB, C++, Spark, Parquet

ML Frameworks & Tools – PyTorch, TensorFlow, HuggingFace, Langchain, Scikit-learn, OpenCV

MLOps Toolkit: Databricks, AWS, Bash, Docker, Kubernetes, Git, Airflow, CircleCI, Helm, Flask, JIRA

ACHIEVEMENTS

- Top 2% (16/829) in Kaggle Team competition, WiDS Hackathon 2022 | Forecast Energy Consumption Mar 23
- Top 3% (87/3537, Solo Silver) in Kaggle competition, PetFinder.my | CNN | Predict Image Popularity Jan 23
- Received DST-INSPIRE scholarship (Top 1% in India ISC Exams) from Govt. of India 2018

ADDITIONAL LEARNING COURSES

- **Certifications:** Databricks Generative AI Fundamentals, Databricks Certified Machine Learning Associate, Databricks Lakehouse Fundamentals, SQL Intermediate (HackerRank), Deep Learning Specialization (Coursera), Building Transformer-Based NLP Applications (NVIDIA)
- Graduate Certificate of Advanced Public Policy with Merit, The Takshashila Institution, Bengaluru, India Aug 21
- Completed 100 hours of French language course run by Ambassade de france en inde equivalent to A1 Apr 17