

Sourabh R. DATTAWAD

✉ sourabhdattawad@gmail.com | ☎ (+49)152-313-88640 | in sourabhdattawad | 🌐 sourabhdattawad |
🌐 sourabhdattawad | 📍 Stuttgart, BW

OBJECTIVE: I am a Machine Learning Engineer with over 4 years of work experience. Presently, I am enrolled in a master's program in computational linguistics (NLP) at the University of Stuttgart. I am seeking part-time working student opportunities, particularly in the fields of Large Language Models (LLMs) and Generative Artificial Intelligence (GenAI).

EDUCATION

Computational Linguistics (NLP) — *Master of Science* APR 2023 - CURRENT
Institute for Natural Language Processing GPA: 2.1
Universität Stuttgart, Stuttgart, BW

Computer Science — *Bachelor of Science* AUG 2014 - MAY 2018
Visvesvaraya Technological University (VTU), Karnataka, India GPA: Distinction

SKILLS

- Programming Languages: Python
- Data Engineering: Apache - Pyspark, Flink, Hive, AirFlow
- Cloud: Amazon Web Services (AWS) - S3, Lambda, Kinesis, SageMaker
- Databases: SQL/No-SQL/Vector Databases, Weaviate, Elasticsearch, PostgreSQL, MongoDB, InfluxDB
- NLP & Machine Learning: Large Language Models (LLMs), OpenAI API, HuggingFace, LangChain, Prompt Engineering, RAGs, Spacy, ScikitLearn, Numpy, Tensorflow, Pytorch, NLTK, Pandas
- Visualization & Monitoring: Grafana, Metabase, Kibana, Plotly, Matplotlib, Weights & Biases
- CI/CD: Git, Gitlab

WORK EXPERIENCE

Cisco Systems — *Machine Learning Engineer - 3* OCT 2020 - MAR 2023

- Developed an algorithm to classify user devices into groups for efficient network management at a scale of 2M+ devices per week.
- Designed a deep learning based spoofing detection model with accuracies ranging from 80%-92% across different device classes like phones, printers, cameras etc.
- Deployed ETL jobs to process over 20 Million data points every day using Pyspark and Airflow.
- Built grafana dashboards to visualize predictions, user interactions and tracking model quality in production.
- Improved unit test coverage by 25% in Pytest and further implemented integration tests to check the correctness of ML systems.
- Implemented Hive queries to monitor customer data quality in Looker, resulting in data quality improvement in the range of 20-50% for 1000+ customers.

Mercedes-Benz R&D India — *Machine Learning Engineer* AUG 2018 - OCT 2020

- Developed a machine learning algorithm for identifying high business impact incidents with an accuracy of 90+%, reducing production downtime by at least 50 hours per month.
- Created context-based embeddings for 2000+ entities involved in the production of automobiles to compare and prioritize critical systems.
- Implemented topic modelling and issue extraction using data mining and NLP techniques involving Mercedes-Me app review data and consumer complaint data for efficient problem management.

Indian Council of Medical Research — *Machine Learning Intern* JAN 2018 - JULY 2018

- Analyzed dengue disease trend by identifying the top 3 environmental factors that influence the growth of the virus.
- Developed an ensemble of models to predict the type of cancer with a 90% accuracy(F-beta) score along with model explainability using LIME.

Sportzify — *Software Engineer Intern* OCT 2015 - MAR 2017

- Developed and integrated a real-time notification system using Socket.io to streamline communication and enhance user experience.
- Integrated payment gateway and developed a comprehensive service for managing referrals, coupons, discounts, refunds, etc., facilitating over 1000 monthly transactions.

TECHNICAL PROJECTS

Personal Project — *Expressive Resume* NOV 2021 - DEC 2021

- Created a **LaTeX** class that makes it easy for anyone to quickly create a beautiful resume and cover letter.

PUBLICATIONS & CONFERENCES

Deep learning for Computer Vision and Natural Language Processing — *FOSSCON* - JULY 2019
Gogte Institute of Technology, Belagavi, India

How vectorizing business-critical entities help solve operational problems — *BAICONF* - DEC 2019
Indian Institute of Management, Bengaluru, India

AWARDS

Innovation Award at Mercedes-Benz Analytics — *Mercedes-Benz R & D India* - OCT 2019

Outstanding Performer at Mercedes-Benz Analytics — *Mercedes-Benz R & D India* - DEC 2019

You Amaze Award for Outstanding Performance — *Cisco Systems* - OCT 2021

LANGUAGE SKILLS

English (C1) — Hindi (C1) — Kannada (C2) — German (A1) — Sanskrit (Hobby)