# SWATI GUPTA

## DATA SCIENTIST | ML ENGINEER | AI ENGINEER

London, United Kingdom | +44 7771084480 | swatibadrigupta@gmail.com | LinkedIn | GitHub

---

## PROFESSIONAL SUMMARY

As a Data Scientist, I specialize in leveraging AI/ML and large language models (LLMs) to drive impact at scale. In my current role at STOP THE TRAFFIK, I apply these skills to address one of society's most critical challenges – combating human trafficking and modern slavery. I have end-to-end experience taking AI projects from idea to proof of concept (PoC) and through to deployment in the cloud. With a strong focus on optimizing machine learning models, scaling global data pipelines, and building AI-powered solutions, I've led initiatives that improved data accuracy by 26% and partnered with leading organizations such as AWS and IBM.

I bring hands-on expertise in Generative AI and NLP, specialising in translating complex unstructured regulatory, legal, and fraud-related data into actionable models to support business decision-making. I have delivered GenAI and NLP solutions across finance, intelligence, tech, and operations teams, driving real-world business impact through AI integration, structured and unstructured data modelling, and cross-functional collaboration. Passionate about uncovering hidden patterns in data and translating them into strategic action, I am now seeking opportunities to apply my expertise to help organizations solve high-impact problems through data-driven innovation.

---

## EDUCATION

**University of Sussex | Brighton, UK | Sep 2022 - Nov 2023**
MSc in Data Science | Distinction
- Key Modules: Machine Learning, Natural Language Processing, Algorithmic Data Science, Web Applications
- Dissertation: Empowering Sign Language Communication Through Image Based Recognition, Domain Specific Language Modelling

**R.C.O.E.M. Nagpur University | Nagpur, India | Aug 2014 - Aug 2018**
B.E. in Information Technology | Distinction
- Key Modules: Database Management System, Web Technologies, Object Oriented Programming, Data Structures

---

## SKILLS

- **Programming Languages and Microservices :** Python, Flask, Django
- **Database Management :** SQL, MongoDB
- **Machine Learning :** Supervised and Unsupervised Learning, Classification, Regression, Numpy, Pandas, Matplotlib, Seaborn, OpenCV, Prompt Engineering, Web Scraping, NLP, Gensim, SpaCy
- **Deep Learning :** Neural Networks (CNN, RNN, LSTM)
- **Natural Language Processing :** Transformers, Text Classification, Named Entity Recognition, LLM, GEN AI
- **Libraries:** Scikit-Learn, PyTorch, Tensorflow, Keras, Hugging Face
- **Cloud Platforms :** AWS, IBM Cloud
- **Version Control and Tools :** Git, JIRA, Swagger, Postman, Jupyter Notebook, Pycharm, VS Code
- **Data Analysis & Statistics :** Data Cleaning, Wrangling, Visualisation, Pipelines, Power BI
- **Soft Skills :** Presentation, Collaboration, Communication Skills, Strong Interpersonal Skills, Data Storytelling

---

## EXPERIENCE

**Data Scientist / AI Engineer**
**STT - Projects | Stop The Traffik | London, UK | Feb 2024 - Present**

- Led 5 AI/ML projects powering the Traffic Analysis Hub to combat human trafficking globally, including entity extraction, LLM-based classification, and multilingual data processing.
- Increased multilingual data coverage by 33% by integrating 15+ new regions and languages, leveraging LLMs to translate and extract insights from Human Trafficking and Modern Slavery (HTMS) datasets.

- Delivered GenAI and NLP solutions across the organization, translating complex unstructured regulatory, legal, and fraud-related data into structured, task-based AI models to support business decision-making; collaborated cross-functionally with finance, intelligence, tech, and operations teams to implement scalable real-world solutions.
- Boosted model accuracy by 26% by optimizing classification pipelines, feature engineering, and retraining ML models on structured and unstructured datasets.
- Designed and deployed LLM-powered NER pipelines using SpaCy and transformer models, enhancing recognition of trafficking-related entities across multiple languages.
- Built 6+ interactive dashboards in Power BI to visualize trafficking patterns, helping stakeholders identify risks, trends, and actionable interventions.
- Managed a team of 10+ across data annotation, web scraping, and data engineering projects—ensuring timely delivery of high-quality data assets.
- Automated annotation processes using LLMs, leading to a 40% reduction in manual effort and consistent high-quality training data.
- Developed AI chatbots and agents with AWS, enabling real-time data interaction and reporting for NGOs and partner organizations.
- Established gold-standard annotation guidelines and rule-based systems to structure open-source data for machine learning pipelines.
- Collaborated with global partners (including AWS & IBM) to co-develop scalable, ethical AI solutions and intelligence-sharing frameworks.
- Supported 4+ technical grant proposals and solution architectures, advocating LLM integration to solve real-world social impact challenges.
- Contributed to NLP models detecting fraudulent job ads in partnership with IBM, using named entity recognition and labour exploitation markers.
- Led talent acquisition and onboarding, expanding the data science and AI engineering team to meet growing project demands.

**Data Scientist**
**Wipro IP - Webinar BOT | Wipro | Bangalore, India | May 2021 - Feb 2022**

- Built a classification model to categorize web-scraped data, saving 1,460 hours annually and enhancing decision-making. Leveraged Flask microservices, REST APIs, and NLP to process insights from webinars, delivering valuable data to marketing teams.

**Data Scientist**
**Wipro IP - IntelliAssure ML Classifier for Defect Log Categorization | Wipro | Bangalore, India | April 2020 - April 2021**

- IntelliAssure is a Wipro IP having ML Classifier to Categorize Defect Logs using Machine Learning Classification techniques like Logistic Regression, Random Forest, SVM.
- Reduced manual human efforts by 18%, resulting in significant time and cost savings.
- Conducted rigorous experimentation to optimize the performance of the algorithms.

**Project Engineer**
**Wipro IP - Intelligent Talent Engineering Platform | Wipro | Pune, India | June 2018 - March 2020**

- Developed a Python script which resulted in enhanced workforce management, enabling the matching of future demand and planned ramp downs based on specified band criteria.
- Empowered decision-makers with accurate skill requirement forecasting and demand analysis 30 to 90 days in advance, leveraging historical data and understanding influencing factors. Facilitated strategic resource allocation and optimised workforce planning, enabling proactive matching of resources for future demands and planned ramp downs.

**ACCOMPLISHMENTS**

- Customer Centricity Award by Wipro Technologies for extra ordinary work
- Certificate of Awesomeness for contribution of work in Make a Difference