

# Arijit Gupta

Website | [arijitgupta2000@gmail.com](mailto:arijitgupta2000@gmail.com) | +447342887938 | Github | LinkedIn

## EDUCATION

### THE UNIVERSITY OF SHEFFIELD

MSc. COMPUTER SCIENCE WITH SPEECH AND LANGUAGE PROCESSING  
September 2024 - Present | Sheffield, UK

Classification: Distinction (Predicted)

### BITS PILANI

BACHELOR OF ENGINEERING ELECTRONICS AND INSTRUMENTATION

August 2018 - May 2022 | Goa, India

Classification: First Class

## EXPERIENCE

### ZS ASSOCIATES | SENIOR DATA SCIENCE ASSOCIATE

Jul 2022 - Aug 2024 | Bangalore, India

- Built scalable NLP pipelines to extract AI-driven insights from large unstructured textual corpora, saving 100+ man-hours per week
- Developed a segmentation pipeline using Haystack, implementing clustering, extractive QA, and fine-tuned few-shot relation extraction to analyze 10K+ physician notes and identify key decision drivers
- Built and optimized a PyTorch-based LLM solution with Flan-T5 to analyze drug side effects, uncovering co-occurrence patterns across 5K+ patient reports and identifying socio-economic factors affecting patient reluctance

### MICROSOFT | SOFTWARE ENGINEER INTERN

May 2020 - Jul 2020 | May 2021 - Jul 2021 | Hyderabad, India

- Worked on the Microsoft Search, Assistance, and Intelligence team, adding voice skills to a state-of-the-art chatbot. Built new features using TypeScript and Express for bot communication, and YAML for training data
- As part of the Skype team, created a pipeline to automatically migrate a static website from C# to a dynamic ReactJS-based site with real-time component updates via systematic API polling, increasing future conversion efficiency by 50%

## RESEARCH

### THE UNIVERSITY OF EDINBURGH | VISITING RESEARCH STUDENT

Aug 2021 - May 2022 | Edinburgh, UK

- Worked with Prof. Kenny Smith at The Center for Language Evolution on grammatical cues in language for noun gender prediction
- Experimented extensively on the usability of contextual and formal types of vector word embeddings in a task of noun gender prediction

## SELECT PROJECT

### RASOEE | PYTHON, DJANGO, PYTORCH | 🍌

Aug 2020

- Rasoe is an application for food dish identification using image classification.
- Performed extensive model comparison to find the most efficient and lightweight architecture
- Scraped and cleaned data for 308 unique food classes. The Kotlin-based mobile app runs entirely on-device without an internet connection.
- The website is built with Django and hosted on Heroku

## SKILLS

### PROGRAMMING

Python • C++ • C# • JavaScript •  $\LaTeX$  • MATLAB • TypeScript • Shell



### FRAMEWORKS

LangChain • PyTorch • TensorFlow • Keras • Angular • ReactJS

### TOOLS

Git • NumPy • OpenCV • Pandas • Scikit-Learn

## PUBLICATIONS

- Transformer-based Models for Supervised Monocular Depth Estimation | 
- Towards Using Diachronic Distributed Word Representations as Models of Lexical Development | 

## SCHOLARSHIPS AND COMPETITIONS

2017 UPSTSE Scholarship - Govt. of India  
- Top 1000 students from my home state

2020 PyTorch Global Summer Hackathon  
- First Runners-Up out of 2,500 participants from 114 countries

2021 OpenCV Spatial AI Competition -  
Top 210 finalists from 1500+ worldwide

## COURSEWORK

ML and Adaptive Intelligence  
Natural Language Processing  
Speech Technology  
Scalable Machine Learning  
Text Processing  
Speech Processing  
Team Software Project  
OOP Programming  
Data Structures and Algorithms  
Discrete Mathematics

## CONFERENCES

ACL 2022 - Student Research Workshop  
Research Week with Google  
International CCCP Symposium - Cognition, Communication, Computation, and Perception