Shataxi Dubey

✉ dubey.shataxi3010@gmail.com   🔗 website   in shataxi-dubey   ○ shataxiDubey

Education

M.Tech Computer Science & Engineering   Jul 2023 – June 2025
- Advisor: Prof. Nipun Batra
- Indian Institute of Technology Gandhinagar
- CPI: 9.71/10

B.Tech Computer Science & Engineering   Jul 2018 – June 2022
- University of Allahabad
- CPI: 9.43/10

Experience

IndiaMart InterMesh Ltd, Associate Engineer   Noida, Uttar Pradesh
- Developed a Flask API to load model pickle files, enabling the generation of smart replies for user interactions.   May 2022 – Jun 2023
- Implemented concurrent execution of processes using threading, reducing the time from 1 sec to 600ms.
- Generated smart replies using TF-IDF and Kmeans & glove and cosine similarity for word embeddings & clustering.
- Integrated API with PostgreSQL database for suggesting text messages based on user reply patterns.
- Analyzed the data of user replies by writing complex SQL queries and created views for data analysis on datamarts.

Indian Institute of Technology Gandhinagar, Teaching Assistant   Gandhinagar, Gujarat
- **Data Structures and Algorithms I:** Guided a group of 30 students on C programming assignments, evaluated their work, and responsibly managed invigilations to ensure the smooth operation of the course.   Jul 2023 – Jun 2025
- **Data-Centeric Computing:** Evaluated assignments and resolved doubts related to numpy, pandas and SQL.
- **Machine Learning:** Conducted viva and tutorial session on convex functions for a class of 150 students.
- **Probability Statistics & Data Visualisation:** Conducted viva, evaluated assignments, and solved students' doubts related to numpy, pandas, torch.

Publications

Space to Policy: Scalable Brick Kiln Detection and Automatic Compliance Monitoring with Geospatial Data

Zeel B Patel, Rishabh Mondal, **Shataxi Dubey**, Suraj Jaiswal, Sarath Guttikunda, Nipun Batra

Accepted in ACM Journal of Computing and Sustainable Societies 2025

10.1145/3736649 ↗

Research Experience

Exploration of Vision Language Models for Brick Kilns detection from Satellite Imagery   Jan 2025 - Jun 2025
M.Tech Thesis
- Implemented GeoChat, PaliGemma, Florence2, Qwen2.5 VL, Yoloe for open vocabulary detection task.

- Achieved consistent performance with Yoloe in domain shift from Lucknow to West Bengal region.
- Implemented Low Rank Adaptation for finetuning the VLMs.
- Made open source contribution by integrating LoRA configuration in Maestro ↗ library developed by Roboflow.
- Curated dataset using Geemap for Lucknow and West Bengal region with Planet imagery.

Space to Policy: Scalable Brick Kiln Detection and Automatic Compliance Monitoring with Geospatial Data   Aug 2024 - Dec 2024

- Successfully detected 30638 kilns in the Indo-Gangetic Plain covering 5 states using Yolo11-l model.
- Achieved 0.94 correlation between detected brick kiln count and the state government survey counts.
- Extracted District-wise count of brick kiln using shapefiles of the state where no survey is conducted till date.
- Implemented the Faster-RCNN detection model and non-max suppression for removal of multiple detections.

Eye in the Sky: Detection and Compliance Monitoring of Brick Kilns using Satellite Imagery   Jan 2024 - Apr 2024
ACM COMPASS 2024

- Presented the paper through poster presentation at ACM COMPASS 2024, held at IIIT Delhi.
- Implemented Yolov8m model on high resolution satellite imagery and successfully identified 19579 brick kilns in the Indo-Gangetic Plain using iterative fine-tuning, conducted compliance monitoring as well.
- Analysis using KD trees shows that 50% kilns violate inter-brick kiln and 16% violate brick kiln-hospital distance.

Projects

Image Caption Generator   Mar 2022- Apr 2022

- Built an LSTM-based model for image caption generation, used pretrained Xception for image feature extraction.
- Used tokenizer for vectorization of captions of images, trained the model on Flickr dataset of 800 images and captions.
- Measured the accuracy of the model using BLEU score,the BLEU score of the model was 0.36.

Diabetes Detection using Machine Learning Techniques   Nov 2021 - Dec 2021

- Performed data preprocessing and performed feature selection for minimum redundancy and maximum relevance.
- Applied supervised machine learning algorithms Logistic Regression, SVM, KNN, Random Forest and XGBoost.
- Achieved accuracy of 82% with logistic regression when relevant features are selected & 80% with Random Forest.

Technologies

Languages: Python, SQL, Fundamentals of C, Javascript, CSS.

Frameworks/Tools: Numpy, Pandas, geopandas, ultralytics, geemap, Unsloth, opencv, NLTK, Sklearn, PyTorch, Git.

ML Algorithms: Machine Learning, Deep Learning, Natural Language Processing, Computer Vision.

Courses: Data Structures and Algorithms, Database Management System, Operating System, Computer Networks.

Achievements

- Awarded third prize for presenting the research work in Simply Speaking 4.0, 2025 at IIT Gandhinagar.
- Awarded Emerging Star of the Month twice, in September 2022 and February 2023 at IndiaMart Intermesh.
- Awarded Star Intern of the Week Award in Rehabilitation Society of the Visually Impaired in June 2021.
- Qualified Bronze Level of Teco Booster track of Code Predators organized by Ramarao Adik Institute of Technology and Mentor Buddy in June 2020.