Purva Parmar

Nationality: Indian   **Gender:** Male   📞 **Phone number:** (+91) 6354943839

✉ **Email address:** thereconpilot@gmail.com   in **LinkedIn:** https://www.linkedin.com/in/purvaparmar/

👤 **GitHub:** https://github.com/TheReconPilot

👤 **Scholar:** https://scholar.google.co.in/citations?hl=en&user=NFANvm4AAAAJ

🌐 **Website:** https://thereconpilot.github.io/

ABOUT ME

MTech AI Student at IISc Bengaluru with a strong foundation in Data Science. Willing to relocate or work remotely.

PUBLICATIONS

[2022]

Object Detection in Indian Food Platters using Transfer Learning with YOLOv4 Object detection is a well-known problem in computer vision. Despite this, its usage and pervasiveness in the traditional Indian food dishes has been limited. Particularly, recognizing Indian food dishes present in a single photo is challenging due to three reasons: 1. Lack of annotated Indian food datasets 2. Non-distinct boundaries between the dishes 3. High intra-class variation. We solve these issues by providing a comprehensively labelled Indian food dataset- IndianFood10, which contains 10 food classes that appear frequently in a staple Indian meal and using transfer learning with YOLOv4 object detector model. Our model is able to achieve an overall mAP score of 91.8% and f1-score of 0.90 for our 10 class dataset. We also provide an extension of our 10 class dataset- IndianFood20, which contains 10 more traditional Indian food classes.

2022 IEEE 38th International Conference on Data Engineering Workshops (ICDEW)

WORK EXPERIENCE

Data Science Intern

AlgoAnalytics Pvt. Ltd. [1 Jul 2022 – 30 Apr 2023]

- Conducted in-depth research on **Semantic Search and Question-Answering Systems** during an internship at AlgoAnalytics Pvt. Ltd. This formed my **Master's Thesis** on the same topic at IISER Pune as well. I primarily worked in the field of **Natural Language Processing (NLP)**.
- Demonstrated **strong initiative and self-reliance**, taking on significant responsibilities and gaining expertise in search systems implementation and deployment.
- Developed the backend for a critical financial product, leveraging NLP framework library Haystack and various **Transformer models** (both Sparse and Dense Retrievers) from **HuggingFace**.
- Conducted **extensive experiments** with databases including Elasticsearch and FAISS, and created a **custom annotated dataset** to evaluate model performance.
- Selected final models based on a thorough assessment considering model scores and evaluation speed.
- Collaborated with a mentor Data Scientist to architect a search application serving a **search API, involving data retrieval, embedding computation, and Semantic Search using Dense Passage Retrievers**.
- Employed **Cloud services** (Azure/AWS) for central database hosting and utilized FAISS for application-specific vector storage.
- Implemented the **search API** with FastAPI, tailored to the front-end requirements
- Maintained **version control with Git** on the company's GitHub repository, engaging in **code review sessions with the mentor** for continuous improvement.
- Implemented **testing with pytest** and **logging**, and conducted **performance profiling** with multiple parallel requests.
- Assumed the role of **product owner** for the Search application, overseeing its complete maintenance.

- Effectively communicated complex technical concepts to non-technical stakeholders, including project managers, through clear and concise **presentations**.
- Conducted **demonstrations** of the API, showcasing its functionality and value proposition to both technical and non-technical team members, fostering a shared understanding of the project's impact and potential.

Summer Research Intern

Indraprastha Institute of Information Technology, Delhi [1 Jun 2020 – 31 Jul 2020]

- Participated in a high-impact Summer Research Internship under the guidance of **Dr. Ganesh Bagler**, renowned for pioneering **Computational Gastronomy**, an innovative fusion of Data Science techniques and the culinary world.
- Collaborated as a key member of a **multidisciplinary team** comprising interns from diverse institutes across India, focusing on implementing **Object Detection with YOLOv4 using Transfer Learning on Indian Food Platter Images.** This enriched my experiences in Computer Vision and Data Science in general.
- Developed a **Python-based web scrape**r leveraging Requests and Selenium libraries to collect images from Google and Instagram, while also conducting data analysis to identify popular Indian food-related hashtags on social media platforms.
- Compiled a list of 10-20 popular and commonly recognized Indian foods and utilized the web scraper to acquire several thousand images for each food category.
- Conducted meticulous **manual image curation and annotation**, collaborating with the team to identify and label relevant bounding boxes for over 12,000 images, using the annotation tool MakeSense.ai.
- Pioneered the creation of the **IndianFood10 dataset**, the largest dataset of its kind, containing thousands of images for 10 distinct Indian food classes, facilitating breakthrough research in Indian cuisine recognition.
- Implemented the **YOLOv4 Object Detection** model with Transfer Learning on the IndianFood10 dataset, achieving remarkable results.
- Co-authored a research paper titled [Object Detection in Indian Food Platters using Transfer Learning with YOLOv4](#), which was presented at the **5th International Workshop on Data Engineering meets Intelligent Food and COoking Recipes** (DECOR@ICDE2022), receiving the prestigious **Best Student Paper Award**.
- The research addressed challenges in Indian cuisine recognition, including a lack of annotated datasets, indistinct boundaries between dishes, and high intra-class variation, achieving an overall mAP score of 91.8% and an f1-score of 0.90 for the 10-class dataset.
- Additionally, contributed to the extension of the IndianFood10 dataset to IndianFood20, incorporating 10 more traditional Indian food classes, further enriching the research community's resources.

Semester Project Student - Mathematics Department

Indian Institute of Science Education and Research, Pune [1 Aug 2021 – 31 Dec 2021]

- Collaborated with **Dr. Anindya Goswami**, an Associate Professor in both Mathematics and Data Science departments at IISER Pune, for an intensive project focused on the **fundamentals of Reinforcement Learning**
- Acquired a comprehensive knowledge of key Reinforcement Learning concepts, including Value Functions, Reward Signals, and Tabular Methods like Cross Entropy, demonstrating proficiency in foundational techniques within the field. Explored Generalized Policy Iteration methods, and also Approximate Solution methods, which can also involve neural networks.
- Explored the critical aspect of Reward Design, recognizing its pivotal role in achieving desired RL agent behavior and understanding potential pitfalls that can arise from improper design choices.
- **Implemented various RL concepts in Python**, demonstrating practical application and coding proficiency in complex machine learning algorithms.
- Maintained digital notes in Markdown and LaTeX using a documentation tool, and hosted the project site on GitHub.
- **Project Site:** https://thereconpilot.github.io/rl-notes/

Semester Project Student - Data Science Department
Indian Institute of Science Education and Research, Pune [1 Jan 2022 – 31 May 2022]

- Collaborated with **Dr. Leelavati Narlikar**, Associate Professor and Deputy Chair of Data Science at IISER Pune, on a comprehensive project focused on **Gaussian Mixture Models (GMMs)**. GMMs can be used for soft **clustering**, a part of **Unsupervised Learning** in **Machine Learning**.
- Partnered with a colleague to explore distinct implementations of GMM clustering, with a focus on **Expectation-Maximization (EM)** method and **Gibbs Sampling** method.
- Mastered the underlying theory and intricacies of GMMs, translating this knowledge into robust **Python implementations from scratch**, showcasing expertise in implementing complex algorithms central to modern Data Science and Machine Learning workflows.
- Conducted **extensive experimentation on various cluster configurations**, including univariate and multivariate clusters, as well as scenarios with unique conditions (e.g., clusters with identical means but different variances).
- Produced **insightful data visualizations** in 2 and 3 dimensions to enhance comprehension and interpretation of clustering results.
- Maintained **meticulous documentation**, encompassing both **theoretical concepts and code implementations**, and utilized **version control with Git** for efficient project management, reflecting **industry-standard practices** essential for collaborative Data Science projects.
- Showcased the project and its findings through a dedicated GitHub site, ensuring accessibility and transparency for fellow researchers and interested parties.
- **Project Site:** https://thereconpilot.github.io/gaussian-mixture-models/

EDUCATION AND TRAINING

MTech in Artificial Intelligence
Indian Institute of Science, Bengaluru [22 Jul 2024 – Current]

City: Bengaluru | **Country:** India | **Website:** https://iisc.ac.in

Bachelor of Science - Master of Science Dual Degree (Mathematics)
Indian Institute of Science Education and Research, Pune [1 Aug 2018 – 31 May 2023]

Address: IISER, Dr. Homi Bhabha Road, Ward No. 8, NCL Colony Pashan, Pune, Maharashtra, India, 411008 Pune (India) | **Website:** https://www.iiserpune.ac.in/ | **Field(s) of study:** Natural sciences, mathematics and statistics:
- Mathematics | **Final grade:** CGPA of 8.1 out of 10 | **Thesis:** Semantic Search and Question-Answering Systems

Academics

- The BSMS Dual Degree Program provided a comprehensive and integrated five-year curriculum encompassing a diverse range of natural science disciplines. This program is equivalent to a traditional 5-year integrated Bachelor of Science - Master of Science program.
- First two years we had compulsory courses from all major natural science disciplines, including some inter-disciplinary courses
- Specialized in Mathematics during the third and fourth years, focusing on advanced courses including Statistical Inference, Graph Theory, Algorithms, Numerical Analysis, Random Processes and Queueing Models, Mathematical Optimization, alongside interdisciplinary subjects like Data Science, Bioinformatics, Non Linear Dynamics, Mathematical and Computational Biology, etc.
- Dedicated the fifth year to an intensive Master's Thesis Project conducted as an intern at AlgoAnalytics Pvt. Ltd., an AI solutions company in Pune, India. Thesis title: "Semantic Search and Question-Answering Systems".

Volunteering and Leadership Experience

- Co-Organizer, TEDxIISERPune 2020
- Organizer, World Cube Association (WCA)-approved Cubing Competitions
    - IISER Pune Cube Open 2019
    - IISER Pune Cube Open 2022

- IISER Pune Summer Open 2023
- Member of the Disha Club, the social services club which helps the community around the institute with educational programs

Engaged in various leadership roles and volunteer activities, demonstrating strong organizational and communication skills. These experiences provided opportunities to work with a diverse student population, fostering cultural understanding and enhancing leadership capabilities.

LANGUAGE SKILLS

Mother tongue(s): Gujarati | Hindi

Other language(s):

English

LISTENING C1 **READING** C2 **WRITING** C2

SPOKEN PRODUCTION C1 **SPOKEN INTERACTION** C1

Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user

DIGITAL SKILLS

Technical Skills

Basic Web Development / Natural language Processing / Machine Learning / Python (advanced - pandas, spacy, transformers, sklearn) / LaTeX / Git / Documentation / Deep Learning / Docker / Command Line / Windows / Cloud Services / Hugging Face / PyTorch / GitHub / Amazon AWS, Azure, OVH / Python / Linux

Soft Skills

Problem Solving / Critical Thinking / Creative Thinking / Data Storytelling / Teamwork & collaboration / Research and Analysis / Adapatability / Project Management / Attention to Detail