

Jinen Setpal

+1 (765)-490-1435 | jsetpal@cs.purdue.edu | github.com/jinensetpal | jinensetpal.github.io

EDUCATION

Purdue University

Bachelor of Science in Data Science

West Lafayette, IN, USA

Aug. 2021 – May 2024

Relevant Coursework: Deep Learning (Graduate), System Security (Graduate), Data Mining & Machine Learning, Embedded Systems, Data Structures & Algorithms, Linear Algebra, Multivariate Calculus, Statistical Programming

EMPLOYMENT

Machine Learning Engineer

DagsHub

Jun. 2022 – Present

Tel Aviv, Israel

- Implemented and deployed open-source data science projects reproducing and extending past research. Examples: [CheXNet](#), [Point-E](#), [YOLOv6](#).
- Developed a data streaming client by monkeypatching Python's `open()` and extending FUSE to lazily pull files from a specified remote using DagsHub's web APIs.
- Building trainer integrations (automatic model, experiment and artifact logging) with [HuggingFace's Transformers](#) library and the [PyCaret](#) framework.

Student Researcher

ARKaNLU @ Purdue University

Jan. 2023 – Present

West Lafayette, IN, USA

- Authored a research proposal which earned a \$250,000 research grant from Amazon Prize.
- Developing BoilerBot, a task-oriented conversational chatbot as part of the Alexa Prize TaskBot Competition.

Systems Developer

Teachiq AB / exam.net

Sep. 2020 – Jul. 2021

Stockholm, Sweden

- Packaged custom security implementations by forking open source `xmodmap(.c)` utility to a node.js module.
- Critical Vulnerability Disclosure for the assessment kiosk on [exam.net](#)'s web client.

TEACHING

Course Instructor

CS 39000 – Web Application Development @ Purdue University

Aug. 2022 – Dec. 2022

West Lafayette, IN, USA

- Curriculum design and course instructor for a two-credit course. Enrollment: 70 students.
- Covered HTML/CSS, JavaScript, React, Node.js, Express.js, MongoDB, Web Security & Cloud Hosting.

Undergraduate Teaching Assistant

STAT 190 – Topics in Statistics for Undergraduates @ Purdue University

Feb. 2022 – May 2022

West Lafayette, IN, USA

- Lab instructor for Purdue's Corporate Partner MISO, developing industry solutions using Data Science.
- Graded assignments, held office hours, conducted code review. Taught classes on git, CI/CD & Data Mining.

PUBLICATIONS

CutLang V2: Advances in a runtime-interpreted analysis description language for HEP data

Frontiers in Big Data, 4, 27, Dr. Gökhan Ünel, et al.

Jul. 2021

CERN, Switzerland

- Setup CI/CD Scripts w/ Automated Email Delivery using GitHub Actions & SendGrid.
- Developed Interpreter Functions through lexical analysis using Flex & Bison (.cpp).

ArchiMeDe @ DankMemes: A New Model Architecture for Meme Detection

7th Evaluation Campaign, Final Workshop, EVALITA 2020. Jinen Setpal, Gabriele Sarti

Dec. 2020

Turin, Italy

- Achieved .7664 F1-Score on test dataset (+.2466 baseline) w/ Video Presentation during final workshop.
- Built a multimodal ensemble using transfer learning through AlexNet, DenseNet & ResNet pre-trained networks.

CONFERENCE PRESENTATIONS

Interpretability Tools as Feedback Loops

Toronto Machine Learning Summit (TMLS) 2022

30th Nov. 2022

Toronto, Canada

PENDING PATENTS

Semi-Supervised Class Activation Mappings for Target Localization & Super-Resolution

Sep. 2021 – Apr. 2022

Final Presentation, TE AI Cup 2022. Jinen Setpal, et al.

TE Corporate, UK

- Won the **Best Innovation Award**, developing subclassed TensorFlow layers for accurate, efficient prediction over classes with minute differences.
- By evaluating feature vectors from the model's penultimate convolutional layer over a dynamic weight threshold, we generate a bounding box to localize the region of the image critical to the final classification.

Leveraging Latent Features for Modular Multiclass Classification

Sep. 2021 – Apr. 2022

Final Presentation, TE AI Cup 2022. Jinen Setpal, et al.

TE Corporate, UK

- Designed & developed a novel modular, scalable architecture for classification achieving .99846 real-data classification accuracy over a +.2466 baseline.
- Exponential Energy, Time and Environmental impact savings with over 195.5 hours and \$1,501.23 USD savings for training the 5,000th classification target.

FUNDED RESEARCH

BoilerBot: Amazon Alexa TaskBot Challenge 2

Jan. 2023 – Present

Amazon Research Grant – PI: Prof. Julia Rayz

Purdue University

- By extending the Open Assistant Toolkit (OAT) developed in TaskBot Challenge 1, we aim patch existing components crucial to customer satisfaction, while building solutions that enhance user experience.
- We aim to focus on novel contributions within: preference and skill elicitation, upstream task resolution, intrastage user interaction as well as building multimodal approaches for incorporating visual aids.

Drone Video Object Recognition

Jan. 2022 – Present

NSF Award 21204301 – PI: Prof. Yung-Hsiang Lu

Purdue University

- Team lead over the Spring 2023 Semester. Our team is leveraging Gazebo, ROS2 & the previous year's scoring function to develop a multi-agent reinforcement learning approach to the sample solution. Won 2nd Place for the **Undergraduate Research Expo Award** under Purdue College of Science.
- Developed an architecture for split-confidence resolution, achieving .9937 test accuracy as part of the reference solution made for the IEEE international autonomous UAV competition. Bootloader patching and setup for linux-based drones with OpenVINO accelerated IoT.

Identifying Cryptographic Functions from Pre-Compiled Binaries

Aug. 2021 – Dec. 2022

NSF Award 2047991 – PI: Prof. Christina Garman

Purdue University

- Employing rudimentary techniques within NLP to establish a baseline approach for reconstructing cryptographic functions from disassembler code used to generate corresponding binaries.
- Evaluating current state-of-art classification tools against rigorous benchmark scripts.

TECHNICAL SKILLS

Languages: Python, C, C++, x86 Assembly, Java, Kotlin, Bash, JavaScript, MATLAB, R, SQL, ROS2

Frameworks: TensorFlow, PyTorch, Keras, NumPy, Pandas, Pillow, ROOT, Matplotlib, FUSE, Node.js, Vue.js

Tools: Git, MLFlow, DVC, Docker, Radare2, Ghidra, TravisCI, GitGuardian, Kubernetes, Gazebo

Cloud Utilities: Google Cloud Console (Compute, Networking, Storage), Amazon Web Services (Redshift, EC2, ECR, S3, Sagemaker, CodePipeline, CodeCommit, CloudWatch), Azure Pipeline, GitHub Actions