

SIDDHI POTDAR

✉ siddhi.potdar2013@gmail.com ☎ +14128165433 💻 [siddhipotdar](https://siddhipotdar.github.io) 🌐 [siddhi-potdar](https://siddhi-potdar.github.io) 🐙 siddhi-potdar.github.io

SUMMARY

Passionate and driven full-stack machine learning engineer with a strong background in computer science and experience in developing, optimizing, and deploying ML models. Seeking an opportunity to contribute to innovative projects and further develop expertise in the field of machine learning.

EDUCATION

University of Texas at Dallas, Richardson, USA Aug 2022 - May 2024
Masters of Science - Computer Science GPA: 3.8/4.0
Savitribai Phule Pune University, Pune, India Aug 2017 - Jul 2021
Bachelor of Engineering - Computer Science GPA: 3.8/4.0

EXPERIENCE

Hewlett Packard Enterprise, Houston, USA May 2023 - May 2024
AI & Machine Learning Engineer Intern

- Brainstormed and prototyped **innovative AI/ML features** for large-scale HPC monitoring systems, enhancing performance and reliability.
- Designed and implemented a cutting-edge visualization prototype for **pinpointing network bottlenecks**, therefore optimizing HPC infrastructure efficiency.
- Leveraged diverse data sources to construct intuitive dashboards for **real-time job execution** trends, empowering system administrators with actionable insights.
- Engineered a robust job failure prediction model with **80% accuracy**, significantly reducing failures and maximizing resource utilization in HPC environments.

PubMatic, Pune, India Apr 2021 - July 2022
Software QA Automation Engineer

- Contributed to the ML team at the programmatic advertising company, serving over **500 global clients** on its platform.
- Constructed and enhanced Spark and Hadoop pipelines to execute **automation test suites for machine learning** models involved in header bidding.
- Generated comprehensive reports for **three machine learning projects** focusing on **header bidding and ad-monetization**.
- Conducted thorough unit testing and regression testing, achieving an impressive **code coverage of over 80%**, across various ad-bidding projects. These efforts helped identify data drift issues that led to model decay in production.

Persistent Systems, Pune, India Jan 2021 - Mar 2021
Software Developer Intern

- Underwent training and assessments in industry technologies like Java, RDBMS, Maven, and Git at the multi-national software product company.

myMO IT Solutions, Remote, India Jan 2020 - Mar 2020
Python Developer Intern

- Developed a web app that extracts the table of contents of a documentation website, whilst preserving its structure.
- The deployment showed a **90% accuracy on standard documentation websites** and created a browsable index per site.

PUBLICATIONS & PATENTS

Artificial Vision for Visually Impaired (Patent: July 2021) (Application No.: 202121028865)

- Developed an end-to-end voice-controlled android app that assists the visually impaired by answering questions about their surroundings. It gives **near-instant and accurate responses** to users by using web APIs for speech-to-text and text-to-speech and **lightweight machine learning libraries** for computer vision.

SELECTED PROJECTS

Predicting fares for NYC taxis using ML (Tech: Tensorflow, Scikit-learn, Python, Pandas, Numpy) Sep 2023

- Developed a TensorFlow and Scikit-learn model to predict NYC taxi fares using Python. Engineered features, optimized hyperparameters, compared model performances and integrated real-time prediction for enhanced taxi service efficiency. The best model showcased a **2.95 MSE**.

Generating Counter Narrative for Hate Speech (Tech: Python, Tensorflow) Apr 2023

- Utilized **BERT model** to perform hate speech classification on different narratives. Fine-tuned **GPT-2 and XLNet** using multi-conan dataset for NLG task. Evaluation resulted in XLNet achieving a **BLEU score of 0.598**.

arXiv Paper Tagger (Tech: Python, scikit-learn) Sep 2022

- Used the scikit-learn library to solve a **multi-label classification** problem to tag arXiv papers with relevant categories. Performed preprocessing of the data using word embeddings and sentence embeddings. Assessed the performance of different supervised machine learning algorithms, performed hyperparameter tuning and achieved **96% accuracy**.

Revenue Forecasting (Tech: Python, Prophet) Jul 2022

- Used the Prophet library from Facebook to **forecast** the daily and quarterly revenue of PubMatic using real-time data. Insights and trends obtained helped the company **maximize their profits** by scaling efforts in peak holiday season.

Video Ads Categorization (Tech: Python, TensorFlow, Keras, Hadoop & OpenCV) Dec 2021

- Used the **EfficientNet model** to predict the metadata from video ads to help their IAB categorization. The implementation showcased an **accuracy of 72%** on test data, and received **special recognition in PubMatic's 2021 ML Hackathon**.

SKILLS

Computer Languages: Python, C++, HTML, Java
Frameworks: scikit-learn, TensorFlow, Flask, Bootstrap, Git, Spark, Hadoop, Jupyter, Matplotlib, Plotly
Databases: RDBMS, MongoDB

HONORS AND AWARDS

First place at the Revenue Forecast Competition held over 2 quarters at PubMatic with more than 10 teams.
Received **Teamwork Award** at PubMatic for reducing automation backlogs and enhancing frameworks.
First place at the IET-Hackathon at PCCOE, among more than 50 participating teams.
Research on job failure prediction at Hewlett Packard Enterprise, **accepted into CUG 2024** conference in Perth, Australia.