Manish Gawali

Portfolio Google Scholar

Email : gawali@usc.edu
Mobile : +1-(213)-574-6099

Education

- **University of Southern California** — Los Angeles, CA
  Master of Science in Computer Science; GPA: 3.86/4.0 — *May 2022 – Dec 2023*
- **Pune Institute of Computer Technology** — Pune, India
  Bachelor of Engineering in Computer Science; GPA: 9.20/10.0 — *Aug. 2014 – May 2018*

Experience

- **Amazon** — New York City, NY
  Applied Scientist Intern — *Sept 2023 - Dec 2023*
  - **Product Relationship**: Use of Large Language Models (LLMs) to determine relationships.
- **Amazon** — Denver, CO
  Software Development Engineer Intern — *May 2023 - Aug 2023*
  - **UI Page**: Creating a Forecast History Page which enables the operations team to view history of uploaded forecast files along with the statistics and metadata about that forecast.
  - **Customer Impact**: This page allows the user to export the excel for failed purchase orders along with the failure reason for it. This improved operations team's user experience by ensuring that they don't have to manually search for emails to find information regarding forecast uploads.
  - **AWS Lambdas Step Function**: Modified five lambdas in two step functions to support this feature and created a new lambda for exporting the excel file for failed purchase orders to S3.
  - **API and Coral API**: Created new APIs and Coral APIs to support fetching data on the UI through a gateway service and service which owns data for forecast history.
  - **Database Schema**: Created a new DynamoDB table for forecast metadata.
  - **Design Document**: Created a design document after gathering requirements for the feature from product owner and the customer [Operations Team Lead].
- **vITAL Lab, University of Southern California** — Los Angeles, CA
  Graduate Research Assistant — *May 2022 - Present*
  - **Effective Aggregation of Weights in Federated Learning (FL)**: Research on a novel aggregation method in FL which uses information from client model updates and determines aggregation weight for that client.
  - **Semi-supervised Federated Learning**: Research on FL where only few silos have ground truths.
  - **FedML**: Open Source Contribution for the feature FLamby in FedML. This feature enabled users to enable open source access to large X-ray, CT, and MRI datasets and baseline FL code for training models on these datasets.
  - **Bone Metastasis Segmentation**: Research on a challenging generic segmentation approach for bone metastasis for CT scans. Challenges: Small dataset, Noisy annotations. Semantic Segmentation of lesions using ViT (Vision Transformer).
- **DeepTek Medical Imaging Pvt Ltd.** — Pune, India
  Senior Data Scientist (Tech Lead) — *Aug 2020 - Mar 2022*
  - **Research**:
    - Published 5 research papers in reputed conferences and journals. 2 more are in review process.
    - Conducted research, contributed to the ***Privacy-Preserving Distributed Deep Learning*** field in the form of a research paper. Proposed a novel distributed deep learning method called ***SplitFedv3*** and a novel distributed model training approach known as ***Alternate mini-batch training***. Moreover, comparison of **Federated Learning**, **Split Learning**, and **SplitFed** variants was done for healthcare domain.
  - **Development**:
    - Developed 10+ deep learning models for detection of pathologies like Pleural Effusion, Fractures, Lung Mass, Covid-19 and Nodules etc. from chest X-rays, CT scans, and MRI scans and productionized them.
    - Designed pipelines for image classification, segmentation (UNET), object detection (Mask-RCNN, YOLO architectures), anomaly detection algorithms (OneClassSVM, AutoEncoders), generative algorithms (GANs, VAEs).
    - Built an UI tool using p5.js and node.js and carried out controlled experiments with 5 radiologists to measure if AI helped improve the productivity and diagnostic accuracy. Statistical analysis and hypothesis testing was done.
    - Automated AI validations for clients.
    - Built an image level (present in dicoms) anonymization software to remove textual patient information from images present in endoscopy and ultrasound scans.
  - **Leadership**:
    - Led the research team of 3 data scientists and 2 researchers for the collaborative project: Privacy-Preserving Distributed Deep Learning between DeepTek and SBIC (Singapore Bioimaging Consortium) - A*Star.

    - Led a team of 3 data scientists for developing end-to-end ML workflows for pathology detection in CT/MRI model research, optimization, and engineering activities.
    - Mentored 3 interns.

- **AlgoAnalytics Pvt. Ltd.** Pune, India
  Associate Consultant (Deep Learning) *Jan 2020 - Jul 2020*
    - **Signature Generation**: Developed a novel variant of DCGAN (Deep Convolutional Generative Adversarial Network) architecture for generating handwritten signatures using the GPDR-960 handwritten dataset.
    - **Virtual Try-On**: Implemented end-to-end client-facing Fashion AI application for Virtual Try-On.
- **Veritas Technologies LLC** Pune, India
  Associate Software Engineer *Jul 2018 - Jul 2019*
    - **Enterprise Vault & Compliance/Discovery Accelerator**: Developed multiple features like Auditing Enhancements, Cloud-SQL, Hotword Statistics, Hotword, and Hotword Set Facet.
    - **POCs, Bug Fixes, Optimization**: Contributed to 8 winning POC's and rectified 30+ bug fixes.Optimized legacy code to increase the speed of operations like archiving e-mails and file shares by 20%
- **GSLab** Pune, India
  Machine Learning Intern *Aug 2017 - Mar 2018*
    - Developed Chatbot for Automated Customer Service for cutting costs by eliminating need of human intervention.
    - Built the backend machine learning & natural language processing pipeline using concepts like TF-IDF, Clustering, and Cosine similarity to help agents deliver better customer support achieving an accuracy of 82%.
    - Deployed ML model as REST API using Flask as a backend service.

Projects

- **Human Activity Recognition | Technologies: Python, Pandas, Keras**: Built a deep learning model that predicts the human activities such as walking, walking upstairs, walking downstairs, sitting, standing or laying in a video.
- **eYantra | Technologies: C++, Embedded C, Python**: Co-ordinated the synchronous motion of two robots by building path planning algorithms, to match the musical notes extracted from audio files by striking steel rods at specific nodes of the graph.

Programming Skills

- **Languages**: Python, Java, Javascript, Typescript, C, C++, C#, Embedded C, SQL
- **Libraries**: Tensorflow, Keras, PyTorch, OpenCV, Pandas, NumPy, Scikit-Learn, Gensim, spaCy, NLTK, p5.js, React.js, Redux
- **Frameworks**: Langchain, Hugging Face, Flask, .NET
- **Developer Tools**: Git, Docker, GCP, AWS, IntelliJ, VS Code, Visual Studio, PyCharm, Eclipse, Android Studio

Publications

- **Springer Publication**: "Comparison of Privacy-Preserving Distributed Deep Learning Methods in Healthcare". Presented the paper at the MIUA 2021 conference.
- **IEEE Publication**: "Deep Learning Models for Calculation of Cardiothoracic Ratio from Chest Radiographs for Assisted Diagnosis of Cardiomegaly", icABCD 2021 conference.
- **JMIR Publication**: "Key Technology Considerations in Developing and Deploying Machine Learning Models in Clinical Radiology Practice", JMIR Medical Informatics Journal.
- **Springer Publication**: "Vulnerability Due to Training Order in Split Learning". ICT4SD 2021 conference.
- **Springer Publication**: "A deep learning approach for automated diagnosis of pulmonary embolism on computed tomographic pulmonary angiography".
- **Arxiv Pre-Print**: "Application of Federated Learning in Building a Robust COVID-19 Chest X-ray Classification Model". (In Publication Process).
- **Medrxiv Pre-Print**: "Automated assessment of chest CT severity scores in patients suspected of COVID-19 infection". (In Publication Process).

Research Services

- **Reviewer**: ML4H: Machine Learning for Health 2023.
- **Reviewer**: ML4H: Machine Learning for Health 2022.
- **Reviewer**: ML4H: Machine Learning for Health 2021.
- **Reviewer**: Secure and Privacy-Preserving Machine Learning for Medical Imaging: MICCAI 2021 Workshop and Tutorial (Springer).
- **Reviewer**: 2021 International Conference on Artificial Intelligence and its Applications (icARTI 2021) (ACM).