Ankur Singh

+1 (408) 941 - 5818 | ankur.singh01@sjsu.edu | ankur-singh.github.io

Summary

With over 4+ years of experience specializing in the domains of Product Management, Deep Learning, Computer Vision, and Software Development. I am passionately curious in solving business problems by building products and services from the ground up. A lifelong learner, tinkerer, and a team builder. In my recent stint, I built a phenomenal team of 6 members and deployed 4 ML services in production in just 10 months. Before that, I developed 5 full-fledged courses on topics around ML/DL and mentored over 600+ students.

Skills

Languages	Python, SQL, Shell Scripting
Databases	Postgres, SQLite3, MongoDB, Redis
Web Development	Flask, FastAPI, Rest APIs, Web scraping
Time Series Forecasting	Stats Model, GluonTS, Prophet
Software Development	CI/CD, Git, Docker, Pytest, Python Packaging
ML/DL Frameworks, Libraries & Git repos	Scikit-Learn, Pytorch, Pandas, Keras, Fastai, Pytorch Lightning, YOLOv5, YOLOR, U2net, Timm, Torchvision, XGboost, LightGBM, Catboost, Spacy, NLTK.
Deep Learning Problems	**CV:** Classification, Regression, Detection, Segmentation, Keypoint detection. **NLP:** Classification, Text Generation, NER, POS tagging, Translation. **Forecasting:** ARMA, ARIMA, SARIMA, RNN, LSTM, 1D convolution, **Others:** OCR, Face Detection & Recognition; Recommender System, Multi-Label, Similarity Search, Unsupervised Learning, Tabular Data, , Multimodal Architecture.

Work Experience

San Jose State University

San Jose, CA

GRADUATE RESEARCH ASSISTANT

Sept. 2017 - present

Building an Edge AI platform to deploy and manage ML/DL models at scale on Edge devices for real-time inference.

Zoop.One

Pune, India

TEAM LEAD - MACHINE LEARNING

Sep. 2021 - Jul. 2022

As a Founding Member of the ML team, built a phenomenal team of 6 ML engineers and led a large effort to scale up the use of ML across the company. I improved the state of MLOps (from level-0 to level-1) by setting up annotation tools, model registry, monitoring, automating model training, and establishing other best practices.

- **OCR Service**
  - Developed an OCR service to extract relevant information from Identity Cards like PAN, Aadhaar, Driving License & Passport. It served as the backend for multiple Sign-up & Verification journeys in several products.
  - The service had more than 7 deep learning models to detect and extract cards, correct card orientation, enhance the image, identify card type, layout analysis, and OCR to extract textual information.
  - Our service was almost 4 to 6 times faster than other players in the market and was much more accurate. Additionally, we provide support for multi-line names & addresses that no other player was offering.
- **Document Scanner Service**
  - Developed Heatmap Regression based extractor services to extract ID cards & documents from images.
  - Compressed and exported the model to TFLite (4.4 MB) for edge deployment.
- **Liveliness Service**
  - This service provided features like face detection, recognition, matching, and other face details like age, gender, race, etc. It also had modules to check for eyeglasses, face mask, liveliness, and spoof detection.
  - Config driven backend, enabling users to dynamically turn on/off any module on per request basis.
  - Low latency (~150 MS for liveliness & spoof detection) allowed us to do real-time inference on video feeds. While other competitors were using active liveliness, we used passive liveliness, with very high accuracy.

- **Other Roles & Responsibilities**
  - Identify areas of improvement in existing services and investigate new technologies to optimize ML workflows.
  - Regularly conducted training sessions to help new and junior team members.
  - Developed end2end frameworks for preparing data, training, model validation, and model deployment.

AiAdventures LLP *Pune, India*

CO-FOUNDER & CEO *Aug. 2018 - Sep. 2021*

We provided AI/ML solutions to businesses, and also trained people in Data Science & Machine Learning.

- **Developed full-fledged courses** for Python, Data Science, Machine Learning, Deep Learning, & Computer Vision. Each course had at least 9 hours of content in the form of Jupyter notebooks, assignments, quizzes, & projects.
- Built the tech-stack required for generation, deployment & distribution of all courses.
- Helped 600+ students get started with python & Data science.
- Conducted 25+ college and corporate workshops on SOTA Machine Learning and Deep Learning systems.
- Wrote 20+ blogs on topics related to Python, Data Science, ML, DL, databases, etc.
- Managing and developing client projects.

Extracurricular

- Built a python package called "**colab-everything**" which lets you run web-apps on any jupyter notebook like environment. The library has over **24K+ downloads**.
- **Won Grand Prize** in "AI for Social Good" hackathon at **Intel Innovation 2022.**
- Have **contributed to** Fastai, MLflow, LazyPredict, Pytorch Lighting, Category Encoding, YOLOv5, etc.
- Build **Bank Cheque service** that allows one to extract bank and account details from canceled cheques.
- Used **stats model** and **prophet** to forecast demand of bakery products to avoid wastage. The final model was an ensemble of 3 different models and was able to address both seasonality and trends in product demands.
- Trained YOLO & Faster-RCNN detection models to detect jewelries in images. Used **unsupervised learning** (triplet & contrastive loss) to find similar images in the database. The application was exposed as Rest APIs. **Technologies:** Pytorch, Faiss, Flask, Rest-API, OpenCV, React, MongoDB, Docker

Kaggle Competitions

Shopee- Price Match Guarantee *Bronze Medal (207th / 2426)*

- Task was to identify similar products using product image, title, description, and phash.
- It was a similarity search problem with new unseen products in the test set.
- Used unsupervised and self - supervised learning techniques (like SIMCLR, SWAV, ArcFace loss), and **multi-modal model** to extract product features, followed by KNN (faiss & rapis) & ranking to find similar products.

Global Wheat Detection *Bronze Medal (191st / 2245)*

- Task was to detect wheat heads from field images across the globe.
- It was an Object detection problem with extremely noisy data and numerous wrong labels.
- Used YOLOv5, EfficientDet-B5, pseudo labeling, knowledge distillation, mosaic augmentation and TTA.

Mechanisms of Action *Bronze Medal (318th / 4373)*

- Task was to measure the effect of drugs on genes and cells.
- It was a **multi-label** classification problem with 897 input features and 207 output labels.
- Used an ensemble of LightGBM, Tabnet and Custom Pytorch models.

Education

San Jose State University (SJSU) *San Jose, California*

MASTERS IN SOFTWARE ENGINEERING *Aug. 2022 - May 2024*

- Major Area of Specialization in **Data Science**.

College of Engineering, Pune (COEP) *Pune, India*

BACHELOR OF TECHNOLOGY IN INFORMATION TECHNOLOGY *Aug. 2014 - May 2018*

- Cumulative Grade Point Average: **7.67**/10