

ABHAY SHASTRY

Gainesville,FL • abhay.s.shastry@gmail.com • +1 (860) 912-6408 • [Website](#) • [Github](#) • [Linkedin](#)

EDUCATION

UNIVERSITY OF FLORIDA

August 2024 -Present

Master's in Computer and Information Sciences

Coursework: Distributed Operating Systems Principles, Advanced Data Structures, Mathematics for Intelligent Systems

DAYANANDA SAGAR UNIVERSITY,INDIA

2018-2022

Bachelor of Technology in Computer Science and Engineering. 8.51 CGPA

PROFESSIONAL EXPERIENCE

NAVANA.AI

Singapore

Machine Learning Engineer

September 2022-June 2024

- Co-led the development of nCloud, a cutting-edge cloud platform for deploying computer vision models, by engineering high-performance inference APIs using FastAPI, Docker, TensorRT, and NVIDIA Triton, ensuring seamless and efficient model deployment.
- Achieved scalable and robust deployments by architecting AWS infrastructure and implementing tiered containerization with Docker Swarm, optimizing resource utilization and system reliability.
- Elevated niia, a generative AI platform by fine-tuning LoRA embeddings and developing the Mannequin2Photoshoot API, enabling precise and realistic clothing visualization on fashion model templates.
- Improved ABinBev's product recognition accuracy by developing computer vision models that achieved over 98% classification accuracy and 96% detection accuracy, enhancing the reliability of their promotional analysis.
- Engineered and deployed a high-throughput inference API, capable of handling 300 requests per second, using NVIDIA Triton and CUDA optimization and created auto deployment CI/CD pipelines on Azure, significantly enhancing real-time processing for ABinBev.
- Streamlined food tray inspection processes for Emirates Airlines by developing segmentation and classification models, automating quality checks with 95% accuracy and improving operational efficiency.

INTEL CORPORATION

Bangalore, India

Software Developer Intern

January 2022-April 2022

- Developed ETL pipelines for procuring Intel Alder Lake production data and fill rates using Apache Airflow and SAP HANA, streamlining data ingestion and significantly enhancing data-driven decision-making within the Supply Chain Team.
- Wrote MapReduce programs for concurrent processing of manufacturing data, reducing ETL time by 43% and improving operational efficiency.
- Designed and implemented automated data validation checks, ensuring data accuracy and integrity, which minimized errors in production reporting by 30%.

SAARAGH

Singapore

Machine Learning Engineer Intern

June 2021-September 2021

- Spearheaded the development of high-precision cancer detection classification models and engineered inference APIs using Flask, significantly enhancing the company's model registry for streamlined deployment and integration.
- Played a key role in designing advanced detection and segmentation models for warehouse management in partnership with Konica Minolta, resulting in a 73% boost in operational efficiency over manual processes.

PROJECTS

- [Music Recommendation System](#)**: Developed an emotion-based music recommender using CNNs, Haarcascades, and K-Means clustering, achieving 86.9% training accuracy with TensorFlow and Pandas.
- [Ecommerce Website Using Blockchain](#)**: Built a full-stack ecommerce platform on Ethereum's EVM, enabling decentralized transactions without a traditional backend database using React.js and Solidity.
- [Food Delivery Website](#)**: Created a responsive full stack food delivery website using React.js, Tailwind and Go
- [Lexical Analyzer](#)**: Implemented a basic lexical analyzer in C using Lex and Yacc for tokenizing and parsing
- [Yolov8 API](#)**: Dockerized API that can be used for inference any yolov8 detection,classification or segmentation onnx file

SKILLS

- Languages**: Python, Vanilla Javascript, CSS, Core Java, GoLang
- Frameworks,Libraries & Tools**: MLFlow, OpenCV, Tensorflow, Keras, Pandas, Numpy, Scikit Learn,Hugging Face, NLTK, Bootstrap, Hadoop, MapReduce, Apache Kafka, Apache Airflow, Jenkins, CUDA, ,PL/SQL, SAP HANA, MongoDB, CockroachDB, CosmosDB, PostgreSQL, RabbitMQ,Amazon Web Services (EC2, ELB, S3, Lambda, SageMaker, EBS, EFS, DynamoDB), Microsoft Azure(AFS, AVM, ALB), GCP(Big Query, Firebase, VertexAI), Docker, Kubernetes