RAAJESH LAGUDUVA RAMESHBABU

Experienced Data Scientist Transitioned to MLOps Engineer

Trying to make the world a better place, one line of code at a time.

https://www.linkedin.com/in/raajeshlr

ML Dev & MLOps - https://github.com/raajeshlr?tab=repositories

raajeshlr2@gmail.com

+91 - 9600745502

Mace Surabi, Chennai.

Capgemini - 10 Years

B.E ECE: 86.3 %

SKILLS

- LLMOps, GCP / AWS / Azure MLOps, Vertex AI, Sage Maker, Azure ML, Docker, Kubernetes, OpenShift, Airflow, Kubeflow, Jenkins, Databricks, Mlflow, Terraform, Cloud Formation, DVC, GitHub Actions.
- Machine Learning, Deep Learning, NLP, Python, RNN, LSTM, Keras, PyTorch, OpenCV, MongoDB, SQL.

MLOPS WORK EXPERIENCE

Capgemini 16th Aug 2022 - Present

On-premises MLOps - https://github.com/raajeshlr/Resume/blob/master/ArchitectureOpenshift.jpg

- **E2E ML Pipeline**: Transformed DS team's notebooks into a robust ML pipeline, managing data loading, processing, training, and report management.
- **Agile Collaboration & Continuous Improvement**: Enforced software engineering best practices, overcoming DS team challenges and promoting continuous improvement.
- **Scalable Deployment**: Orchestrated Fast API-based scoring engine routers, Dockerized, and deployed on OpenShift for efficient production deployment.
- **Production Deployment & Environment Expertise**: Addressed production deployment issues, deploying across various environments using Jenkins and OpenShift clusters, ensuring smooth transitions and optimized performance.

Azure MLOps - https://github.com/raajeshlr/Resume/blob/master/Architecture%20Image%20Azure.png

- **Azure DevOps Pipeline and Infra Management:** Implemented Azure DevOps Pipeline & orchestrated resources such as storage accounts, Databricks workspaces, Kubernetes clusters, Airflow using Terraform.
- **Databricks for ML Pipeline Execution:** Utilized Databricks for ML Pipeline execution, facilitating seamless data operations with storage accounts.
- **Airflow Orchestration on Kubernetes:** Used Airflow on Kubernetes managing DAGs & execute Databricks Jobs, utilizing Kubernetes Pod Operator for running containerized images from Azure Container Registry.
- **End-to-End Azure DevOps Integration:** Integrated Azure DevOps with the project repository for automating Docker image creation, pushing to ACR, and uploading Airflow DAGs to storage account file shares.

AWS MLOps - https://github.com/raajeshlr/Resume/blob/master/Architecture%20Image%20AWS.PNG

- **CI/CD DevOps Pipeline with AWS Resources:** Utilized AWS Code Pipeline for ML model CI/CD, integrating S3, Code Commit, Code Build, and Amazon EKS. Implemented Docker containers for modular ML components.
- **ML Pipeline:** Implemented different steps of Machine Learning Pipelines through AWS Sage Maker.
- **Infra Automation:** Created the Infra through Cloud Formation and automated it using the bash scripts.
- **Deployment:** Deployed containers on Amazon EKS/ECS Clusters using Code Build and Kubernetes commands.
- **Cluster Management:** Experienced on creating Kubernetes Clusters, exposing, and scaling the pods.
- **Sustainability**: Used the code carbon package on the ML Pipeline to measure the CO2 Emissions and performed actionable recommendations to reduce it.

Accenture – 27th July 2021 to 09th August 2022

GCP MLOps - https://github.com/raajeshlr/Resume/blob/master/Architecture%20Image%20GCP.PNG

- **CI/CD DevOps Pipeline with GCP Resources:** Designed and executed CI/CD DevOps Pipeline for Machine Learning models, utilizing GCP Cloud Source Repository, Cloud Build for triggers, and Cloud Function.
- **Machine Learning Pipelines:** Used Kubeflow components and deployed it to GCP Vertex AI.
- **Scheduled Jobs:** Periodic jobs for running pipelines using Cloud-Scheduler, Cloud Pub/Sub, Cloud Function.
- **Infra Provisioning:** Created the Infrastructure through Terraform.
- **Sustainability:** Used the code carbon package on the ML Pipeline to measure the CO2 Emissions and performed actionable recommendations to reduce it.

Mlflow

- **Machine Learning Pipelines with Mlflow:** Created experiments and registered the best models.
- **Deployment and Monitoring:** Deployed models using both real-time inferencing and batch transform methods and developed a model monitoring system integrated with Mlflow, utilizing DVC for version control and GitHub actions for automated Python testing.
- **Production ready ML Models**: Ensuring user authentication through Single Sign-On (SSO) Azure Active Directory. Installed SSL certificates on VMs for secure HTTPS transport and seamlessly integrated with various systems for ML pipelines spanning from data ingestion to model serving.

ML DEVELOPMENT WORK EXPERIENCE **Infosys 03rd Dec 2018 to 23rd July 2021**

Infosys Intelligent Assistant - To automate support projects.

- **Web App Development**: Used Angular (UI), Python on business layer, MongoDB on Data Layer.
- **Machine Learning Ticket Classification:** Pulled tickets history from ISTM and used Algorithms like Logistic Regression, Random Forests, Decision Trees, BERT for ticket category classification.
- **E2E Machine Learning Pipelines:** Implemented for both training and Inferencing.
- **Text Rank and KB Integration:** For related tickets search and providing automatic solutions from KB.

Infosys Fraud Detector

- **Text and Numerical Fraud Detection:** Have used LSTM + CNN based ensemble network for this.
- **Image Tampering Detection:** Using Deep Learning model + NLP techniques Spacy, tesseract, bounding box.
- **Deployment and Architecture:** Used Docker, Azure AD Single Sign On (SSO) with 3-tier architecture.

Infosys Intelligent Chatbot – Understands user screen and provides solution.

- **Implementation:** Developed Azure Chatbot using node.js and used Python for backend Machine Learning.
- **Advanced techniques:** Used One-Shot Learning, NLP techniques, and built with QnA and RPA Services.
- **Deployment with Docker on Azure:** Deployed Python code to Azure as a Docker Image.

Comprehensive cross check for new joiners

- **Document processing and Data comparison:** Developed Python Bots for extracting fields from documents and pdfs, comparing it with SAP DB data.
- **Deployment Pipeline setup:** Created Deployment pipelines of this project in the Virtual Machines.
- **HR Automation:** It's for the HR Team & product is live, it automated the manual efforts and reduced FTE's.

Signature classification using CNN.

- **Data Collection & Training:** Started project from scratch, collected images, labelled it, built with CNN model.
- **Script Development:** Written Python code for cropping, finding the coordinates and completed this project.

JUNIOR ML ENGINEER WORK EXPERIENCE **TCS, 22nd Dec 2014 – 30th Nov 2018**

Home Credit Default Risk

- Our goal is to predict loan defaulters using Logistic Regression, Random Forest, and Light-GBM model.
- Performed EDA, preprocessing done, tried feature engineering, and evaluated with ROC AUC.

Finding Donors for Charity

- Our goal is to predict individuals makes money > $50,000 to appeal donor for a non-profit organization.
- Performed EDA, pre-processing including skewed continuous feature transformation, normalization, encoding.
- Performed Grid Search CV and fine-tuned essential parameters for GBM, which achieved 87% test accuracy.

Densely Connected Convolutional Networks - Dense-Net

- CIFAR10 Dataset: Dense-Net Paper https://arxiv.org/pdf/1608.06993.pdf
- Created Dense-Net architecture with three convolution blocks and two Transition blocks.
- Achieved the max validation accuracy of 88% after fine-tuning and implementing OneCycleLR.

Creating Customer Segments

- Developed Unsupervised Learning, clustering technique, demonstrated feature scaling, dimensionality reduction and feature transformation using PCA, and identify customer segments hidden in the data.
- Developed K-Means clustering algorithm and GMM and measured performance with Silhouette score.

Convolutional Neural Networks using Fashion MNIST Data - No Obsolete method.

- Goal is to achieve 99.2% Val Accuracy with less than 20,000 Hyper-parameters (No Hidden layers should be used.
- Developed the model with a high-level framework Keras, with selection of TensorFlow for backend.
- The model achieved 99.2% validation accuracy in 11 Epochs.

Restaurant Review Management System

- Goal is to classify reviews, performed cleaning, stemming, created corpus and bag of words with 2000 features.
- Implemented Gaussian Naïve Bayes Classifier, and trained and tested the model, evaluated using f1 score.

INTERNSHIP EXPERIENCE https://theschoolof.ai/

Extensive Vision for AI - CNN

- CNN Projects in Keras - https://github.com/raajeshlr/CNN_KERAS_NLP_INTRO_EVA1_INTERNSHIP
- CNN Projects in PyTorch - https://github.com/raajeshlr/CNN_PyTorch_EVA4B2_Internship
- We worked on the concepts like Image Classification, Object Detection using YOLO v3, Transfer Learning, Super Convergence, Landmark's Detection.

Natural Language Processing - NLP

- NLP Projects in PyTorch - https://github.com/raajeshlr/NLP-END
- We worked on the concepts like Text Classification, Sequence to Sequence translation using multi head attention models.

Extensive Machine Learning Operations MLOPS - EMLO

- MLOps - https://github.com/emlopsinfy?tab=repositories
- We worked on Docker Internals, Kubernetes Clusters, Heroku, CI/CD, Model Deployment on AWS EC2.

Certificates

- ML Nanodegree Udacity, Sequence Models - Coursera, Machine Learning - Coursera, Machine Learning A-Z Udemy, LLMOps Specialization.

Honor Awards

- Value creator award from Capgemini
- Impact creator award from Infosys
- Best performer of the year award from TCS
- Service and commitment award from TCS.

Languages

- Saurashtra, English, Tamil, RW Hindi, Learning Kannada.

Achievements

- Secured ‘Certificate A’ Exam under authority of, Ministry of Defense, Government of India.
- Completed Hindi Exams until Praveshika.
- Presented Parallel Parking Robots and Image Processing surveillance system papers during college.