Name	: **Yempati Karthik**
Mobile number	: +91-8008527441
Email	: karthikyempati997@gmail.com
LinkedIn	: https://www.linkedin.com/in/yempati-karthik/

Summary

Having approx **3 years** of industry working experience. For the past **2 years** working in Data Science, Machine Learning, Deep Learning and Computer Vision Development.

Certified Tensorflow Developer recognized by Google.

Professional Synopsis

- Excellent Data Cleaning, Data Exploration Analysis skills, Data Engineering(ETL) skills, Quantitative Analysis and Data Visualization. Maintaining Data Quality.
- Design, Build, Develop and Implementing end-to-end ML pipelines and building both batch and real-time model pipelines with existing applications.
- Create/Maintain CI/CD MLOps pipeline(Continuous Integration/Continuous Delivery/Deployment), Verification/Validation and Monitoring of AI/ML models in Kubernetes(K8s) clusters. Scheduling automated retraining based on data drifting.
- Good knowledge of Advanced Analytics like NLP, Predictive/Prescriptive models, Demand Forecasting(Anomaly Detection), Descriptive modeling and Data Mining.
- Experienced in varied technology environments and platforms such as image processing and image manipulation techniques – image enhancement, contour detection, edge detection, feature point tracking etc. Image Classifiers and Object Detection/Detectors. Developing state-of-the-art (SOTA) data products.
- Proficient in Natural Language Processing (LSA, LDA, TF-IDF, Tokenizer, Bag of Words), Natural Language Understanding (Sentiment Analysis, Word Embedding such as Word2Vec) and Text Mining.
- Experience in Statistical Techniques and ML such as Hypothesis Testing(A/B Testing), Regression, Classification and Time Series Analysis.
- Done custom training tesseract 4.0 to extract Arabic text from documents (Optical Character Recognition [OCR]).

Experience

Data Scientist

Organization : Amnet Digital

Tenure　　　 : May 11, 2021 to Present

Machine Learning Engineer

Organization : Metagogy Learning Systems Pvt Ltd

Tenure　　　 : Jan 08, 2020 to May 08, 2021

Data Science - Intern

Organization : Techolution

Tenure   : Sep 26, 2019 to Jan 07, 2020

IT Analyst
    Organization : Teleperformance
    Tenure   : July 02, 2018 to May 06, 2019

Technical Skills

Machine Learning (Supervised Learning)	Linear Regression, Logistic Regression, Polynomial Regression, KNN, SVM, Naive Bayes, Decision Trees, Ensembling Techniques, Random Forest, AdaBoost, Gradient Boosting, XGBoost
Machine Learning (Unsupervised Learning)	K-means Clustering, Hierarchical clustering, Apriori, ICA, PCA
TimeSeries Forecasting	ARIMA, SARIMA, Facebook Prophet
Deep Learning	Neural Networks, ANN, CNN, RNN, LSTM, GRU, Autoencoders
Programming Languages	Python
Databases	MySQL(RDBMS), PostgreSQL, MSSQL
IDE Tools	Jupyter Notebook, JupyterLab, PyCharm, Databricks
Others	Statistical Modeling, Data Structures and Algorithms, Mathematics, Linear Algebra, Probability
Operating System	Windows, Ubuntu (Linux/Unix)
Version Control System	Git and Github
Orchestration tool	Docker, GKE
Big Data	Spark(PySpark), Apache Hive
Cloud Computing	GCP

Production Projects

Project1):
Product name   : FaceOpen
Company   : Techolution
Details   : Touch Free Door Open and Unlock Technology
Tools and Technique used   : Python, OpenCV, Keras, TensorFlow on GPU, PyTorch, Numpy, Pandas, Sklearn(SciKit-learn), Matplotlib, Seaborn, CUDA, NVIDIA Drivers, Intel Realsense Depth Camera, opencv2, pyrealsense2.
Architectures used   : Transfer Learning, MobileNetV2, DenseNet121, DenseNet169, ResNet50, VGG16, NASNet, Efficient Net, Inception, FeatherNetB, YOLO, ImageNet.

Faceopen is designed to prevent face spoofing. Our 3D face liveness detection technique powered by artificial intelligence rejects printed photographs or images shown on digital devices, reducing security threats and the risk of unauthorized entry.

About Product : Our proprietary and powerful biometric-based user identification system is built on artificial intelligence technologies such as image recognition, deep neural networks, and robotic process automation. This allows each Faceopen system to evolve with every regular user interaction, enabling adaptability in different environments and conditions.

Roles and Responsibilities : Responsible for collecting data (images) from different sources through the internet and outside world.
Collected data using logitech camera (2d) using tensorflow and realsense D435 camera using pytorch framework. Optimizing algorithms for hardware acceleration cores.
Worked on data preprocessing, face detection(MTCNN), feature extraction(FaceNet) and face classification(KNN,SVM). Working with heterogeneous or parallel computing systems.
Built model for anti-spoofing classifier using various architectures.
Built model for detecting blur faces (blur-classifier). Used docker for containerizing the application (DockerFile). Here deployment of the models is done in edge, which is android in the form of TensorflowLite (tflite).
Converted spoofing pytorch model to ONNX format for deployment.

Project2):

Product name : [AI Educator](#)
Company : Metagogy Learning Systems Pvt Ltd
Details : Solving student's knowledge gaps using graph neural networks (Problem solving gap analysis).
Tools and Technique used : Python, Deep Graph Neural Networks, networkx.
Use case : Analyzing student's problem solving skills in assignments which he/she solves. Based on student performance in understanding concepts and solving coding problems the model will suggest where the student is outperforming.
Roles and Responsibilities : Worked on deep graph neural networks such as converting a python script into an abstract syntax tree format.
Worked on analyzing graph data such as finding patterns in the coding assignments.
Built a python function which takes a coding question as input and gives the output such as the user's knowledge score, his/her problem solving score and many more use cases.
This application also recommends a user what he/she needs to study to solve the particular coding assignment.
Based on these scores we are able to find the root cause of why a user is unable to solve a coding problem and fill their knowledge gaps.
This application is deployed using the django framework at the backend of our web application.

About Product : AI Educator is an online learning platform which is enabled by AI technology. The system and method to find the knowledge gap and problem solving gap using the response data from the user. Using this system we can be able to find out why a student is not able to solve a coding question. By using the current system and method we find the reason why an user is unable to solve a coding question and reinforce his knowledge and problem solving skills immediately so that he is able to solve the problem with utmost confidence.

Project3):

Product name   : [Epro](#)
Company   : Amnet Digital
Client   : CDW
Tools and Technique used : Python, Citrix Workspace(Virtual Machine), Microsoft SQL Server(MSSQL).

About Product   : Electronic Procurement, or e-Procurement, is the business to business purchase of products or services through the Internet and other networking systems. ERP/SRM system i.e., E-commerce for business.

About Client   : CDW Corporation is a leading multi-brand provider of information technology solutions to business, government, education and healthcare customers in the United States, the United Kingdom and Canada.

Roles and Responsibilities   : Collecting data from Enterprise Data Warehouse (EDW) using SQL Queries from multiple databases and tables as per business requirements.
Connecting and Extracting data from Apache Hive tables using PySpark and performing transformations using Spark SQL. Handling Imbalanced data/classes using SMOTE.
Implementing End-to-End Data Science pipelines such as Problem scoping, Data Gathering/Collection, Data Analysis, Exploratory Data Analysis (EDA), Model Training/Modelling/Evaluation/Validation, Model Monitoring and maintenance.
Worked on advanced boosting algorithms such as CatBoost, LightGBM etc. Monitoring Model/Data Drift and Debugging Code, Re-training models.
Preparing presentations(PPT) based on insights generated to business clients.

Proof of Concepts (POC) Projects

Project1):
Project details   : Real Time Signature Extraction From Documents
Tools and Technique used   : Python, OpenCV, Numpy, Pandas
Architecture used   : YOLOv4
Data Annotation Tool used : LabelImg
Roles and Responsibilities   : Responsible for collecting document data (images) from different sources through the internet and outside world.
Conversion of pdf files into images as part of preprocessing.
Done data annotation using open source LabelImg tool.
Built model to extract signature from documents using YOLOv4 architecture.

Project2):
Project details   : Real Time Logo and Stamp Extraction From Documents
Roles and Responsibilities   : Responsible for collecting document data (images) from different sources through the internet and outside world.
Conversion of pdf files into images as part of preprocessing.
Done data annotation using open source LabelImg tool.
Built model to extract logo and stamp from documents using YOLOv4 architecture.

Project3):

Project details : Real Time Photos, Cartoons, Logo's, Maps Extraction From Documents

Roles and Responsibilities : Responsible for collecting document data (images) from different sources through the internet and outside world.

Conversion of pdf files into images as part of preprocessing.

Done data annotation using open source LabelImg tool.

Built model to extract photographs, cartoons, logo's, maps from documents using YOLOv4 architecture.

Project4):

Project details : Cricket Analytics using Computer Vision

Roles and Responsibilities : Responsible for data collection, data cleaning, data preprocessing and data labelling.

Built model using SOTA (state-of-the-art) model yolov4-tiny for Object Detection, pose estimation for batsman cricket shot angle calculations.

Education

Qualification	University Name	College Name	Year of Passage	Percentage
B.Tech (Electronics and Communication Engineering)	JNTU Hyderabad	Marri Laxman Reddy Institute of Technology and Management	2018	66.08

Certifications

CLICK HERE : **Google's Tensorflow Developer Professional Certificate**