DINESH KUMAR

Email **dineshkumarbalu07@gmailcom**
Phone **+91 7402282440**
Github **https://github.com/DineshKumar-Balu**
Website **www.linkedin.com/in/dineshkumar-balasubramaniyan**

SUMMARY

Currently pursuing a comprehensive third-year curriculum in Artificial Intelligence and Data Science at Kongu Engineering College, integrating theoretical knowledge with hands-on experience in cutting-edge technologies.

LEADERSHIP QUALITIES

Qualities

As **Secretary** of the AI Department, exhibited strong leadership skills in efficiently managing responsibilities. During my tenure at **Centillion** as an intern, demonstrated effective leade rship by overseeing and contributing to multiple projects, including the successful management of two key initiatives

Secretary | quick Leaner | Team Managment

PROJECT MANAGEMENT

SKILLS

In Project Management, contributed to multiple projects with a focus on GCP cloud management. Demonstrated leadership by taking a separate leadership role for Vertex AI, ensuring distinct client management and project success

Team player | Time Managament | Communication skills

PAPER PUBLISHED

SCOPUS / 2023

CERVICAL CANCER

Published a research paper on cervical cancer using ma chine learning in a recognized Scopus-indexed journal. Shared new insights, showcasing dedication to connect ing healthcare with the latest technology. Recognized for excellence in bridging medicine and machine learnING.

Machine Learning ALgorithms

SPRINGER NATURE / 2023

LEARNING DYSLEXIA

Published a research paper in Springer Nature Journal on Learning Dyslexia, leveraging machine learning tech niques for innovative insights. Contributed to the ad vancement of knowledge in education and cognitive sciences with a focus on data-driven approaches for dyslexia analysis and intervention. Demonstrated research proficiency and commitment to addressing real world challenges in learning disorders.

MACHINE LEARNING

PAPER PRESENTED

KIT COLLEGE-SENTIMENTAL ANALYSIS Got 1st prize in Sentimental Analysis using NLP, demonstrating expertise in Natural Language Process ing and data-driven insights.SENTIMENTAL ANALSYIS USING NLP **KARPAGAM UNIVERSITY-DRESS CODE DETECTION** Awarded 2nd prize in paper presentation Dress Code Detection at Karpagam University,gained experince at different in-sights DRESS CODE DETECTION USING ML

SKILLS

Languages: Python,Java,C,JavaScript,React,Pig,HTML ,CSS

Tool and Framework: PowerBI,TerraForm,Git,Linux, MongoDB ,Tableau,Tensor Flow,Pandas,Flask,Streamlit, Pyspark, LangChain, RAG,cloud craft, Ex- calidraw,Draw.io, Docker

Other: Machine Learninng,Deep Learning,NLP,vision control,Object oriented programs

TECHNICAL SKILLS

Google Cloud Platform(GCP), DataProc, Vertex AI, Cloud Storage ,AWS,EC2,S3

AREAS OF INTEREST

Data Engineer, Machine Learning, Computer Networks

EXPERIENCE IN CLOUD AND NETWORKS

Data Migration for Taboola (Connexity Client):

Spearheaded the migration of business logic written in legacy PIG SCRIPTs for Click Data processing at Taboola, an industry-leading advertising company.

Successfully transitioned from a local server infrastructure to **Google Cloud Platform (GCP)**, leveraging Dataproc for efficient execution of business logic.

Key Contributions:

Cloud Migration:

Orchestrated the migration from on-premise servers to GCP, deploying **Dataproc** as the primary VM for running business logic.

Utilized **Cloud Storage** (bucket) for seamless input data storage and leveraged SSH connections for job execution.

Dataproc Cluster Management:

Engineered a **10-node Dataproc** cluster to parallelize the processing of click data, optimizing job execution and reducing processing times.

Implemented Java files to validate and execute PIG files, ensuring accuracy and efficiency in data processing.

Logging and Monitoring:

Established comprehensive logging using GCP, offering a dedicated dashboard for job monitoring.

Monitored job success rates, warnings, alerts, errors, and failures, ensuring a proactive approach to issues.

Health Checks and Network Connectivity:

Instituted Monitoring in GCP, actively checking cluster health, capacity, and network connectivity.

Scheduler Implementation:

Implemented scheduling using **Apache Airflow**, automating job runs and ensuring timely processing of click data.

Configured and optimized Airflow workflows for seamless integration with GCP services.

Large Language Models (LLM) models:

Advanced AI Solutions Development:

Spearheaded the development of sophisticated AI solutions, integrating **Hugging Face's** large language models (LLM) and fine-tuning techniques.

Implemented cutting-edge language processing algorithms for the **Gmail Summarizer**, ensuring efficient summarization of user emails through optimized transformer models.

GCP Vertex AI-Powered Multi PDF Reader:

Engineered a high-performance Multi PDF Reader utilizing GCP Vertex AI, emphasizing Optical Character Recognition (OCR) capabilities for precise content extraction.

Integrated the **Langchain framework**, **FSSI** vector database, and RAG for superior comprehension, elevating document understanding to unprecedented levels.

Legal and HR AI with GPT-3.5 Turbo:

Developed specialized Legal AI and HR AI solutions, leveraging the GPT-3.5 Turbo as the foundation for large language models (LLM).

Employed advanced techniques from Langchain, including vector databases and grounding strategies, to enhance contextual awareness and provide nuanced responses to legal and HR inquiries.

AWS Deployment Excellence:

Orchestrated a seamless Proof of Concept (POC) and subsequent deployment on AWS, showcasing technical prowess in infrastructure management.

Established separate EC2 instances to ensure resource isolation and scalability, implementing advanced deployment methodologies for optimal system performance.

PROJECTS DONE

CERVICAL CANCER

CERVICAL CANCER using **Machine learning** Applying machine learning algorithms to analyze cervical cancer data for early detection and personalized treat ment strategies.

Machine Learning ALgorithms

LEARNING DYSLEXIA

Done a project in Learning dyslexia to study Uti lizing machine learning for early identification and personalized intervention **strategies in dyslexia.**

Machine Learning ALgorithms

DRESS CODE DETECTION

Implementing machine learning for automated dress code detection, enhancing compliance monitoring in various settings. - **Dress Code Detection**

open CV YOLO V8

DATA MIGRATION

As a Data Engineer at Centillion Labs, spearheaded successful data migration for the Connexity Taboola client to GCP. Proficiently worked on Dataproc, Cloud Storage, VM Instances, Airflow scheduler, logging, monitoring, Hadoop, and Pig for seamless project execution.

GCP DATAPROC VM INSTANCE

SENTIMENTAL ANALYSIS

Applying NLP for sentiment analysis to discern and in terpret emotions expressed in text data for actionable insights. - **Sentimental Analysis**

Sentimental Analysis NLP Hugging Face

MULTI CHAT

Developed and completed the MultiPDF chat appli cation utilizing Vertex AI models in GCP for real-time instance processing. Implemented concepts such as embedding, vector database (FASSI), pincone, vector search,RAG and grounding. Deployed the application on AWS and integrated Streamlit as the user interface for a seamless and interactive experience - **MultiChat PDF**

Vertex AI GCP Vector DataBase FASSI Paincone

VIRTUAL AGENTS

Developed HR AI and Legal AI bots using GPT models from OpenAI. Integrated Streamlit as the user interface for a user-friendly and interactive experience. Imple mented embedding and vector database techniques, leveraging frameworks such as Langchain for efficient development - **HR and LEGAL AI**

GPT Model | Emberding | LangChain | RAG | Grounding

Currency Recognition

Currency recognition using the ResNet model in deep learning is employed to train on real-time data. The system utilizes a Streamlit UI for user interaction, allowing users to upload or show a rupee note in front of the webcam. The model then recognizes the denomination, distinguishing between various rupee notes such as 2000 and 100. A critical aspect of this project is addressing the needs of visually impaired individuals. To achieve this, a small camera is attached to sunglasses, serving as an integrated system. Before giving or receiving money, the system informs the user, specifically those with visual impairments, about the denomination of the currency. This is achieved through Optical Character Recognition (OCR) technology, which reads and interprets the bill, providing an auditory response to convey the exact amount.

Resnet | OCR | currency | AI model

EXPERIENCE

JUNIOR WE DEVELOPER INTERN IN IDOT

During my 6-month internship at **IDOT**, I served as a Junior Developer Intern, contributing to the develop ment team by actively participating in coding, testing, and debugging tasks. This experience enhanced my programming skills and provided hands-on exposure to real-world software development projects.

TBI SITE | HTML | CSS | JAVA SCRIPT

DATA ENGINEER AT CENTIILION LABS

Led the successful migration of Hadoop architecture to GCP Dataproc for a prominent **US Client,Connexity Taboola**, deploying a 10-node high-performance clus ter. Utilized Pig scripting to streamline business logic execution, optimizing data workflows. Orchestrated jobs efficiently using a Java project for **Cluster Management and Integrated Apache Airflow** for sched uled runs. Developed logging and monitoring so lutions, configured dashboards for real-time health checks, and established a dedicated logging dash board for centralized log viewing.

Centillion Labs | GCP | Migration | DataProc | Taboola | Connexity | AirFlow

LEAD AI DEVELOPER AT CENTIILION LABS

Developed AI-based products,including Multi PDF, Gmail Summarizer, HR AI, Legal AI, and JSON AI chat bots. Utilized Vertex AI in GCP and GPT models in OpenAI for implementation. Employed the Langchain framework for building these solutions

GPT | Grounding | RAG | Vertex AI | Gmail Summarizer | HR | LEGAL

EDUCATION

UG- 2021-2025

Kongu Engineering College
Thesis Currently Purusing 3rd Year AI Sixth semester with current **CGPA-7.64**

10TH GRADE 2018-2019

SSV School, ParthasarathiPuram,Tiruppur
Thesis Completed by SSLC with **74 percentage**

12TH GRADE 2021

Veveaham Matric Hr.sec School
Thesis Completed 12 Grade with **85 percentage**.

INOVATION CHALLENGES

- Empowering the blind with Smart Glass and AI pow ered financial Assistance presented - POC Level.
- Light transmitting using AI in Smart India Hack thon(SIH) TRL-2

CLUBS AND MEMBERSHIP

- As the **Secretary** for our **AI Department**, I stream line administrative processes, coordinate departmen tal events, and facilitate effective communication. My role involves meticulous planning and collaboration to support the department's objectives.
- **Active member of CSI** (Computer Society of India), contributing to a dynamic community of technology enthusiasts.