

# Anbu Valluvan Devadasan

[anbu.valluvan@outlook.com](mailto:anbu.valluvan@outlook.com) | 408-646-0130 | [LinkedIn-Profile](#) | [github.com/AV-D](https://github.com/AV-D) | [anbuvalluvan.me](http://anbuvalluvan.me)

Solutions-focused professional combining 10 years of enterprise experience with advanced Data Science expertise to bridge business challenges and AI-powered solutions. Demonstrated success in stakeholder management and process optimization with proven ability to translate complex technical concepts into actionable business strategies, while consistently exceeding performance metrics. Seeking to apply this unique combination of technical depth and business acumen to accelerate AI adoption and create measurable business value.

## Education

M.S., Applied Data Science & AI,  
San Jose State University, Dec 2024

B.S., Computer Science,  
Osmania University, Apr 2011

## Research Projects

Blood Pressure Monitoring using Edge AI & Deep Learning (Proof of Concept) Dec 2024

*Tensorflow, Tensorflow Lite, Arduino Nano, Python, Scikit-learn, Numpy, Scipy*

- Architected deep learning framework using TensorFlow for continuous blood pressure estimation, implementing custom signal processing algorithms for PPG waveform analysis and feature extraction
- Developed novel AttentiveBPNet architecture achieving medical-grade accuracy standards (MAE: 6.36mmHg SBP, 4.09mmHg DBP) on MIMIC IV waveform database
- Optimized neural network for edge deployment using TensorFlow Lite, achieving 90.71% model compression (0.13MB) while maintaining accuracy for Arduino implementation
- Developed and trained SVM and regression-based model as baseline models locally in HP PC using Z by HP AI Studio, leveraging its integrated development environment for model experimentation and training

Retrieval Augmented Generation System with Mistral 7B & LangChain Sep 2024

*Langchain, Mistral 7B, FAISS, PyPDF, HuggingFace, Python, PromptEngineering, ChromDB*

- Engineered RAG system leveraging Mistral 7B LLM, LangChain framework and FAISS vector storage implementing advanced context retrieval and real-time response streaming with call back management.
- Optimized model performance through quantization (Q4\_K\_M) and F16 key-value cache, achieving 4096 token context processing with configurable batch operations, while maintaining high inference accuracy through custom prompt engineering
- Developed comprehensive document processing pipeline integrating PyPDF for extraction, ChromaDB for vector storage, semantic search capabilities and custom chunking strategies, implementing bi-encoder architecture for dense passage retrieval and hierarchical document

Multi Label Toxic Comment Classifier – Natural Language Processing (NLP) July 2024

*PyTorch, LSTM, BERT, NLP, Numpy, Pandas, NLTK, Scikit-Learn*

- Built a NLP classification system using PyTorch and BERT to detect 6 categories of toxic comments, implementing custom loss functions and achieving a 94% accuracy.
- Developed a comprehensive text preprocessing pipeline with NLTK for tokenization, lemmatization and cleaning, featuring dynamic token length computation and vocabulary based embedding dimensioning.
- Implemented and optimized neural network architectures including Bi-LSTM with dropout regularization, embedding layers and attention mechanisms. Experimented with GRU variants and transformer models.

German to English Translation Model – Natural Language Processing (NLP) Apr 2024

*PyTorch, Transformer, Torchtext, NLP, Spacy and Numpy*

- Developed a machine translation system using Pytorch Transformer module to translate German sentences into English, leveraging the Multi30k dataset for training and achieving effective translation performance
- Implemented text preprocessing pipeline with Spacy for tokenization and vocabulary creation. Built and optimized Seq2Seq transformer architecture featuring embedding layers, multi-head attention mechanisms, feedforward networks and dropout regularization. Xavier initialization was applied to stabilize training dynamics.
- Greedy decoding algorithm was used for inference enabling sequential generation of target tokens until the end of sequence token was reached.

Object detection System – Computer Vision – Yolo V8 Dec 2023

*PyTorch, Ultralytics, YOLOv8, OpenCV, Numpy*

- Implemented object detection system using YOLOv8 architecture to detect and classify 5 different classes, achieving 88.9% mean Average Precision.
- Built an end to end training pipeline using the Ultralytics YOLO framework with data augmentation techniques and the model demonstrated robust real time detection capabilities with 0.6ms pre-processing and 15.8ms inference times and strong per class performance across all 5 classes.

Yelp Review Sentiment Analysis streaming pipeline – Big Data Analytics Oct 2023

*Google Cloud Platform (GCP), Apache Spark, MLlib, BigQuery, Pub/Sub, Dataproc, Word2Vec*

- Engineered a scalable sentiment analysis system for YELP reviews using GCP services processing 9GB of historical data in JSON format and ~500 daily streaming reviews (via API calls) through pub/sub achieving 95.7% ROC AUC score using Linear SVM classifier
- Built a Pyspark data preprocessing pipeline on Dataproc clusters for text normalization (tokenization, stopword removal, stemming) and feature engineering including TF-IDF vectorization, Word2Vec embedding and LDA topic modeling
- Implemented a real time streaming architecture integration Pub/Sub for data ingestion, BigQuery for storage and Dataproc for distributed processing. Deployed optimized sentiment model as microservice to analyze streaming reviews and generate insights through BigQuery dashboard.

Facial Recognition System for Missing Person Identification Sep 2023

*PyTorch, OpenCV, Computer Vision, Video Processing, CNN, GAN, Siamese Network*

- Architected facial recognition system utilizing ensemble of deep learning models achieving 93% recall and 89% precision in identifying missing persons from video streams through multi tiered detection, alignment and feature extraction pipeline
- Implemented comprehensive data engineering workflow processing 62000 video frames across LFW and Youtube faces datasets incorporating image standardization, quality analysis and augmentation techniques with PCA based dimensionality reduction for optimal model training.
- Developed scalable video processing pipeline integrating real time face detection using Faster RCNN, YOLO models and CNN based facial landmark regression and feature extraction through DeepID2+/ VGG face networks for accurate face matching and temporal fusion

## Publications

Devadasan, Anbu Valluvan., Veerla, S., Masum, M., Chowdhury, M., Shahriar, H., "E-SMOTE: Entropy based Minority Oversampling for Heart Failure and AIDS Clinical Trails Analysis," in 2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC), Osaka, Japan, 2024.

## Professional Experience

### San Jose State University

May 2023 - Dec 2024

#### Graduate Research Assistant

- Led research initiatives under Prof. Masum in Edge Artificial Intelligence and edge computing, focusing on optimized machine learning algorithms for healthcare monitoring systems. Specifically, concentrated on **blood pressure monitoring applications** supported by the RSCA 2024 research grant, with a specialized emphasis on implementing neural networks for real-time monitoring at Edge devices
- Contributed to pioneering research on **entropy-based oversampling technique (E-SMOTE)** to address imbalanced medical datasets, resulting in a more balanced representation of data and presented research outcomes at the IEEE COMPSAC 2024 International Conference, utilizing Python programming and Machine learning algorithms

### State Bank of India

Dec 2013 – January 2023

#### Business Intelligence Analyst

June 2021 – January 2023

- Led the execution of sentiment analysis implementation across various customer touchpoints in Hyderabad region driving a **21% surge in customer satisfaction** through data driven service enhancement and proactive issue resolutions
- Collaborated with the senior analytics team to conduct market research on retail asset products using data analytics tools providing crucial insights that **doubled the average loan ticket size to \$6000** and paved the way for the successful launch of the innovative YONO Retail Xpress Credit (RTXC) platform
- Developed data analytics reports on retail assets post YONO digital product launches, focusing on Pre-Approved Personal Loans and Pension Loans, utilizing Power BI and Python tools resulting in **13% increase in asset penetration** metrics
- Elevated the YONO Pre-Approved Business Loan platform for Small Enterprises resulting in **8% increase in loan origination volume** by fostering cross functional coordination, aligning with customer needs, and meeting market demands
- Achieved a **23% increase in conversion rates** by conducting targeted analysis of CRM data to identify high potential customers for savings products and mutual funds, utilizing CRMNEXT for data analytics

#### Manager, Branch & Risk Operations

January 2019 – June 2021

- Streamlined the management of a mid-sized urban branch alongside a 20-member team, **bolstering the branch asset portfolio by 38% to \$2.5B** through strategic customer segmentation methodologies driven by analytics data from the bank and CRM insights via CRMNEXT sales
- Conducted Excel based trend analysis using pivot tables for proactive stressed assets account management, resulting in **34% decrease in non-performing assets from \$80M to \$53M**
- Employed excel for data analytics techniques to drive a **40% year-on-year surge** in branch deposits from \$1.2B to \$1.6B alongside integrating market research, cross selling strategies, and customer segmentation
- **Boosted** customer satisfaction scores from **3.1 to 4.2 out of 5** amidst the COVID 19 pandemic while utilizing Excel tools for tracking and analyzing key performance indicators, empowering data-informed decision making

#### Credit Analyst

January 2016 – Dec 2018

- Managed an asset portfolio of \$530M with a special focus on agriculture and Self-Help Group (SHG) advances in rural banking, driving an **18% increase in market share** through innovative sale practices and personalized engagement across 14 villages. Proactively analyzed customer needs and tailored solutions, achieving 110% of the allocated advance sales goals while fostering trust and long-term client relationships in underserved markets.

#### Assistant Manager

Dec 2013 – January 2016

- Managed comprehensive foreign exchange operations by executing financial instruments with strict adherence to FFMC and RBI guidelines, conducting thorough customer due diligence ensuring meticulous regulatory compliance and documentation for audits, while **building strategic client relationships that drove 19% growth** in foreign exchange sales revenue through personalized solutions, customer service and efficient issue resolution

## Skills

### Technical Skills

- Programming Languages (Python and SQL), Libraries (Pandas, Numpy, Seaborn, Matplotlib, Scikit-learn)
- Machine Learning and Artificial Intelligence: Tensorflow, Tensorflow Lite, PyTorch, Keras, MLflow, LLM Framework (Langchain, HuggingFace, Mistral), ML DevOps, RAG Systems, Model Optimization (Q4\_K\_M, KV cache), Prompt Engineering (Few Shot learning, Single Shot learning, Chain of Thought)
- Databases & Storage: MySQL, MongoDB, Snowflake, Vector Databases (FAISS, ChromaDB), Amazon Redshift and Google Big Query
- Cloud and Visualization: Amazon Web Services (AWS), Google Cloud Platform (GCP), Azure Cloud, Tableau, PowerBI, Looker

**Business Skills** : Market Trend Analysis, Strategic planning, Competitive Analysis, Customer Need Assessment, KPI development, Dashboard creation, Agile Methodologies, CRISP-DM Framework, Data Story telling, Stakeholder Management, Consultative Selling, Relationship building, Cross functional Collaboration, ability to work independently, RFPs, Sales Strategy, Revenue Planning and Analysis, Business Sales