

TARUN KUMAR DEGALA

Associate Data Scientist

Email: dvtarun1@gmail.com

Mobile: +91 7799200960

LinkedIn: [linkedin.com/in/tarun-kumar-degala-venkata-7419061a5](https://www.linkedin.com/in/tarun-kumar-degala-venkata-7419061a5)

Career Objective

To apply my deep technical expertise and innovative problem-solving capabilities in a dynamic, forward-thinking organization. Committed to continuous professional development, I excel in collaborative environments and have a proven ability to communicate complex, data-driven insights.

Professional Summary

- GCP Certified Professional Machine Learning Engineer
- GCP Certified Professional Data Engineer
- Proficient in conversational Japanese
- Total 6+ years of experience in AI, ML, and Python.
- 4+ years of experience in Google Cloud Platform.
- 4 Years experience Natural Language Processing (NLP)
- 4 Year Experience in (GenAI) including areas such as prompt engineering, and prompt tuning.
- Extensive Experience on working with Google Gemini API, GPT, Dall-E, and Large Language Models such as BERT and PaLM
- Skilled in utilizing AutoML for automated model development
- Expertise in deploying models using Cloud Functions, App Engine, Kubernetes Engine, and Cloud Run.
- Proficient in leveraging Cloud Storage, BigQuery, Dataflow, Dataproc, and Data Fusion for efficient data engineering and storage
- Hands-on experience with Vertex AI, AI Platform Training, and AI Platform Prediction for building, training, and deploying machine learning models.
- Experience with Cloud Composer, monitoring with Cloud Monitoring, and ensuring security with Cloud Security Command Centre and Cloud IAM.
- Experience in building chat-bots using Google Dialog Flow.
- Expert in Python machine learning and deep learning libraries such as Tensorflow, Keras, NLTK, Spacy, NumPy, Pandas, and Matplotlib.
- Highly Skilled in SQL

Skills

Operating System	Windows
Languages	Python 3.13, C, SQL
IDES	Spyder, Jupyter, Visual Studio Code, and Google Collaboratory (Colab)
Frameworks	TensorFlow, Keras, Kubernetes, Docker
Artificial Intelligence	Generative AI, Prompt Engineering, Prompt Tuning, Hyper Parameter Tuning, Large Language Models(LLM) like BERT, PaLM, Gemini API, Google DialogFlow, Natural Language Processing.
GCP Services	Vertex AI, Cloud AutoML, AI Platform, BigQuery, Cloud Storage, Compute Engine, Cloud Functions, App Engine, Bigtable, Cloud Endpoints, Cloud Firestore, Cloud Logging, Cloud Monitoring, Cloud Natural Language API, Cloud Run, Cloud Scheduler, Cloud Search, Cloud Spanner, Cloud Text-to-Speech API, Cloud Translation API, Compute Engine, Data Catalog, Data Fusion, Dataflow, Dataproc, Deployment Manager, Dialogflow, Endpoint API, Error Reporting, Firebase, Logging

Professional Experience

Project:1 - Company Name: Accenture	
Name of the Project	Transcription Service
Job Title	Ind & Func AI Decision Science Analyst
Tools and Environment	Python 3.13, Speech to Text API, Cloud Storage, PLX Scripts, Google Collaboratory (Colab)

Project Description:

- To provide a unified and highly accurate transcription service empowering businesses to unlock valuable audio insights.
- To build configurable, unified service broadening its potential applications and making it more adaptable to diverse needs.
- To expand language support with better accuracy by model adaptation and other approaches
- To baseline quality metrics for 10+ languages across clients and continuously measure quality, integrating API with ML foundations.
- Incorporate a user-friendly widget providing a more intuitive and insightful user experience, including side by side audio and transcripts.

Roles and Responsibilities:

- Improve application capabilities through transcript translation, multi-language transcription, audio dubbing content filtering, and speaker diarization (detecting different speakers in an audio recording).
- Improving Advanced Language Support:
 - Improving language detection by configuring agent self-language settings.
- Quality Improvements:
 - Adaptation of new models and approaches.
 - Editing Transcripts and refining the model.
- Quality Insights:
 - Automated Transcription Quality Insights.
- Advanced Transcription Support:
 - Worked on bi-directional real time transcript translations.
 - Worked on user UI(widget) for transcribed audio player
- Worked on Advanced Features:
 - Transcript search API's
 - Audio and transcript content filtering
 - Speaker identification
 - Speaker attribution
 - Re-transcription and model switching
 - Automatic punctuations
- Audio
 - Worked on audio dubbing and translation within audio recording player
 - Designed model to use ML foundation capabilities
- Worked on Quality metrics and dashboard.

Project:2 - Company Name: Accenture	
Name of the Project	Niki
Job Title	Ind & Func AI Decision Science Analyst
Tools and Environment	Python 3.13, BERT, Google Colaboratory (Colab), BigQuery, Large Language Models.

Project Description:

- Leveraging email intent detection and instant response technologies for optimizing FinX's operational efficiency and customer deflection. It is a continuously evolving classification model.
- Providing Customer support solution to salesforce where 40% of the cases will be answered quickly out of 600k cases.

Roles and Responsibilities:

- Improved accuracy of the model from **74%** to **87%**.
- Improved deflection rate through model refinements and tailored confidence scores.
- Worked on automatic email generation which provides an automated email response, either a canned response or a link to the portal that enables users to lookup information.
- Extract information from email body to enrich cases
- Continued support for the deployed model.

Project:3 - Company Name: Accenture	
Name of the Project	Data Analytics Solutions
Job Title	Ind & Func AI Decision Science Analyst
Tools and Environment	Python 3.13, Gemini API , Generative AI, Prompt Engineering SQL, Google Cloud Platform (Vertex AI, Big Query), Google Colaboratory (Colab), PLX Dashboard,

Project Description:

Engineered a data quality assurance framework leveraging Gemini API. This solution automates the identification of data quality issues, including uniqueness, data type conformity, completeness, consistency, and timeliness. By analyzing complex datasets and extracting patterns, the framework generates comprehensive reports to empower data-driven decision-making.

Roles and Responsibilities:

- Engineered and deployed a data quality framework using Gemini API to identify and flag violations.
- Leveraged Python libraries for data preprocessing and analysis.
- Implemented prompt engineering techniques to extract precise data quality rules from Gemini.
- Utilized BigQuery for efficient data querying and extraction.
- Fine-tuned machine learning models using advanced techniques like chain-of-thought reasoning.
- Developed interactive PLX dashboards to visualize data quality insights.

Project:4 - Company Name: Accenture	
Name of the Project	Databot
Job Title	Ind & Func AI Decision Science Analyst
Tools and Environment	Python 3.13, Gemini API , Generative AI, Prompt Engineering, Dialog Flow, , Google Cloud Platform (Vertex AI)

Project Description:

Engineered a sophisticated chatbot leveraging **GenAI** and **Dialogflow** to facilitate natural language interactions with complex datasets. The chatbot efficiently processes user input, executes complex queries, and presents relevant information in a user-friendly format.

Roles and Responsibilities:

- Designed and implemented a robust knowledge base and prompt engineering strategies for the chatbot.
- Built a user-friendly Dialog flow chatbot to enable natural language interaction with the data.
- Developed a front-end interface to facilitate seamless user engagement with the chatbot.
- Rigorously tested and chatbot for optimal performance and accuracy.

Project:5 - Company Name: Accenture

Name of the Project	DLP
Job Title	Ind & Func AI Decision Science Analyst
Tools and Environment	Python 3.9,HTML,CSS,Google cloud platform, Google AppsScript, Generative AI, Big Query, Vertex AI, Colab

Project Description:

The client wants to incorporate a tool that conducts sentiment analysis, as well as a Dialogflow chatbot that provides sentiment analysis for user queries.

Improving the performance and accuracy of the system by implementing Gemini API andGenerative AI-related technology such as prompt engineering and prompt tuning.

Working on Large Language Models (LLMs) like BERT, PaLM and fine-tuning them using techniques such as LoRA tuning and hyper-parameter tuning.

Roles & Responsibilities:

- Working on LLM along with prompts
- Prompt Creation
- Prompt tuning
- Designing and implementing Large Language Models
- Preprocessing data
- Improving training methodologies
- Fine-tuning the models to improve their accuracy and precision using techniques like LoRa tuning, soft attention and hyper parameter tuning.
- Creating Dialogflow chatbots to resolve user queries
- Working on labelling and annotating the data provided by the models
- Monitoring the performance of the models

Project:6 - Company Name: Accenture

Name of the Project	Vertex AI Templates
Job Title	Ind & Func AI Decision Science Analyst
Tools and Environment	Python 3.13, BERT, Vertex AI, Jupyter, Cloud Storage, BigQuery

Project Description:

- Engineered Vertex AI-based templates to streamline the development of regression, clustering, and classification models.
- Enhanced semantic analysis capabilities with Distilled BERT-based templates to improve decision-making and customer experience.

Roles & Responsibilities:

- Collaborated with cross-functional teams (e.g., data engineers, product managers) to understand business requirements and translate them into technical solutions.
- Created and refined reusable templates to accelerate model development and deployment.
- Optimized model performance and accuracy
- Cleansed, transformed, and prepared data for model training.
- Conducted exploratory data analysis (EDA) to identify patterns and insights.
- Documented model development processes and best practices.

Project:7 - Company Name: Accenture	
Name of the Project	CORTEX Prediction of lead time for goods
Job Title	Ind & Func AI Decision Science Analyst
Tools and Environment	Python 3.13, Big Query, Kubeflow, AutoML, BQML, Vertex AI, Cloud Endpoints

Project Description:

- Develop a baseline AUTOML regression model on the Google cloud environment to predict the lead time for products in SAP environment.
- Streamline data pipelines using Kubeflow to automate data ingestion and model training, reducing manual effort and increasing accuracy.
- Create Vertex AI Endpoint, Implement batch and online prediction.

Roles & Responsibilities:

- Collaborated with stakeholders to define requirements.
- Developed an AutoML regression model on GCP to predict SAP product lead times.
- Orchestrated data pipelines using Kubeflow to automate data ingestion and model training.
- Implemented model monitoring and alerting mechanisms to ensure model performance.
- Stored batch prediction results in BigQuery for dashboarding in Looker.
- Utilized online predictions to gain insights into feature importance and model behavior.
- Leveraged BigQuery ML to deliver accurate and timely insights, driving data-driven decision-making.
- Utilized BigQuery's powerful querying capabilities to extract valuable information from large datasets.
- Generate Synthetic data using Python libraries.
- Triggering dags in Airflow, pylintng of code, mock data creation.

Project:8 - Company Name: Tech Mahindra	
Name of the Project	Sales Forecasting for Pharma Client
Job Title	Associate Analyst
Tools and Environment	Python 3.13, Jupyter Notebook, Pandas, NumPy, Scikit-learn, Statsmodels

Project Description:

- To accurately forecast sales for multiple products using Arima Model
- Effectively capturing and modelling seasonal trends, the model will enable businesses to optimize inventory levels, improve production planning, and make informed decisions to drive revenue growth.

Roles & Responsibilities:

- Collaborate with business stakeholders to understand their forecasting needs and translate them into actionable insights
- Gather historical sales data for multiple products from various sources (e.g., ERP systems, CRM systems, spreadsheets).
- Cleanse and preprocess the data to handle missing values, outliers, and inconsistencies.
- Implement an ARIMA model to capture the time series components (trend, seasonality, and residual).
- Experiment with different model configurations (e.g., ARIMA(p,d,q)(P,D,Q)s) to find the optimal parameters.
- Evaluate model performance using appropriate metrics (e.g., Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE)).

Project:9 - Company Name: Tech Mahindra	
Name of the Project	Customer Churn prediction for an Insurance company
Job Title	Associate Analyst
Tools and Environment	Python 3.13,

Project Description:

- Develop a KNN-based churn prediction model to identify at-risk customers.
- Utilize K-means clustering to identify distinct customer segments based on behavioural and demographic characteristics.

Roles & Responsibilities:

- Gather and clean customer data from various sources (e.g., CRM, transactional data).
- Feature engineer relevant variables to improve model performance.
- Develop a KNN-based churn prediction model to identify customers at risk of churning.
- Experiment with different hyper-parameters (e.g., number of neighbors, distance metric) to optimize model performance.
- Apply K-means clustering to segment customers based on their behavior and demographics.
- Analyze the characteristics of each customer segment to identify opportunities for targeted marketing and retention strategies.
- Communicate findings and recommendations effectively through clear visualizations and reports.
- Present insights to decision-makers and provide guidance on how to leverage the model to improve customer retention.

Certifications and Achievements

- GCP Certified Professional Machine Learning Engineer
- GCP Certified Professional Data Engineer
- Proficient in Conversational Japanese.
- Won prizes in multiple football tournaments.
- Honored with an award for exceptional athleticism and teamwork in the inter-sports football tournament.

Languages Known

Language	Proficiency
English	Fluent
Hindi	Fluent
Telugu	Native Language
Japanese	Conversational