# Personal Profile

## Introduction

I am Sumukhi Ganesan, a data enthusiast currently enrolled in the last semester of MS in Data Science program. With a background in Electronics and Communication Engineering and Business Intelligence, I am excited to leverage my skills and experience to contribute meaningfully to this project.

## Skills

* Programming Languages: Proficient in Python, SQL. Intermediate in C++ and basics of R.
* Statistical Analysis, Probability models and Linear Algebra
* ML Algorithms and Concepts
* ML Frameworks: Tensorflow, PyTorch, Scikit-learn
* Cloud Infrastructure: Practical experience in building data pipelines on MS Azure
* Data Visualization: Matplotlib, Seaborn, Tableau, Power BI
* Data Modeling for Business Intelligence
* Open CV and Open GL for Computer Vision
* Deep Learning Models and Concepts
* Natural Language Processing

## Work Experience

In my 4+ years of professional experience as a Business Intelligence and Analytics Consultant, I have had the opportunity to work with datasets from various industries including Retail, FMCG, Tobacco and Finance. This has equipped me with practical insights in data engineering, analysis, and modeling, allowing me to apply theoretical knowledge in real-world scenarios.

Prior Projects

* CookMate

A recipe recommendation system that identifies ingredients using a live camera feed and suggests appropriate recipes according to preset user preferences.

Role: CV Data Scientist – Identified and Implemented Vision models for ingredient identification, curated ingredient dataset and achieved 96% accuracy using transfer learning on MobileNetV2.

Technologies Used: OpenCV for image processing, Tensorflow for Deep Learning

* Decoding Dropouts: Understanding the relationship between Curricular Metrics and Student Attrition in CS Education

A research project in the field of Learning Analytics (particularly Curricular Analytics).

Role: Student Researcher – Explored a dataset of 12k+ students to identify patterns of dropouts from an undergraduate CS program. Performed a correlation study to discover an 87% correlation between centrality and attrition.

Technologies Used: Python for heuristic and statistical data analysis

* Implementing CutLER

Reproducing 2023 paper on CutLER by Meta AI. Rebuilt the novel object detection and instance segmentation model from scratch on Tensorflow (original was built using PyTorch).

Role: ML Engineer – Interpreted the “Cut-and-Learn” approach and translated the model to Tensorflow; Tested the efficiency of the rebuilt model on KITTI dataset for traffic sign detection

Technologies Used: ML Frameworks – Tensorflow and PyTorch

* SQLearn

A SQL Auto Grader that uses a hybrid approach to grade SQL queries based on its structure. It employs Abstract Syntax Trees (ASTs) to decode the SQL query structure and uses distance metrics to compare instructor and student queries.

Role: Team Contributor

Technologies Used: Python for control flow and method testing

## Interesting Topics

I am passionate about the applications of data science and machine learning in Education. I am particularly intrigued by Learning Analytics, TinyML, Applied Computer Vision. I am excited to explore these further during the capstone project.

## Capstone Project Interests

I am eager to apply my skills and knowledge to the Data Science Capstone Project, contributing to its success and gaining valuable insights into challenges in Education. I am currently considering a Handwritten Mathematical Expression Recognition (HEMR) System for auto grading K-12 Math assignments.