# Project Initiation

This document presents the project ideas and descriptions for DS 5500 Capstone Project. I, Sumukhi Ganesan, will be the sole contributor in this project.

As stated in the personal profile document, my topic of choice is Auto-Grading of K-12 math assignments. The aim of this project is to facilitate the auto-grading of K-12 math assignments that predominantly consists of basic algebra, geometry, and word problems. This process presents three major sub-problems –

1. Mathematical Expression Recognition (MER) – recognizing handwritten/printed mathematical characters and images.
2. Mathematical Language Processing (MLP) – processing the semantics behind mathematical text (like NLP). This is needed specifically for word problems.
3. Multi-modal math reasoning – Common K-12 math problems and solutions are a combination of text and images/diagrams. The system needs not only to recognize them, but also understand them so that multiple solutions / partial grading is possible.

Upon reviewing the existing literature on these sub-problems, I have identified several works that demonstrate the applications of deep NNs in MER. MLP is approached like NLP and popular LLMs have been evaluated on mathematical text. Lastly, multi-modal math reasoning is handled by LLM/LVM combined models. Research on math auto grading revolves around the challenges in parsing responses to open-ended questions in math and partial grading.

For this capstone project, I aim to start with an effective OCR + image interpretation system that can parse math questions and answers that contain text and image inputs. CROHME is a publicly available dataset for MER. It contains over 11,000 math expressions (some of them include images) handwritten by hundreds of writers from different countries along with the ground truth in LaTeX for evaluation. A second dataset called MathVista was published recently to study multi-modal reasoning and can also be used for MER.

After MER, the grading of the recognized solutions needs to be approached with either a Large Multi-Modal Model (LMM) or a specifically crafter deep neural network for this specific task. The literature study for the auto-grading part is in progress and I believe that I will be able to define the scope of this part before the next submission.