

CSC4220-001
Data Mining and Machine Learning:
Hypertuning a Medical LLM

Group Members: Edward Gannod
Tania Perdomo Flores
Sharon Colson

- Topic: Fine-tuning a Medical LLM
- Problem Description: American health care services are in crisis. Healthcare professionals are overworked and understaffed. Data science has the opportunity to aid in relieving this crisis through the development of assistive models.
- Objectives: The purpose of this project is to explore and fine tune a Medical LLM.
- Dataset: <https://www.kaggle.com/datasets/jpmiller/layoutlm/data> This dataset includes data and NLP-type models, including LLMS, to accomplish a variety of tasks including sequence to sequence, summarizations, q & a, token labeling, and text classification.