

# 11-711 Assignment 2 : Retrieval Augmented Generation System

**Rajeev Veeraraghavan** \*

rveerara@andrew.cmu.edu

## Abstract

This report summarizes the work done to build a stand alone Retrieval augmented generation system to answers questions on CMU (Carnegie Mellon University). The system uses publicly available data to build a knowledge base, and LLama2-7b model to generate answers to questions.

## 1 Introduction

## 2 Data Creation

To produce a PDF file, pdfL<sup>A</sup>T<sub>E</sub>X is strongly recommended (over original L<sup>A</sup>T<sub>E</sub>X plus dvips+ps2pdf or dvi2pdf). XeL<sup>A</sup>T<sub>E</sub>X also produces PDF files, and is especially suitable for text in non-Latin scripts.

## 3 Model Details

## 4 Results

## 5 Analysis

## Acknowledgements

---

\*This work was done as part of the 11-711 course at Carnegie Mellon University. The assignment started off with a group of 3 members, but after 21st February, I had to complete the assignment individually. Permission to work alone was taken from the course instructor. Permission to use material that the group had worked on together, prior to splitting up, was also taken from the course instructor. Shared work has been attributed to contributors.