Title
AlphaChef - Build a transformer-based model to generate recipes from food videos.

Who
ChenHao Lu (clu63)
Nan Chen (nchen40)
PingYao Shen (pshen6)
Ziao Zhang

Introduction
We are trying to build a model that auto-generates recipes from food videos. While food video is a great way to learn cooking, some videos are too long and may be difficult to memorize all the steps mentioned in the video. We believe it will be useful if we can generate a precise text recipe that gives the learner an overview of the cooking steps.

Related Work
We found an existing paper that are related to our goal:
[UniVL: A Unified Video and Language Pre-Training Model for
Multimodal Understanding and Generation] (https://arxiv.org/pdf/2002.06353v3.pdf)

The paper will only be our reference and starting point.

Data
Youcook2
http://youcook2.eecs.umich.edu/.
Contains 2000 cooking videos on 89 recipes with 14K video clips.

Methodology
We will be using the already preprocessed video data from the YC2 dataset, and build our own preprocessing method for the text data. They will then be encoded separately with two encoders. These encodings will go through a transformer to generate the output text descriptions.

Metrics
We will use common NLP metrics like perplexity, BLEU, etc., to evaluate the performance of our model.

Ethics
1.
We decide to use **Youcook2** as our dataset. We notice that most of the videos are about western dishes. Given that all the videos and recipes in our dataset are in English, this uneven distribution is understandable. Nevertheless, this might affect the model's performance when we try to generate recipes for certain types of dishes.

2.
The stakeholders of this project might be those who are trying to learn to cook a new recipe. Unclear instructions in our generated text recipe may lead to waste of food ingredients, or more severely, kitchen accidents.

Division of labor
ChenHao Lu: encoder and decoder
Nan Chen: preprocess text and video data
PingYao Shen: preprocess text and video data
Ziao Zhang: encoder and decoder