

# Philosophy Professors

Predictive analysis of UMD faculty reviews

# Problem Overview

## 1. What matters in a philosophy class?

Exploring this question can teach us if open debate, better readings, or if the professor themselves is a critical of a critical-thinking oriented class.

## 2. Create a means of determining which matter

We can break this problem down into determining which traits of a class/professor lead to having a higher rating for a particular class

## **3. Develop a predictive model to predict the ratings from a given students reviews**

This will be our solution

# Data preparation



First, review data was collected on 5 philosophy professors through the planet terp API.



Then, the text and ratings of each review was extracted, then converted into a pytorch dataset to later then be tokenized with Hugging Face's Auto Tokenizer package

Now the data was in a flexible format that could be used to fine tune a variety of models!

# Fine tuning (Model Selection)

After experimenting with several models, Hugging Face's 'distilbert-base-uncased' was selected as the foundation of the transformer due to it being the one that worked naturally with Hugging Face's Auto Encoder tokenizer, and had the least compiler hiccups.



# Fine tuning (Training Process)

When provided a default learning rate of  $e^{-5}$  and 3 training epochs, the model significantly struggled to differentiate ratings between positive and negative sentiment reviews, and provided a rating from 0.3-1.2 consistently for any response given with little variance.

Example:

Review	Rating
"The workload was quite heavy, really didn't like this professor. Don't take him."	0.981
"This professor was amazing! Loved the way he taught and made the class engaging"	1.012

When given an extreme higher learning rate of 1, and 10 epochs, a complete collapse of predictions caused all values to flatten to a 4.5647, meaning all tested inputs result in the same positive verdict no matter the sentiment

Gradually decreasing the learning rate from 1 to  $e^{-1}$ ,  $e^{-2}$ , so on and so forth down to  $3e^{-3}$  resulted in a more contrast of ratings between the positive and negative sentiment reviews, with  $3e^{-3}$  being the best, and the one that was decided on, resulting in the following results on the previous example

Review	Rating
"The workload was quite heavy, really didn't like this professor. Don't take him."	1.141
"This professor was amazing! Loved the way he taught and made the class engaging"	4.554

# Analysis of results

To determine which factors matter most in a philosophical class, it was decided to use a *silver lining* approach.

The idea is if it can determine what positive trait of class has the highest influence in terms of increasing the score of a negative review, then we will determine what factor(s) matter most in a class.

From the results of this analysis posted on the right, we can make either one of two conclusions.

**1:** There is no significant influence on the factors in a class in and of themselves if a student has already decided that they do not like the class. From the chart we see that students produce a rating of 2.1-2.3 regardless of the reason of why they produced that rating, and the difference is arguable negligible

**2:** Class culture and community is more important than any other feature, even workload. For the results that did end up being higher, they were based around people: the professor, the discussion, the way the student interacted with others in the class. That mattered most to them, and resulted in a higher predicted rating.

Best Factor: Negative Review Rating Distribution	
Review	Rating
"The workload was quite heavy, really didn't like this professor. Don't take him."	1.141
"I Really didn't like this professor, but there was plenty of open discussion though that I thought was good."	2.388
"I Really didn't like this professor, but there were pretty interesting readings though that I thought was good."	2.157
"I Really didn't like this professor, but there was very little homework though that I thought was good."	2.168
"I Really didn't like this class, but the professor was really engaging though and I thought was good."	2.303

# Bonus question for extra credit?

Does a more verbose review with identical sentiment result in a higher or lower predicted rating?

Using Chat GPT to increase text length while keeping sentiment the same, we obtain the following results

Review	Ratings
"The workload was quite heavy, really didn't like this professor. Don't take him."	1.145
"The workload was pretty heavy, and I honestly didn't enjoy this professor's teaching style at all. The class felt overwhelming most of the time, and I really wouldn't recommend taking him if you can avoid it."	3.464
"The workload in this class was extremely heavy, and I really struggled to stay on top of all the assignments because the professor's teaching style just wasn't helpful or engaging. Most of the time the class felt overwhelming, frustrating, and unnecessarily difficult, so I genuinely wouldn't suggest taking him unless you have absolutely no other option available."	1.895

Has diminishing returns, but seems to increase ratings quite a bit!

# Link to code

[https://drive.google.com/drive/folders/1A0wYrJTqvKk33MCSdPjUORPDYvSEull9?usp=drive\\_link](https://drive.google.com/drive/folders/1A0wYrJTqvKk33MCSdPjUORPDYvSEull9?usp=drive_link)