WEBVTT

1
00:00:00.840 --> 00:00:04.860
Hello and thank you for clicking on my
talk. My name is Jillian Mellen. And

2
00:00:04.860 --> 00:00:09.340
today I will be speaking to you about
predictors of faculty sentiment on

3
00:00:09.340 --> 00:00:13.400
their transition to online teaching.
And if you scan the QR code on the top

4
00:00:13.400 --> 00:00:17.180
right of the slide, you'll find this
talk in pdf form in a google drive

5
00:00:17.180 --> 00:00:22.660
folder along with a transcribed copy of
the talk, thanks to Zoom.

6
00:00:24.140 --> 00:00:31.860
So for this talk, for this project,
rather, we looked at sentiment of

7
00:00:31.870 --> 00:00:37.230
responses to an open ended question
that was part of a large national

8
00:00:37.230 --> 00:00:44.170
survey of physics faculty. For this
portion, I used data from 364

9
00:00:44.170 --> 00:00:48.830
participants, because that is how many
answered the open ended response. I

10
00:00:48.830 --> 00:00:55.740
also used their answers to 14 other

survey questions, including their

11
00:00:55.750 --> 00:00:59.920
institution, department, and teaching
loads, their prior experience with

12
00:00:59.930 --> 00:01:05.620
online instruction, preparation in
transition to online instruction, and

13
00:01:05.620 --> 00:01:06.660
comfort in teaching.

14
00:01:09.240 --> 00:01:15.170
So when when looking at this, um, we
were wondering, first of all, what was

15
00:01:15.170 --> 00:01:18.910
the overall sentiment of their
experience? Did they have good

16
00:01:18.910 --> 00:01:23.850
experiences? Did they have negative
experiences? Did it vary widely? What

17
00:01:23.850 --> 00:01:28.560
was the overall feeling amongst the
people who answered this open ended

18
00:01:28.560 --> 00:01:34.340
question? And to do that, we performed
a sentiment analysis. Also, we wanted to

19
00:01:34.340 --> 00:01:42.510
look at their um other answers, the
answers to the other questions and see

20
00:01:42.510 --> 00:01:48.060
if there was any way to predict the
sentiment that was found in the first

21
00:01:48.740 --> 00:01:54.120
portion of this. So we trained a
machine learning model to generate

22
00:01:54.130 --> 00:01:59.180
sentiment score predictions using
their answers to the other survey

23
00:01:59.180 --> 00:01:59.960
questions.

24
00:02:01.940 --> 00:02:06.560
For those who are not familiar
sentiment analysis is a way of

25
00:02:07.140 --> 00:02:15.140
looking at the sentiment, favorability,
type um, attribute of a body of text.

26
00:02:15.150 --> 00:02:19.970
It can be as low as a single word, and
that's generally how these libraries

27
00:02:19.980 --> 00:02:26.970
are created. You can look at a phrase,
sentence, paragraph, entire document,

28
00:02:26.970 --> 00:02:31.620
probably. But I imagine that would be
fairly complicated for my portion. I

29
00:02:31.630 --> 00:02:38.220
used the responses from the survey
participants and these ranged from a

30
00:02:38.220 --> 00:02:44.440
sentence to a small paragraph. I used a
python library called TextBlob, which

31
00:02:44.450 --> 00:02:50.690
takes apart the individual words and
analyzes them for their sentiment

32
00:02:50.690 --> 00:02:58.460
values that are found in a lexicon that
has stored away within the code. So the

33
00:02:58.470 --> 00:03:05.650
overall sentiment for the responses was
actually slightly positive, 0.1. They

34
00:03:05.660 --> 00:03:13.030
ranged from -1, to positive 0.65 where
-1 is the most negative

35
00:03:13.030 --> 00:03:19.340
possible and 1 is the most positive
possible. And I have an example of the

36
00:03:19.340 --> 00:03:24.060
highest sentiment response and one with
the lowest sentiment response.

37
00:03:25.940 --> 00:03:27.350
I'll let you pause and read those.

38
00:03:29.040 --> 00:03:35.230
So for the predictions and analysis, I
I split my data into training and

39
00:03:35.230 --> 00:03:43.380
testing sets of 75% and 25%, or 273 and
91 participants, meaning 273

40
00:03:43.380 --> 00:03:47.620
participants’ data went into the
training set to train the model. And the

41

00:03:47.620 --> 00:03:52.260
model was tested on 91 participants’
data. I used a Keras sequential

42
00:03:52.260 --> 00:03:56.540
model, which was trained and then used
to generate the score predictions.

43
00:03:56.550 --> 00:04:02.900
So there were 99 predictions for
91 participants, for the 14 survey

44
00:04:02.900 --> 00:04:07.260
questions. And if this this sounds like a
lot of data, it ended up being a lot of

45
00:04:07.270 --> 00:04:12.120
data, which is why I ended up using the
medians of the predictions for each

46
00:04:12.120 --> 00:04:19.890
participant, for each question. From
there, I found correlation values for

47
00:04:19.890 --> 00:04:25.520
the predictions and the original
sentiment scores that were found during

48
00:04:25.520 --> 00:04:27.160
the sentiment analysis.

49
00:04:28.440 --> 00:04:33.770
So the results so far are that the
sentiment is overall positive with a

50
00:04:33.770 --> 00:04:37.910
mean of 0.1 and as far as the
correlations go, I thought this was

51
00:04:37.910 --> 00:04:41.840

fairly interesting - question seven,
which is the size of the largest class

52
00:04:41.840 --> 00:04:51.150
taught online, was very slightly
positive at 0.147 and for question 41,

53
00:04:51.150 --> 00:04:55.450
which was perceived job security, there
was there was a very slight negative

54
00:04:55.450 --> 00:04:59.160
correlation at -0.179.

55
00:05:00.240 --> 00:05:04.330
I do have other correlation values for
the other questions. However, they were

56
00:05:04.340 --> 00:05:05.460
much smaller.

57
00:05:07.140 --> 00:05:13.270
So there were a number of things that
put limitations on this project. First

58
00:05:13.270 --> 00:05:20.200
of all, the data size, 364,
participants really is not a large

59
00:05:20.200 --> 00:05:26.220
amount of people, but it's what I had
and it was still interesting. So I went

60
00:05:26.220 --> 00:05:27.360
with it. And

61
00:05:28.540 --> 00:05:33.990
TextBlob, the lexicon that it uses
for the sentiment value, it is

62
00:05:33.990 --> 00:05:39.220
actually fairly small. It's mainly
comprised of adjectives and there are

63
00:05:39.220 --> 00:05:47.250
about 2900 of them line by line, but
many are repeats as they have different

64
00:05:47.250 --> 00:05:49.960
sentiment scores when used in different
contexts.

65
00:05:51.040 --> 00:05:57.240
And the physics, there's no physics
corpus for training a model like this,

66
00:05:57.240 --> 00:06:01.530
for training the sentiment analysis
type portion, meaning

67
00:06:02.540 --> 00:06:08.180
this this TextBlob library was
actually trained on movie reviews with

68
00:06:08.190 --> 00:06:15.470
an accuracy score of about 75%, which
I understand is not great. Hopefully in

69
00:06:15.470 --> 00:06:19.630
the future there will be some kind of a
physics corpus for this sort of thing

70
00:06:19.640 --> 00:06:23.500
and the sentiment analysis and
qualitative analysis of this type of

71
00:06:23.500 --> 00:06:30.170
thing will improve. In addition,
machine learning is generally kind of a

72
00:06:30.180 --> 00:06:36.120
black box sort of thing where you can
you can put things into it and you can

73
00:06:36.120 --> 00:06:39.330
get things out of it and they look
interesting, but you don't necessarily

74
00:06:39.340 --> 00:06:40.760
know what's going on inside.

75
00:06:42.540 --> 00:06:47.440
And that's not great. Also, it would have been great if there

76
00:06:47.440 --> 00:06:51.750
were more sentiment analysis model
alternatives that I could have used in

77
00:06:51.750 --> 00:06:56.710
addition to this, in order to be able to
use different scores for the

78
00:06:56.710 --> 00:07:00.340
correlation values or for the
predictions and correlation values.

79
00:07:00.350 --> 00:07:07.960
Unfortunately, there are not many that
are open access or easy to use.

80
00:07:10.140 --> 00:07:16.750
So with this in mind, the future work
that I am working on now is looking at

81
00:07:16.750 --> 00:07:22.440
model accuracy. I'm going to be
performing model evaluation firstly

82
00:07:22.450 --> 00:07:29.000

using linear regression and depending
how that works, I may branch out into

83
00:07:29.010 --> 00:07:33.460
other types of regression depending on
what we feel is appropriate.

84
00:07:35.740 --> 00:07:36.760
Thank you very much.

85
00:07:38.240 --> 00:07:43.320
And here's that QR code again along
with a really fantastic reference that

86
00:07:43.320 --> 00:07:47.150
I'm able to use for this especially for
the upcoming evaluation portion.