

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Chat Log Analysis in an Online CS1 Course

Permalink

<https://escholarship.org/uc/item/9z18q934>

Author

Zhao, Stanley Jialiang

Publication Date

2022

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Chat Log Analysis in an Online CS1 Course

A Thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science

in

Computer Science

by

Stanley Zhao

March 2023

Dissertation Committee:

Dr. Frank Vahid, Chairperson

Dr. Paea LePendu

Dr. Mariam Salloum

Copyright by
Stanley Zhao
2022

The Thesis of Stanley Zhao is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

Firstly, I give thanks to my mother, Ming Bao, and father, David Zhao, for tirelessly working and laboring to provide me with love, care, safety, and resources in support of me. Thank you for being there and making my journey possible in the first place.

To my friends and companions of many years: Brad Duran, Krishna Patel, and Bryant Ta. Without our weekly tabletop and board game nights, I would not have been able to pull through with this one.

To all my peers and my cohort of six years in the crucible of higher education: David, Keerthi, Elijah, Varun, Carson, Elise, Carissa, Sam, Janine, and Christian. Without the unflappable might of the computer science study group, I would have faltered and failed many years earlier.

To Dr. Frank Vahid, my advisor and mentor, for your unwavering guidance and steady patience, without either of which I would not have been able to reach this point in my development as both a student and an academic. I am unfathomably grateful.

Finally, I thank the University of California, Riverside (UCR) for providing me with an education and opportunities such as this Ph.D., all my colleagues over the years that have knowingly and unknowingly supported me, and anyone else I forgot to mention or do not have room to mention

TABLE OF CONTENTS

Acknowledgements.....	iv
Table of Contents.....	v
List of Figures.....	ix
List of Tables.....	xii
Chapter 1: Introduction.....	1
1.1 Online Classes.....	1
1.2 Automated Analysis.....	4
Chapter 2: Chat Analysis Tool.....	7
2.1 Introduction.....	7
2.1.1 The Chat Log and it's Qualities.....	9
2.1.2 Making Use of the Chat Analysis Tool (CAT).....	10
2.2 Table Outputs.....	11
2.2.1 Messages Per User Tables.....	12
2.2.2 Chat Log Metrics Tables.....	13
2.2.3 Participation Points Tables.....	13
2.2.4 Poll Statistics Tables.....	14
2.3 Conclusion.....	15
Chapter 3: Participation Points Calculation.....	16
3.1 Introduction.....	16

3.2 Methodology.....	17
3.2.1 Unadjusted Participation Score.....	17
3.2.2 Factoring Message Spread Into Participation.....	18
3.2.2.1 Participation Breakdown for Student A.....	19
3.2.2.2 Participation Breakdown for Student B.....	20
3.2.2.3 Full Breakdown for Student C.....	21
3.2.3 Reasoning for Default Participation Thresholds.....	22
3.2.4 Non-class Related Messages.....	23
3.4 Conclusion.....	25
Chapter 4: Identification of Polls and Poll Messages.....	26
4.1 Introduction.....	26
4.2 Methodology.....	27
4.2.1 How Do We Detect Polls in the Chat Log?.....	27
4.2.2 Poll Message Detection Accuracy.....	28
4.3 Discussion.....	29
4.4 Conclusion.....	29
Chapter 5: Predictors of Course Outcomes.....	30
5.1 Introduction.....	30
5.2 Methodology.....	32
5.2.1 Participation Points as a Predictor.....	32
5.2.2 Poll Messages as a Predictor.....	33
5.2.3.1 Early Participation Points.....	34

5.2.3.2 Early Poll Messages.....	35
5.2.3.3 Earnestness.....	37
5.2.3.4 Challenge Activity Struggle.....	38
5.2.4 Impact of Previous Experience.....	39
5.3 Discussion.....	40
5.4 Conclusion.....	40
Chapter 6: Summary.....	41
References.....	43

LIST OF FIGURES

Figure 1.1.1: Percentage of students in the United States taking distance learning courses from 2012 to 2020 [3].....	2
Figure 1.1.2: Increase in online enrollments to large public institutions between 2019 and 2022 [5].....	3
Figure 1.2.1 An example of a built-in polling system.....	5
Figure 1.2.2 The chat analysis tool (CAT) described in this thesis.....	6
Figure 2.1 Extracting chat logs from Zoom.....	8
Figure 2.2.1 An example raw chat log from a CS1 lecture.....	9
Figure 2.1.2 Flow chart depicting the process of converting recorded class chat logs into analysis results.....	10
Figure 2.2.1 A chart outlining the outputs.....	11
Figure 2.2.1.1 An output table from the chat log in Fig. 2.2 showing number of messages per unique user.....	12
Figure 2.2.1.2 An output table showing every message by timestamp from each unique use.....	12

Figure 2.2.2.1 An output table from the chat log in Fig. 2.2.1 showing number of messages per minute of the recorded chat log.....13

Figure 2.2.3.1 An output table from the chat log in Fig. 2.2.1 depicting the participation point calculation process for each unique user in the uploaded chat logs.....13

Figure 2.2.4.1 An output table from the chat log in Fig. 2.2.1 depicting, for each poll detected in the chat log, the number of messages and total runtime.....14

Figure 3.1 A flowchart depicting how the participation point algorithm generates scores for each user present in a chat log.....16

Figure 3.2.1 A line chart showing the scale between number of messages a student posted and the resulting unadjusted participation score, before penalties, where the set threshold is 5 messages.....17

Figure 3.2.2 An example of the time spread of student A (red) and student B (blue).....18

Figure 3.2.2.1 An example of a weak spread and low message count, student A’s first and last message cover 15% of the total class time.....19

Figure 3.2.2.2 An example of an average spread and message count, student B’s first and last message cover 30% of the total class time.....20

Figure 3.2.2.3 An example of a strong time spread and message count, student C’s first and last message cover 78% of the total class time.....21

Figures 3.2.4.1 Excerpts from a CS1 class chat log showing the non-class related messages.....	23
Figure 3.3.1 A pie chart depicting student opinion on how much the participation points system kept them engaged with the class.....	24
Figure 3.3.2 A pie chart depicting student opinion on how much the participation points system caused extra stress.....	24
Figure 4.1 Excerpt of chat log showing an example of a poll.....	26
Figure 4.2.1 A flowchart depicting the process used to calculate which sections of the chat log are polls.....	27
Figure 4.2.2 Chart showing number of messages per minute, poll messages marked in orange.....	28
Figure 5.1 Graph showing the relationship between number of messages and course grade in a CS1 class.....	31
Figure 5.2.1 Graph showing the relationship between student lecture participation and course grade in a CS1 class.....	32
Figure 5.2.2 Graph showing the relationship between student poll messages and course grade in a CS1 class.....	33

Figure 5.2.3.1 Graph showing the relationship between the first four weeks of lecture participation and midterm score in a CS1 class.....	35
Figure 5.2.3.2 Graph showing the relationship between the first four weeks of recorded poll messages and midterm score in a CS1 class.....	36
Figure 5.2.3.3 Scatterplot showing relation between first four weeks of earnestness data and midterm score in a CS1 class [7].....	37
Figure 5.2.3.4 Scatterplot showing relation between first four weeks of earnestness data and midterm score in a CS1 class [7].....	38
Figure 5.2.4 Scatterplot showing previous experience with programming with regards to participation predicting course total.....	39

LIST OF TABLES

Table 3.2.3 A table showing the change of average number of messages per user.....	22
Table 4.2.2 An accuracy table for the poll detection algorithm.....	28

Chapter 1: Introduction

1.1 Online Classes

As there's been an increase in the number of students taking online classes at universities across the country, there has been a greater need for instructors to be able to track and understand the level of engagement students have with online lectures. Figure 1.1.1 is a Statista figure that shows the massive increase in students who take at least one online class and students who take exclusively online classes between 2019 and 2020.

Particularly in introductory classes, students who take an online course more commonly suffer from a few well-known issues: low levels of engagement [6], high levels of passivity [4], shortening attention spans [2], and the tendency to skip class or to minimize the online lecture and ignore it [1]. This paper seeks to provide instructors with more tools to identify which students in their classes are experiencing these issues, in order to address them earlier in the term.

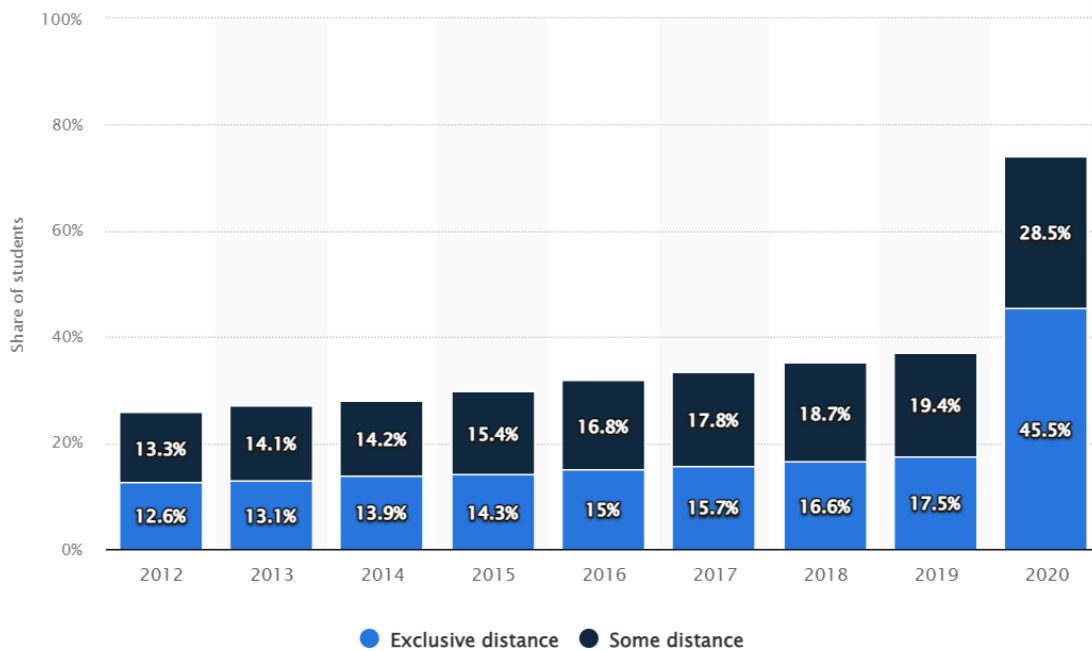


Figure 1.1.1 Percentage of students in the United States taking distance learning courses from 2012 to 2020, [3]

We make the observation that for many new university students going forward, there is an increased likelihood an online introductory class will be their first experience with higher education. Figure 1.1.2 below showcases the large jump in distance education enrollments from 2019 to 2020 at SARA-participating institutions. SARA – the State Authorization Reciprocity Agreement – is an agreement between member states, territories and districts of the United States of America. As of July 2019, 49 states (all but California), the District of Columbia, Puerto Rico and the U.S. Virgin Islands have joined SARA.

93% ↑ Increase in distance education enrollments at SARA-participating institutions between Fall 2019 and Fall 2020

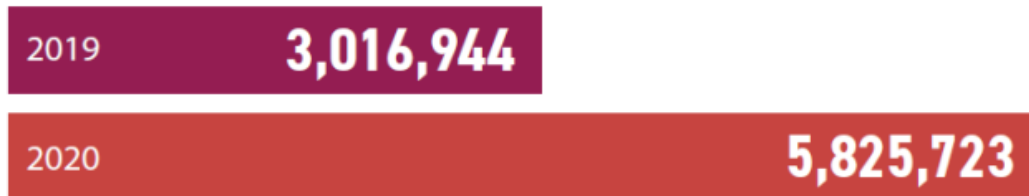


Figure 1.1.2 Increase in online enrollments to large public institutions between 2019 and 2022, [5]

The quantity of online courses continues to grow in the post-pandemic era of higher education. As more students are attending classes remotely, instructors are finding ways to adapt their learning environments for a virtual medium. We will focus specifically on the aspect of student activity and participation in the class, through their usage of the public class chat box. The importance of the chat is in its function as the analogue for students interacting with each other and with the instructor in a physical environment, such as talking questions or having open discussion periods.

1.2 Automated Analysis

Many instructors already have methods of assigning credit to students for participation, whether through self sign-in sheets, clicker questions, or having more informal manual checks. Many teleconferencing apps used for online classes, such as Zoom, already have built-in class survey and polling features that can be used as a way to track who is present and get dynamic feedback from the class. Figure 2.1 below is the default poll on the Zoom mobile application. We found these default options too constricting and simplified to get the level of insight we were looking for. Students using online courses were much more likely to be highly active in the public chat box, using it to interact with other students and to ask questions about the lecture to the professor and the TAs.

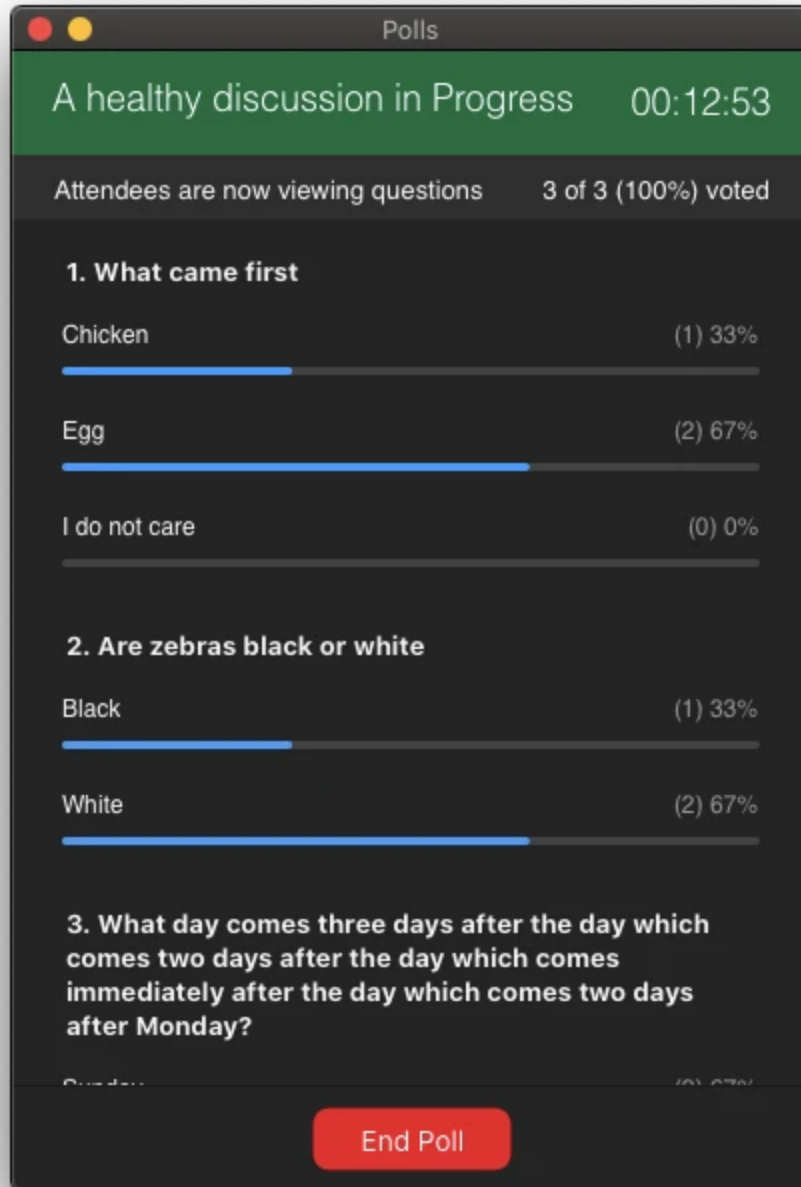


Figure 1.2.1 An example of a built-in polling system

In this thesis, we will describe a tool that analyzes the chat logs of each online class to check for student presence, assign participation points for different levels of engagement, and identify where there were peaks in student activity. Our tool is a lightweight, automated alternative to the more involved manual methods of tracking student participation that are more commonly in use. We will discuss the benefits and drawbacks of using automated participation algorithms in online courses, methods to identify naturally occurring polls within the chat log, and present metrics on how chat log data can be used to predict student performance.

UCR - Zoom Chat Log Analysis (alpha)

We know many classes are being run online. Having students participate via chat is an effective way to run an online class. This tool helps instructors gain insight into class behavior by analyzing zoom chat logs.

[Need help downloading your zoom chat logs?](#)

Questions? Email: analysis-support@cs.ucr.edu

NOTE: This is a beta version with use by invitation only. Please do not circulate this URL without permission. This tool comes with no warranties, express or implied. Use entirely at your own risk. No chat log files are stored. Created by [Joe Michael Allen](#), [Stanley Zhao](#) and [Frank Vahid](#) at UC Riverside

Upload Zoom chat log:

Choose Files

 No file chosen

Use [sample_1.txt](#) ([View](#)) ([Download](#))

Figure 1.2.2 The chat analysis tool (CAT) described in this thesis

Chapter 2: Chat Analysis Tool

2.1 Introduction

Our large public university has been teaching online sections of CS1 quarterly since 2013. The volume of students who take CS1 courses in online sections has only risen in the post-2019 environment. Many instructors who teach online sections make use of a built-in chat feature inherent in most teleconferencing software to interact with students during lectures. Zoom, which is most commonly used at our university, allows these chats to be downloaded as a chat log in a text file format. We use these chat logs to gain more insight on the characteristics of the class and course overall.

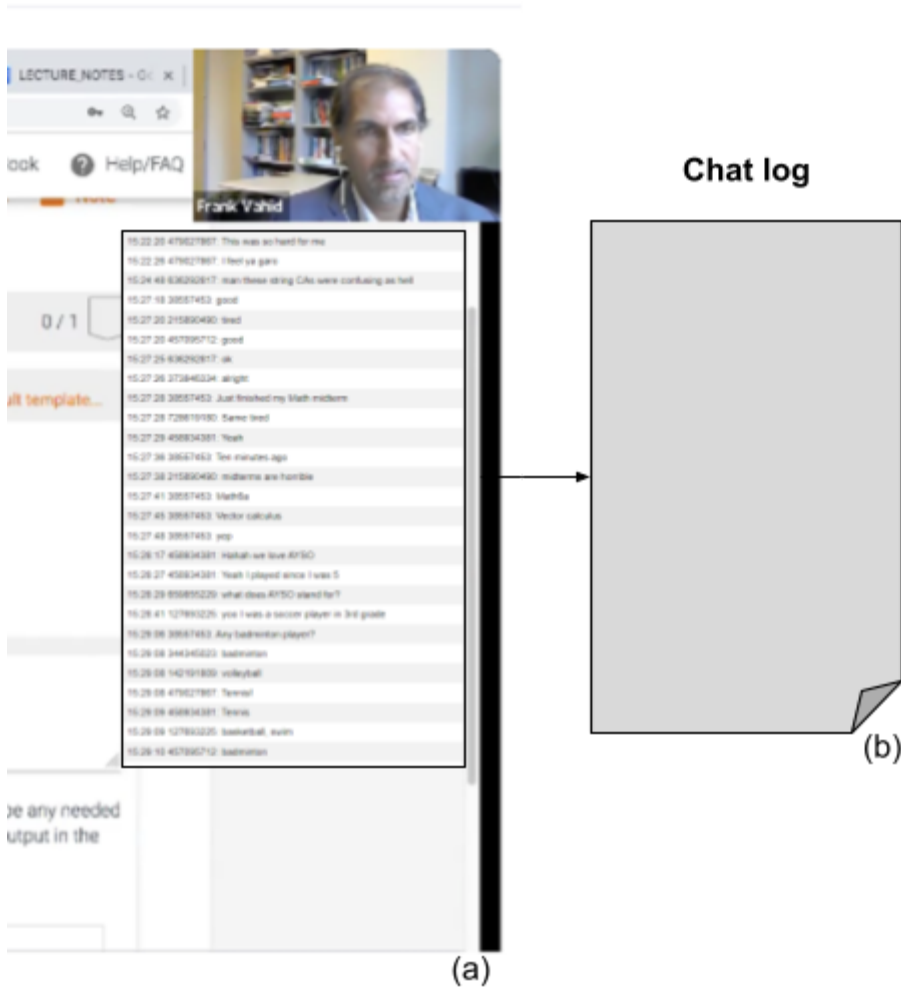


Figure 2.1 Extracting chat logs from Zoom

The record of the active chat log shown in Figure 2.1 (a) is downloadable as a text file, represented by the image in Figure 2.1 (b). The text file containing the stored chat log is used as the input into the tool we developed to analyze the activity levels of the class and its individual students.

2.1.1 The Chat log and it's Qualities



Figure 2.2.1 An example raw chat log from a CS1 lecture

Chat logs contain a recording of the entirety of the messages sent into the public chat box during an online class. Figure 2.2.1 above shows an excerpt from a chat log of a CS1 class, with anonymized student names. Each line of the downloaded chat contains a student's name, their message, and the timestamp of the message. We use the chat log as input to the chat analysis tool (CAT) described below.

2.1.2 Making Use of the Chat Analysis Tool (CAT)

The CAT automates much of the process of tracking and rewarding student participation in online classes. Previously, TAs of the CS1 courses at our university would manually look over chat logs and assign participation points to individual students.

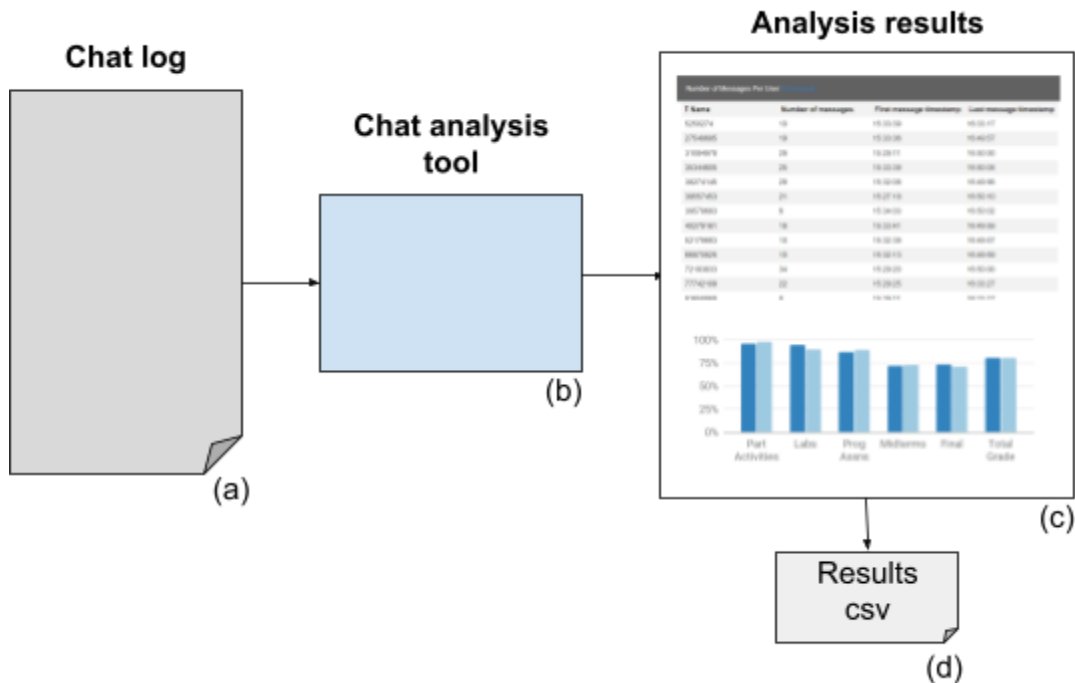


Figure 2.1.2 Flow chart depicting the process of converting recorded class chat logs into analysis results.

The CAT in Figure 2.1.2 (b) takes chat logs as input, and outputs a multitude of analysis tables, all of which can be downloaded as csv files. The CAT has the option to take multiple chat logs as input, and outputs tables containing summarized information from all uploaded chat logs.

2.2 Table Outputs

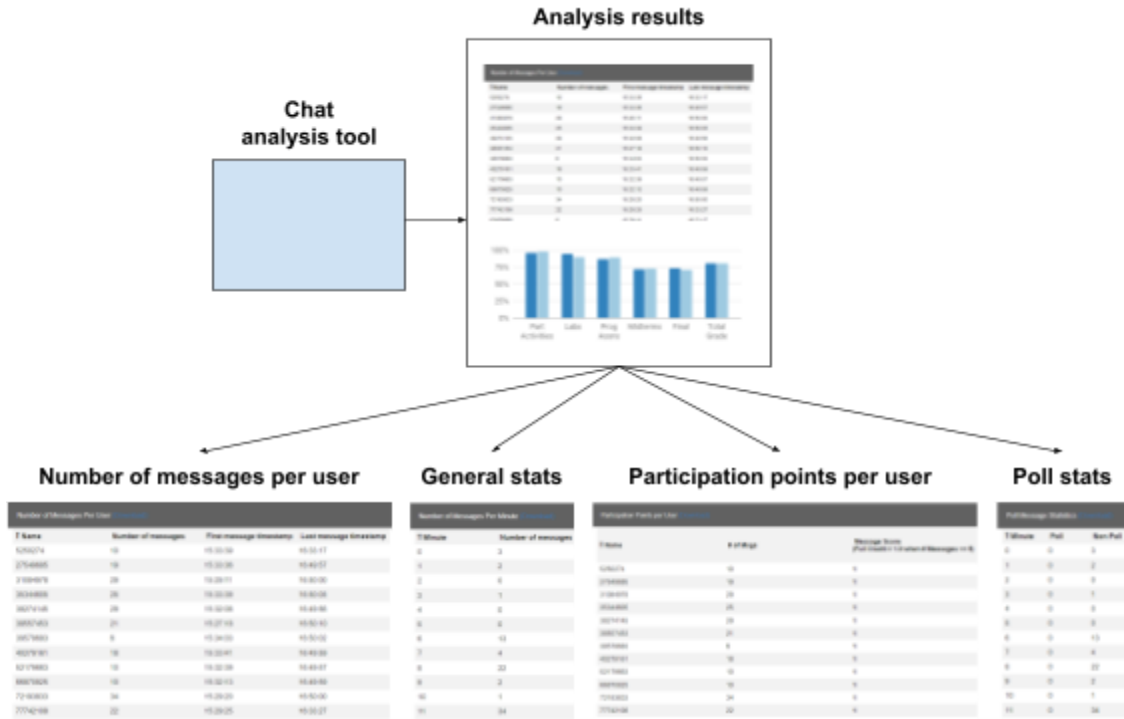


Figure 2.2.1 A chart outlining the outputs

Figure 2.2.1 showcases the four categories of analysis results outputted by the CAT for each chat log. The CAT allows a quick overview of the general class activity level as well as deeper insight on individual students. Calculation of participation points is an automated process.

2.2.1 Messages Per User Tables

Number of Messages Per User (Download)			
↑ Name	Number of messages	First message timestamp	Last message timestamp
ADURG003	6	16:23:12	16:49:51
AKATK003	8	16:23:09	16:49:38
ANGUY295	13	16:23:05	16:50:00
APISU001	5	16:28:52	16:41:22
ASALA135	8	16:22:53	16:49:59

Figure 2.2.1.1 An output table from the chat log in Fig. 2.2 showing number of messages per unique user.

Messages by User by Timestamp (Download)		
User	Message	Timestamp
ADURG003	hii	16:23:12
ADURG003	no	16:25:07
ADURG003	yes	16:28:27
ADURG003	oh wow	16:34:33

Figure 2.2.1.2 An output table showing every message by timestamp from each unique user.

The two messages per user tables allows us to see each student's number of messages and personal history of messages in the chat log. Figure 2.2.1.1 is the main table, which gives us the compiled list of all unique users, their number of messages, and the timestamps of their first and last messages. Figure 2.2.1.2 shows the history of each individual user's activity in the chat log. These analysis results allow instructors to easily tally the activity levels of each of the students in their class.

2.2.2 Chat Log Metrics Tables

Number of Messages Per Minute (Download)		Number of Messages and Unique Users (Download)	
↑ Minute	Number of messages	Number of messages	Unique users
0	2	462	89
1	34		
2	9		
3	32		
4	1		

(a)

(b)

Figure 2.2.2.1 An output table from the chat log in Fig. 2.2.1 showing number of messages per minute of the recorded chat log

The chat log metrics tables gives us general information on the activity levels of class as a whole. Figure 2.2.2.1 (a) is the number of messages recorded each minute in the chat log, and (b) in the same figure is the sum of all messages and sum of all unique users in the chat log. These analysis results are used to calculate local minimums and maximums in the frequency of messages throughout the chat log. We use this data in our participation point and poll identification algorithms, described below.

2.2.3 Participation Points Tables

↑ Name	# of Msgs	Message Score (Full Credit = 1.0 when # Messages >= 5)	First Timestamp	Last Timestamp	Time Spread (Must be over 30.0 percent of total class time)	Penalty Multiplier based on Time Spread (Time Spread / Total Class Time) / (0.30) 1.0 = No Penalty)	Final Score (Message Score * Penalty Multiplier)
ADJUR003	6	1	16:23:12	16:49:51	0:26:39	1	1
AKATK003	6	1	16:23:08	16:49:38	0:26:29	1	1
ANGLY295	13	1	16:23:05	16:50:00	0:26:55	1	1
APRBU001	5	1	16:28:52	16:41:22	0:12:30	1	1
ASALA135	6	1	16:22:53	16:49:59	0:27:06	1	1
ASHLEY PANG	2	0.4	16:23:52	16:51:24	0:27:32	1	0.4
ATONG027	4	0.8	16:23:11	16:49:41	0:26:30	1	0.8
BUJUNG020	1	0.2	16:49:49	16:49:49	0:01:00	0.1	0.0
BVAR003	1	0.2	16:49:46	16:49:46	0:01:00	0.1	0.0

Figure 2.2.3.1 An output table from the chat log in Fig. 2.2.1 depicting the participation point calculation process for each unique user in the uploaded chat logs.

The participation point table contains all of the data used in our automated algorithm, and is adjustable by the instructor using the CAT. Each user's number of messages and message timestamps are compared to the general chat log statistics for each chat log.

Each user is assigned a participation point score for each uploaded chat log, and the CAT also outputs a summarized table of each user’s score across all uploaded chat logs. The participation point system is explained fully in Chapter 3.

2.2.4 Poll Statistics Tables

Number of Messages Per Poll Download				Poll Message Statistics Download			Poll Messages per User Download	
† Poll	Messages	First Timestamp	Last Timestamp	† Minute	Poll	Non-Poll	† Name	Poll Messages
1	29	16:23:06	16:23:25	0	0	2	MSUN042	8
2	33	16:24:41	16:25:23	1	29	5	DTINO006	6
3	30	16:28:16	16:28:27	2	2	7	ASALA135	3
4	40	16:29:20	16:29:39	3	31	1	HSUN096	8
5	24	16:30:24	16:31:16	4	0	1	MMOHA020	5
6	43	16:41:11	16:41:29				ZKXAO037	3

Figure 2.2.4.1 An output table from the chat log in Fig. 2.2.1 depicting, for each poll detected in the chat log, the number of messages and total runtime.

We define a poll as a period in the online class where the students are responding to an instructor’s question to class, or are otherwise invited by the instructor to use the chat log en masse. The CAT looks at the number of messages per minute and identifies sections where the frequency peaks, marking them as possible polls. The results allow instructors to see how many polls they had conducted, and each individual student’s number of poll messages. Tracking poll messages is a more specific method of measuring student engagement and attentiveness.

2.3 Conclusion

Using the CAT is a highly lightweight and automated alternative to extract data on student engagement with online classes compared to manually examining chat logs, or relying on sign-in forms. The algorithms that automatically calculate student participation points can be adjusted to each instructor's preference, and the downloadable results can be examined further to return insightful information about the level of activity in the class and the future performance of the students.

Chapter 3: Participation Points Calculation

3.1 Introduction

Participation points for each user is calculated based on multiple metrics and two adjustable variables. As shown in Figure 3.1 (a), each user’s number of messages in a specific chat log must surpass the minimum message threshold set by the instructor in order to qualify for full credit. We then apply a penalty to this “unadjusted score” received for the student, by calculating the difference between their first message timestamp and last message timestamp, and comparing that “time spread” duration to the total class runtime in minutes.

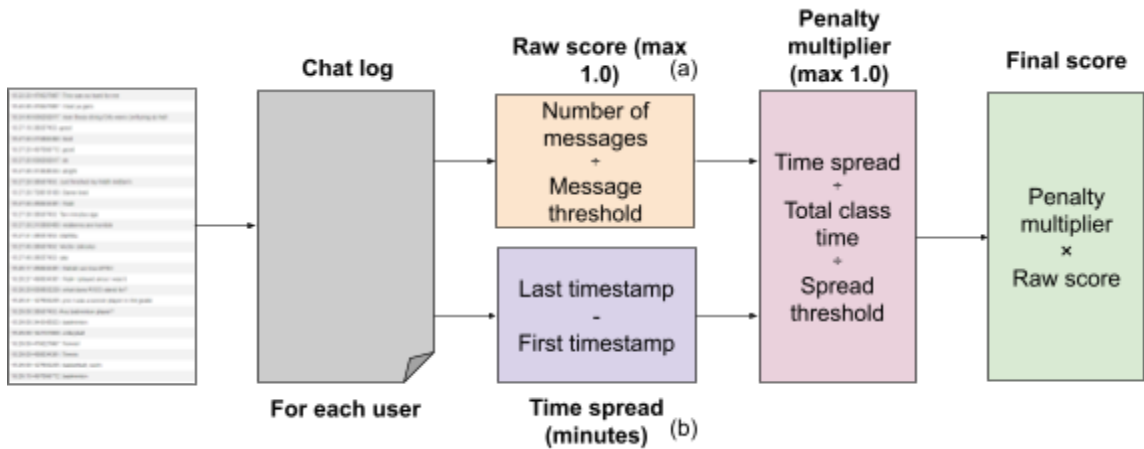


Figure 3.1 A flowchart depicting how the participation point algorithm generates scores for each user present in a chat log.

Figure 3.1 (b) above shows the time spread calculation process. If the student's time spread does not surpass the adjustable spread threshold set by the instructor, then the CAT applies a negative multiplier to their participation score.

3.2 Methodology

3.2.1 Unadjusted Participation Score

Unadjusted score from number of messages

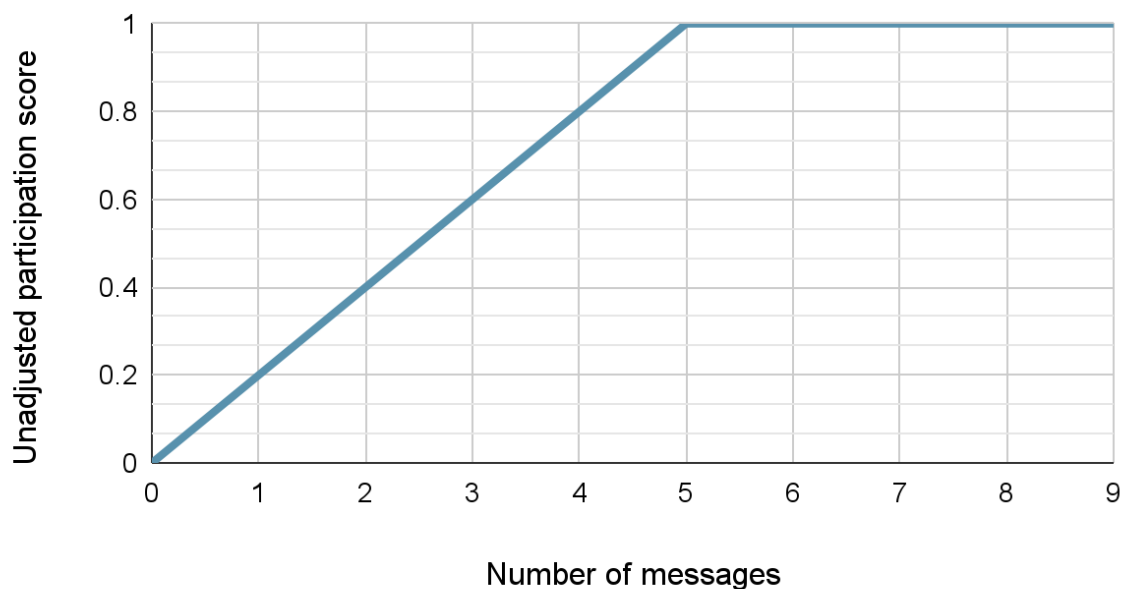


Figure 3.2.1 A line chart showing the scale between number of messages a student posted and the resulting unadjusted participation score, before penalties, where the set threshold is 5 messages.

The unadjusted score for participation is based on the number of messages per student and is calculated as a fractional percentage. The CAT algorithm does not reward students for messages submitted after reaching the instructor's set message threshold. The default value used in our CS1 class was 5 messages. Students can earn partial credit on their

unadjusted score, as seen on the vertical axis in Figure 3.2.1 above. This partial credit is awarded on a linear scale, set according to the adjustable message threshold.

3.2.2 Factoring Message Spread into Participation

In the initial stages of developing the CAT, we relied on only the unadjusted score to assign participation points. We noticed that students had the tendency to submit a large number of messages in a short span of time, usually clustered near the start of class. The time spread variable, threshold, and associated penalty multiplier was introduced to the algorithm to account for this behavior.

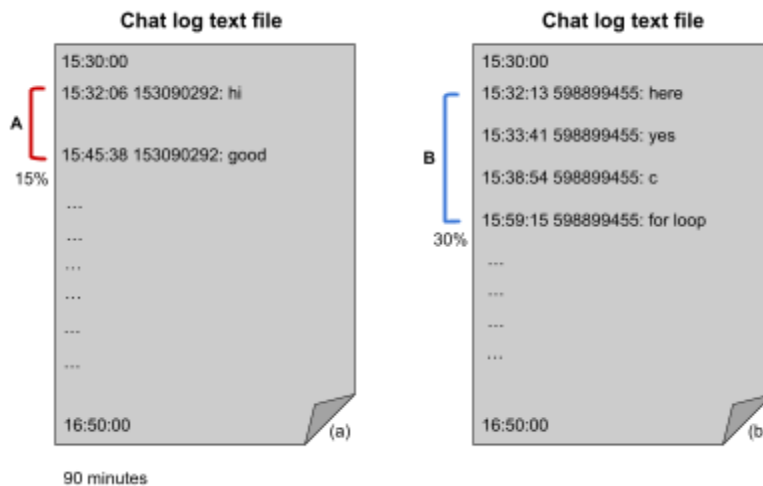


Figure 3.2.2 An example of the time spread of student A (red) and student B (blue).

Student B reaches the minimum value of 30% of the total class length, student A's time spread only reaches 14.8%. The spread penalty multiplier is calculated as a fractional percentage. We take the time in minutes between a student's first and last timestamp, and

divide the value by the total class runtime in minutes. We then take the resulting dividend and divide it by the adjustable spread threshold. The default used in our CS1 class was 30%, expressed in the formula in Figure 3.2.2.1 below as the decimal value 0.30.

3.2.2.1 Participation Breakdown for Student A

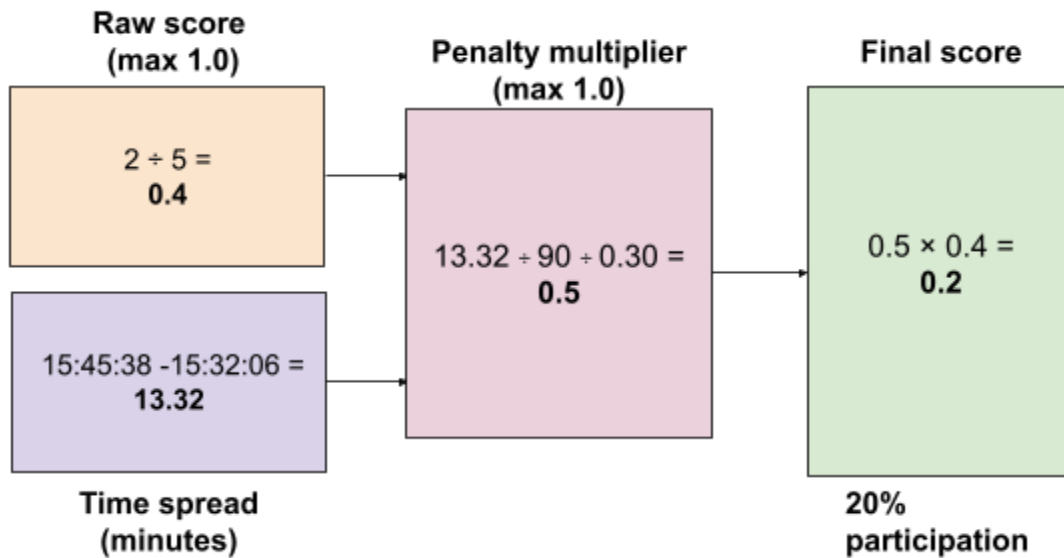


Figure 3.2.2.1 An example of a weak spread and low message count, student A's first and last message cover 15% of the total class time

In Figure 3.2.2.1, student A (red) only submitted 2 total messages, giving a unadjusted score of (0.4) and achieved a spread of 15%, in which case our algorithm described in the earlier chapter will give a penalty multiplier of 0.5, reducing student A's unadjusted score by half. Their final score would be 0.2, 20% participation credit for the lecture.

3.2.2.2 Participation Breakdown for Student B

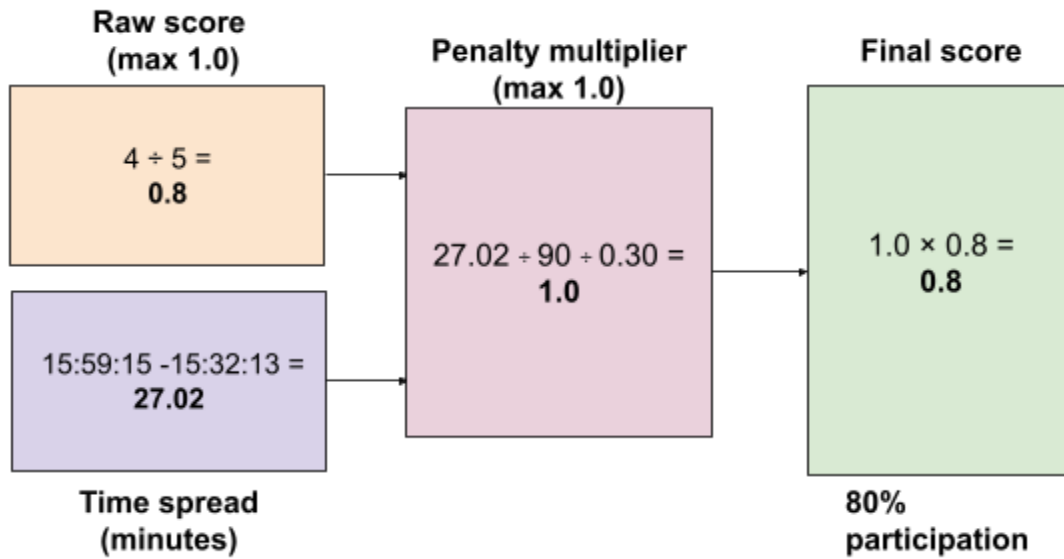


Figure 3.2.2.2 An example of an average spread and message count, student B's first and last message cover 30% of the total class time

In Figure 3.2.2.2, student B (blue) submitted 4 total messages, giving an unadjusted score of (0.8) and achieved the minimum spread of 30%, meaning the history of all messages they had sent in the chat log covers at least 30% of the total chat log length. They receive a penalty multiplier of 1 (no penalty), which is the maximum multiplier. The final score would be 0.8, 80% participation credit for the lecture.

3.2.2.3 Full Breakdown for Student C

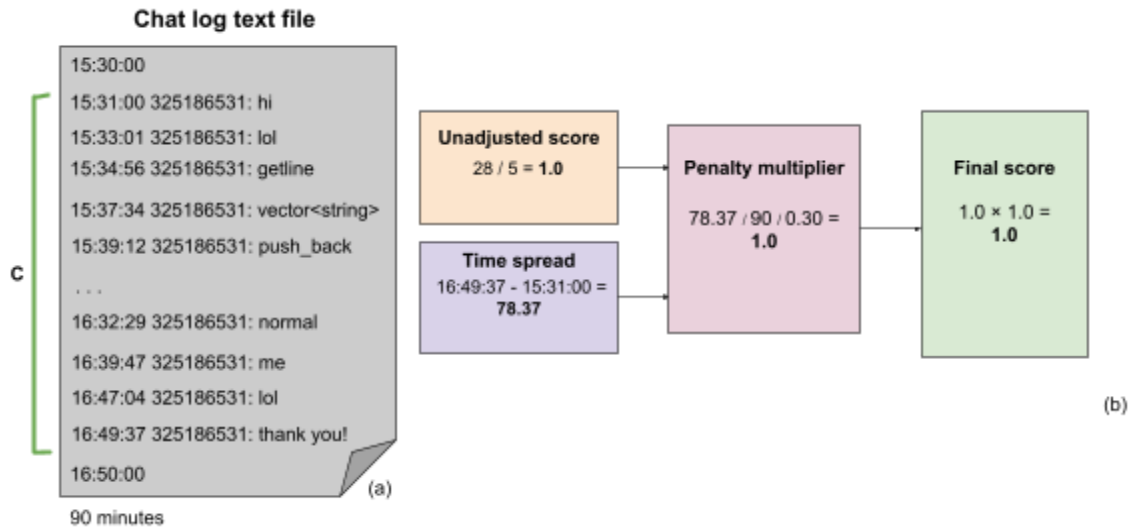


Figure 3.2.2.3 An example of a strong time spread and message count, student C's first and last message cover 78% of the total class time.

In Figure 3.2.2.2, student C submitted 28 total messages in the chat, giving them full unadjusted points (1.0) and also achieved a spread of 78%, in which case our algorithm described in Figure 3.1 would give a multiplier of 1 (no penalty). They would receive 100% participation credit for the lecture.

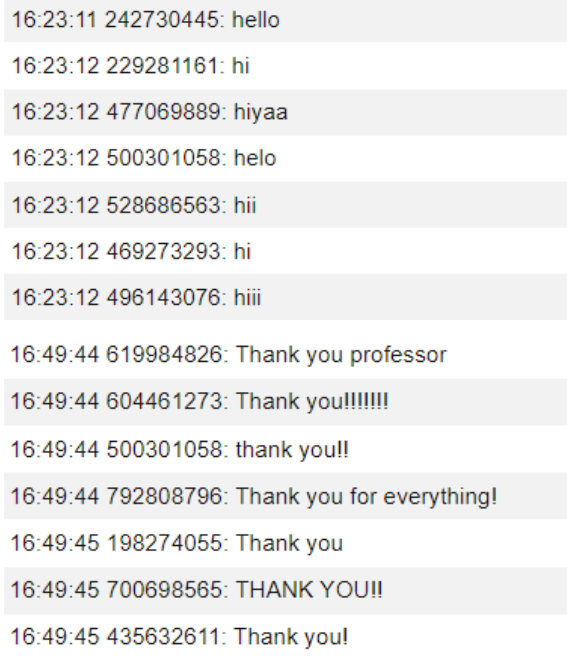
3.2.3 Reasoning for Default Participation Thresholds

	Average Number of Messages Per User
3/31/22 (Week 1)	7.916
4/21/22 (Week 5)	20.717
6/2/22 (Week 10)	5.128

Table 3.2.3 A table showing the change of average number of messages per user

The number of messages that a student must surpass to qualify for full participation credit can be decided by the instructor and adjusted for each class in the CAT, but we found that the minimum average of messages per student over the course of a standard 10 week quarter never fell below 5 messages.

3.2.4 Non-class Related Messages



The image shows a list of chat messages from a CS1 class chat log. Each message is on a separate line, with the timestamp and sender ID on the left, and the message content on the right. The messages are as follows:

- 16:23:11 242730445: hello
- 16:23:12 229281161: hi
- 16:23:12 477069889: hiyaa
- 16:23:12 500301058: helo
- 16:23:12 528686563: hii
- 16:23:12 469273293: hi
- 16:23:12 496143076: hiii
- 16:49:44 619984826: Thank you professor
- 16:49:44 604461273: Thank you!!!!!!
- 16:49:44 500301058: thank you!!
- 16:49:44 792808796: Thank you for everything!
- 16:49:45 198274055: Thank you
- 16:49:45 700698565: THANK YOU!!
- 16:49:45 435632611: Thank you!

Figure 3.2.4.1 Excerpts from a CS1 class chat log showing the non-class related messages

The CAT currently does not look at the contents of the student messages in the chat log, only username and timestamp. Figure 3.2.4.1 shows common non-class related messages present in many chat logs, where students greet the instructor at the start of class and thank them at the end. As we do not differentiate between messages based on content, these messages are included in our raw message count and time spread calculations.

3.3 Discussion

The lecture participation points keeps me more engaged

125 responses

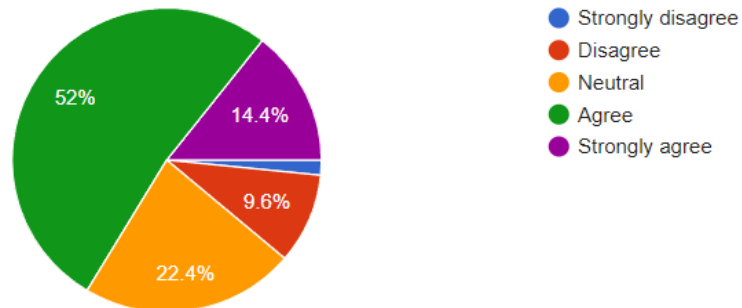


Figure 3.3.1 A pie chart depicting student opinion on how much the participation points system kept them engaged with the class

The lecture participation points stress me out

125 responses

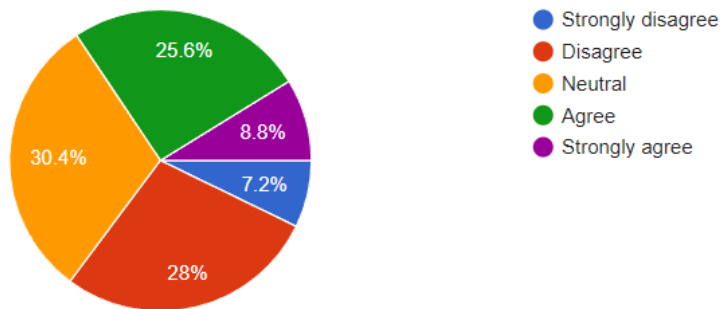


Figure 3.3.2 A pie chart depicting student opinion on how much the participation points system caused extra stress

The automated algorithm that calculates participation points is a large part of why the CAT is a very lightweight option for instructors to assign credit for class attendance and engagement, as well as track the performance of both the overall class and individual students. Surveys conducted on the students in our CS1 class revealed that the system can

also cause undue strain in a subset of the student population. In Figure 3.3.2 above, 34% of the students in the class answered with “Agree” or “Strongly Agree” when asked whether the automated participation points caused stress.

3.4 Conclusion

The usage of participation points as a way to incentive student attendance and track engagement is already in common use in both in person classes and in pre-2019 online and hybrid courses. Sign-in sheets, clicker questions, and discussion boards are all used as ways to foster attentiveness in university students. Using the CAT allows us to have greater insights into the engagement level of students in an online class, using automated methods. The major parameters taken into account for deciding which students are active participants are their total number of messages and the amount of time that those messages cover across the duration of each class. The CAT does not look at message contents, only the message’s timestamp and the frequency of each user’s submissions to the chat log.

Chapter 4: Identification of Polls and Poll Messages

4.1 Introduction

We defined polls as sections of the chat log where students are responding to a question from the instructor asked to the class, or are otherwise invited to be active in the chat for a period of time. Figure 4.1 below is an excerpt of a manually identified poll in a CS1 chat log, a response to an instructor asking the question “What type of variable should we make on line 10?”. The CAT uses a poll detection algorithm that automatically looks through uploaded chat logs and flags messages that are considered possible polls.

```
15:37:18 831774504: vector <string> chatheaders
15:37:18 410255600: Vector<string>
15:37:18 163321719: vector<string>
15:37:19 105286020: vector<string> chatHeaderLines,
15:37:19 704551312: vector<string>
15:37:19 700639285: vector<string>
15:37:19 404189179: vector<string>
15:37:20 489453371: vector<string>
15:37:20 140600331: vector<string>
15:37:21 781802844: vector <string>
15:37:21 312302536: vector<string>
15:37:22 815609803: vector<string>
15:37:22 604534608: vector<string>
15:37:23 615772639: vector<string>
15:37:23 719875311: vector <string>
15:37:24 549618247: vector <string>
15:37:26 61353637: Vector<string>
15:37:26 5795039: Vector<string>
```

Figure 4.1 Excerpt of chat log showing an example of a poll

4.2 Methodology

4.2.1 How Do We Detect Polls in the Chat Log?

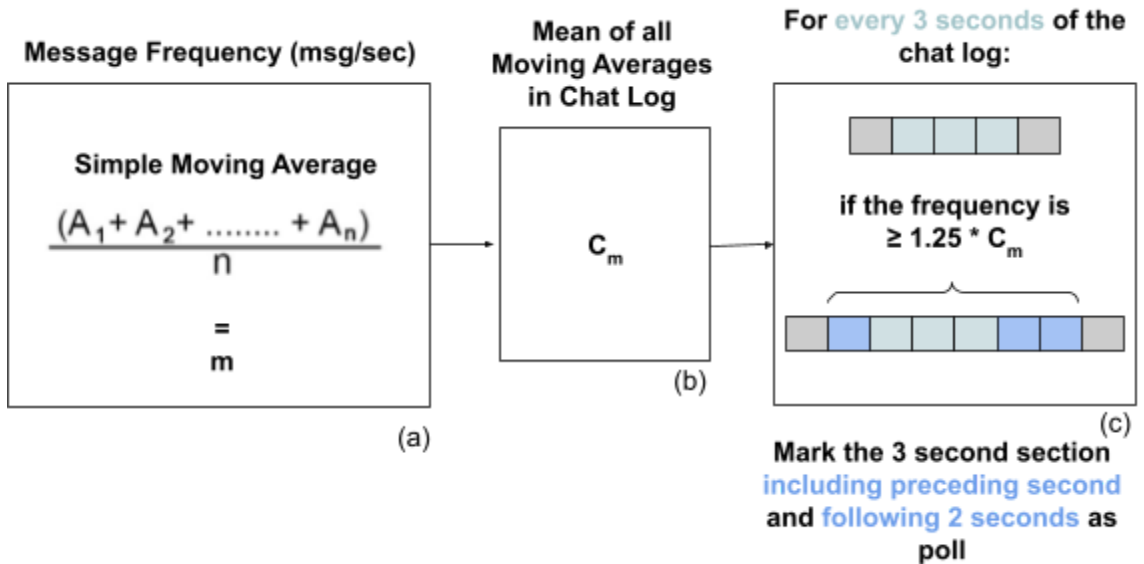


Figure 4.2.1 A flowchart depicting the process used to calculate which sections of the chat log are polls

We identify polls in the chat log by isolating local maximums in the messages per second (MPS) frequency throughout the chat log. We start by calculating the simple moving average of the MPS, shown in Figure 4.1 (a), where $n = 3$. Once the entirety of the chat log is broken down into windows of moving averages, we calculate the mean of all the averages, C_m . We then look over each 3-second block of the chat log and compare the MPS value to check if it surpasses the poll threshold, 125% of C_m . The messages contained in that block are then flagged as poll messages. We also noted a tendency of students to “straggle” when answering polls, a behavior noticed during manual analysis of chat logs. The majority of students will respond to the poll after the local maximum in MPS frequency, and the first few students who respond will submit before the local maximum. To account for this, each flagged block of messages also includes the

preceding second and the following two seconds as part of the poll. Multiple polls can be adjacent to each other, and when the CAT flags blocks that overlap, they are counted as one poll.

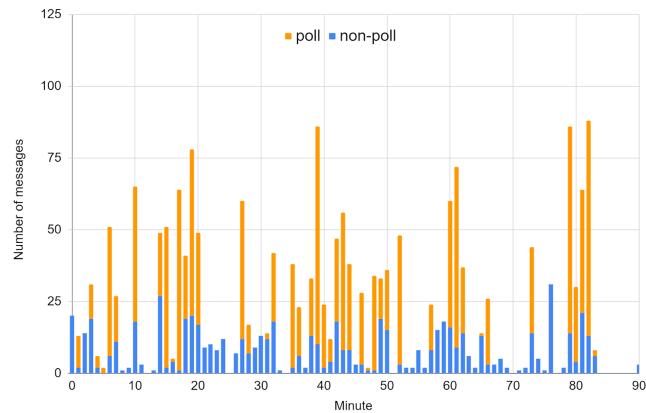


Figure 4.2.2 Chart showing number of messages per minute, poll messages marked in orange.

4.2.2 Poll Message Detection Accuracy

n = 1981 messages	Poll message	Non-poll message
Positive	1212	99
Negative	247	423

Table 4.2.2 An accuracy table for the poll detection algorithm

Table 4.2.2 above is an accuracy table generated after comparing the results of the CAT automatically flagged poll messages to the manually flagged poll messages in the same chat log. We notice that our algorithm has an average specificity of 0.81 and average sensitivity of 0.83, and that the CAT generally ends up with more false negatives than false positives, due to students who reply to polls too late or too early to be caught in the windows described in the previous section.

4.3 Discussion

The poll detection algorithm of the CAT does not look at message content, only examining message frequencies in order to decide where polls are located. The way we flag polls involves setting a starting and ending timestamp that establishes the boundaries of the poll, and all messages contained within are considered poll messages, which contributes to the false positive count. We also noticed the tendency of multiple polls in quick succession being detected as one long poll by the CAT. Polls that are very short, consisting of only 5-10 messages or fewer, are often missed entirely by the detection algorithm.

4.4 Conclusion

We use poll messages in order to identify the peaks of instructor-student interaction during online classes, where students are directly interfacing and responding with the instructor. Knowing when and how often students were reacting to the material taught in online classes allows us to examine a specific facet of class activeness, and may be used to further refine the participation point algorithm in the preceding chapter. Poll messages can also be used as a predictor of future student performance in the class, even when looking at only the first few weeks of collected data.

Chapter 5: Predictors of Course Outcomes

5.1 Introduction

The data gathered from the CAT can be used in conjunction with the gradebook data from our online CS1 sections to create predictors of student performance over the course of the term, across a variety of course outcomes. We define two main categories of course outcomes: course totals, which are the weighted letter grades students achieve, and exam performance, which are midterm and final scores. We specifically look at examining only the first 4 weeks of data from a standard term, which gives us early insight on a student's trajectory and likelihood of achieving certain levels of course outcome.

Course Total vs. Raw Messages

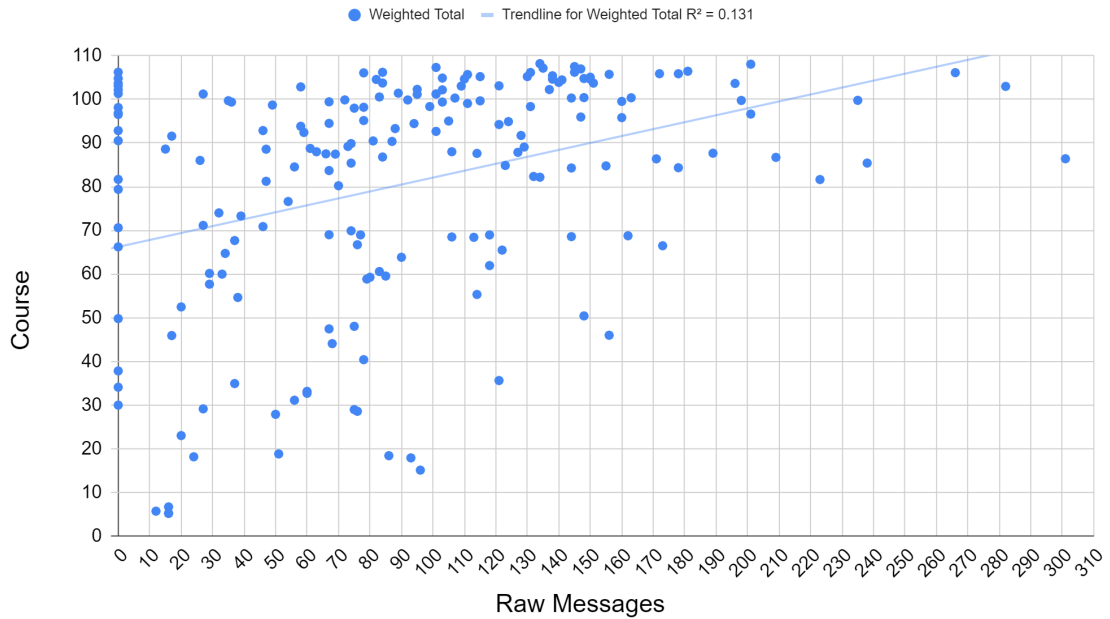


Figure 5.1 Graph showing the relationship between number of messages and course grade in a CS1 class.

Figure 5.1 compares student's raw number of messages to their course total. We note that the majority of students who earned above 80% in the class submitted more than 60 messages during the term. We put forth the hypothesis that other metrics gained from chat logs would have correlation with course outcomes, and examined them further.

5.2 Methodology

5.2.1 Participation Points as a Predictor

Participation points, despite being only 5% of the course grade in our CS1 class, had a very positive correlation with course outcomes. The results discussed below were gained over the course of the entire term, and use the default 5 message threshold and 30% spread penalty threshold discussed in Chapter 3.

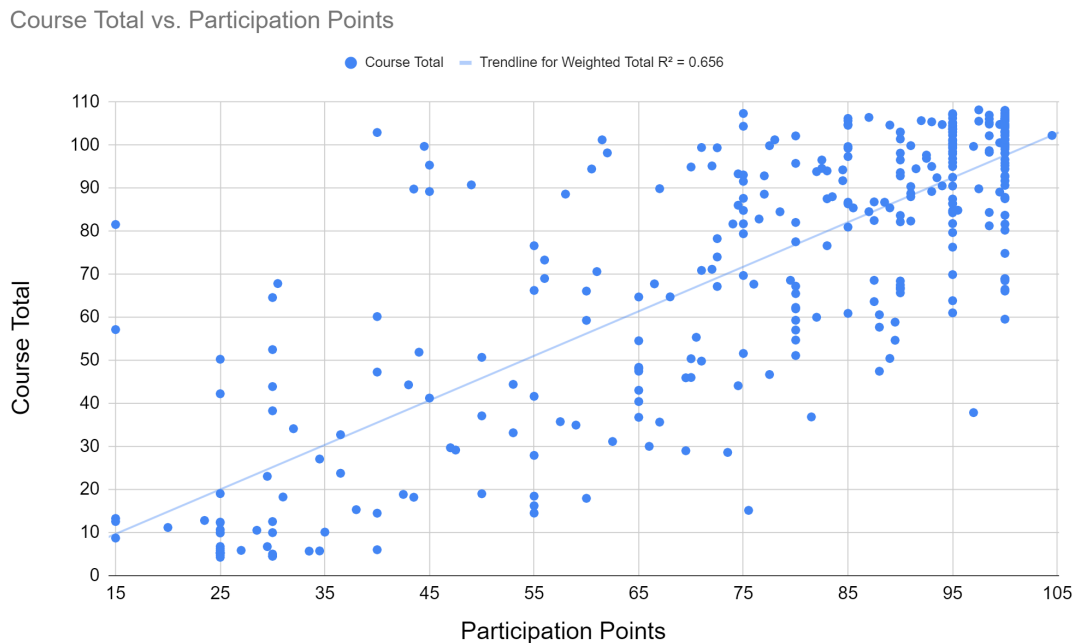


Figure 5.2.1 Graph showing the relationship between student lecture participation and course grade in a CS1 class.

The CAT's automatically assigned participation points are significantly correlated to course totals in a CS1 class. The weighted R-squared value is 0.65 when graphing participation from the entire term against course total. Students with higher participation will do well in the class, and on average students that earned above 70% of the available

participation points finish the class with a 75% course total or higher. The data is somewhat unreliable, as the set of students examined includes students who dropped the class, which skews our correlation in the positive direction.

5.2.2 Poll Messages as a Predictor

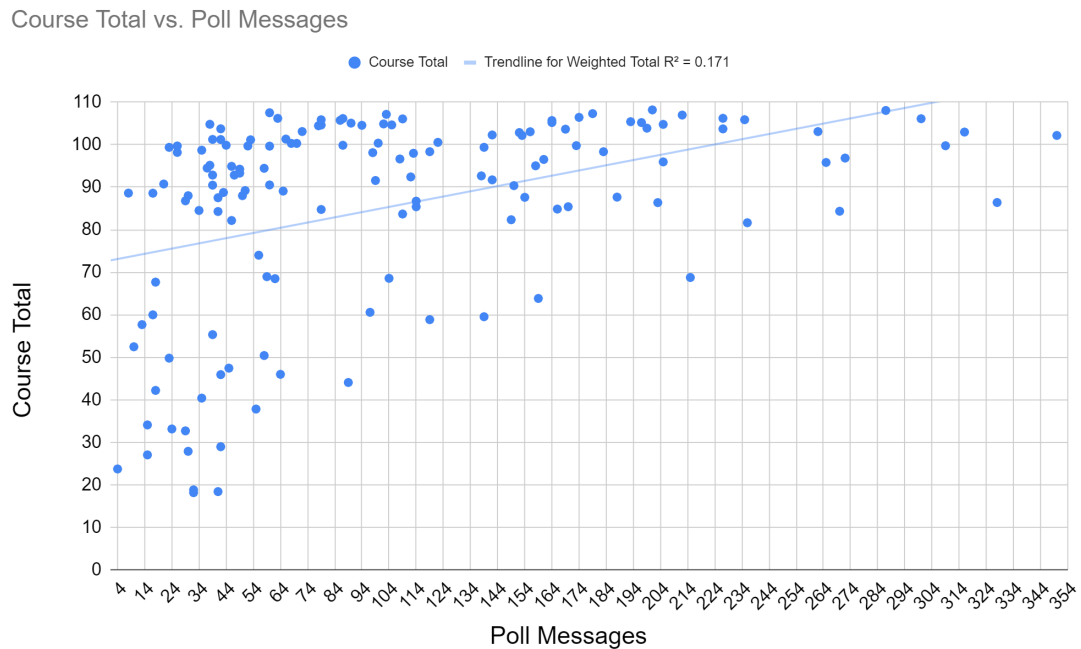


Figure 5.2.2 Graph showing the relationship between student poll messages and course grade in a CS1 class.

The automated poll detection algorithm returns data that can be used as a predictor for course totals. The correlation of student's total poll messages to course total grade is not as strong, but we do see interesting patterns emerge in the graph shown in Figure 5.2.2. The majority of students who did less than 70% in the class had 60 or less poll messages.

5.2.3 Early predictors

After establishing that there existed positive correlations between chat log metrics and course outcomes, we sought to extrapolate useful data about student performance earlier in the term. The specific goal was to be able to find strong dynamic predictors of student performance in the midterm by the fourth week of the term.

5.2.3.1 Early Participation Points

The first four weeks of participation data allows us to see the likelihood of a student to earn above 70% in the midterm.

Midterm vs. Weeks 1-4 Participation

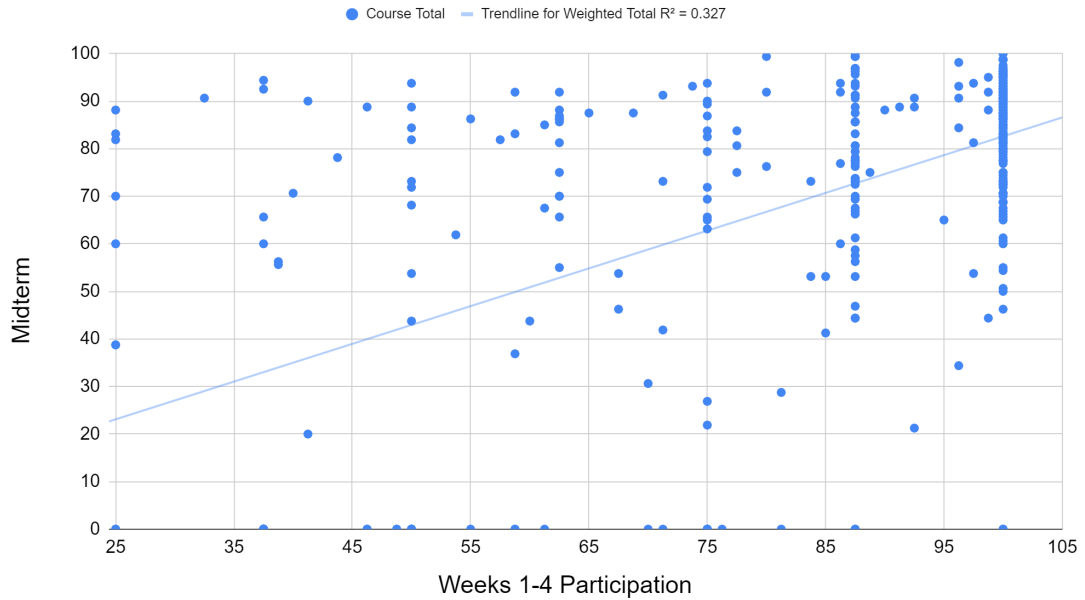


Figure 5.2.3.1 Graph showing the relationship between the first four weeks of lecture participation and midterm score in a CS1 class

Students who earned 85% or higher on their early class lecture participation are much more likely to do well in the midterm. We note that higher class participation in the first few weeks of class is impactful on a student's midterm grade.

5.2.3.2 Early Poll Messages

Early poll message data is used to further refine our predictors of student midterm performance. Figure 5.2.3.2 shows student performance in our CS1 class clusters around 85% exam score and 45 recorded poll messages in the first four weeks of the term.

Midterm vs. Weeks 1-4 Poll Messages

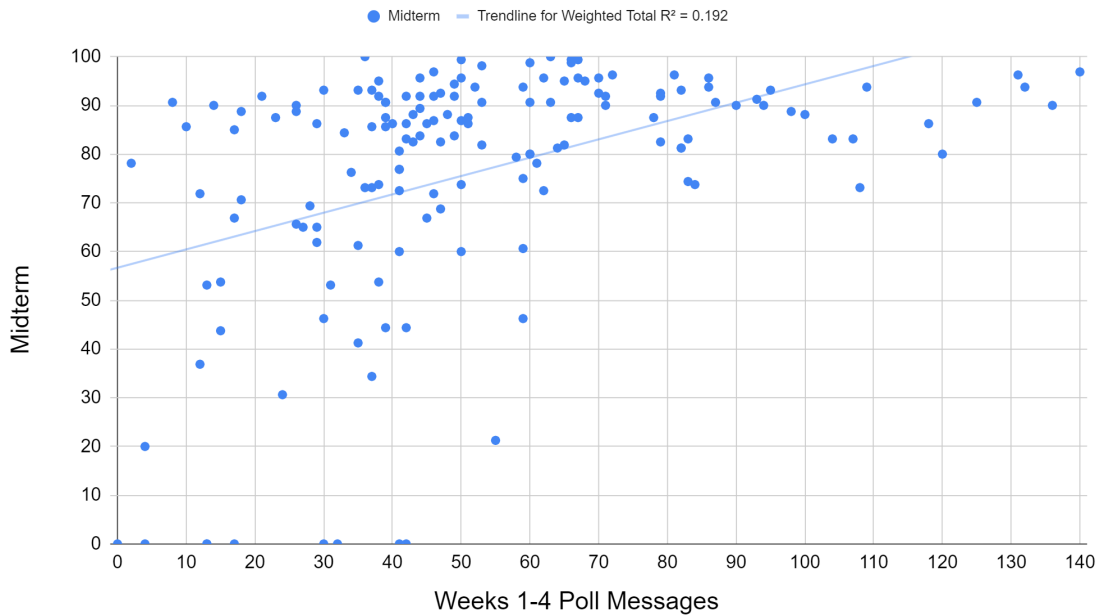


Figure 5.2.3.2 Graph showing the relationship between the first four weeks of recorded poll messages and midterm score in a CS1 class

Poll messages are a better visualizer of student performance than only raw message counts or participation points. We can also state that students who are active enough to participate in the majority of polls in the early weeks will have a high likelihood to earn above 80% in the midterm.

5.2.3.3 Earnestness

We found that using only chat log metrics were not as accurate as including data gained from zyBooks, the learning platform used by our CS1 class.

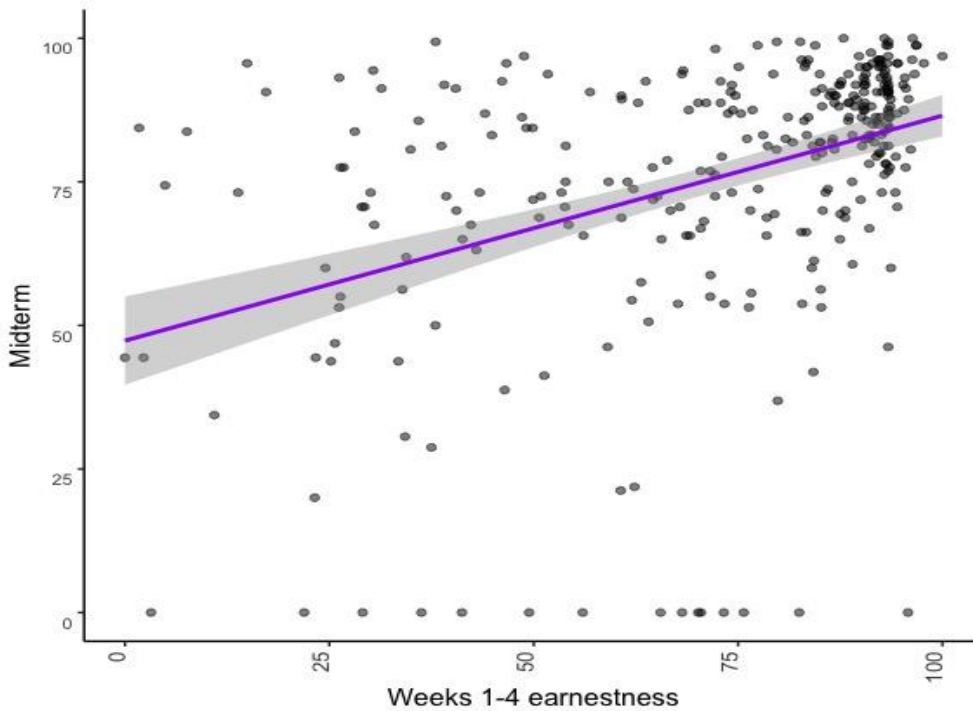


Figure 5.2.3.3 Scatterplot showing relation between first four weeks of earnestness data and midterm score in a CS1 class [7]

Many zyBooks activities require typing a short answer in a text box, which will include a clickable “show answer” button, which reveals the correct answer. We define an “earnestness” metric as the proportion of short answer questions that students earnestly attempted to complete. An earnest attempt is when a student attempts to provide a response to the question, whether incorrect or correct, prior to clicking "show answer".

Figure 5.2.3.3 shows that there is a positive correlation of early earnestness to midterm performance, with an R-squared value of 0.44.

5.2.3.4 Challenge Activity Struggle

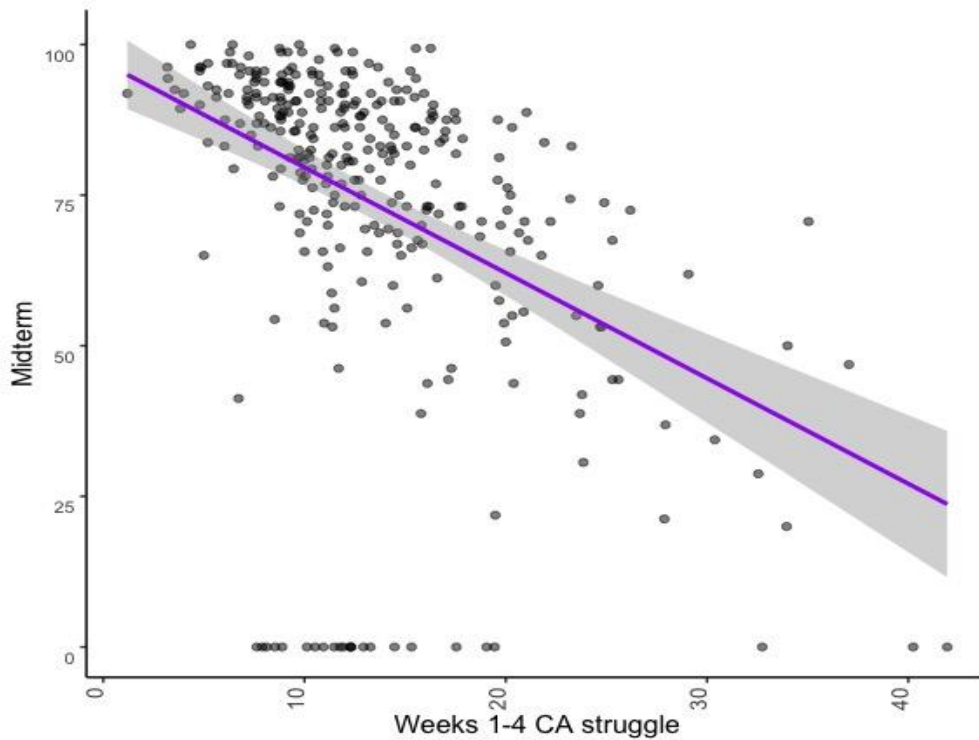


Figure 5.2.3.4 Scatterplot showing relation between first four weeks of earnestness data and midterm score in a CS1 class [7]

The second zyBook metric we use is challenge activity (CA) struggle. A zyBooks CA represents a small homework problem, each requiring perhaps between 1 to 4 minutes, involving typing the output of a small program, or typing code to complete a small program to achieve a task, such as "Complete the program to output the maximum of the three input integers". We define CA struggle using two parameters: number of attempts and time spent trying to solve the CA. Students may submit multiple submissions before

achieving the correct solution to a CA. We define a struggling student for a particular CA as a student who has spent more than twice the average time and attempted more than double the average attempts.

5.2.4 Impact of Previous Experience

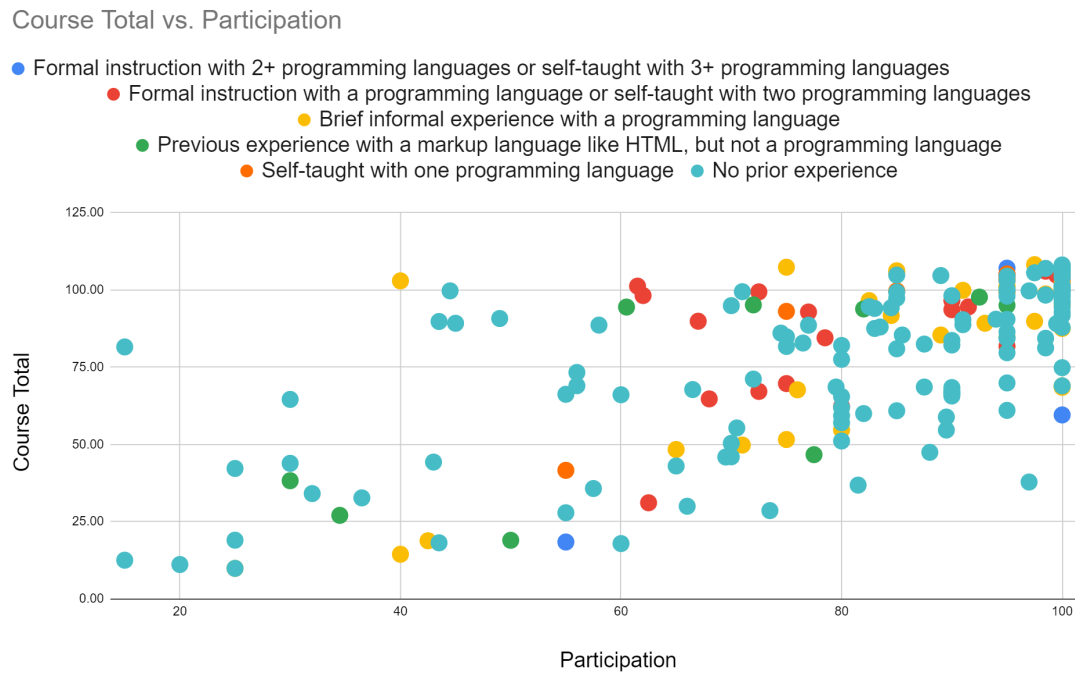


Figure 5.2.4 Scatterplot showing previous experience with programming with regards to participation predicting course total.

The majority of students with no prior experience ended the course with high overall participation, and a strong correlation between high participation and high course totals. Students with formal previous experience or instruction with multiple programming experience tend to outperform the other categories of students, and in general do not slack on participation in class.

5.3 Discussion

The data used for the predictor graphs showcased in the figures above are skewed by the inclusion of students who dropped the class over the course of the term.

Many of our predictors are not as accurate as some pre-existing methods for dynamic prediction of student performance. Both the PreSS algorithm and other ultra-lightweight methods outstrip the methods discussed in this chapter. When gathering data for previous experience, many students did not respond. Student awareness of the participation algorithm in our CS1 class may also compromise the validity of the results.

5.4 Conclusion

Chat log metrics can be used to predict the success rates of students early in the term.

Using participation, poll message, and general gradebook data collected from both the early weeks and over the course of the quarter allows us to make predictions on student behavior and outcomes.

We were not able to achieve the level of efficacy with our predictors we had aimed for when starting the project. The algorithms used for calculating participation points can be further improved, by examining the contents of the messages or by smarter adjustment of the threshold values used. The poll detection methods used by the CAT can also be refined by including message content and better identification of when students are reacting to the same prompt from the instructor. Improvements to the way we collect data to predict student behavior may yield more accurate results.

Chapter 6: Summary

The fundamental importance of the chat log is its multipurpose usefulness as a way to more seamlessly allow student-instructor interaction, while also being a source of data on student engagement [8]. In addition, we aim to facilitate further interaction between isolated online students and both their class, professor, and fellow peers. The chat analysis tool described in this thesis (CAT), shown in Figure 1.2.2, to help instructors and TAs automatically grade their online classes for participation and to give insight on the activity and engagement levels of the students in those classes. Currently, the tool is only available for instructor use by invitation, and has been used by about 5 instructors at 3 different institutions. We plan to eventually release a stable public version for any class to use.

In Chapter 4.2.2 we discuss how parts of the CAT are not completely accurate, specifically poll detection which averaged about 82% in overall accuracy. Chapters 5.3 and 5.4 further mention that the graphed data is affected by the inclusion of dropped students. Our correlation values are positively inflated due to this.

We have tried to improve our online CS1 course, and online courses in general, through the development and use of the CAT. Many instructors already track class participation as part of course credit, and the CAT automates much of this process. The CAT allows users to customize the algorithms used to calculate student participation and poll detection.

This paper also discusses the possibility of using the data gleaned from the chat log to make predictions about student performance and course outcomes, such as midterm scores. We found that there was a positive correlation between student participation in the chat log to success in exams and in the overall class grade, and that the correlation was present even when we only examined the first few weeks of the term. However, the results were not strong enough to form a definite statement on how much participation impacts increased student success. Finally, we explored various other metrics gathered from student activity data recorded in the learning platform used by our CS1 class, zyBooks, as well as examining the impact of previous experience with coding. There is still space for further improvement to be made on both the CAT, such as adding the ability to process and categorize the contents of each student's message. Finding new approaches to including the chat log as part of courses will allow instructors to create a much more engaging environment for online classes.

REFERENCES

- [1] Adnan, M. and Anwar, K., 2020. Online Learning amid the COVID-19 Pandemic: Students' Perspectives. Online Submission, 2(1), pp.45-51.
- [2] Baum, S. and McPherson, M., 2019. The human factor: The promise & limits of online education. *Daedalus*, 148(4), pp.235-254.
- [3] Duffin, E., 2022. "U.S. Student Distance Learning Enrollment 2020." Statista, 18 May 2022, <https://www.statista.com/statistics/944245/student-distance-learning-enrollment-usa/>.
- [4] Jona, K., 2000, December. Rethinking the design of online courses. In *Ascilite* (pp. 14-19).
- [5] Lederman, D., 2021. New Data Offer Sense of How COVID Expanded Online Learning, <https://www.insidehighered.com/news/2021/09/16/new-data-offer-sense-how-covid-expanded-online-learning>.
- [6] McPherson, M.S. and Bacow, L.S., 2015. Online higher education: Beyond the hype cycle. *Journal of Economic Perspectives*, 29(4), pp.135-54.
- [7] Vahid, F., Zhao, S. and Gordon, C., 2023, March. Ultra-Lightweight Early Prediction of At-Risk Students in CS1. In *2023 ACM SIGCSE Technical Symposium*.
- [8] Vahid, F., Zhao, S. and Allen, J., 2022, August. Automated Zoom Chat Analysis Including Chat-Based Polls for an Online Introductory Programming Course. In *2022 ASEE Annual Conference & Exposition*.