DEC2014-34797

KNOWLEDGE DISCOVERY OF STUDENT SENTIMENTS IN MOOCS AND THEIR IMPACT ON COURSE PERFORMANCE

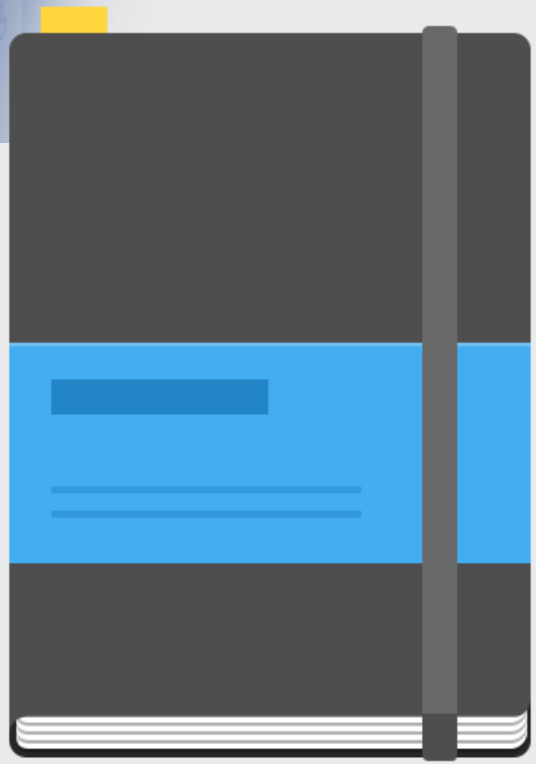
Conrad S. Tucker, Engineering Design and Industrial Engineering

Bryan Dickens, Computer Science and Engineering

Anna Divinsky, School of Visual Arts, World Campus

{ctucker4, bid5098, axd289}@psu.edu

PRESENTATION OVERVIEW



- Background
- Motivation
- Methodology
  - Mining MOOC data
  - Performance Correlation
- Case Study
- Results
- Conclusions
- Future Work

In-class student engagement is readily observable…

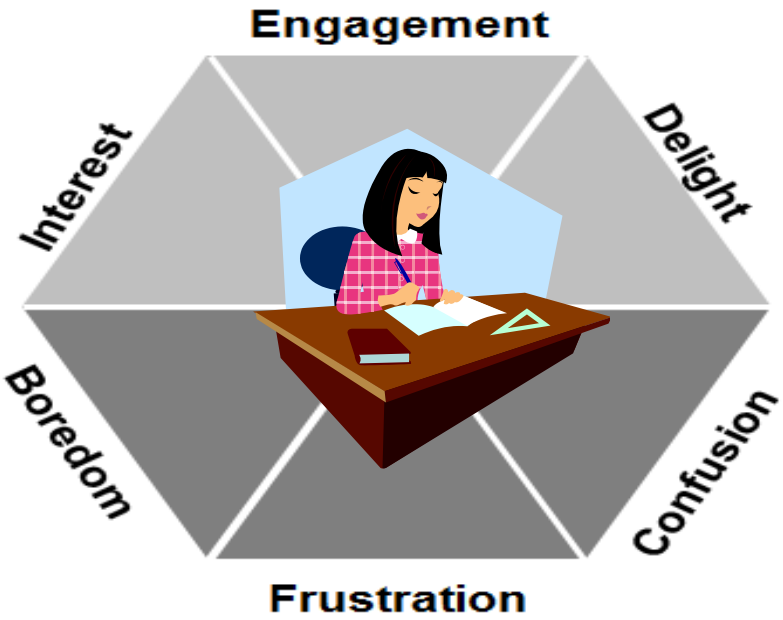
EMOTIONS IN THE CLASSROOM

Emotional State	Learning Gains Impact	References
Engagement/ Interest	Positive	[21,7]
Frustration	Negative	[25,9]
Boredom	Negative	[26]
Confusion	Positive	[9,25,27]
Delight	Positive	[9]

Students **experience positive mental states while minimizing those mental states associated with negative connotations**.

EMOTIONAL STATES AND LECTURES

In-Class Learning



Online Learning

Hypothesis: Students' emotional states (quantified through textual data) are correlated with their performance on assignments

-towards individually customized learning

The Challenge in Massively Open Online Courses (MOOCS)

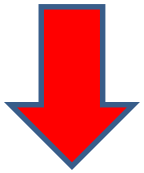
Educational Data Mining

Educational Systems
("Brick and mortar", MOOCs, laboratories, etc.)

Mostow et al. (2005), Vialardi et al. (2009), García-Saiz et al. (2014)

data

Data Mining Algorithms
(classification, clustering, etc.)



knowledge

Educators

Students

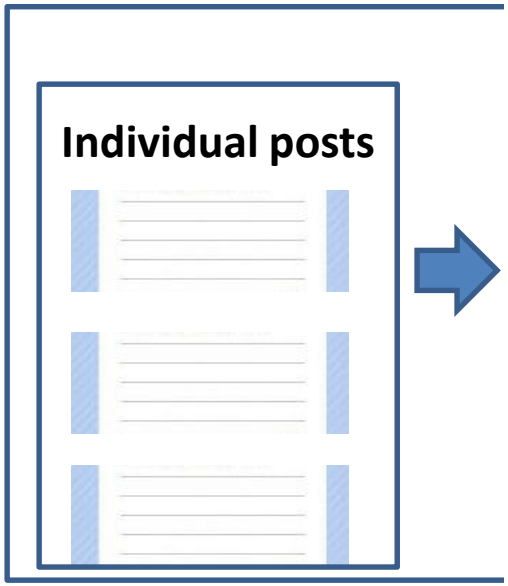
Research Objectives

- Quantify students' sentiments in MOOCs by mining student-generated textual data

- Determine whether students' sentiments correlate with course performance

Research Methodology

Forums



Sentiment Scoring Algorithm

$$Sentiment(post_i) = \sum_{n=1}^{k} map\_lookup(post[n])$$

```
procedure map_lookup( ArrayList<String> post )
1    Sentence Sentiment Score = 0;
2    Clean Sentences (post);
3    while post not empty
4      String temporary word = post.remove(post.size() – 1);
5       if (temporary word abbreviated)
6           Substitute full word for temporary word;
7       if( word is emoticon)
6         Sentence Sentiment Score += emoticon sentiment;
7       if( word is in general word list)
6           Sentence Sentiment Score += word sentiment;
7    end
7    return Sentence Sentiment Score;
end map_lookup.
```

Sentiment Scoring Example

Post Text	Sentiment of Post
Hello everyone!! im from Cyprus! i was born with this talent and i cant stop drawing on walls,papers or whatever i can from an early age! :) im so enthusiastic about art and expressing our selfs! Good luck everyone and i hope this new beggining of ours comes to a success!! Have a great week!!!! :))	
i will confess that nothing I create for this class I will consider art. I will consider it studies.	

CORRELATION WITH PERFORMANCE

r	Type of relationship	Color Code
± [0.0 to 0.2]	Weak or no relationship	
± [0.2 to 0.4]	Weak relationship	
± [0.4 to 0.6]	Moderate relationship	
± [0.6 to 0.8]	Strong relationship	
± [0.8 to 1.0]	Very strong relationship	

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

n: *Sample size*

X_i: *Value of* $i-$ *th observation from sample* $X, i: 1$ *to* n

$\bar{X}$: *Average value of all observations from sample* X

Y_i: *Value of* $i-$*th observation from sample* $Y, i: 1$ *to* n

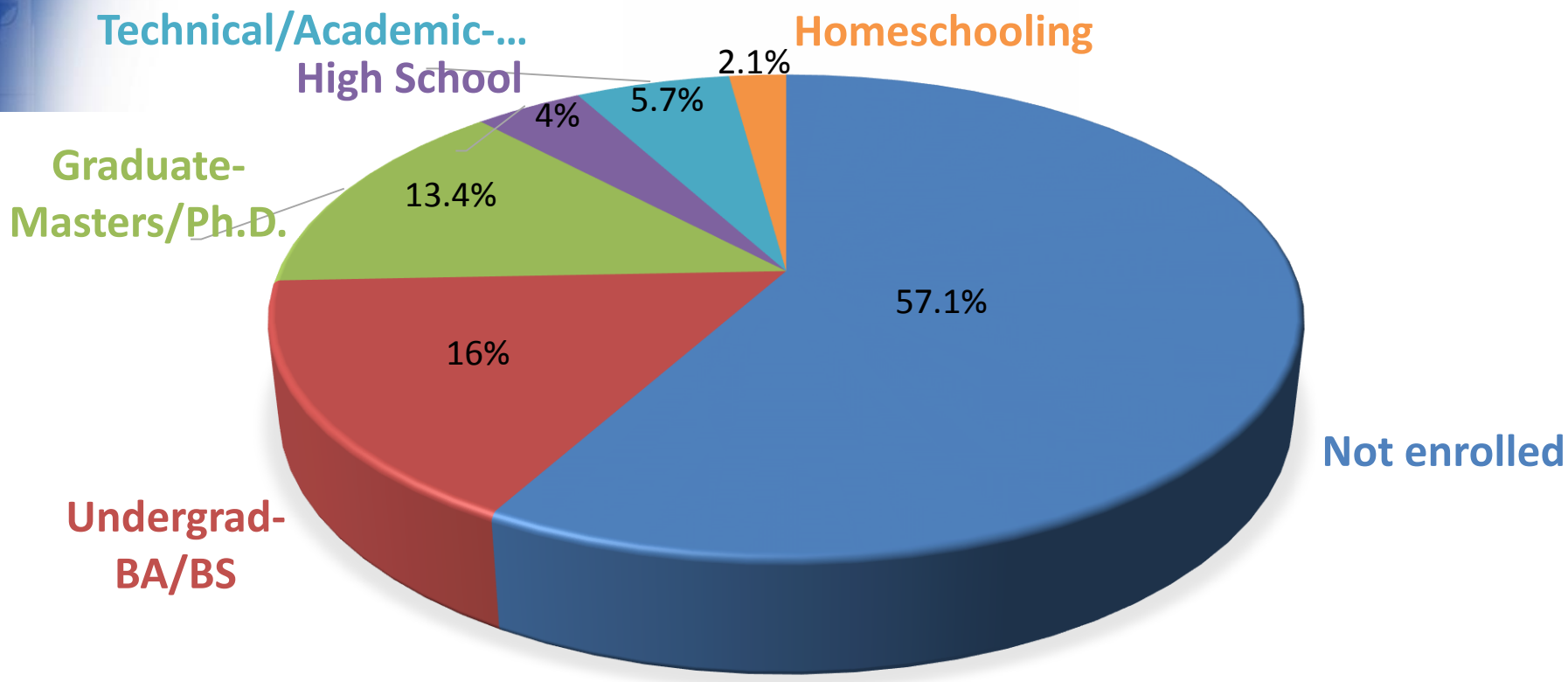
$\bar{Y}$: *Average value of all observations from sample* Y

CASE STUDY

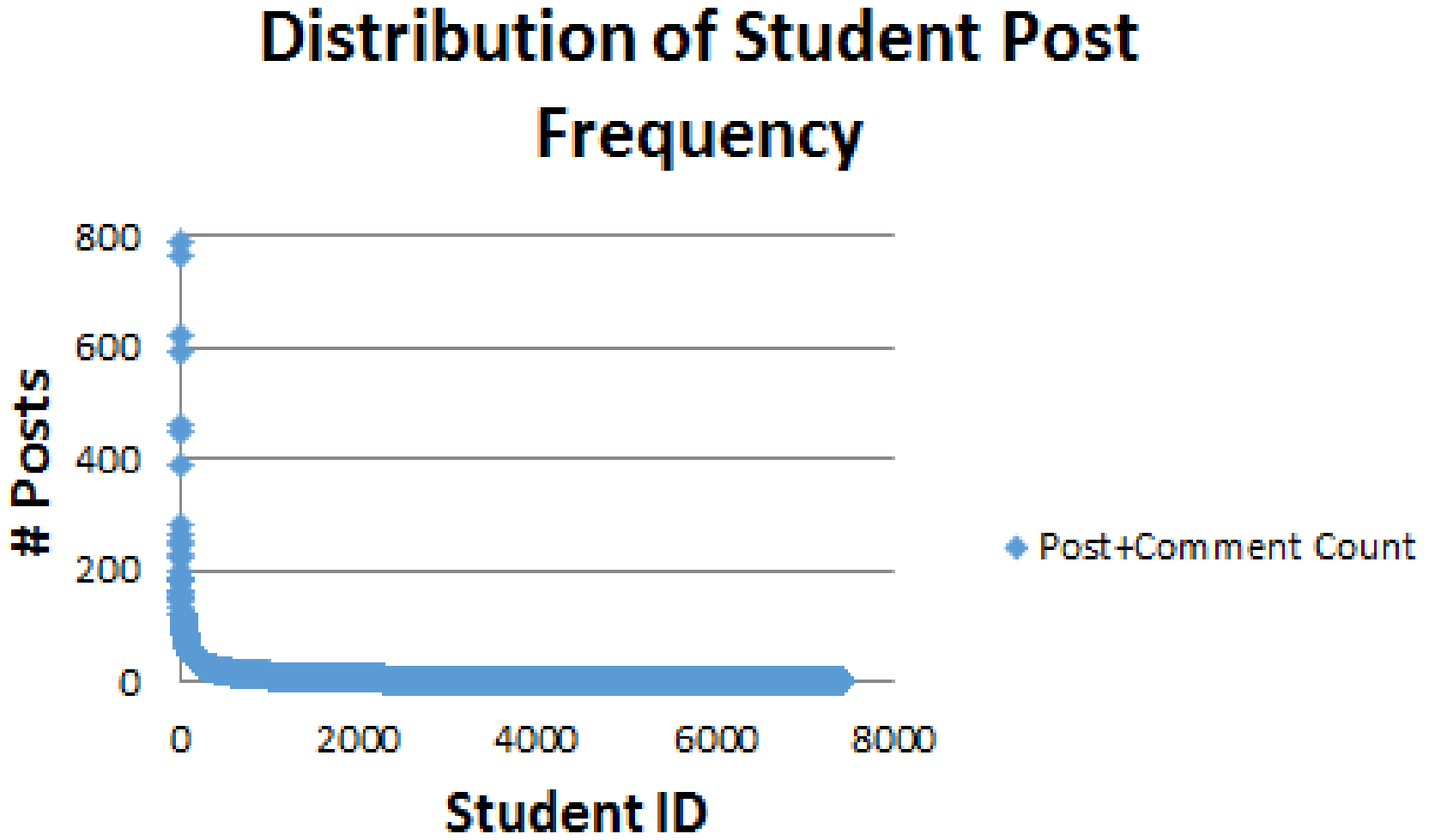
Introduction to Art: Concepts and Techniques

- Penn State MOOC
- was offered through Coursera: Spring of 2013.
- modeled after an online art course "Art 10: Introduction to Visual Studies" taught by Anna Divinsky.

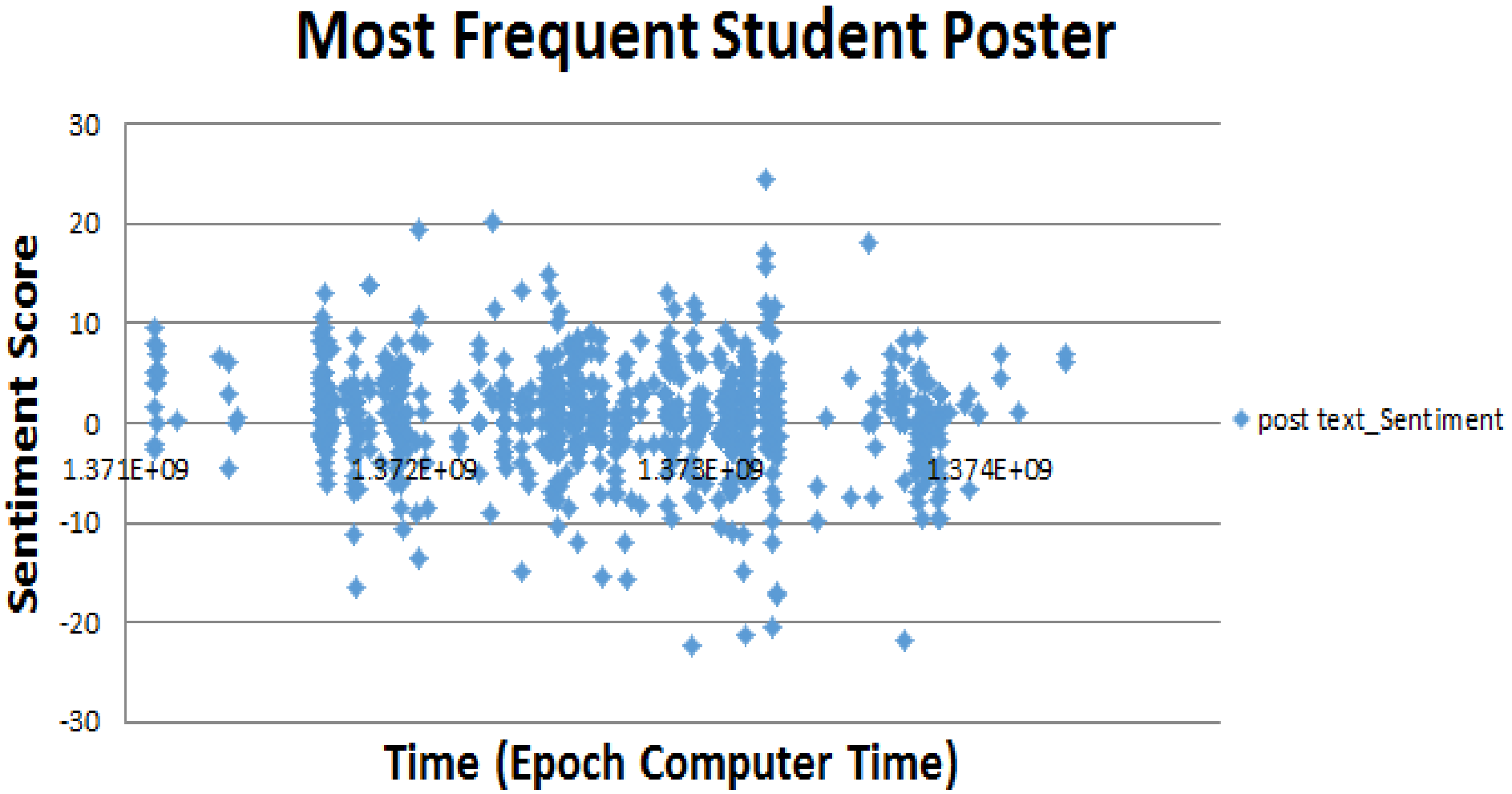
Introduction to Art: Concepts and Techniques



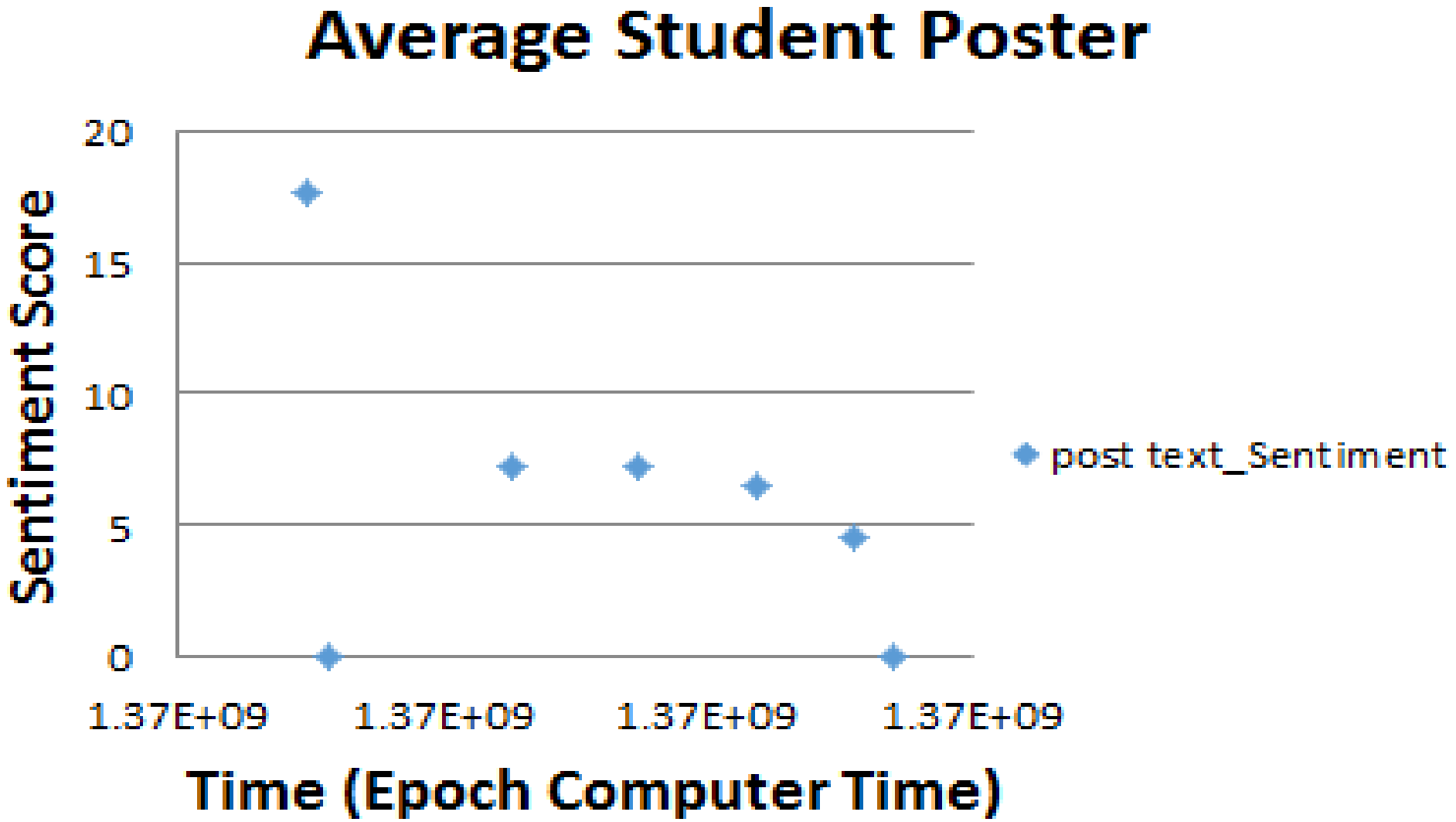
Results


Distribution of Student Post Frequency
800
600
400
200
0
Posts
0
2000
4000
6000
8000
Student ID
Post+Comment Count

Results

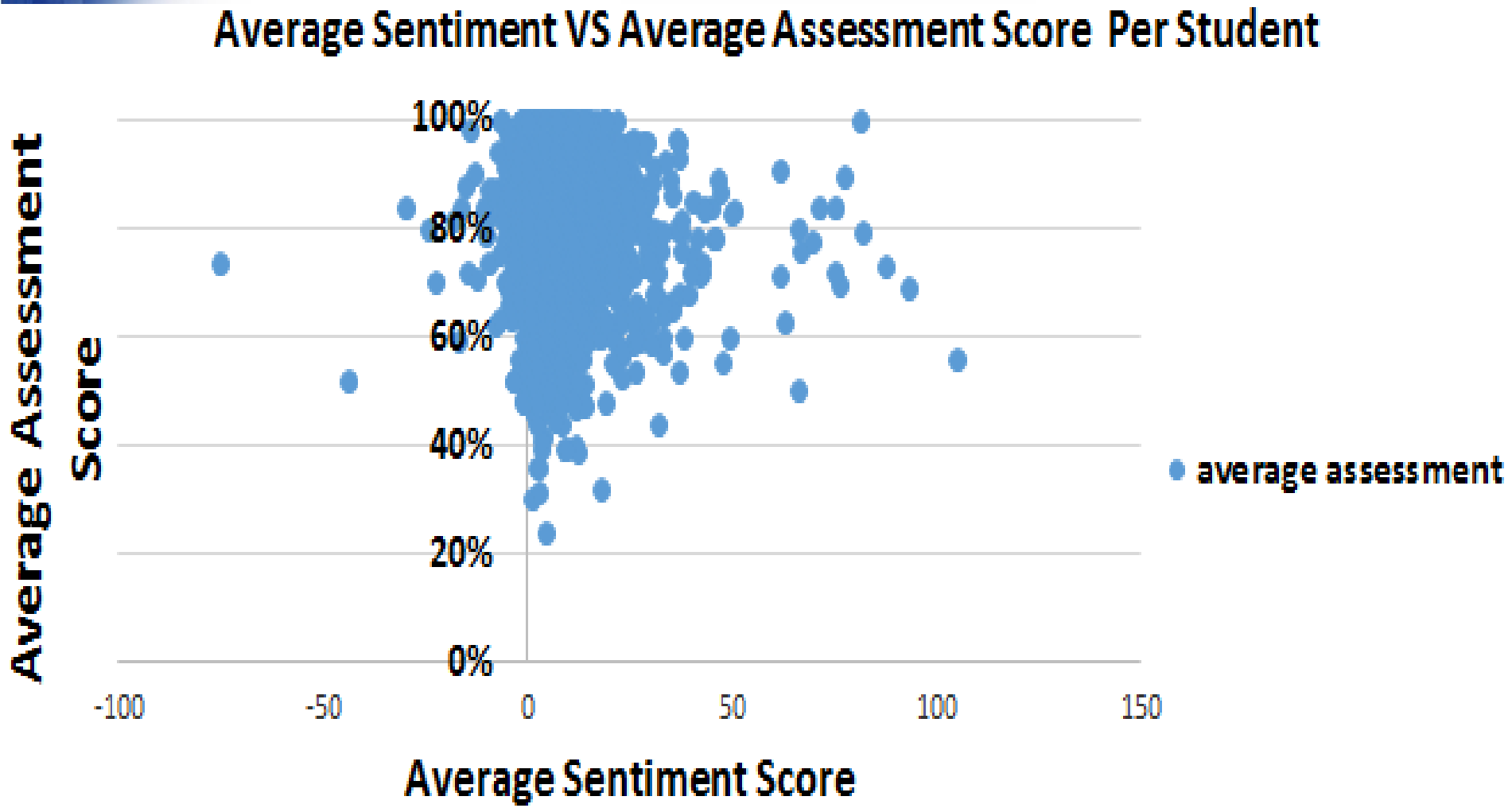

Most Frequent Student Poster
30
20
10
0
-10
-20
-30
Sentiment Score
1.371E+09
1.372E+09
1.373E+09
1.374E+09
Time (Epoch Computer Time)
post text_Sentiment

Results


Average Student Poster
Sentiment Score
20
15
10
5
0
1.37E+09
1.37E+09
1.37E+09
1.37E+09
Time (Epoch Computer Time)
post text_Sentiment

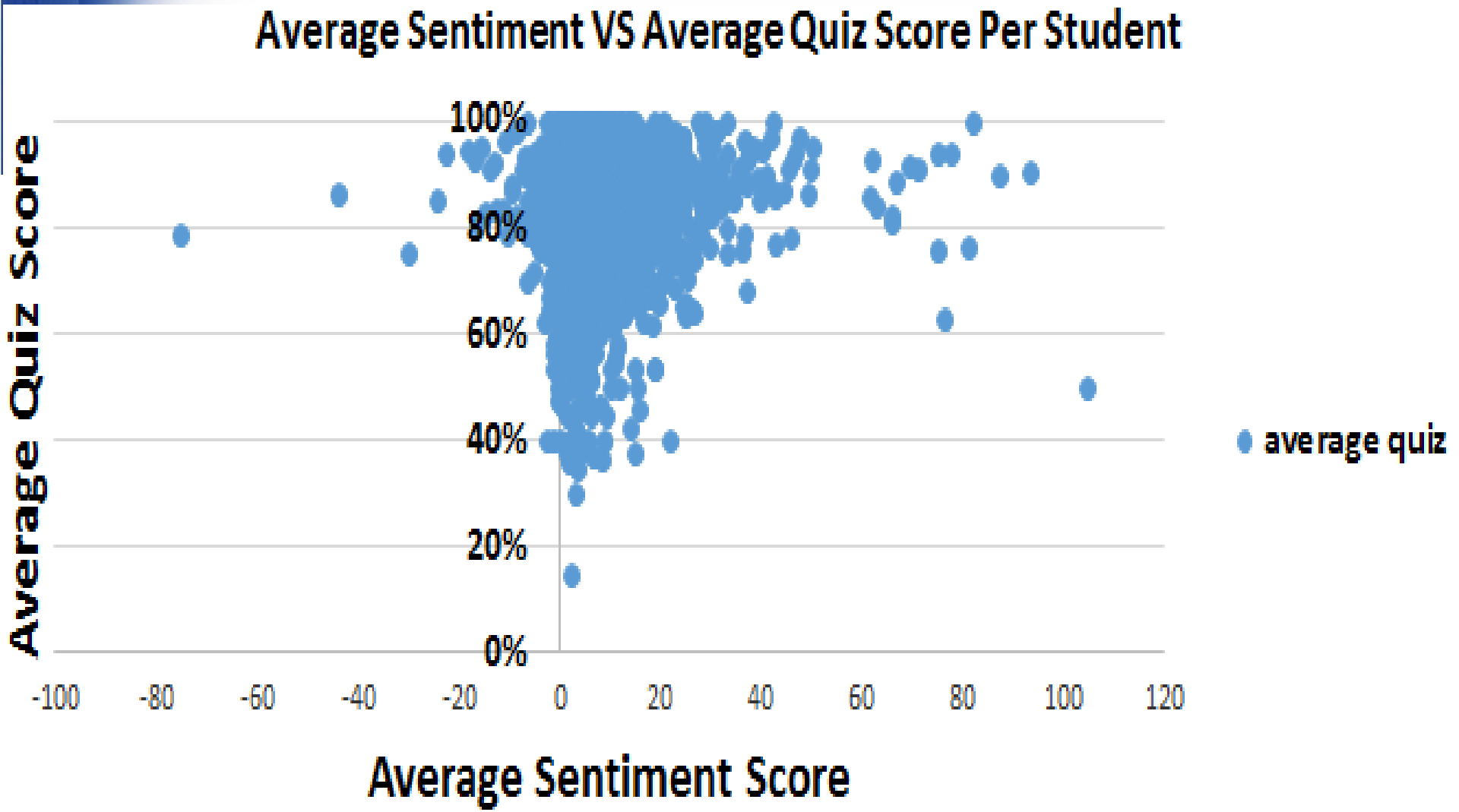
Results

r= -0.016



Results

r= 0.04



FEEDBACK TO EDUCATORS

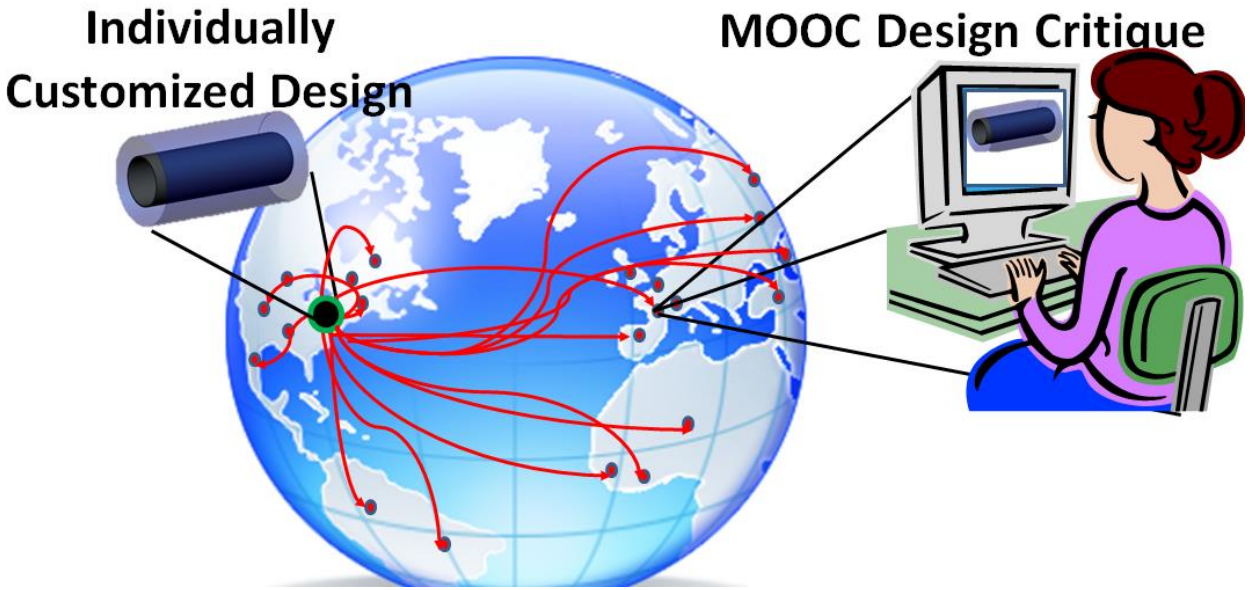
Task	Approach
Grouping Students	• Clustering/Classification • Social network analysis with visualization
Detecting Undesirable Student Behaviors	• Outlier Detections
Predicting Student Performance	• Feature selection • Classification/Clustering • Network Mining

FEEDBACK TO STUDENTS

Task	Approach
Recommending courses to Students	• Relationship mining • Classification/Clustering
Encouraging participation online	• Text mining • Relationship mining

Conclusion and Path Forward

- MOOCs as “Design Critique” Platforms
- Expand to other MOOCs
- Ground truth data, based on students’ surveys
- Data beyond textual (e.g., image, geospatial, etc.)



Acknowledgement & References

Contributors:

Conrad Tucker (Pennsylvania State University, University Park)

Bryan Dickens (The Pennsylvania State University),

Anna Divinsky (The Pennsylvania State University)

References

- Jia, L., & Wang, R. (2010, October). The application skills of body language in teaching. In *Artificial Intelligence and Education (ICAIE), 2010 International Conference on* (pp. 18-21). IEEE.
- Antes, T. A. (1996). Kinesics: The value of gesture in language and in the language classroom. *Foreign Language Annals*, *29*(3), 439-448.
- Corts, D. P., & Pollio, H. R. (1999). Spontaneous production of figurative language and gesture in college lectures. *Metaphor and Symbol*, *14*(2), 81-100.
- Kellerman, S. (1992). 'I see what you mean': The role of kinesic behaviour in listening and implications for foreign and second language learning. *Applied Linguistics*, *13*(3), 239-258.