2023 Annual Conference & Exposition
Baltimore Convention Center, MD | June 25 - 28, 2023

The Harbor of Engineering
Education for 130 Years

ASEE

Paper ID #38728

Work in Progress: Using Machine Learning to Map Student Narratives of Understanding and Promoting Linguistic Justice

Harpreet Auby, Tufts University

Harpreet is a graduate student in Chemical Engineering and STEM Education. He works with Dr. Milo Koretsky and helps study the role of learning assistants in the classroom as well as machine learning applications within educational research and evaluation. He is also involved in projects studying the uptake of the Concept Warehouse. His research interests include chemical engineering education, learning sciences, and social justice.

Dr. Milo Koretsky, Tufts University

Milo Koretsky is the McDonnell Family Bridge Professor in the Department of Chemical and Biological Engineering and in the Department of Education at Tufts University. He received his B.S. and M.S. degrees from UC San Diego and his Ph.D. from UC Berkeley,

WIP: Using Machine Learning to Map Student Narratives of Understanding and Promoting Linguistic Justice

Introduction

This work-in-progress paper expands on a collaboration between engineering education researchers and machine learning researchers to automate the analysis of written responses to conceptually challenging questions in statics and dynamics courses [1]. Using the Concept Warehouse [2], written justifications of challenging conceptual questions, called ConcepTests (CTs), were gathered from a diverse set of two- and four-year institutions. Written justifications for CTs have been used to support active learning pedagogies [3], [4], making it essential to investigate how students assemble their problem-solving narratives of understanding. However, despite the considerable benefit that analysis of student written responses may provide to instructors and researchers, manual review of responses is cumbersome, limits analysis, and can be prone to human bias.

In efforts to improve the analysis of student written responses, machine learning has been used in various educational contexts to analyze short and long texts [5], [6]. Natural Language Processing (NLP) uses machine learning methods like transformer-based machine learning models [7], [8], which can be used through fine-tuning or in-context learning methods. NLP can be used to train algorithms that can automate the coding of written responses. Only a few studies for educational applications have leveraged transformer-based machine learning models, further prompting an investigation into its use in STEM education. However, since language analysis is challenging to automate because of its complexity, NLP has been criticized for increasing the possibility of perpetuating and amplifying harmful stereotypes and implicit biases [9], [10].

This study details preliminary results to plan for using NLP for linguistic justice. Linguistic justice is defined as equitable access to political or social life through language [11]. Through text summary and topic modeling utilizing machine learning tools like Bag-of-Words (BoW) and latent Dirichlet allocation [12], we identify critical aspects of student narratives of understanding in written responses to statics and dynamics CTs. We seek to use machine learning to identify different ways students talk about a problem. Through this process, we hope to help reduce human bias in the classroom and through technology by giving instructors and researchers diverse narratives that include insight into their students' histories, identities, and understanding. These can then be used to connect technological knowledge to students' everyday lives.

Background

Ways of understanding that deviate from normative Western discourses have been historically excluded from schooling [13], [14], [15]. These narratives of understanding give us ideas regarding the sensemaking and processing students undertake when learning. During problem-solving processes, such as answering complex concept questions, we risk losing their narratives because their everyday language may not fit into the standard accepted by the majority [15]. Students and teachers unconsciously form conceptions of performance-based expectations due to status characteristics like gender, race, class, etc. How others express their ideas significantly impacts which students are taken seriously and who is given access to the conversational floor [16], [17]. Multiple solutions have been proposed to alleviate these inequities in education.

Culturally responsive [18] and sustaining pedagogies [19] aim to ensure that students' histories and identities are sustained in the classroom. However, in larger classrooms, as with the mechanics courses studied here, these pedagogies can be challenging to implement.

As students have different histories and identities that they draw upon to formulate their academic discourses to write these written responses, analysis using NLP becomes tricky as most data don't include culturally diverse language and narratives within them [9], [10]. Thus, existing algorithms will recognize students who express their ideas in specific ways more often. Linguistic justice in NLP aims to create algorithms that can effectively analyze large amounts of data without leaving out the voices of non-dominant discourses.

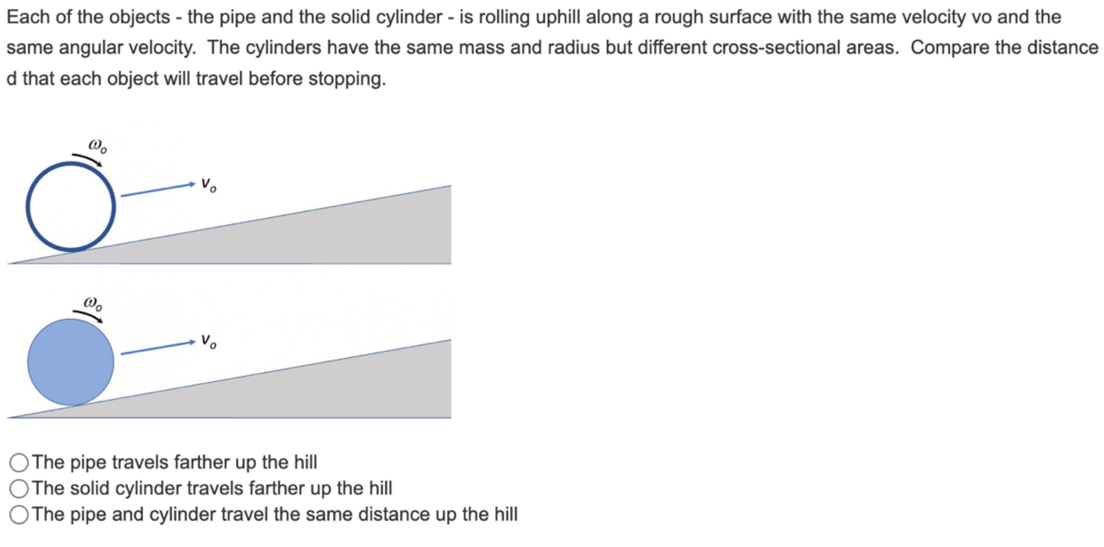
Methods

Context and Setting

This study emerges from a larger study investigating the uptake of the Concept Warehouse [2] in mechanics courses at diverse two- and four-year institutions. Eight common concept questions (four in statics and four in dynamics) were given at all institutions. All instructors used the written short answer follow-up, which asks, "Please explain your answer in the box below," to a conceptually challenging multiple-choice question.

Qualitative Coding

Emergent and *a priori* coding methods [20] were used to analyze CW Question 6141 (shown in Figure 1), as detailed previously [1]. Similar themes emerged from the analysis of this question.



Each of the objects - the pipe and the solid cylinder - is rolling uphill along a rough surface with the same velocity vo and the same angular velocity. The cylinders have the same mass and radius but different cross-sectional areas. Compare the distance d that each object will travel before stopping.

○ The pipe travels farther up the hill
○ The solid cylinder travels farther up the hill
○ The pipe and cylinder travel the same distance up the hill

Figure 1. CW Question 6141

Students use the processes of **identification, comparison, and inference** to navigate through answering this problem. This cognitive process sets the basis of a narrative of understanding. Students **identify** concepts like the moment of inertia and other physical properties of the system. Students use these concepts to **compare** initial energy, angular momentum, and kinetic energy. Finally, students use these concepts and comparisons to make **inferences** about translational and rotational motion. Although students use these common themes in their narratives, they convey knowledge differently.

Machine Learning

Text summary, topic modeling, and the Naïve Bayes Classifier were used as exploratory methods to analyze 106 written responses. Text summary and topic modeling are unsupervised machine learning methods, while the Naïve Bayes Classifier is a supervised machine learning method.

Text Summary

Bag-of-Words modeling compiles words and phrases from the data set to condense large amounts of text. More formally, this kind of test feature extraction is done by splitting all words in the data set and creating representations of the word in a vector format. This can be used alongside sentence- and word-tokenizing tools to summarize the text.

Text Modeling

Methods 1 and 2 of text modeling were completed using the same pre-processed data. Method 1 and Method 2 used the same algorithm but extracted different numbers of keywords. They both utilized latent Dirichlet allocation (LDA), which is a Bayesian model that makes documents "random mixtures over latent topics, where a distribution over words characterizes each topic" [12, p. 996]. More simply, LDA sees topics as probability distributions, and to generate these, LDA finds patterns of words that repeat together, occur frequently, or are similar. Then the algorithm will tag documents with these topics.

Naïve Bayes Classifier

To see the potential loss of narratives of understanding, a multinomial Naïve Bayes Classifier was used to see how well a computer could predict, based on a response, whether it was correct. We must emphasize that this **is not** to develop a tool that could predict correct answers of a response but to see how many narratives of understanding **we could lose**. Naïve Bayes is just one metric to see the potential loss of student narratives of understanding. Naïve Bayes in NLP uses the principles of Bayesian thinking to predict a posterior probability. In this case, we predict if a response is correct based on a topic (a set of words), as seen in the equation below.

$$P(Correct|Topic) = \frac{P(Topic|Correct) * P(Correct)}{P(Topic)}$$

Since multinomial Naïve Bayes assumes words are independent of one another, these probabilities for each word being correct in a topic are multiplied to calculate the final probability. For example:

$$\begin{aligned} &P(Topic|Correct) \\ &\qquad = P(Topic_{word\ 1}|Correct) \times P(Topic_{word\ 2}|Correct) \times \ldots \\ &\qquad \times P(Topic_{word\ n}|Correct) \end{aligned}$$

This can then be used to predict if a response will be correct based on these probabilities.

Preliminary Findings

Text Summary

Extensive explanations can be condensed into shorter explanations, as shown in Table 1. N-grams (shown in Appendix A.1) are examples of words and phrases that the BoW method uses to create the vectorized list of words. This can then be used to summarize text as shown in Table 1.

Table 1. Text summary example results.

	Student Response	**Machine Output**
Example 1	The mass moment of inertia for the ring is I=mr^2 and for the solid disk, its I=1/2mr^2. So the MMOI of the ring cross section is larger, so the initial kinetic energy of the ring is larger. Therefore, it has more initial energy. All of this energy will be converted to potential energy. Since the ring cross section has more initial energy, it will travel higher before stopping because all kinetic energy is converted to potential energy. This correlates to it going farther up the ramp before stopping.	The solid cylinder has a smaller mass moment of inertia, so it will lose less kinetic energy over time compared to the pipe.
Example 2	Although they have the same angular velocity the moment of inertia of the ring is greater. The ring will have more energy going into rotational kinetic energy and they have the same linear kinetic energy. Therefore the first system has more energy and will go further up the hill.	Therefore the first system has more energy and will go further up the hill.

Text Modeling

Through Method 1, the following keywords were determined.

[['energy', 'pipe', 'cylinder'], ['inertia', 'moment', 'larger']]

Method 2 obtained the following 10 topics. Each topic lists related words as determined by LDA. As the increase in topics happens, the terms get less specific regarding the problem topic. Sticking with a smaller set of topics in Method 1 presents a more accurate set of keywords.

Topic 0: ['bit', 'b.', 'assumed', 'compressed', 'bd', '4', 'ac']

Topic 1: ['assuming', 'answers', 'at least', 'a-c', 'chose', 'confusing', 'believe']

Topic 2: ['bars', 'called', 'angles', 'and', 'act', 'change', 'a.']

Topic 3: ['acting', 'coming', 'central', 'concluded', 'clear', '(', 'cd=tension']

Topic 4: ['compresses', 'balancing', 'center', 'clicking', '45deg', 'basically', 'because']

Topic 5: ['analyzing', '@', 'bc', 'causes', '100', 'cb', 'b-']

Topic 6: ['approach', '?', '0.667p', 'able', 'asking', 'cancel', '"']

Topic 7: ['cd', 'adjacent', 'connected', '2p', 'b', 'apart', 'cause']

Topic 8: ['certain', 'calculations', 'based', 'bodies', '>', 'al', 'caused']

Topic 9: ['answer', 'balanced', 'actually', ')', 'approaching', '1.89p', 'completely']

Naïve Bayes Classifier

Using a multinomial Naïve Bayes classifier with a 70/30 train-test split, a **70.00%** accuracy was obtained.

Implications and Future Directions

This exploratory data analysis has shown that machine learning has potential applications for promoting linguistic justice in NLP. These methods help us understand the different language students may use to answer questions. As NLP can help us shorten and gain knowledge about common concepts that students think about in a problem, it can help researchers and instructors determine what topics or patterns to look for in the data. This is especially true from text summarization and topic modeling, as these allow instructors to discover topics students may discuss. Once instructors and researchers have ways to understand possible trends, they can search for non-conventional ways students may choose to convey these topics. For researchers, this can provide ways to iterate the codebook and look for ways to improve the training set for machine learning. For instructors, this provides interesting information regarding patterns and trends in their classes to draw conversation upon. We summarize these potentials as the following:

- **Text summary using BoW:** Human coders can compare the shortened response to investigate how the machine condenses responses and then make decisions to improve how the machine creates these shortened responses.
- **Text modeling using LDA:** Researchers and instructors can learn about common language students use when answering conceptual questions using the generated topic lists.
- **Naïve-Bayes Classifier:** Researchers can look back at responses tagged as incorrect to see why the machine may have tagged them as incorrect to investigate the possible error in the machine's interpretation of the language.

We intend to use these tools to build a **partnership** between the human coder and automated machine coding. Humans and machines can both be biased. In other small-group collaborative learning settings, if those biases are different, then reconciling those differences can help with promoting linguistic justice. With further investigation, these methods above can help researchers explore what can be done to reconcile the biases between human and machine coders. Additionally, this work applies to the professional formation of engineers since engineers must be able to communicate with many audiences. Using machine learning tools like those described above, instructors and researchers can learn more about the language students use and can emphasize and attend to the diversity of language that can be used to answer conceptually challenging problems.

Moving forward, we hope to:

- Further understand what narratives of understanding are excluded from analyzing student written responses to conceptually challenging problems.
- Gather more text samples that center written responses to conceptually challenging problems from underrepresented groups to adequately train algorithms.

Acknowledgments

We acknowledge support from the National Science Foundation (NSF) through the NRT DGE 2021874 and DUE 2135190. Any opinions, findings, conclusions, or recommendations expressed are those of the authors and do not necessarily reflect the views of the NSF.

References

[1] H. Auby, N. Shivagunde, A. Rumshisky, and M. D. Koretsky, “WIP: Using machine learning to automate coding of student explanations to challenging mechanics concept questions,” presented at the American Society for Engineering Education, Minneapolis, MS, Jun. 2022.

[2] M. D. Koretsky, J. L. Falconer, B. J. Brooks, and Silverstein, “The AIChE Concept Warehouse: A web-based tool to promote concept-based instruction,” *Advances in Engineering Education*, vol. 4, no. 1, p. 27, 2014.

[3] M. D. Koretsky, B. J. Brooks, R. M. White, and A. S. Bowen, “Querying the questions: Student responses and reasoning in an active learning class,” *Journal of Engineering Education*, vol. 105, no. 2, pp. 219–244, 2016, doi: 10.1002/jee.20116.

[4] M. D. Koretsky, B. J. Brooks, and A. Z. Higgins, “Written justifications to multiple-choice concept questions during active learning in class,” *International Journal of Science Education*, vol. 38, no. 11, pp. 1747–1765, Jul. 2016, doi: 10.1080/09500693.2016.1214303.

[5] J. Burstein *et al.*, Eds., *Proceedings of the 15th Workshop on Innovative Use of NLP for Building Educational Applications*. Seattle, WA, USA → Online: Association for Computational Linguistics, 2020. [Online]. Available: https://aclanthology.org/2020.bea-1.0

[6] J. Burstein *et al.*, Eds., *Proceedings of the 16th Workshop on Innovative Use of NLP for Building Educational Applications*. Online: Association for Computational Linguistics, 2021. [Online]. Available: https://aclanthology.org/2021.bea-1.0

[7] T. B. Brown *et al.*, “Language models are few-shot learners.” arXiv, Jul. 22, 2020. Accessed: Apr. 03, 2023. [Online]. Available: http://arxiv.org/abs/2005.14165

[8] C. Raffel *et al.*, “Exploring the limits of transfer learning with a unified text-to-text transformer.” arXiv, Jul. 28, 2020. Accessed: Apr. 03, 2023. [Online]. Available: http://arxiv.org/abs/1910.10683

[9] K.-W. Chang, V. Prabhakaran, and V. Ordonez, “Bias and fairness in natural language processing,” in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): Tutorial Abstracts*, Hong Kong, China: Association for Computational Linguistics, Nov. 2019. [Online]. Available: https://aclanthology.org/D19-2004

[10] E. Mayfield *et al.*, “Equity beyond bias in language technologies for education,” in *Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications*, Florence, Italy: Association for Computational Linguistics, Aug. 2019, pp. 444–460. doi: 10.18653/v1/W19-4446.

[11] J. Nee, G. M. Smith, A. Sheares, and I. Rustagi, “Linguistic justice as a framework for designing, developing, and managing natural language processing tools,” *Big Data & Society*, vol. 9, no. 1, p. 20539517221090930, Jan. 2022, doi: 10.1177/20539517221090930.

[12] D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent Dirichlet allocation,” *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[13] L. D. Delpit, “The politics of teaching literate discourse.,” in *Language and Linguistics in Context*, H. Luria, D. M. Seymour, and T. Smoke, Eds., 1st ed.Routledge, 2005.

[14] J. P. Gee, *Social linguistics and literacies: ideology in discourses*, 3rd ed. London ; New York: Routledge, 2008.

[15] K. D. Gutierrez, "Unpackaging academic discourse," *Discourse Processes*, vol. 19, no. 1, pp. 21–37, Jan. 1995, doi: 10.1080/01638539109544903.

[16] R. A. Engle, J. M. Langer-Osuna, and M. McKinney de Royston, "Toward a model of influence in persuasive discussions: negotiating quality, authority, privilege, and access within a student-led argument," *Journal of the Learning Sciences*, vol. 23, no. 2, pp. 245–268, Apr. 2014, doi: 10.1080/10508406.2014.883979.

[17] L. A. Kurth, C. W. Anderson, and A. S. Palincsar, "The case of Carla: Dilemmas of helping all students to understand science," *Sci. Ed.*, vol. 86, no. 3, pp. 287–313, May 2002, doi: 10.1002/sce.10009.

[18] G. Ladson-Billings, "Toward a theory of culturally relevant pedagogy," *American Educational Research Journal*, vol. 32, no. 3, pp. 465–491, Sep. 1995, doi: 10.3102/00028312032003465.

[19] D. Paris and H. S. Alim, "What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward," *Harvard Educational Review*, vol. 84, no. 1, pp. 85–100, Mar. 2014, doi: 10.17763/haer.84.1.982l873k2ht16m77.

[20] J. W. Creswell and C. N. Poth, *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, 4th ed. 2018.

Appendix A: Machine Learning Additional Results

A.1 N-grams

Digram and trigram detailing common phrases found in summarization. The x-axis represents the frequency of the sets of words.

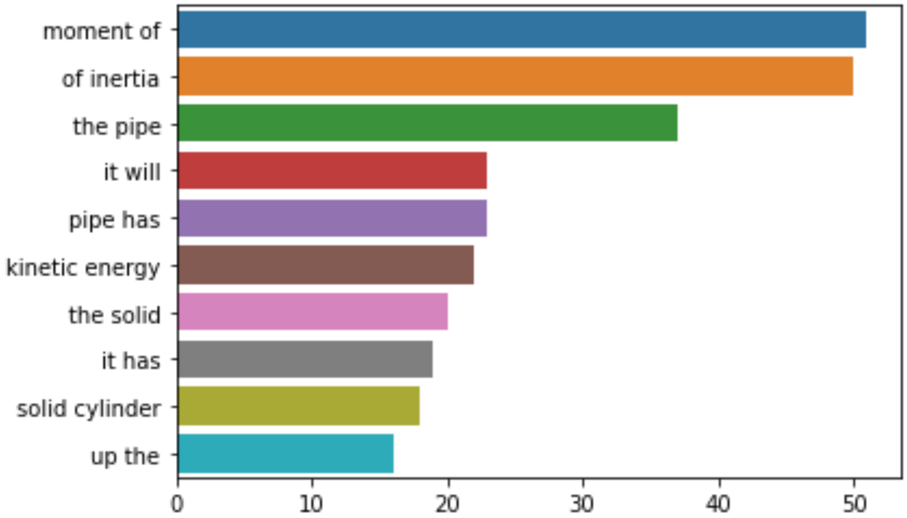


Figure A1. Digram frequencies for CT 6141.

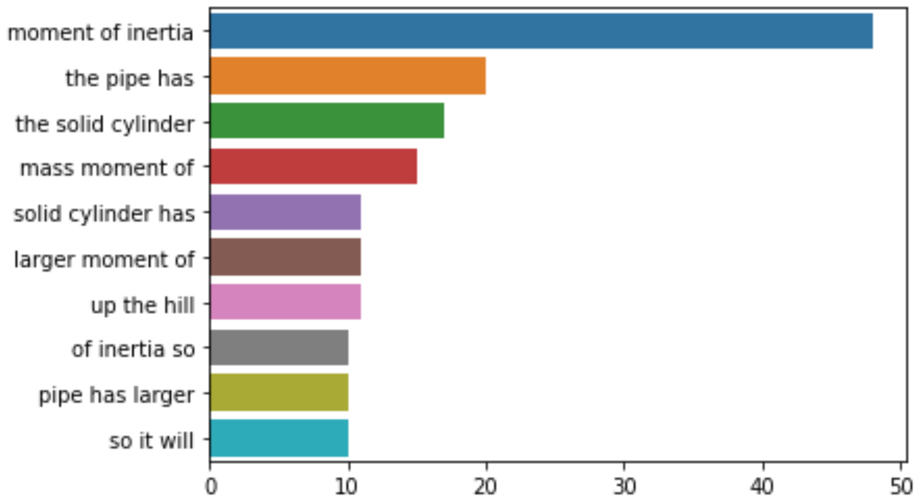


Figure A2. Trigram frequencies for CT 6141.