ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022 ASEE

Paper ID #36551

Work-in-Progress: Understanding learners' motivation through machine learning analysis on reflection writing

Elizabeth Pluskwik

Elizabeth facilitates project-based and co-op based engineering education in the Iron Range Engineering program, Minnesota State University, Mankato. Her specialties leading entrepreneurial mindset in engineering, engineering management, accounting, product-costing, and lean six sigma. Her research interests include motivation to persist in engineering, emotional intelligence, and industry 4.0.

Yuezhou Wang

Dr. Yuezhou Wang is an associate professor at Minnesota State University, Mankato. After receiving his Ph.D. from University of Minnesota in 2017, he works for the Iron Range Engineering, a project-based learning program. His teaching interests are in areas of materials science, structural analysis, finite element modeling and dynamic systems. His technical research focuses on multiscale modeling on mechanical behavior of nano and granular materials.

Lauren Singelmann

Lauren Singelmann earned her Ph.D. from North Dakota State University in Electrical and Computer Engineering and STEM Education in 2022. She is a faculty member for Iron Range Engineering through Minnesota State University, Mankato, and she supports instruction of Innovation-Based Learning courses at multiple institutions. Her research interests include learning analytics, experiential learning, and equitable grading and assessment.

Work-in-Progress: Understanding learners’ motivation through machine learning analysis on reflection writing

Abstract

Educational data mining (EDM) is an emerging interdisciplinary field that utilizes a machine learning (ML) algorithm to collect and analyze educational data, aiming to better predict students' performance and retention. In this work-in-progress paper, we plan to report our methodology and preliminary results from utilizing an ML program to assess students' motivation through their upper-division years in the Iron Range Engineering project-based learning (PBL) program. ML, or more specifically, the clustering algorithm, opens the door to processing large amounts of student-written artifacts, such as reflection journals, project reports, and written assignments, and then identifies keywords that signal their levels of motivation (i.e., extrinsic vs. intrinsic). These results will be compared against other measures of motivation, including student self-report, faculty observation, and externally validated surveys. As part of a longer-term study, this preliminary work sheds light on the key question for student success and retention: how does student motivation evolve through the 3rd and 4th years in college?

The purpose of this research project is to gain insights into learners' motivation levels and how it evolves during the last two years in college, as well as to extend current Educational Data Mining research and Machine Learning analysis described in the literature. It is significant on two fronts: 1) we will extend the ability of ML in analyzing reflective written artifacts to explore student physiological and emotional development; 2) the longitudinal study will help monitor the progressive change of motivation in college students in a PBL environment.

Preliminary results from an initial preliminary study are promising. By analyzing written reflection journal entries from previous students, the ML algorithm has differentiated keywords into three student motivation levels: "high", "neutral" and "low". Using supervised classes, for example, the ML algorithm differentiated words in the highly motivated student text such as "team" and "learning", while the text coded as low motivation included "use", "pushed" and "nothing".

For our future research, we aim to create a dictionary that identifies words/phrases related to positive/negative motivation. We will extend the preliminary study to a longitudinal evaluation of student motivation over the four semesters of engineering education as well as prediction of student success in a PBL environment.

Introduction

Machine Learning (ML) is a branch of computer science that focuses on developing automatic self-improving algorithms based on data input and experience [1]. As a highly impactful tool to analyze big data in manufacturing, engineering, and even social and political sciences, ML has experienced rapid growth and gone far beyond software development itself. In education, for instance, the fusion of ML and pedagogy becomes an emerging field of study known as educational data mining (EDM). The International Educational Data Mining Society defines EDM as "a developing method for exploring unique and increasing large scale data that comes from an educational setting to better understand the students" [2]. In other words, EDM converts educational information into intelligent action, aiming to improve learning outcomes [3]. Major methods of EDM include regression, associating rules, sequential pattern analysis, and clustering

[4]. These methods are generally borrowed from data science/mining. For example, associating rules are used to identify the correlation between data sets and clustering is to divide the dataset into small groups based on their similarities.

The popularity of EDM stems from the efficiency of ML to process a large amount of student data as well as the low cost for maintenance due to its automatic improvement. Not only can a prediction be made, but also an individualized recommendation would be generated by the algorithm, allowing a personalized education approach, also known as "precision education" [5]. More specifically, educators focused on forecasting students' academic performance (i.e., GPA and grades). Beck and Wolf [6] created a learning agent that determined the likelihood for a student to answer a question correctly. The agent was trained with the student's previous efforts and feedback was generated to the instructor. The agent gradually evolves and was built into a platform known as the "Intelligent Tutoring System". Alkhasawneh and Hobson [7] developed neural network models to predict the retention rate of more than 300 incoming freshmen students at the Virginia Commonwealth University based on their GPAs. The model classified students as at-risk, intermediate, and advanced, and had an overall accuracy of 70%. Similarly, Lakkaraju et al. [8] designed an ML framework to identify students at risk of not graduating high school among more than 200,000 students. Their pioneer work in the K-12 setting promoted a data-driven approach to combat academic adversity. Huang and Fang [9] predicted student learning outcomes in an engineering dynamics course by comparing four commonly used ML models, namely, multiple linear regression perception network, function network, and support vector machine model. They concluded that the selection on the type of model would depend on whether the assessment was done on average or individual performance, meaning the ML model may have a certain bias if not employed appropriately. At the program level, Xu et al. [10] recognized the dynamic progress a student could make through the years in college. Their model weighted in student's major and course relevance information and tracked student's performance over the next three years. Many university administrators employed EDM in student recruitment and placement. For example, the development of placement predictor systems was used in the admission process to evaluate students' programming, analytical and communication skills [11, 12].

Despite the achievement of ML in education, with a shift to a more student-centered learning environment, such as project-based, inquiry-based, and innovation-based learning, the traditional approach of analyzing GPA and grades may not accurately reflect students' creativity and critical thinking skills. A more advanced ML application that can process qualitative data is thus needed. Gobert et al. [13] launched a web-based learning environment, Inquiry Intelligent Tutoring System, in which middle schoolers conducted inquiries on an aquatic ecosystem through simulations and animations. Being tasked with these inquiries, students then were asked to make hypotheses, run simulations with multiple trials, analyze the data and finally communicate their findings. The system was able to distinguish an effective learning process from a poor one with 79% accuracy. Inspired by the well-known "Bloom's Taxonomy", Singelmann et al. applied ML in assessing the depth of learning in an innovation-based learning classroom [14, 15]. In addition to quantitative metrics (e.g., frequency of login, number of learning objectives and deliverables, etc.), text artifacts were also clustered into four categories: surface level, surveyors, learners, and innovators. Keywords were selected by the ML algorithm to signify the category. For example, frequent keywords related to class assignments, such as "website, presentation, and review", were more associated with surface learners, since they might care about the basic requirement of a passing grade.

Knowing the rapid change of higher education, we pursue the direction of extending ML capability in assessing college students' cognitive process. Starting within our Program, we implement an ML algorithm to analyze students' perception and motivation under project-based learning. Besides the technical development of the ML program, this study will also help gain insight into learners' motivation levels and how it changes during college.

Preliminary Study: Research Questions

The goal of this study is to explore the application of ML in evaluating student motivation levels. As can be seen in Figure 1, we plan to measure learners' motivation levels during their 3rd and 4th years of college, using ML to differentiate keywords and phrases used by students in their learning journals (i.e., written artifacts of weekly reflection). The results of the ML analysis will be compared to other measures of motivation, including student self-report, faculty observation, and externally validated surveys.

Our preliminary study (P) aims to answer the following questions:

<u>*P1.*</u> Can a machine learning classifier be trained to analyze learning journal text and differentiate between low and high motivation with sufficient agreement to the research team?

<u>*P2.*</u> If the classifier model has sufficient performance, what types of words/phrases are most likely to differentiate between low and high motivation?

<u>*P3.*</u> Which type(s) of learning journals most highly correspond with overall student motivation (as reported by faculty observation)?

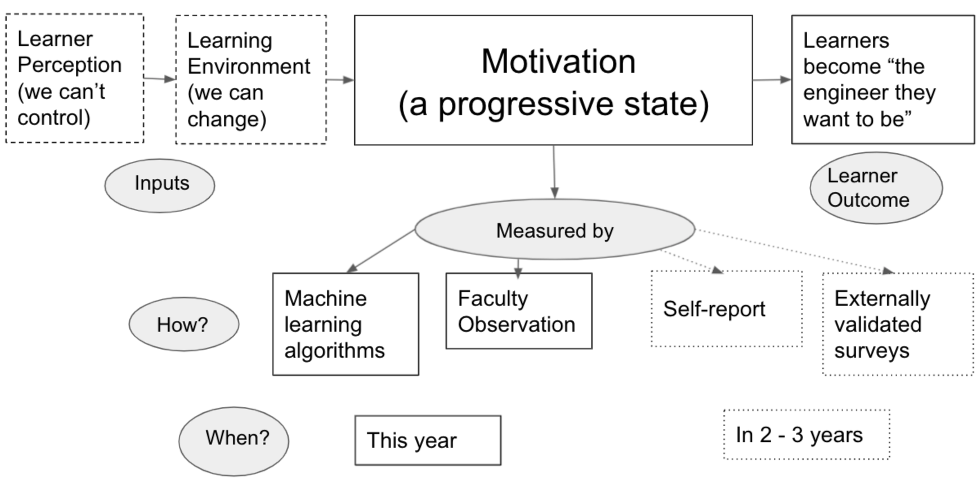


Figure 1 The progression of learner motivation, its inputs, and the plan to validate the results of the machine learning algorithm with other methods.

Research Design and Significance

The research is planned by three preparation actions. First, we identify the connection between reflection and motivation and discuss the significance of understanding the learner's motivation, especially under the PBL environment. Second, based on this connection, we select appropriate content from the learning journals that can potentially indicate the learner's motivation. Third, these journals will feed the existing ML algorithm for revision and training that can better capture the student's choice of words.

We follow Moon's definition of "reflection" as: "A form of mental processing with a purpose and/or an anticipated outcome." [16] It is an internal self-engaged conversation that often reveals perceived value on certain things. [17, 18] "Motivation", on the other hand, is the desire for one individual to perform a task itself. In short, reflection can manifest a purpose, perception, and meaning, and motivation carries them into a desire for action.

By further assessing the reflection details, such as mindset (positive vs. negative), depth, and identification of action steps, we can further evaluate the learner's motivation level through the "continuum of the motivation" shown in Figure 2. This motivation continuum is based on the work of [19], in framing motivation along a continuum of amotivation to extrinsic to intrinsic motivation. They state that people vary in both the level of motivation (how much), as well as the orientation (what type) of motivation. Both measures change over time and environment.

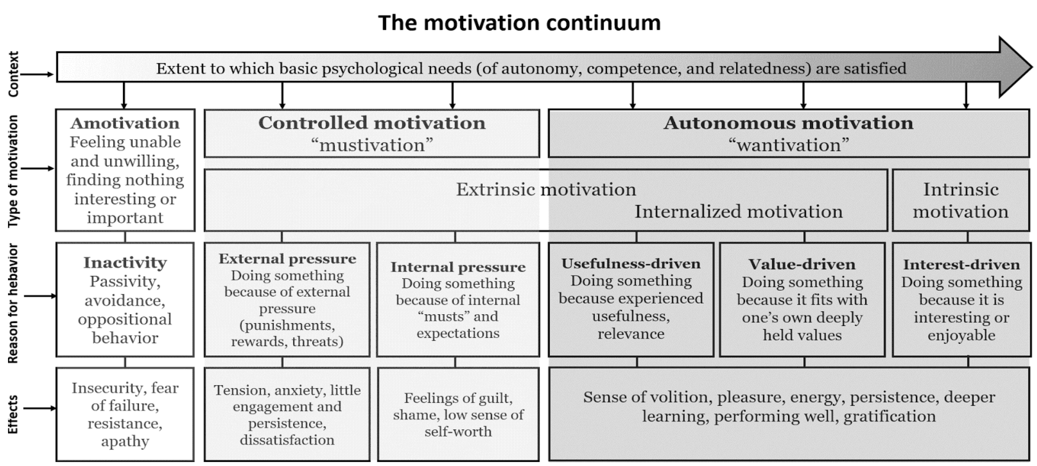


Figure 2 A continuum of learner motivation in self-determination theory [20]. Our study aims to understand the placement and movement of project-based engineering learners through these various levels of motivation to persist.

We hypothesize that this continuum map will hold true in our sample group of learners. While the extrinsic motivation factors such as grades, graduation credit requirement, and salaries, remain a driving force, more students may have started to develop rationale behind their learning activities and discover the inherent value of being an engineer. These intrinsically motivated students are often empowered by their self-efficacy and self-expectancy [21-23]. Again, perception matters. If one believes the ability to reach the goal and expects to create great value, then this individual is motivated to act. Also, classroom environments can facilitate or harm intrinsic motivation, curiosity, and the desire for challenge according to how autonomy-supportive they are perceived to be [19].

Therefore, by understanding perception, not only can educators improve the curriculum and teaching methods for higher retention [24, 25], but also students can also have a positive identity to make career decisions [26] as well as achieve better mental wellbeing [27]. One of the pedagogical goals of the Iron Range Engineering (IRE) program is to create a project-based learning environment that promotes intrinsic motivation among upper-division college students. Unlike traditional lecture-based classes, IRE students engage in open-ended problem solving by working on industry client projects [28, 29]. This innovative approach to engineering education also lets us ponder if our students are truly intrinsically motivated. Surprisingly, our previous study on students' resistance to active learning indicated that a number of students did not perceive they are learning effectively [30]. In fact, according to a report published by Wijnia et al. [31], students in PBL environment may not always be better motivated compared to their peers in the lecture-based environment. Their perception of the learning environment and autonomy played an essential role in their motivation. It is thus interesting to assess the student's perception of the learning environment as well as themselves with the ML approach. Our research will also further the understanding of where project-based learners are generally landing on the continuum (Figure 2) and how their motivation to persist in "becoming an engineer" changes during their 3rd and 4th years of PBL engineering education.

Preliminary Study: Data Collection

The preliminary study was conducted with the written reflection documents of a convenience sample of 28 on-campus students enrolled in the IRE program, an ABET-accredited program of Minnesota State University, Mankato from a course in the Fall of 2017. This program is delivered at an off-campus location, and was designed to educate engineering students in the local region. These students were all participating in industry-sourced design projects and were enrolled in the course which required reflection writing as weekly assignments. Co-op students were not included in the study.

As shown in Table 1, the sample size for the preliminary study was 28 undergraduate engineering students. Sixty four percent were male, and 36% were female. Seventy nine percent were traditional college-age learners (20 - 25 years old), and 21% were non-traditional age (mostly in their 30s and 40s). Half of them were juniors (3rd year) and half were seniors (4th year).

Table 1 The demographics of the participants (N = 28) in the preliminary study.

Demographic Information	**Categories**	**N**	**Percentage**
Gender	Male	18	64.30%
	Female	10	35.70%
Age	Traditional college age	22	79%
	Non-traditional	6	21%
Level in college	Juniors	14	50%
	Seniors	14	50%

Table 2 Selected questions from student learning journals that are used in the ML analysis

Categories	List of Reflection Questions
Design Process (a)	1. What are the ways that you think experiment results should be reported?
	2. Which step in the design process do you consider to be the most important? Why?
	3. Describe how the interactions are going with your client?
	4. Briefly describe your personal model for solving an open-ended problem.
Technical Competency (b)	1. How do you feel about your progress in your technical competencies to this point in the block?
	2. Which competency is going worst? Why?
Professionalism (c)	1. What are three important aspects of interpersonal communication?
	2. How might you work to minimize your own unconscious biases?
	3. What are the essential elements of leadership?
	4. In 2-3 paragraphs, describe your own personal mission statement with regards to making ethical decisions every day.
	5. How do you see mindfulness as being of value to you in the context of your engineering team?
General Experience (d)	1. Describe the most impactful experience you have had in the past 8 weeks in the IRE program?
	2. How can you practice happiness during your college experience?
Teamwork (e)	1. In what ways does your team's current performance hinder team success?
	2. What actions would improve your team's performance?
	3. What is the one thing your team should do to improve the project at this point?
	4. How did individual contributions impact project completion?
Extending to Future (f)	1. How will the jobs package experience guide your future?
	2. What does lifelong learning mean to you?
	3. How does self-directed learning take a role in your lifelong learning?
	4. What steps are you taking to develop your leadership skills for the future?
	5. How is contemporary issue knowledge pertinent to your career as an engineer?
	6. How can you further develop your capacities for empathy and care?

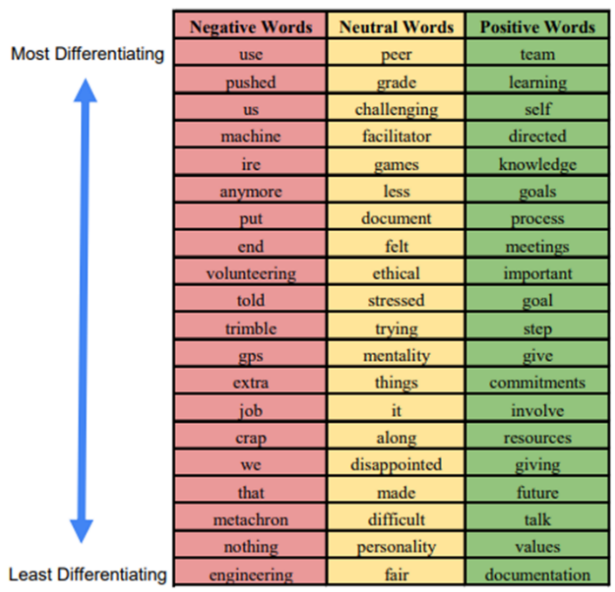
Regarding the reflection practice, in IRE, students write learning journals that are assigned weekly (14 weeks in total) and will count for credit in student Professionalism coursework. The journal is both progressive and reflective, with topics ranging from a reflection on a career fair to a discussion on project team performance. In general, there are three steps of reflection, problem definition, analysis, and generalization [14]. An effective reflection under the engineering education context will also add on open-mindedness, responsibility, and wholeheartedness [14]. Based on this theoretical framework, we select six major categories (Table 2), containing a total number of 23 related questions. Student responses are documented after their identification information is coded.

Preliminary Study: Results and Discussion

To answer P1 and assess preliminary feasibility of a machine learning classifier that differentiates between low and high motivation in learning journals, a machine learning classifier was trained using learning journal text labeled by the instructional team as "positive", "neutral" and "negative" by the instructional team according to the motivation continuum (Figure 2). The classifier was a support vector machine with a linear kernel, a model appropriate for text classification with small sample sizes. In order to assess the performance of the classifier, ten-fold cross-validation accuracy, Cohen's Kappa, precision, recall, and F1 score were calculated. Rather than having one training set and one testing set, ten-fold cross-validation splits the data into ten subsets. A model is trained using nine of the sets and tested on the tenth, the performance is calculated, and this process is completed for each of the ten folds. This method prevents overfitting while still considering all data available. The performance metrics for the optimized model were calculated and listed in Table 3. The ten-fold cross-validation accuracy, Cohen's Kappa, precision, recall, and F1 score all showed sufficient performance above baseline, supporting the hypothesis that the text that students use when writing learning journals can differentiate between low and high motivation at a level of sufficient agreement with a human evaluator.

Table 3 Performance metrics comparing the trained classifier and the random baseline classifier

Performance Metric	**Classifier Performance**	**Random Classifier (Baseline)**
Accuracy	0.900	0.500
Cohen's Kappa	0.750	0.000
F1	0.866	0.424
Recall	0.875	0.500
Precision	0.871	0.368



Negative Words	Neutral Words	Positive Words
use	peer	team
pushed	grade	learning
us	challenging	self
machine	facilitator	directed
ire	games	knowledge
anymore	less	goals
put	document	process
end	felt	meetings
volunteering	ethical	important
told	stressed	goal
trimble	trying	step
gps	mentality	give
extra	things	commitments
job	it	involve
crap	along	resources
we	disappointed	giving
that	made	future
metachron	difficult	talk
nothing	personality	values
engineering	fair	documentation

Figure 3 "Most differentiating" words classified by ML algorithm as negative, neutral, and positive. These words were extracted from students' written reflection journals in the small initial data set from the Spring, 2017 semester. Words that were specific to the program were redacted.

To answer P2, the most important features were extracted from the trained classification algorithm. Faculty members qualitatively coded each learning journal as having "high", "neutral" or "low" levels of motivation. The ML algorithm then used chi-2 values to determine which features were most highly differentiating for each of the classes, and these results can be seen in Figure 3. Negative words such as “pushed”, “nothing” and “anymore” can be distinguished from positive words such as “goal”, “commitment” and “values”. For future work, we propose aligning words and phrases with the various types of motivation included in Ryan and Deci’s motivation continuum [19]. For example, the word “goal” could align with “usefulness-driven” motivation.

Finally, to answer P3, the classifier predicted motivation levels for learning journals in the categories listed in Table 2, and these predictions were compared to instructor observation of overall motivation. This question was explored to see if motivation in learning journal categories corresponds to overall motivation as identified by instructor observations. For example, if the algorithm detects low motivation in a learning journal about a technical competency that the student is taking, does that correlate with low motivation overall? For each student, their learning journal responses in each of the six categories were given a score between 0 and 1. If the score is less than 0.5, the journal would be classified as negative, and if the score was greater than 0.5, the journal would be classified as positive. The difference between these scores and the instructor observation (negative=0, neutral=0.5, positive=1) was calculated, and these error measurements were averaged for each category. For a baseline, a majority class classifier was used and compared to instructor observations; because most students were identified to be positively motivated, a majority classifier scores all students with a 1. The error levels of a random classification, a majority class baseline classification, and each of the learning journal categories can be seen in Figure 4.

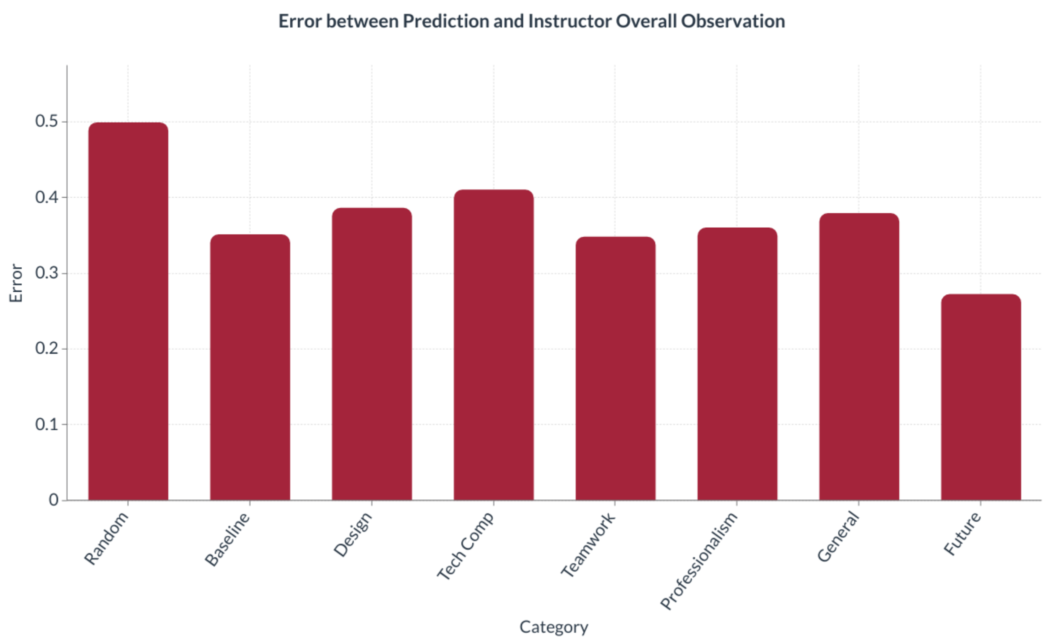


Figure 4 Average error between the classifier score for each category and the instructor observed overall motivation. Low error shows that there was a higher correlation between the classification of that learning journal type and the instructor observation.

These results show that most learning journal categories did not give an overall representation of a students’ motivation. However, when they were writing responses to the “Extending to Future” category, the learning journal classification was more aligned with instructor observations. More data will need to be analyzed to make stronger conclusions, but these results suggest that there is more work to be done to understand "micro-motivation" versus "macro-motivation". A student may be writing about a negative experience or frustrations in a specific learning journal ("micro-motivation"), but that does not necessarily align with having negative motivation throughout the entire engineering program ("macro-motivation"). It is interesting to note the close alignment between students reflecting about their future (an instance of "micro-motivation") and their overall motivation as noted by instructors ("macro-motivation"). More work should be done to see if these measures align in a larger dataset.

Limitations

These preliminary results have shown promise, but the conclusions that can be drawn are still limited. The sample size is small, and questions still remain about how we measure and understand student motivation. For example, how are the ideas of "micro-" and "macro-" motivation related? In other words, what do learning journals (which are single snapshots of motivation in time) tell us about overall motivation? Even more broadly, what does it mean to have motivation as an engineering student? Does motivation equate to perseverance in the program? How does it relate to engineering identity? By combining careful qualitative analysis and the scalable speed of algorithms, we hope to gain a greater understanding of how each of these factors interact.

Conclusion and Future Work

Since the time of the preliminary study, Institutional Review Boards (IRB) approval was sought and granted – allowing for further analysis with a larger dataset. With this dataset, we plan to explore the following broader research questions for the future work (F):

F1. How do the results of a machine learning classifier compare to other measures of student motivation (such as self-report, faculty observations, or externally validated instruments)?

F2. How do students' levels of motivation change during their 3rd and 4th year of college in this particular PBL program?

F3. How are ideas written in learning journals connected and communicated, and how do the learning journals compare to the existing literature about student motivation?

To answer F1, future work will involve collecting a variety of other measures of student motivation including self-report, faculty observations, and other external validated instruments. During the preliminary study, only data from past students was accessible. IRB approval allows for the collection of current student data and other measures of motivation. These results will be compared with the goal of validating the machine learning method and/or identifying and analyzing its limitations.

Similarly, with the approval of the IRB, longitudinal data can be collected to analyze the trajectory of student motivation, answering F2. By performing time series analysis, these results may give further insight about how and why motivation changes during the junior and senior year of college, as well as a better understanding of the static or dynamic characteristics of motivation.

Finally, we propose using epistemic network analysis [32] to further understand how ideas and themes are connected in the learning journals and how these ideas contribute to engineering identity. By coding learning journal responses with characteristics and features presented in motivation models such as Ryan and Deci's Motivation Continuum (in Figure 2 above), epistemic network analysis can be used to identify which of these characteristics and features are most closely related for sets and subsets of students. By using existing literature, a framework can be created that allows for an understanding of student motivation at a deeper level. For example, codes such as "anxiety" and "grades" may suggest controlled motivation whereas codes such as "excitement" and "value" may suggest autonomous motivation.

Acknowledgment

The authors would like to acknowledge the support from the Faculty Research Grant provided by Minnesota State University, Mankato.