Learning Analytics in Outer Space: A Hidden Naïve Bayes Model for Automatic Student Off-task Behavior Detection

Wanli Xing
University of Missouri-Columbia
School of Information Science & Learning Technologies
wxdg5@mail.missouri.edu

Sean Goggins
University of Missouri-Columbia
School of Information Science & Learning Technologies
gogginss@missouri.edu

ABSTRACT
Learning analytics (LA) has invested much effort in the investigation of students' behavior and performance within learning systems. This paper expands the influence of LA to students' behavior outside of learning systems and describes a novel machine learning model which automatically detects students' off-task behavior as students interact with a learning system, ASSISTments, based solely on log file data. We first operationalize social cognitive theory to introduce two new variables, *affect states* and *problem set*, both of which can be automatically derived from the logs, and can be considered to have a major influence on students' behavior. These two variables further work as the feature vector data for a K-means clustering algorithm in order to quantify students' different behavioral characteristics. This quantified variable representing student behavior type expands the feature space and contributes to the improvement of the various model performance compared with only time- and performance-related features. In addition, an advanced Hidden Naïve Bayes (HNB) algorithm is coded for off-task behavior detection and to show the best performance compared with traditional modeling techniques. Implications of the study are then discussed.

Categories and Subject Descriptors
J.1 [**Data Processing**]: Education; K.3.1 [**Computers and Education**]

General Terms
Algorithms, Measurement, Design, Assessment, Human Factors, Theory.

Keywords
Learning Analytics; ITS; Social Cognitive Theory; Off-task Behavior; Hidden Naïve Bayes

1. INTRODUCTION
The field of learning analytics (LA) has invested considerable energy in modeling, understanding and assessing the behavior and performance of students while they use learning systems [17][37][41][42]. Most of these documented studies center specifically on the interaction between students and learning environments. However, students' behavior outside of the system may also influence how well the students learn. One such type of behavior is off-task behavior in which students' attention becomes lost and disengaged from the learning environment and activities. Incorporating off-task behavior detection into the scope of LA can extend the application of LA and enhance the utility of LA in technology-mediated education, such as online courses and MOOCs. Similarly, introducing LA, especially its methodological perspective, to behavioral investigation has the potential to assist teachers to provide timely intervention and guide system designers to develop learning environments that respond to off-task behavior.

Off-task behavior is defined as behavior not directly related to the learning activities in a course, and has been recognized as a problem by both researchers and practitioners for over a century [6][10][3][25][32]. It can take a number of forms: talking with other students without any learning aims, surfing the web, disrupting other students, etc. [18]. Systematic reviews have found a negative impact of off-task behavior on student learning outcomes [15][19]. Therefore, it is important to detect off-task behavior and design an effective way to reduce it.

One possible way to solve the problem is to detect off-task behavior using data from students' interaction with the learning system. Such a detector can be built into the learning system and discretely signal off-task behaviors in real time. In turn, it has the potential to improve students' learning experiences and outcomes. Research on automatic detection of off-task behaviors is limited and some exemplar studies are examined here. Baker [2] applied a latent response model algorithm using only time and performance features, derived from user-system interaction logs by fast correlation-based filtering and forward selection method with a model performance of 0.55 with 10-fold cross-validation. Similarly, Pardos et al. [19] built the selected features from the logs using forward selection and compared different machine learning algorithms such as J48, Naïve Bayes, and K*, and with F-measure reaching 0.69 using 5-fold cross validation.

Time and performance features are often used to detect off-task behaviors, and are usually generated in the interaction log data; however, the majority of previous work focused on low-level time and performance features only and the performance of the off-task detection model is inconsistent and relatively low. This paper aims to develop a more reliable model with improved performance in detecting off-task behaviors. To accomplish this goal, this paper employed social cognitive theory to contextualize data that influence students' behavior, and then applied K-means cluster analysis to quantify the inter-user difference, together with the time and performance features from log data, consisting of the feature space for the model building. Then to address data collinearity and data interdependency problems that often take place in educational contexts, an advanced machine learning algorithm – Hidden Naïve Bayes (HNB) – is coded to develop the detection model. The proposed model is compared with other traditional modeling techniques provided that they 1) only utilize time and performance features; 2) utilize time and performance features as well as the quantified student behavior type variable. It is expected that both the quantified student behavioral variability and HNB can improve the model performance of off-task behavior detection. The paper is organized as follows: Section 2 shows the context of the study and data format. Section 3 describes overall methodology. Sections 4 and 5 present experimental results and analysis. Section 6 discusses results. Section 7 summarizes this study and points out future research directions.

2. RESEARCH CONTEXT & DATA

In this paper, we attempt to detect off-task behavior in the context of the intelligent tutor system ASSISTments. ASSISTments is a web-based platform for tutoring K-12 school mathematics, which provides scaffolding, help and feedback as students complete math problems [21][24]. Question items in ASSISTments are designed to hone students' skills in corresponding state standardized examinations. Figure 1 demonstrates how the system decomposes a problem into steps when a question is answered incorrectly. Hints are provided for each step and students can ask for a bottom-out hint that shows the final answer. Both teachers and the system can assign questions to students. In this current study, a diverse population in race and SES of 229 8th graders are involved and assigned math test prep questions from the system in 2010.

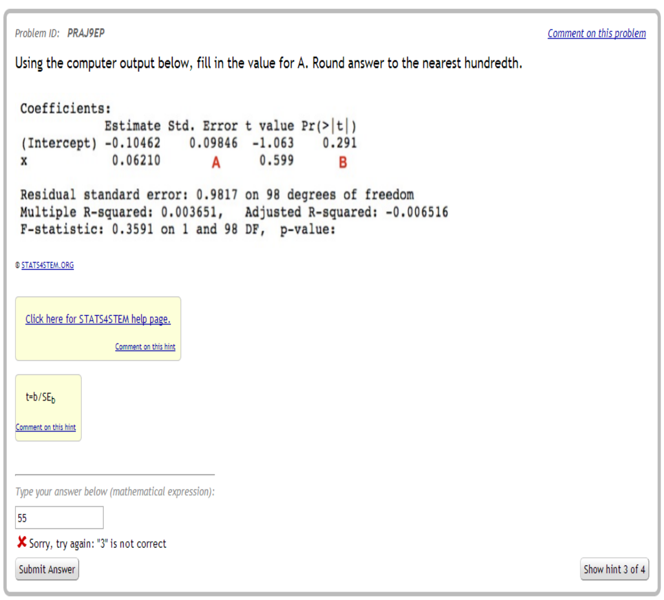

Figure 1. A sample ASSISTment lesson where the student answers incorrectly and asks for the first hint.

In order to build a model to detect off-task behavior in students, two sources of data on student behavior occurring during learning sessions are employed. The first source of data is the interaction log gathered by ASSISTment when the students use the system. It logs time and performance information on each action that a student makes in the environment such as attempts to respond, results, requests for hints, time taken to complete actions etc. In addition, some distilled features are about the past actions such as number of attempts and number of incorrect actions the student had made for this problem. In sum, 40 features are recorded for each action made on ASSISTments. Those seeking specific information about the features can refer to [2][29]. In total, 88,179 data points are collected from 229 students while they use the ASSISTments system.

The second data source used is the ground truth data to represent whether a student is on-task or off-task coded by a pair of expert field observers as students use ASSISTments. While the on and off-task ground truth data was coded, the observers also coded the student affect states from categories of boredom, frustration, engaged concentration, confusion, and any other states, though these affect data can be inferred automatically from the log data [27][29]. The coders applied the Baker-Rodrigo Observation Method Protocol (BROMP) to code affective and behavioral data in the learning settings [28][31][34] using a synchronization software for android handheld computers. Both coders received extensive training and reached Cohen's Kappa of 0.86 for inter-rater reliability and exceeded the Fleiss threshold value 0.75 to refer as "excellent" for field observation [13]. For specific criteria and procedure of judgments on affect state and behavior, refer to Baker et al. [4]. In total, 915 observations were taken for the period of observation across the students with 260 (28.4%) coded as off-task and 655 coded as on-task (71.6%). To aggregate individual student actions and synchronize them with field observations is fairly easy because both the field observation data and log data contain the problem number, student id, time and date and then a series of sum, minimum, maximum and average values of students actions were calculated for the corresponding observation.

3. METHODOLOGY

The first objective of this work is to differentiate students' behavior characteristics for a new feature construction. Cluster analysis grounded in social cognitive theory is then applied to determine types of similar students, so that a quantified difference variable manifesting student behavior type can be added to the feature space. Partitioning students reflecting types of students captures varying levels of strain and influence on students' behavior beyond the log data effect of ASSISTment system.

The ultimate objective is to build a reliable method to detect off-task behaviors using the cluster-generated feature and the available features in the log data and especially to correlate changes in detection rate with the addition of the built feature. Furthermore, in addition to the traditional modeling building techniques, another advanced machine learning algorithm, HNB, is designed to build the detection model due to its power in dealing with collinearity and efficiency for model building. As stated, a binary variable (the ground truth data), generated from the observation data is used to express whether the student is on task or off-task and the proposed algorithm is employed to detect whether the student is off-task as using ASSISTment. Other traditional modeling techniques such as Logistic Regression and Naïve Bayes are also coded to serve as the baseline algorithms to benchmark the performance of proposed

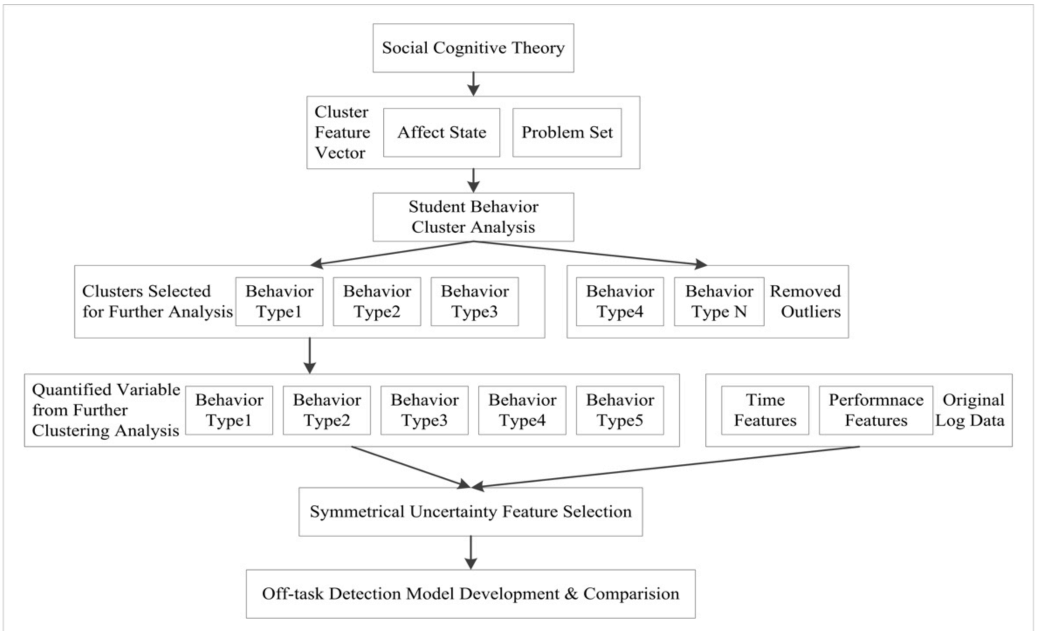

Figure 2. Methodology overview.

HNB model in both the original feature space (log data only) and the constructed feature space (the generated variable and log data). The flowchart in Figure 2 outlines our methodology and how the learning objectives are related.

4. CLUSTERING FOR VARIABLE CONSTRUCTION

Variations in off-task student behaviors are affected by factors beyond those recorded in the interaction logs, such as previous experience, background, cognitive abilities etc. The present research employs a clustering technique that can be applied to evaluate the different types of students. Students with similar variations in cognitive capability, background and other possible variables outside of the logs are identified thereby improving the likelihood that these clustered students have similar behavioral characteristics and are confined by a similar set of constraints. Student types are considered as a quantified variable, and are added into the feature space generated by the log data.

4.1 Social Cognitive Theory and Cluster Vector

Social Cognitive Theory [5] relies on the premise of triadic reciprocal causation, which shows that cognitive factors, environmental factors, and human behavioral factors interact with and influence with each other as shown in Figure 3. Cognitive factors represent personal cognition, emotion, efficacy and biological events. Environmental factors refer to the social and physical environments. Human behavioral factors can include both cognitive and social changes. Since the aim of the current research is to investigate students' behavior, focus is placed on the ways in which cognitive and environmental factors impact individual behavior.

Clustering can be employed to identify groups of similar students through unsupervised classification, subject to specified clustering constraints. These constraints are often referred to as feature vectors. According to social cognitive theory, cognitive factors and environmental factors have direct influence on students' behavioral changes. Therefore, in our context, in searching for the feature vectors, we focused on factors related to cognitive and environmental effects located outside of the interaction log data. In searching for the cognitive related factors, data concerning students' affective state are a good candidate as social cognitive models have been criticized for not considering human affective states sufficiently [22][30], and have supported that affective states can influence the human decision-making process [14]. In addition, affect data is not collected through interaction log data. Introducing the outside factor into the clustering and further detector model can expand the data space and has the potential to improve the explained variance. Affect state data is obtained through field observation, but the automatic detection of affect data have been successfully developed by [27][29]. Therefore, the current study has the potential to be fully automated using log data even without the observed affect data.

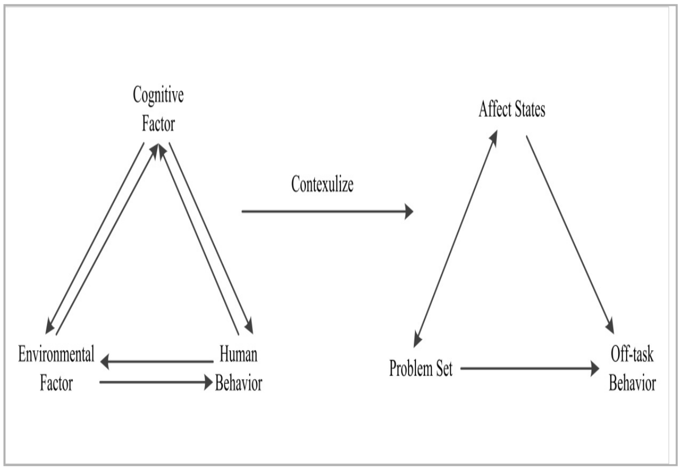

Figure 3. Social cognitive theory contextualization in ASSISTment .

In terms of environmental factors, many studies have been conducted to demonstrate that learning environments influence students' behavior and performance. Wu et al. [35] categorized how the technology mediated environment influence students' behavior within functional environment and social environments. Since students usually work on their own in ASSISTment, social factors have no obvious effects on student behavior. The interaction log data have recorded many possible system functional factors. However, these factors are rather low-level action information, and it is difficult to reflect the overall and accumulated effect of environmental factors on students' behavior. Therefore, a deductive variable – problem set – is constructed from the log data to represent the diversity and variance of problems assigned to the students by either teachers or systems [34]. The problem set is defined as the unique number of problems students encountered during the observation. As a result, each student is represented by a six-dimensional vector, affect states and problem set, for clustering analysis.

4.2 Two-Stage K-means Cluster Analysis

In this study, students were classified into several categories by K-means clustering analysis using the feature vector constructed in the above section. K-means clustering [20][36][37][38] has been widely used to separate data into groups that are relatively homogeneous within themselves and heterogeneous among each other on the basis of the defined cluster vector data (variables). This clustering process is sensitive to the scale between different variables [40][41]. Therefore, each cluster variable is standardized to prevent large magnitude from disproportionately influencing the centroid calculations. The K-means algorithm requires (k) number of clustering centers to be pre-specified and iterates case allocation to the nearest center point until the cluster centers no longer change [12]. Therefore, this work completed two steps of cluster analysis: to remove outliers, and to select intuitive cluster groups for K (number).

The first stage of clustering is assessed over a range of possible K. When many cluster partitions are chosen (K>15), the minimum membership is very small and may indicate the existence of extreme outlier students. The cluster solutions for K = 16, 17, 18, 19, 20 selected most of the same outliers in their respective low membership clusters, suggesting that the selection of outliers is robust for high values of K. Therefore, students with highest frequency appearing in those low membership clusters under the different conditions of K were removed from the dataset. The second stage K-means analysis is performed with the standardized data with the outliers removed. There is no definitive number of types of students, so multiple k is chosen with results for selecting the best K value shown in Table 1. The Pseudo F measure [7] is a reliable method for interpreting the optimal number of clusters. In terms of the pseudo value, the optimal number of clusters is found at the elbow of the pseudo curve, where the gap between the pseudo value for K and for K-1 is relatively large [7]. The decreasing pseudo value has a local maximum in the elbow of the curve when K = 5 (Table 1). As a result, each student has an assigned quantified variable 1 – 5 reflecting their behavioral type.

Table 1.Stage 2 Cluster analysis for K selection.

K Number	Pseudo F Value	Minimum Membership	Maximum Membership
3	26.2956	17	132
4	161.6596	9	103
5	17.7084	17	87
6	55.0533	9	76
7	108.4264	12	60
8	69.0984	9	48
9	63.334	9	47
10	47.9475	5	36
16	78.4309	3	31
17	71.8805	2	41
18	14.0211	3	41
19	38.6701	1	71
20	11.9383	1	38

5. HNB MODEL

5.1 Feature Selection

There are 40 features in the interaction log data with one additional feature constructed from the cluster analysis. However, not all of these features have influence on students' off-task behavior, and some may have overlapping influence on off-task behavior. Therefore, before building the model; feature selection is conducted on the dataset to remove redundant and/or irrelevant attributes. Specifically, the symmetrical uncertainty (SU) method, derived from information theory and numerical recipes, is applied as a fast correlation measure to evaluate the relevance of individual features and to put the most relevant features at the beginning of the list [1]. The symmetric nature of this method reduces the number of comparisons, that is $\mathrm{SU}(i,j)$ is the same with $\mathrm{SU}(j,i)$, where i and j are different features. In addition, SU is not affected by multivalued attributes as is the case of information gain [1], and the values are normalized. Symmetrical uncertainty (SU) is calculated as:

$$\mathrm{SU}(X,Y) = 2 \times \frac{\mathrm{InfoGain}(X \mid \mathrm{Y})}{\mathrm{Ent}(X) + \mathrm{Ent}(\mathrm{Y})}$$

Where $\mathrm{InfoGain}(X \mid \mathrm{Y})$ is the information gain of variable X as an independent attribute and Y is the class attribute. $\mathrm{Ent}(X)$ and $\mathrm{Ent}(\mathrm{Y})$ are the respective entropy of features X and Y. The value of SU ranges from 0 to 1. An SU value of 0 indicates that the feature X has no relation with the off-task behavior and Value 1 indicates the variable X can completely predict off-task behavior. After performing 10-fold cross validation, 25 features are chosen. Other sample features commonly used to detect off-task behaviors and are related to time and performance, e.g. time taken to complete the problem, percentage of correct answers so far, number of hints used, total number of attempts made, whether the bottom-out hint was requested, etc. Specifically, the 25 features with manually assigned on task or off-task labels for each student become the dataset and input for the various machine learning models, where the 25 features are independent variables/features and are used to predict the dependent variable/class, whether the student is on-task or off-task.

5.2 Model Development

This section describes how we use chosen features (independent variables) to detect the off-task behavior (dependent variable). A Bayesian model is an attractive option in educational domains, where an element of uncertainty is always involved. General Bayesian networks are usually too complex for high dimensional data to learn, which may easily be a NP-hard problem [23]. Therefore, the Naïve Bayes model is often a preferable option by virtue of its simplicity. In this study, an instance E of whether a student is off-task or not is represented by a vector $\{x_1, x_1, x_1, ..., x_n\}$, $n = 25$, and x_i is the value of attribute X_i. Let C denote the binary class on-task or off-task and c as the value of C. As a Naïve Bayes model, it is assumed that all attributes are independent of each other [26]. Then according to Bayesian theorem,

$$p(E \mid c) = p(x_1, x_2, x_3, ..., x_n) = \prod_{i=1}^{n} p(x_i \mid c)$$

So the resulting classifier is a Naïve Bayes classifier:

$$\mathrm{c(E)} = \arg\max_{c \in C} \left(p(c) \prod_{i=1}^{n} p(x_i \mid c) \right)$$

This is the basic reasoning behind the Naïve Bayes model. Compared to a Bayesian network, this model only has two layers, the class variable c in the root node, and all the other variables X_i in the leaf nodes, as shown in Figure 4. The Naïve Bayes model assumes that all leaf nodes are conditionally independent, given the class value. This assumption is rarely true in an educational context, where all variables have some dependency on each other. This is especially the case in the current study in which many features are relative features or constructed features from other more basic features.

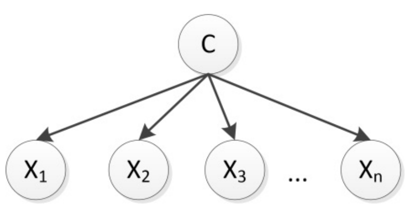

Figure 4. Naïve Bayes Model

We aim to relax the assumption on Naïve Bayes while maintaining a desirable level of simplicity such as TAN [16]. The TAN model expands the structure of Naïve Bayes by adding additional dependencies among nodes (features). However, only one parent is allowed per node in TAN, although several attributes may have similar influences on the model. For these reasons we propose to use Hidden Naïve Bayes (HNB) model [43] in this work. The idea is to build a hidden parent for each attribute that combines the influences from all other attributes (Figure 5). This not only avoids the intractable computational complexity of Naïve Bayes or TAN, but also takes influences from all attributes into account.

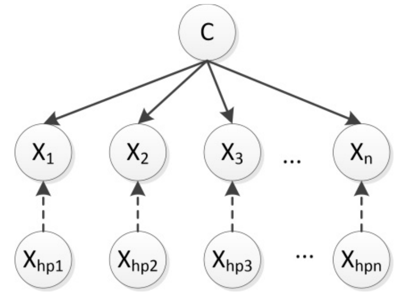

Figure 5. HNB

In the HNB model, each attribute X_i has a hidden parent X_{hpi}, where $i \subset [1, \dots, n]$, representing the weighted influence from all other attributes as shown with the dashed circles in Figure 5. Therefore, similar to Naïve Bayes based on the Bayesian theorem, the HNB model can be defined as follows:

$$\mathrm{c(E)} = \arg\max_{c \in C} \left(p(c) \prod_{i=1}^{n} p(x_i \mid x_{hpi}, c) \right)$$

$$\text{Where } p(x_i \mid x_{hpi}, c) = p(\mathrm{c}) \sum_{j=1, j \neq i}^{n} W_{ij} * p(x_i \mid x_j, c).$$

As shown in the formula, the hidden parent X_{hpi} is basically a mixture of the weighted influences from all other features. There are multiple approaches for determining the weights W_{ij} such as an expectation maximization algorithm (EM) or conditional mutual information. For this work, mutual information between attributes X_i and X_j is applied. Naïve Bayes and Logistic Regression are also coded for comparison with the proposed HNB model.

5.3 Model Performance

This section presents the performance of the Naïve Bayes, Logistic regression, and HNB models in detecting students' off-task behavior. We examine simple models including time and performance measures, as well as more expansive feature spaces, which include a cluster-generated measure. All models are conducted with 10-fold cross validation with specific results shown in Table 2 and built from the electronic trace data alone.

As we can see from Table 2, the performance of all models is improved when the cluster-based variable is added to the feature space for precision, recall and F-measure or using the HNB model. The F-measure is a more comprehensive evaluation criteria for model performance; therefore, the explanation for the results will focus on F-measure. The F-measure increases from 2.9% (Logistic Regression) to 20.4% (Naïve Bayes) when the cluster-generated variable is incorporated in the feature space. On the hand, HNB has the best performance both with the original feature sets and with the new constructed feature sets. Specifically, on the interaction only log feature sets, the improvement of HNB ranges from 3.5% to 23.9% in comparison with Logistic Regression and Naïve Bayes respectively. The enhancement of HNB on F-measure for the new constructed feature space ranges from 3.9% to 6.8% in comparison with Logistic Regression and Naïve Bayes algorithms. The performance of HNB as whole reaches 80.2% for precision, 81.4% for recall and 79.6% for F-measure on student off-task behavior detection and is much higher than the previous modeling techniques.

Table 2 Results of comparison of the model performance in different feature spaces.

	Time and Performance Features			Time, Performance and Cluster Generated Features		
	Pre.	Rec.	F	Pre.	Rec.	F
NB	0.72	0.501	0.524	0.747	0.717	0.728
Logit	0.743	0.772	0.728	0.754	0.776	0.757
HNB	0.763	0.784	0.763	0.802	0.814	0.796

6. DISCUSSION

Learning analytics has historically focused primarily on students' behavior and performance within the interactive learning systems [44][45]. The proposed study aims to expand the scope of LA into outer space – to detect the behavior outside of the learning systems. Specifically, students' off-task behavior as they use learning systems has been found to closely relate with students' learning and is very difficult to monitor and respond to for a teacher in a classroom full of students [2][7]. This work accordingly attempts to build a model to automatically detect students' off-task behavior while they are interacting with the system. More importantly, the built model only used the data available in the log files or data possible automatically collected (the cluster-based variable); therefore, the model does not require any sophisticated equipment (e.g. cameras, eye-trackers) that are inaccessible to most schools for student off-task behavior detection.

Compared with other preliminary studies on detecting off-task behavior [2][29], this study introduces the constructed variable (cluster-based feature) to the modeling space for the detector instead of variables solely based on the time and performance related attributes collected by the ASSISTment log files. This cluster-based feature relies on the premise that students may have different types of behaviors and then the proposed method automatically quantifies the difference for us from students' affect and environmental influence factors. The proposed method increases the explained variance for the students and the resulting measures such as precision, recall, and F-measure all show enhancements after incorporating the new cluster-based variable.

Furthermore, previous work has concentrated on methodological and algorithmic exploration, often overlooking the educational context [2][8][29]. Their processing of data is purely mathematical rather than accounting for student behaviors. The current work aims to obtain deep understanding of students' behavior by looking from the lens of social cognitive theory, investigating how to use behavior-oriented theory to inform measure selection and construction. This theory-informed variable is then connected with advanced computational algorithms and showed an improved result. Analytics will be more powerful if the data is structured for interpretation using theory because theory is understandable by users interpreting the resulting analytics , allowing learning analytics system designers to think and reflect on the fundamental factors influencing student behavior. Of course, social cognitive theory is not the only available theory [41]. We choose it because this theory fits our context well. Other research contexts can totally consider other theories for their research. For example, in CSCL research, researchers may choose activity theory or group cognition theory to connect with their algorithms.

In addition to the contribution to feature space and linking theory with computation, this work went beyond traditional modeling techniques to employ HNB for off-task behavior detection. Bayesian algorithms have always appealed to educational researchers in general, especially considering their capabilities in dealing with uncertainties and their stable modeling performance. However, there are two common difficulties encountered in the educational field when applying the traditional Bayesian techniques: first the number of variables affecting students' performance and behavior can become large especially in technology-mediated environment, resulting in high computational costs; second, the collinearity and codependency among those variables often breach the assumption of the Bayesian algorithms and compromise the their performance [43]. The proposed HNB is especially good at dealing with these two disadvantages of traditional Bayesian algorithms by limiting its usage of computational power and building a hidden parent to compound the influence from other attributes for each attribute (variable). The result of this experiment also demonstrates the optimal modeling performance for off-task behavior detection compared with other machine learning algorithms. This is not necessarily a claim that HNB has the best performance in all situations, but rather to suggest HNB as a starting point for researchers pursuing similar goals in educational and learning analytics research.

Now that we have a reliable system to detect students' off-task behavior, the next concern is how we respond to it. After all, the ultimate goal is to reduce such behavior, increasing students' learning and performance. One possible way to do this is to aggregate the percent of time students are found to be off-task during the session and inform the teachers so that they can talk with the students and provide appropriate intervention. However, this approach has heavy time-lag between students' off-task behavior and intervention. Another possible way is to design a pedagogical agent and integrate it into the learning system to alert the students' when they go off-task. This must be carefully done because responding to off-task immediately and in a heavy-handed way may have counterproductive effect such as irritate the student, and irritating them further when the detection model inaccurately diagnoses them as off-task [2]. It might be more appropriate to deal with students' off-task behavior with a long-term view and less severely, especially considering that off-task behaviors tend to associate with poorer learning at an aggregate level [9]. A constructive method to address off-task behavior is through self-monitoring which has demonstrated effectiveness in reducing students' off-task behavior in traditional classroom [11]. Learning analytics designers can further investigate how to incorporate self-monitoring mechanism into the learning systems when the student is detected as off-task.

7. CONCLUSION

This paper describes a novel machine learning model to automatically detect students' off-task behavior as students are interacting with a learning system based solely on the automatically gathered data. We operationalize social cognitive theory to introduce computationally represented states and indicators that enhance K-means clustering techniques to quantify students' different behavioral characteristics. This cluster-based variable expands the feature space and contributes to the improvement of the model performance as compared with models using solely time and performance related features. We also introduce an advanced HNB algorithm to detect off-task behavior and show the best performance compared with traditional modeling techniques and explore the approaches to reduce off-task behavior. Future studies can explore other possible variables and test its effects on students' off-task detection and its practicality of incorporating the variable into the whole automatic detection system. It is also valuable for future research to design interventions based on the proposed method to reduce student's off-task behavior and investigate how to integrate the interventions into the learning system.

ACKNOWLEDGMENTS

Withhold for Review

REFERENCES

[1] Ali S. I. and Shahzad W. A feature subset selection method based on symmetric uncertainty and ant colony optimization. *International Conference on Emerging Technologies (ICET)*, 2012. IEEE, 2012: 1-6.

[2] Baker R. S. J. Modeling and understanding students' off-task behavior in intelligent tutoring system. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2007: 1059-1068.

[3] Baker, R. S. J., Corbett, A. T., Roll, I., Wagner, A. Z., and Koedinger, K.R. The Relationship Between Gaming the System and Learning in Cognitive Tutor Classrooms. Manuscript Under Review.

[4] Baker R. S. J., D'Mello S. K., and Rodrigo M. M. T., et al. Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. In *International Journal of Human-Computer Studies*, 2010, 68(4): 223-241.

[5] Bandura A. Social foundations of thought and action. Prentice Hall.: Englewood Cliffs, NJ, 1986.

[6] Berliner, D. C. What's All the Fuss About Instructional Time? In *The Nature of Time in Schools: Theoretical Concepts, Practitioner Perceptions*. New York: College Press. 1990.

[7] Caliński T. and Harabasz J. A dendrite method for cluster analysis[J]. In *Communications in Statistics-theory and Methods*, 1974, 3(1): 1-27.

[8] Cetintas S., Si L., Xin Y. P., et al. Automatic detection of off-task behaviors in intelligent tutoring systems with machine learning techniques. In *Learning Technologies*, IEEE Transactions on, 2010, 3(3): 228-236.

[9] Cocea M., Hershkovitz A., and Baker R. S. J. The impact of off-task and gaming behaviors on learning: immediate or aggregate?. 2009.

[10] Currie, J. The principles and practice of common school education. Cincinnati, OH: R. Clarke, 1884.

[11] Dalton, T., Martella, R. C., and Marchand-Martella, N. E. (1999). The effects of a self-management program in reducing off-task behavior. In *Journal of Behavioral Education*, *9*(3-4), 157-176.

[12]Duda R. O., Hart P. E., and Stork D. G. Pattern classification. John Wiley & Sons, 1999.

[13] Fleiss, J. L. Statistical Methods for Rates and Proportions (2nd Edition), John Wiley and Sons, New York, 1981.

[14] Forgas J. P. Affect and cognition. In *Perspectives on Psychological Science*, 2008, 3(2): 94-101.

[15] Fredrick W. C. and Walberg H. J. Learning as a function of time. In *The Journal of Educational Research*, 1980: 183-194.

[16] Friedman N., Geiger D., and Goldszmidt M. Bayesian network classifiers. In *Machine learning*, 1997, 29(2-3): 131-163.

[17] Gašević D., Mirriahi N., and Dawson S. Analytics of the effects of video use and instruction to support reflective learning. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*. ACM, 2014.

[18] Godwin K. E., Almeda M. V., and Petroccia M., et al. Classroom activities and off-task behavior in elementary school children. Cognitive Science Society, 2013.

[19] Goodman L. Time and learning in the special education classroom. SUNY Press, 1990.

[20] Guo R., Zhang Y. Identifying Time-of-Day Breakpoints Based on Nonintrusive Data Collection Platforms. In *Journal of Intelligent Transportation Systems*, 2014, 18(2): 164-174.

[21] Heffernan N. T., Turner T. E., and Lourenco A. L. N., et al. The ASSISTment Builder: Towards an Analysis of Cost Effectiveness of ITS Creation. FLAIRS Conference. 2006: 515-520.

[22] Keer M., Putte B., and Neijens P. The role of affect and cognition in health decision making. In *British Journal of Social Psychology*, 2010, 49(1): 143-153.

[23] Koc L., Mazzuchi T. A., and Sarkani S. A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier. *Expert Systems with Applications*, 2012, 39(18): 13492-13500.

[24] Koedinger, K. and Corbett, A. Cognitive tutors: Technology bringing learning sciences to the classroom. In K. Sawyer (Eds.), Cambridge handbook of the learning sciences, 61 – 78. New York: Cambridge University Press, 2006.

[25] Lemov D. Teach Like a Champion: 49 Techniques that Put Students on the Path to College (K-12). John Wiley & Sons, 2010.

[26] McCallum A. and Nigam K. A comparison of event models for naive bayes text classification. AAAI-98 workshop on learning for text categorization. 1998, 752: 41-48.

[27] Ocumpaugh J., Baker R., and Gowda S., et al. Population validity for educational data mining models: A case study in affect detection. In *British Journal of Educational Technology*, 2014, 45(3): 487-501.

[28] Ocumpaugh, J., Baker, R. S. J. D., and Rodrigo, M. M. T. Baker-Rodrigo Observation Method Protocol (BROMP) 1.0. Training Manual version 1.0. Technical Report. Manila, Philippines: Ateneo Laboratory for the Learning Sciences, 2012.

[29] Pardos Z. A., Baker R. S. J. D., and San Pedro M. O. C. Z., et al. Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. In *Proceedings of the Third International*

Conference on Learning Analytics and Knowledge. ACM, 2013: 117-124.

[30] Pligt J. and De Vries N. K. Expectancy-value models of health behaviour: The role of salience and anticipated affect. In *Psychology and Health*, 1998, 13(2): 289-305.

[31] Roberge D., Rojas A., and Baker R. Does the length of time off-task matter? In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. ACM, 2012: 234-237.

[32] Rodrigo M. M. T., Baker R. S. J. D., and Rossi L., et al. Student Off-Task Behavior in Computer-Based Learning in the Philippines: Comparison to Prior Research in the USA. In *Teachers College Record*, 2013, 115(10).

[34] Rowe J. P., McQuiggan S. W., Robison J. L., et al. Off-Task Behavior in Narrative-Centered Learning Environments. AIED. 2009: 99-106.

[34] Woodburn A., Ryerson M. Airport Capacity Enhancement and Flight Predictability. In *Transportation Research Record: Journal of the Transportation Research Board*, 2014, 2400(1): 87-97.

[35] Wu J. H., Tennyson R. D., Hsia T. L. A study of student satisfaction in a blended e-learning system environment. In *Computers & Education*, 2010, 55(1): 155-164.

[36] Xing W., Guo R., Lowrance N., et al. Decision Support Based on Time-Series Analytics: A Cluster Methodology. Human Interface and the Management of Information. Information and Knowledge in Applications and Services. Springer International Publishing, 2014: 217-225.

[37] Xing W., Wadholm B., Goggins S. Learning analytics in CSCL with a focus on assessment: an exploratory study of activity theory-informed cluster analysis. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*. ACM, 2014: 59-67.

[38] Xing W., Wadholm B., Goggins S. Group Learning Assessment: Developoing a Theory-Informed Analytics. Educational Technology & Society. In Press, 2014.

[39] Xing, W., Guo, R., Fitzgerald, G., & Xu, C. Google Analytics based Temporal-Geospatial Analysis for Web Management: A Case Study of a K-12 Online Resource Website. International Journal of Information Science and Management (IJISM), 2014, 13(1).

[40] Guo, R., & Zhang, Y. Identifying Time-of-Day Breakpoints Based on Nonintrusive Data Collection Platforms. In *Journal of Intelligent Transportation Systems*, (2014), 18(2): 164-174.

[41] Xing W., Guo, R., Petakovic E., Goggins, S., Participation-based Student Final Performance Prediction Model through Interpretable Genetic Programming: Integrating Learning Analytics, Educational Data Mining and Theory. In *Computers in Human Behavior*. In Press, 2014.

[42] Goggins, S., Xing W., Chen, X, Chen, B., Wadholm, B. Learning Analytics at “Small” Scale: Exploring a Complexity-Grounded Model for Assessment Automation. In *Journal of Universal Computer Scien*ce, In Press, 2014.

[43] Jiang L., Zhang H., Cai Z. A novel Bayes model: Hidden naive Bayes. In *Knowledge and Data Engineering*, IEEE Transactions on, 2009, 21(10): 1361-1371.

[44] Xing, W., Wu, Y., Ma, X. Ontology construction and mathematical modeling for the LAS Engines. In *Journal of East China Normal University (Natural Science)*, 2014 no. 6: 126-139.

[45] Xing, W. L. Wadholm, B. & Goggins, S. Assessment Analytics in CSCL: Activity Theory based Method. In *Proceedings of the International Conference on Learning Sciences'14*. 2014: 1535-1536.