Enriching the Student Model in an Intelligent Tutoring System

Ramkumar Rajendran

Supervisors
Sridhar Iyer    Campbell Wilson
Sahana Murthy    Judithe Sheard

IITB-Monash Research Academy, IIT Bombay, Monash University

Aug 22, 2014
Outline

1 Introduction
   - Intelligent Tutoring System
   - Affect Recognition

2 Related Work
   - Predicting Affective States
   - Addressing Affective States

3 Theory-Driven Approach

4 Predicting Frustration using Mindspark Log Data
   - Human Observation
   - Results
   - Discussion

5 Addressing Frustration
   - Strategies to Address Frustration
   - Algorithm
   - Data Collection
   - Results

6 Generalizing Theory-Driven Approach
   - Applying Theory-Driven Approach to Model Boredom
   - Data Collection
   - Results
Objective

To create a model to detect and respond to affective states of the students when they interact with an Intelligent Tutoring System (ITS).
Intelligent Tutoring System (ITS)

ITS dynamically adapts the learning content based on learner’s needs and preferences.
Affective components in Student Model

- The learning process involves both cognitive and affective processes and the consideration of affective processes has been shown to achieve higher learning outcomes [29].

- The importance of the students’ motivation and the affective component in learning has led adaptive systems such as ITS to include learners’ affective states in their student models.

- Affective states used in affective computing research: Frustration, Boredom, Confusion, Engaged Concentration, Delight, and Surprise.
Methodology

**Phase I**
- Definition of frustration
- Operationalize for ITS

**Phase II**
- Log data
- Model to predict frustration

**Phase III**
- User Interface (System)
- Messages to handle frustration
- Motivation Theory

**Phase IV**
- If frustrated
- Reasons for frustration

**Student**

Motivation Theory

Log data  Model to predict frustration

Reasons for frustration.

If frustrated
To include affective states in the student model, students’ affective states should be identified and responded to, while they interact with the ITS.

In affective computing, detecting affective states is a challenging, key problem as it involves emotions—which cannot be directly measured; it is the focus of several current research efforts [32], [9].
Affect Recognition

In order to respond to students’ affective states, the following methodologies are employed to identify affective states of students while they interact with ITS.

1. Human observation [18], [47], [4]
2. Learner’s self reported data [5], [6]
3. Using sensing devices such as physiological sensors [7], [8], [83], [84]
4. Face-based emotion recognition systems [29], [102], [79], [80], [81], [82]
5. Mining the data from the student log [30], [31], [27], [46]
6. Modeling affective states [6], [10]
Affect Recognition

- Identifying affective states using the sensor signals is possible in laboratory settings, but difficult to implement at a large scale. Also, the physiological sensors are intrusive to the users.

- Facial analysis methods use a web-cam to analyze the facial expressions of the users. In the real-world scenario, keeping the camera in the right position, and expecting users to face the camera all the time is not feasible.

- Voice and text analysis methods can only be used in the ITS that considers voice and subjective answers as an input from the users.
Our Context

**System:** Mindspark, a commercial ITS implemented in large scale.

**Affective State:** Frustration.

**Method:** Modeling the data from student log.
- A commercial mathematics ITS developed by Educational Initiatives India (EI-India)
- Incorporated into the school curriculum for different age groups (grade 3 to 8) of students [21].
- Mindspark is currently being implemented in more than hundred schools and being used by 80,000 students across India.
- Mindspark adaptation logic is based on student’s response to the question, question’s difficulty level and student’s education background.
- Sparkies are the reward points to motivate the students.
### Ref Number | ITS/Game used                        | Features used                                  | Method of selecting the feature | Detection Accuracy | Classifiers used                          
---|-------------------------------------|-----------------------------------------------|---------------------------------|---------------------|-----------------------------------------
[30] | AutoTutor                           | Data from students’ interaction               | Correlation analysis            | 78%                 | 17 classifier like NB, DT from Weka[50]  
[46] | Crystal Island                      | Data from students’ interaction and Physiological sensors | All features                   | 88.8%               | NB, SVM, DT                              
[31] | Introductory Programming Course Lab | Data from students’ interaction               | Correlation analysis            | Regression coefficient $r=0.3168$ | Linear regression model            
[10] | Crystal Island                      | Students’ learning pattern and data from questionnaires | All features                   | 28%                 | DBN                                      
[6]  | Prime Climb                         | Students’ learning pattern and data from questionnaires | All features                   | For joy = 69% and for distress $= 70%$ | DDN                                      

*NB- Nave Bayes, SVM- Support Vector Machine, DT - Decision Tree, DBN - Dynamic Bayesian Network, DDN - Dynamic Decision Network, $= this system was not detecting frustration*
Related Work - Predicting Affective States

- Crystal Island [10], and Prime Climb [6] creates a Dynamic Bayesian Network (DBN) model to capture the users’ affective states.
- The users’ affective states are predicted by applying the theory.

The reason identified by the system helps to respond to user’s affective state based on the reasons for it.

Discussion

- Accuracy in data-mining approaches is in the range of 77% to 88%.
- Accuracy for emotions reported by using DBN and DDN model is comparatively less, 28% to 70%.
- Affective state modeling captures not only the affective states but also why the user is in that state.
## Related Work - Addressing Affective States

### Table: Related Research Works to Respond to Student’s Affective States along with the Theories used, Experiment Method and Results

<table>
<thead>
<tr>
<th>Ref Number</th>
<th>ITS/Game used</th>
<th>Theory used to respond to frustration</th>
<th>Experiment Method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>[52]</td>
<td>Affect-Support computer game</td>
<td>Active listening, emotional feedback, sympathy statement [181]</td>
<td>Factorial study, 2 (level of frustration) x 3 (interactive design), N = 71. Self reporting using questionnaire</td>
<td>On an average the affect support group played more minutes compared to non-affect support group.</td>
</tr>
<tr>
<td>[4]</td>
<td>Scooter the Tutor</td>
<td>Agents were given emotions</td>
<td>Control-experiments group study. N = 60. Human observation</td>
<td>Reduction in frustration instances. There is no significant difference in observed affect between control and experimental group.</td>
</tr>
<tr>
<td>[19]</td>
<td>Wayang Outpost</td>
<td>Agent to reflect student’s affective states and messages based on Dweck’s messages [78], [77]</td>
<td>N = 34, physiological sensor data to detect affective states</td>
<td>Initial studies results that students change their behavior based on digital interventions</td>
</tr>
</tbody>
</table>

\( N = \text{Number of participants} \)
The theory-driven approach to detect affective states is given below:

1. Operationalize the theoretical definition of affective state for the system under consideration.
2. Construct features from the system’s log data; based on the theoretical definition of affective state.
3. Create a model using the constructed features to detect the affective state.
4. Conduct an independent method to detect affective state and use the data from independent method to train the weights of model.
5. Validate the performance of the model by detecting the affective state in the test data and compare the results with the data from independent method.
Definitions of Frustration

The following factors of frustration are considered in our research to model the student’s frustration.

- Frustration is the blocking of a behavior directed towards a goal [25].
- The distance to the goal is a factor that influences frustration [88].
- Frustration is cumulative in nature [146].
- Time spent to achieve the goal is a factor that influences frustration [55].
- Frustration is considered as a negative emotion, because it interferes with a student’s desire to attain a goal [88], [146].
1. Define Frustration: An emotion caused by interference preventing/blocking one from achieving the goal

2. Identify the students’ goals while they interact with the system (goal1, goal2,...,goaln)

3. List the blocking factors of each identified goal (goal1_bf, goal2_bf, ..., goaln_bf). Operationalize it for the system using log data

4. Create a linear regression model for frustration index (Fi) with the blocking factors identified

5. Learn the weights of the linear regression model using labeled human observation data

6. Validate the performance of model with test data and compare the results with labeled human observation data
We formulate a linear function $F_i$, as the frustration index at $i^{th}$ question based on the blocking behaviour of student’s goals.

Linear regression formulation of frustration

$$F_i = \alpha [w_0 + w_1 \times \text{goal1.bf} + w_2 \times \text{goal2.bf} + \ldots + w_n \times \text{goaln.bf} + w_{n+1} \times t_i] + (1 - \alpha) [F_{i-1}]$$

$W_0, W_1, \ldots W_n$ are weights, will be determined during training.

$\propto$ is to accommodate the cumulative nature of frustration.

$t_i$ is the response time at $i^{th}$ question.
Human Observation & Data Collection

- Independent method to identify the student’s frustration while they interact with Mindspark

**Figure:** Facial Action Coding System (FACS) [62]
Human Observation & Data Collection

- Students’ facial expressions during the interaction with Mindspark is recorded using a web camera.
- The student’s interaction with Mindspark is recorded using Camstudio\(^1\), open source free streaming video software.
- 932 facial expression form the 27 student’s interaction video.
- Based on guidelines given in [48] and [47] the student’s facial expressions such as outer brow raise, inner brow raise, pulling at her hair, statements like “what”, “this is annoying”, and so on are considered as frustration.
- 80% of time observers agree to other observers facial expression coding and Cohen’s \(\kappa\) was found to be 0.74, a substantial agreement.

We recorded 932 observations from 27 students. Among those, 137 observations were classified as frustration (Frus) and remaining as non-frustration (Non-Frus).

\(^1\)www.camstudio.org
# Metrics

<table>
<thead>
<tr>
<th>Model</th>
<th>Human Observation</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Frustrated</strong></td>
<td><strong>Non-Frustrated</strong></td>
</tr>
<tr>
<td>Frustrated</td>
<td>True Positive (TP)</td>
<td>False Positive (FP)</td>
</tr>
<tr>
<td>Non-Frustrated</td>
<td>False Negative (FN)</td>
<td>True Negative (TN)</td>
</tr>
</tbody>
</table>

**Precision** = \( \frac{TP}{TP + FP} \)

**Recall** = \( \frac{TP}{TP + FN} \)

**Accuracy** = \( \frac{TP + TN}{TP + FP + FN + TN} \)

F1 score and Cohen’s kappa are measured to check the performance of our model compared to random guess.
### Table: Student Goals and Blocking Factors for Mindspark

<table>
<thead>
<tr>
<th>Student Goal</th>
<th>Blocking factor</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>goal</em>1: To get the correct answer to the current question</td>
<td><em>goal</em>1.bf: Answer to the current question is wrong</td>
</tr>
<tr>
<td><em>goal</em>2: To get a Sparkie (answer three consecutive questions correctly)</td>
<td><em>goal</em>2a.bf: Answers to two previous questions are correct and to the current question is wrong</td>
</tr>
<tr>
<td></td>
<td><em>goal</em>2b.bf: Answer to the previous question is correct and to the current question is wrong</td>
</tr>
<tr>
<td><em>goal</em>3: To reach the Challenge Question (answer five consecutive question correctly)</td>
<td><em>goal</em>3a.bf: Answers to four previous questions are correct and to the current question is wrong</td>
</tr>
<tr>
<td></td>
<td><em>goal</em>3b.bf: Answers to three previous questions are correct and to the current question is wrong</td>
</tr>
<tr>
<td><em>goal</em>4: To get the correct answer to the Challenge Question</td>
<td><em>goal</em>4.bf: Answer to the Challenge Question is wrong</td>
</tr>
</tbody>
</table>
Frustration Model for Mindspark Log Data

\[ F_i = \alpha[w_0 + w_1 \times goal1.bf + w_2 \times goal2.bf + w_3 \times goal3.bf + w_4 \times goal4.bf + w_5 \times t_i] + (1 - \alpha)[F_{i-1}] \]
Solving Linear Regression Model

Human Observation, $B_i$ at the $i^{th}$ instance, $B_i = 0$ for non-frustration and $B_i = 1$ for frustration.

Predicted frustration $P_i$, $P_i = 0$ if $F_i < 0.5$ and $P_i = 1$ if $F_i > 0.5$, 0.5 - threshold.

Our Goal:

$$\min(P_i - B_i)^2$$

by varying $w_0, w_1, w_2, w_3, w_4, w_5$

GNU Octave\(^2\) is used to solve the above optimization problem. We used gradient decent algorithm with step size = 0.001.

\(^2\)http://www.gnu.org/software/octave/
Results

Table: Contingency Table

<table>
<thead>
<tr>
<th>Pred Result</th>
<th>Frustrated</th>
<th>Non-Frustrated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frustrated</td>
<td>45</td>
<td>12</td>
</tr>
<tr>
<td>Non-Frustrated</td>
<td>92</td>
<td>783</td>
</tr>
</tbody>
</table>

Table: Performance of our Approach

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>88.84%</td>
</tr>
<tr>
<td>Precision</td>
<td>78.94%</td>
</tr>
<tr>
<td>Recall</td>
<td>32.85%</td>
</tr>
<tr>
<td>Cohen’s kappa</td>
<td>0.41</td>
</tr>
<tr>
<td>F1 Score</td>
<td>0.46</td>
</tr>
</tbody>
</table>
### Performance of Related Data-Mining Approaches Applied to the Data from Mindspark Log File

<table>
<thead>
<tr>
<th>System</th>
<th>Classifiers</th>
<th>Accuracy in %</th>
<th>Precision in %</th>
<th>Recall in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>AutoTutor</td>
<td>Logistic Model Tree</td>
<td>88.63</td>
<td>65.97</td>
<td>46.71</td>
</tr>
<tr>
<td>Crystal Island</td>
<td>Decision Tree</td>
<td>86.05</td>
<td>52.63</td>
<td>51.09</td>
</tr>
<tr>
<td>Programming lab</td>
<td>Linear regression</td>
<td>$r = 0.583$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our Approach</td>
<td>Linear Regression</td>
<td>88.84</td>
<td>78.94</td>
<td><strong>32.85</strong></td>
</tr>
</tbody>
</table>

Our approach performed comparatively better than other approaches in precision of 79.31%
## Performance of Theory-Driven Features using Different Classifiers

<table>
<thead>
<tr>
<th>Order of Polynomial Model</th>
<th>Precision</th>
<th>Recall</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>78.94%</td>
<td>32.85%</td>
<td>88.84%</td>
<td>0.41</td>
</tr>
<tr>
<td>Second</td>
<td>85.1%</td>
<td>29.2%</td>
<td>88.84%</td>
<td>0.3889</td>
</tr>
<tr>
<td>Third</td>
<td>82.4%</td>
<td>30.7%</td>
<td>88.84%</td>
<td>0.3989</td>
</tr>
<tr>
<td>Fourth</td>
<td>77.4%</td>
<td>29.9%</td>
<td>88.4%</td>
<td>0.3808</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Classifiers</th>
<th>Precision</th>
<th>Recall</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes</td>
<td>55.24%</td>
<td>57.66%</td>
<td>86.91%</td>
<td>0.4873</td>
</tr>
<tr>
<td>Logistic</td>
<td>77.94%</td>
<td>38.69%</td>
<td>89.38%</td>
<td>0.4649</td>
</tr>
<tr>
<td>Bagging Pred</td>
<td>60.18%</td>
<td>49.64%</td>
<td>87.77%</td>
<td>0.4741</td>
</tr>
<tr>
<td>Logistic Model Tree</td>
<td>79.69%</td>
<td>37.23%</td>
<td>89.38%</td>
<td>0.4566</td>
</tr>
<tr>
<td>Decision Table</td>
<td>68.97%</td>
<td>43.80%</td>
<td>88.84%</td>
<td>0.4759</td>
</tr>
</tbody>
</table>
Discussion

- The advantage of the theory-driven approach is that the features identified provides the reasons for students’ frustration.
- The reason for frustration provides information on which variables to control while responding to students’ frustration.

Limitations:

- The frustration model is specific to Mindspark.
- To apply our theory-driven approach to other systems, careful thought is required to operationalize the blocking factors of goals.
- The goals of the students when they interact with the system should be captured; this is a limitation in the scalability of our approach.
- The results of the theory-driven approach are dependent on how well the goals are captured and how well the blocking factors of the goals are operationalized.
Methodology

User Interface (System)

Defintion of frustration
Operationalize for ITS

Log data

Model to predict frustration

If frustrated

Messages to handle frustration

Reasons for frustration.

Motivation Theory

Phase I

Phase II

Phase III

Phase IV
Our Approach to Respond to Frustration

1. Detect frustration with its reasons
2. Create motivational messages to respond to frustration
3. Develop the algorithm to show messages
4. Collect data for validation
5. Validate the impact of motivational messages on students’ frustration

Figure: Steps of our Approach to Respond to Frustration
Strategies

- Create motivational message to attribute the students’ failure to achieve the goal to external factors [76].
- Create messages to praise the students’ effort instead of outcome [77].
- Create messages with empathy, which should make the student feel that s/he is not alone in that affective state [52].
- Create message to request student’s feedback [121].
- Display messages using an agent [182], [121].
Algorithm 2 To display motivational messages

Require: Res Time, FrusInst, Question Type.

return Message

if FrusInst = 1 & Question Type is Normal then
    Create Message: Based on the response time, concatenate the messages from Table and display it to the students.
else if FrusInst = 1 & Question Type is Challenge then
    Create Message: Based on the response time, concatenate the messages from Table and display it to the students.
else if FrusInst = 2 & Question Type is Normal then
    Message: It is okay to get the wrong answer sometimes. You may have found the question hard, but practice will make it easier. Try again
else if FrusInst = 2 & Question Type is Normal then
    Message: Don't worry, this is a tough question for many of your friends too. You can attempt it again.
else if FrusInst = 3 then
    Message: Would you like to give your feedback?
end if
It is okay to get the wrong answer sometimes. You may have found the questions hard, but practice will make it easier. Try again.

It’s incorrect!

Correct answer:
\[ w = \frac{(u + 5v)}{5} \]

One method to make \( w \) the subject is shown below:

To isolate \( w \), it is first required to separate the term containing \( w \):

\[ u = 5w - 5v \]

Now, \( 5v \) can be added to both the sides to isolate \( 5w \):
Don’t worry, this is a tough question for many of your friends too. You can attempt it again.

Sorry, that’s incorrect!

Correct answer:

\[
a = \frac{(3b - 1)}{5} = \frac{3b - 10}{5}
\]

Given, \(5a + 10 = 3b\)

Or, \((5 \times a) + 10 = (3 \times b)\)

To isolate the term containing \(a\), we can subtract both the sides by 10.

\(5a = 3b - 10\)

To separate \(a\) from \(5a\), we can multiply both the sides by the reciprocal of 5.

\(a \times 5 \times \frac{1}{5} = (3b - 10) \times \frac{1}{5}\)
Data Collection - Methodology

Select three ICSE board schools. School ID: 1752, 153271, 420525

Collect class 6 student’s log data for one week.

Remove the sessions with no of questions < 10

Remove the sessions with average time spent to answer the questions < 11 seconds

Select the unique user ID and corresponding data

In the following week, implement addressing frustration algorithms for same schools.

Collect class 6 student’s log data for one week.

Remove the sessions with no of questions < 10

Remove the sessions with average time spent to answer the questions < 11 seconds

Select the unique user ID and corresponding data

Calculate number of frustration instances per session for the *identical* students
Table: Details of the data collected from three schools to measure the impact of motivational messages on frustration

<table>
<thead>
<tr>
<th>School Code</th>
<th>Number of students in Class 6</th>
<th>Mindspark topic in first week (Without motivational Messages)</th>
<th>Mindspark topic in second week (with motivational messages)</th>
<th>Number of matching students' sessions considered for analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1752</td>
<td>326</td>
<td>Integers</td>
<td>Integers</td>
<td>54</td>
</tr>
<tr>
<td>153271</td>
<td>279</td>
<td>Decimals</td>
<td>Decimals</td>
<td>72</td>
</tr>
<tr>
<td>420525</td>
<td>164</td>
<td>Algebra</td>
<td>Geometry</td>
<td>62</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>188</td>
</tr>
</tbody>
</table>
**Results**

**Table:** Median and Median Absolute Deviation (MAD) of number of frustration instances from the Mindspark session data from three schools

<table>
<thead>
<tr>
<th>Number of Mindspark Sessions</th>
<th>Median of Frustration Instances</th>
<th>MAD of Frustration Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>188 sessions without motiva-</td>
<td>2</td>
<td>2.1942</td>
</tr>
<tr>
<td>tional messages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>188 sessions with motivatio-</td>
<td>1</td>
<td>1.4628</td>
</tr>
<tr>
<td>nal messages</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure: Box plot of Frustration instances from 188 sessions without and with motivational messages. Box = 25th and 75th percentiles; bars = minimum and maximum values; center line = median; and black dot = mean.
Results

Number of frustration instances is reduced in from very high to less due to the motivational messages.
## Results

**Table: Impact of motivational messages on frustration in three schools**

<table>
<thead>
<tr>
<th>School Code</th>
<th>Number of Sessions</th>
<th>Without Motivational Message</th>
<th>With Motivational Messages</th>
<th>Mann-Whitney’s Significance Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sum of Frustration instances</td>
<td>Median</td>
<td>Sum of Frustration instances</td>
<td>Median</td>
</tr>
<tr>
<td>1752</td>
<td>54</td>
<td>92</td>
<td>1</td>
<td>57</td>
</tr>
<tr>
<td>153271</td>
<td>72</td>
<td>212</td>
<td>3</td>
<td>148</td>
</tr>
<tr>
<td>420525</td>
<td>62</td>
<td>130</td>
<td>2</td>
<td>72</td>
</tr>
</tbody>
</table>
Validation of Impact of Motivational Messages

![Box Plot of Frustration instances](image)

<table>
<thead>
<tr>
<th>School Code</th>
<th>Number of Sessions</th>
<th>First Week Data</th>
<th>Second Week Data</th>
<th>Mann-Whitney's Significance Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sum of Frustration instances</td>
<td>Median</td>
<td>Sum of Frustration instances</td>
</tr>
<tr>
<td>1752</td>
<td>99</td>
<td>215</td>
<td>2</td>
<td>203</td>
</tr>
</tbody>
</table>
Analysis on Ordering Effects - Removal of Motivational Messages

Figure: Box plot of frustration instances from 42 session in each week. First week without motivational messages, second week with motivational messages and third week without motivational messages.

Figure: Box plot of Frustration instances from 42 session in each week. First week without motivational messages, second week with motivational messages and third week without motivational messages.
Discussion

- From the histograms, the frustration instances of students are reduced in the sessions with motivational messages.
- There is a statistically significant reduction in the number of frustration instances per session due to the approach to respond to frustration.
- The significant reduction in the frustration instances is independent of the schools analyzed and topics used in the Mindspark sessions.
- The approach to respond to frustration has a relatively higher impact on the students whose performance in the sessions is low.
- The approach to respond to frustration has a relatively higher impact on the students who spend more time to answer the questions in Mindspark session.
Approach to Detect Boredom

The theory-driven approach to model boredom

1. Define Boredom: State of low arousal and dissatisfaction due to repeated situation and more/less challenging activity.
2. Identify the repeated activities in the ITS, and students' ability to solve the questions while student’s interaction.
3. List the repeated activities, student’s performance in a last three questions. Operationalize it for ITS.
4. Create a logistic regression model for Boredom Index (Bi) with the repeated activities identified.
5. Train the logistic regression model weights using student’s self reported data.
6. Validate the performance of model with test data compare the results with student’s self reported data.
Definition of Boredom Used in Our Research

The most common feature in all existing work on boredom is repetitiveness and monotonous stimulation [189], [191]. The other key features of boredom are

1. Conflict between whether to continue the current situation or not due to lack of motivation [190].

2. The student is forced to do an uninteresting activity. Non-interest occurs when the student is not challenged enough [37], [194].

3. The student is prevented from doing a desirable action or forced to do an undesirable action [191].

4. The student lost the interest in outcome of the event [193].
Boredom Model

The logistic regression model to detect boredom is given below:

\[ B_i = w_0 + w_1 \times f1 + w_2 \times f2 + w_3 \times f3 + \ldots + w_n \times fn \]  

(1)
Figure: EmotToolbar integrated with Mindspark user interface to collect students’ emotions. The emote bar is in right side of the figure.
The emotToolbar consists of six options for the students to choose from as

![The EmotToolbar](image)

**Figure:** The EmotToolbar
We collected 1617 instances of student’s answering the questions in Mindspark from 90 students.

Out of 1617, 442 instances are self reported as boredom (Bored) by students, the remaining instances are marked as (Non-Bored).

The dataset is stratified at questions (instances) level. Unit of analysis is the instances where students respond to questions in Mindspark.
Table: Results of Boredom Model when Applied to Mindspark Log Data

<table>
<thead>
<tr>
<th>Pred</th>
<th>Bored</th>
<th>Non-Bored</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bored</td>
<td>98</td>
<td>46</td>
</tr>
<tr>
<td>Non-Bored</td>
<td>344</td>
<td>1129</td>
</tr>
</tbody>
</table>

The values from Table 9 are used to calculate the performance of our model. The results are given in Table 10.

Table: Performance of our Approach Shown Using Various Metrics when Applied to Mindspark Log Data

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>75.88%</td>
</tr>
<tr>
<td>Precision</td>
<td>68.1%</td>
</tr>
<tr>
<td>Recall</td>
<td>22.22%</td>
</tr>
<tr>
<td>Cohen’s kappa</td>
<td>0.23</td>
</tr>
<tr>
<td>F1 Score</td>
<td>0.33</td>
</tr>
</tbody>
</table>
Major Contributions

- **Theory-driven Approach:** We developed an approach to detect affective states using data from the students’ interaction with the system. Our approach uses only the data from log files, hence, it can be implemented in the large scale deployment of ITS. We have tested our approach on a math ITS to detect frustration. Moreover, we validated the likelihood of generalizing the theory-driven approach to detect other affective states by creating a model to detect boredom in an ITS.

- **Frustration Model:** We developed a linear regression model to detect frustration in a math ITS – Mindspark, using the theory-driven approach. The detection accuracy of our model is comparatively equal to the existing approaches to detect frustration. Additionally, our model provides the reasons for the frustration of the students.

- **Respond to Frustration:** We provided an approach to avoid the negative consequences of frustration, such as dropping out, by using the motivational messages. The messages to respond to frustration are created based on the reasons for frustration. The impact of motivational messages was analyzed and it was found that our approach significantly reduced the number of frustrations per session.
Publications Arising Out of this Thesis


- Responding to Students’ Frustration while Learning with an ITS, To be submitted to the IEEE Transactions on Learning Technologies.


Thank You
Zhao Ying, Quan Bingzhe, and Jin Chunzhao.  
On intelligent tutoring system.  

Chung Leung, Elvis Wai, and Qing Li.  
An experimental study of a personalized learning environment through open-source software tools.  

Elizabeth J. Brown, Timothy J. Brailsford, Tony Fisher, and Adam Moore.  
Evaluating learning style personalization in adaptive systems: Quantitative methods and approaches.  

The effects of an interactive software agent on student affective dynamics while using an intelligent tutoring system.  

Sensors model student self concept in the classroom.  
C. Conati and H. Maclaren.  
Empirically building and evaluating a probabilistic model of user affect.  

K. Brawner and B. Goldberg.  
Real-time monitoring of ecg and gsr signals during computer-based training.  

M. S. Hussain, H. Monkaresi, and R. A. Calvo.  
Categorical vs. dimensional representations in multimodal affect detection during learning.  

Z. Zeng, M. Pantic, G.I. Roisman, and T.S. Huang.  
A survey of affect recognition methods: Audio, visual, and spontaneous expressions.  

Jennifer Sabourin, Bradford Mott, and James C. Lester.  
Modeling learner affect with theoretically grounded dynamic bayesian networks.  

Elvira Popescu.  
Evaluating the impact of adaptation to learning styles in a web-based educational system.  


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.  

Affect-Aware Tutors: Recognising and Responding to Student Affect.  

Angel de Vicente and Helen Pain.  
Motivation diagnosis in intelligent tutoring systems.  

Suchismita Srinivas, Muntaquim Bagadia, and Anupriya Gupta.  
Mining information from tutor data to improve pedagogical content knowledge.  

Ivon Arroyo and Beverly Park Woolf.  
Inferring learning and attitudes from a bayesian network of log file data.  

Modeling student affect and behavior using biometric readings, log files and low fidelity playbacks.  
Cynthia D. Fisherl.
Boredom at Work: A Neglected Concept.

Clifford T. Morgan, Richard A. King, John R. Weisz, and John Schopler.
*Introduction to Psychology*.

Cristina Conati and Heather Maclare.
Evaluating a probabilistic model of student affect.

Ashish Kapoor, Winslow Burleson, and Rosalind W. Picard.
Automatic prediction of frustration.

Viewing student affect and learning through classroom observation and physical sensors.

Predicting affective states expressed through an emote-aloud procedure from autotutor’s mixed-initiative dialogue.
Automatic detection of learner’s affect from conversational cues.

Coarse-grained detection of student frustration in an introductory programming course.

Rafael A. Calvo and Sidney K D’Mello.
Affect detection: An interdisciplinary review of models, methods, and their applications.

Peter Brusilovsky and Eva Milln.
User models for adaptive hypermedia and adaptive educational systems.

Jeff Johns and Beverly Woolf.
A dynamic mixture model to detect student motivation and proficiency.

C. Romero and S. Ventura.
Educational data mining: A review of the state of the art.
Huey-Ing Liu and Min-Num Yang.
Qol guaranteed adaptation and personalization in e-learning systems.

William L. Mikulas and Stephen J. Vodanovich.
The essence of boredom.

Reed W. Larson and Maryse H. Richards.
Boredom in the middle school years: Blaming schools versus blaming students.

Ulrike E. Nett, Thomas Goetz, and Nathan C. Hall.
Coping with boredom in school: An experience sampling perspective.
Students' Emotions and Academic Engagement.

Richard Farmer and Norman D. Sundberg.  
Boredom proneness—the development and correlates of a new scale.  

Marvin Zuckerman, Sybil B. Eysenck, and H.J. Eysenck.  
Sensation seeking in england and america: Cross-cultural, age, and sex comparisons.  

Anna Sierpinska, Georgeana Bobos, and Christine Knipping.  
Sources of students' frustration in pre-university level, prerequisite mathematics courses.  

Markku S. Hannula.  
Affect in mathematical thinking and learning: Towards integration of emotion, motivation, and cognition.  

Markku Hannula, Jeff Evans, George Philippou, and Rosetta Zan.  
Affect in mathematics education-exploring theoretical frameworks. research forum.  
Scott W. Mcquiggan, Sunyoung Lee, and James C. Lester.
Early prediction of student frustration.

Affect and usage choices in simulation problem-solving environments.

Sidney K. D’Mello, Scotty D. Craig, Barry Gholson, and Stan Franklin.
Integrating affect sensors in an intelligent tutoring system.

M. Grant and S. Boyd.
CVX: Matlab software for disciplined convex programming, version 1.21.

Hall Mark, Frank Eibe, Holmes Geoffrey, Pfahringer Bernhard, Reutemann Peter, and Ian H. Witten.
The weka data mining software: An update.
Reed, Lawson.  
*Frustration: the development of a scientific concept.* 

J Klein, Y Moon, and R W Picard.  
This computer responds to user frustration: Theory, design, and results. 

Naveen Nair, Ganesh Ramakrishnan, and Shonali Krishnaswamy.  
Enhancing activity recognition in smart homes using feature induction. 

M. S. Hussain, Omar AlZoubi, Rafael A. Calvo, and Sidney K. D’Mello.  
Affect detection from multichannel physiology during learning sessions with autotutor. 

Jonathan Lazar, Adam Jones, Mary Hackley, and Ben Shneiderman.  
Severity and impact of computer user frustration: A comparison of student and workplace users. 

Paul Ekman.  
Strong evidence for universals in facial expressions: a reply to Russell’s mistaken critique. 
Hillary Anger Elfenbein and Nalini Ambady.
On the universality and cultural specificity of emotion recognition: A meta-analysis.

Similarities and variations in facial expressions of emotions: Cross-cultural evidence.

K. Gottlicher, S. Stein, and D. Reichardt.
Effects of emotional agents on human players in the public goods game.

Sidney K D’Mello, Scotty D Craig, Karl Fike, and Arthur C Graesser.
Responding to learners’ cognitive-affective states with supportive and shakeup dialogues.

Takeo Kanade, Yingli Tian, and Jeffrey F. Cohn.
Comprehensive database for facial expression analysis.
Yan Tong, Wenhui Liao, and Qiang Ji.
Inferring facial action units with causal relations.

Sandra Graham, Bernard Weiner, DC Berliner, and RC Calfee.
Theories and principles of motivation.
*Handbook of educational psychology*, 4:63–84, 1996.

Mental states estimation with the variation of physiological signals.

A. Amsel.
Arousal, suppression, and persistence: Frustration theory, attention, and its disorders.

A. Ortony, G. L. Clore, and A. Collins.
The cognitive structure of emotions.

S. Geisser.
*Predictive Inference: An Introduction*.
*Machine learning, neural and statistical classification.*  

G. Salton.  
*Automatic Text Processing; the Transformation, Analysis, and Retrieval of Information by Computer.*  
Addison-Wesley, 1989.

Y. Yang.  
An evaluation of statistical approaches to text categorization.  

N. L. Stein and L. J. Levine.  
Making sense out of emotion.  

Robert E Thayer.  
*The biopsychology of mood and arousal.*  
Oxford University Press on Demand, 1989.

James Paul Gee.  
What video games have to teach us about learning and literacy.  
James Paul Gee.  
Learning and games.  

F. Heider.  
*The Psychology of Interpersonal Relations*.  

Bernard Weiner.  
An attributional theory of achievement motivation and emotion.  

Carol S. Dweck.  
Messages that motivate: How praise molds students' beliefs, motivation, and performance (in surprising ways).  

Carol S Dweck.  
Motivational processes affecting learning.  

Facial features for affective state detection in learning environments.  
Marian Stewart Bartlett, Gwen Littlewort, Mark Frank, Claudia Lainscsek, Ian Fasel, and Javier Movellan.
Fully automatic facial action recognition in spontaneous behavior.

Maja Pantic and Ioannis Patras.
Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences.

Hatice Gunes and Massimo Piccardi.
Bi-modal emotion recognition from expressive face and body gestures.

Rafael A Calvo, Iain Brown, and Steve Scheding.
Effect of experimental factors on the recognition of affective mental states through physiological measures.

Changchun Liu, Karla Conn, Nilanjan Sarkar, and Wendy Stone.
Physiology-based affect recognition for computer-assisted intervention of children with autism spectrum disorder.
Ashish Kapoor and Rosalind W Picard.
Multimodal affect recognition in learning environments.

Abram Amsel.
Frustration theory: Many years later.

Reinhard Pekrun.
The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators.

Paul E Spector.
Organizational frustration: A model and review of the literature.

Bernard Weiner.
Theories of motivation: From mechanism to cognition.
1972.

Sandra Graham.
A review of attribution theory in achievement contexts.
Charles Darwin.  
_The expression of the emotions in man and animals._  

Paul Ekman.  
Universals and cultural differences in facial expressions of emotion.  
_In Nebraska symposium on motivation._ University of Nebraska Press, 1971.

Paul Ekman and Wallace V Friesen.  
_Unmasking the face: A guide to recognizing emotions from facial clues._  

Carroll E Izard.  
_The face of emotion._  

Carroll E Izard.  
Innate and universal facial expressions: evidence from developmental and cross-cultural research.  

Craig A Smith and Phoebe C Ellsworth.  
Patterns of cognitive appraisal in emotion.  
_Journal of personality and social psychology, 48_(4):813–838, 1985._
Stanley Schachter and Jerome Singer.
Cognitive, social, and physiological determinants of emotional state.

Ira J Roseman.
Cognitive determinants of emotion: A structural theory.

An approach for identifying affective states through behavioral patterns in web-based learning management systems.

Rosalind W Picard.
*Affective computing*.

Classifying facial actions.

Rana El Kaliouby and Peter Robinson.
*Real-time inference of complex mental states from facial expressions and head gestures*.
Michel F Valstar and Maja Pantic.
Fully automatic recognition of the temporal phases of facial actions. 

Chul Min Lee and Shrikanth S Narayanan.
Toward detecting emotions in spoken dialogs.

Diane J Litman and Kate Forbes-Riley.
Predicting student emotions in computer-human tutoring dialogues.

Cecilia Ovesdotter Alm, Dan Roth, and Richard Sproat.
Emotions from text: machine learning for text-based emotion prediction.

Taner Danisman and Adil Alpkocak.
Feeler: Emotion classification of text using vector space model.
Weiyuan Li and Hua Xu.
Text-based emotion classification using emotion cause extraction.
*Expert Systems with Applications*, (0):–, 2013.

Ivon Arroyo, David G Cooper, Winslow Burleson, Beverly Park Woolf, Kasia Muldner, and Robert Christopherson.
Emotion sensors go to school.

Mark Coulson.
Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence.

Selene Mota and Rosalind W Picard.
Automated posture analysis for detecting learner’s interest level.

Sidney K D’Mello and Art C Graesser.
Automatic detection of learner’s affect from gross body language.
Nadia Bianchi-Berthouze and Christine L Lisetti.
Modeling multimodal expression of users affective subjective experience.

Johannes Wagner, Jonghwa Kim, and Elisabeth André.
From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification.

Kyung Hwan Kim, SW Bang, and SR Kim.
Emotion recognition system using short-term monitoring of physiological signals.

Tanja Bänziger, Didier Grandjean, and Klaus R Scherer.
Emotion recognition from expressions in face, voice, and body: the multimodal emotion recognition test (mert).

Multimodal affect recognition in spontaneous hci environment.
Henry A Feild, James Allan, and Rosie Jones.
Predicting searcher frustration.

Rich Caruana and Alexandru Niculescu-Mizil.
An empirical comparison of supervised learning algorithms.

Automatically recognizing facial expression: Predicting engagement and frustration.

Kate Hone.
Empathic agents to reduce user frustration: The effects of varying agent characteristics.

Saul Rosenzweig and Gwendolyn Mason.
An experimental study of memory in relation to the theory of repression.

Rosalind W Picard.
Affective computing.
ELENA Verdu, LUISA M Regueras, MARÍA JESÚS Verdu, JUAN PABLO de Castro, MA Pérez, Qing Li, SY Chen, and Anping Xu.
Is adaptive learning effective? a review of the research.

Peter Brusilovsky, Elmar Schwarz, and Gerhard Weber.
Elm-art: An intelligent tutoring system on world wide web.
In Intelligent tutoring systems, pages 261–269, 1996.

Susanne P Lajoie and Alan Lesgold.

Chris Eliot and Beverly Park Woolf.
Reasoning about the user within a simulation-based real-time training system.

Erica Melis and Jörg Siekmann.
Activemath: An intelligent tutoring system for mathematics.
Steven Ritter, John R Anderson, Kenneth R Koedinger, and Albert Corbett.
Cognitive tutor: Applied research in mathematics education.

Cognitive tutors: Lessons learned.

Mia Stern, Joseph Beck, and Beverly Park Woolf.
Adaptation of problem presentation and feedback in an intelligent mathematics tutor.

CR Beal, JE Beck, and BP Woolf.
Impact of intelligent computer instruction on girls math self concept and beliefs in the value of math.

Intelligent tutoring goes to school in the big city.
Kurt VanLehn, Pamela W Jordan, Carolyn P Rosé, Dumisizwe Bhembe, Michael Böttner, Andy Gaydos, Maxim Makatchev, Umarani Pappuswamy, Michael Ringenberg, Antonio Roque, et al.
The architecture of why2-atlas: A coach for qualitative physics essay writing.

Autotutor: A simulation of a human tutor.

Steven Ritter, John Anderson, Michael Cytrynowicz, and Olga Medvedeva.
Authoring content in the pat algebra tutor.
*Journal of Interactive Media in Education*, 98(9), 1998.

Albert T Corbett, Kenneth R Koedinger, and John R Anderson.
Intelligent tutoring systems.

Arthur C Graesser, Kurt VanLehn, Carolyn P Rosé, Pamela W Jordan, and Derek Harter.
Intelligent tutoring systems with conversational dialogue.

Arthur C Graesser, Patrick Chipman, Brian C Haynes, and Andrew Olney.
Autotutor: An intelligent tutoring system with mixed-initiative dialogue.
More accurate student modeling through contextual estimation of slip and guess probabilities in bayesian knowledge tracing.

Stacy Marsella, Jonathan Gratch, and Paolo Petta.
Computational models of emotion.

Phoebe C Ellsworth and Klaus R Scherer.
Appraisal processes in emotion.

Ira J Roseman, Martin S Spindel, and Paul E Jose.
Appraisals of emotion-eliciting events: Testing a theory of discrete emotions.

Richard S Lazarus.
*Emotion and adaptation*.

Saul Rosenzweig.
A general outline of frustration.
Frustration and aggression. 
1939.

Valerie J Shute and Joseph Psotka. 
Intelligent tutoring systems: Past, present, and future. 

Etienne Wenger. 
Artificial intelligence and tutoring systems. 

JR Hartley and Derek H Sleeman. 
Towards more intelligent teaching systems. 

William J Clancey. 
Intelligent tutoring systems: A tutorial survey. 

Derek Sleeman and John Seely Brown. 
*Intelligent tutoring systems.* 


Richard R Burton and John Seely Brown.
An investigation of computer coaching for informal learning activities.

Edward Hance Shortliffe.
*Computer-based medical consultations, MYCIN*.

William J Clancey.
Tutoring rules for guiding a case method dialogue.

John Seely Brown, Richard R Burton, and Alan G Bell.
Sophie: A step toward creating a reactive learning environment.

James D Hollan, Edwin L Hutchins, and Louis Weitzman.
Steamer: An interactive inspectable simulation-based training system.

Roger C Schank.
*Dynamic memory: A theory of reminding and learning in computers and people*.
Chris Dede, Marilyn C Salzman, and R Bowen Loftin.
Sciencespace: Virtual realities for learning complex and abstract scientific concepts.

Sandra Katz and Alan Lesgold.
The role of the tutor in computer-based collaborative learning situations.

Nicola Capuano, Marco Marsella, and Saverio Salerno.
Abits: An agent based intelligent tutoring system for distance learning.

Peter Brusilovsky.
Adaptive hypermedia.

Detecting student misuse of intelligent tutoring systems.

Ling-Hsiu Chen, Yi-Chun Lai, and Yi-Hsu Weng.
Intelligent e-learning system with personalized misconception diagnose and learning path guidance.
Barry Kort, Rob Reilly, and Rosalind W Picard. 
An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion.

Edward Vockell. 
Educational psychology: A practical approach. 
Purdue University, 2004.

Sadia Batool, Muhammad Imran Yousuf, and Qaisara Parveen. 
A study of attribution patterns among high and low attribution groups: An application of weiners attribution theory. 

Jyotirmay Sanghvi, Ginevra Castellano, Iolanda Leite, André Pereira, Peter W McOwan, and Ana Paiva. 
Automatic analysis of affective postures and body motion to detect engagement with a game companion. 
Mohammad Adibuzzaman, Niharika Jain, Nicholas Steinhafel, Munir Haque, Ferdaus Ahmed, Shiekh Iqbal Ahamed, and Richard Love.
Towards in situ affect detection in mobile devices: a multimodal approach.

Sidney DMello, Rosalind Picard, and Arthur Graesser.
Towards an affect-sensitive autotutor.

Blair Lehman, Melanie Matthews, Sidney DMello, and Natalie Person.
What are you feeling? investigating student affective states during expert human tutoring sessions.

Allan Wigfield and Jacquelynne S Eccles.
Expectancy–value theory of achievement motivation.

Marylène Gagné and Edward L Deci.
Self-determination theory and work motivation.

Edward L Deci and Richard M Ryan.
Self-Determination.
Wiley Online Library, 2010.
Thomas Gordon. 
*PET: Parent Effectiveness Training.*

William R Nugent and Helene Halvorson. 
Testing the effects of active listening. 

Helmut Prendinger and Mitsuru Ishizuka. 
The empathic companion: A character-based interface that addresses users’ affective states. 

Timo Partala and Veikko Surakka. 
The effects of affective interventions in human–computer interaction. 

Craig C Pinder. 
*Work motivation in organizational behavior.* 

Encouraging girls in math and science. 
2007.
Allan Wigfield and Kathryn R Wentzel.
Introduction to motivation at school: Interventions that work.

Edwin A Locke and Gary P Latham.
What should we do about motivation theory? six recommendations for the twenty-first century.

Friedrich Försterling.
Attributional retraining: A review.

Richard P Smith.
Boredom: A review.

Joseph E Barmack.
The effect of benzedrine sulfate (benzyl methyl carbinamine) upon the report of boredom and other factors.

James F O'Hanlon.
Boredom: Practical consequences and a theory.

