# Modelling Activities of Self Regulated Learners as Contextualized Action Sequences

**Research** · August 2015

DOI: 10.13140/RG.2.1.1010.3525



1 author:

Jordan Aiko Deja
University of Primorska

# MODELLING ACTIVITIES OF SELF-REGULATED LEARNERS AS CONTEXTUALIZED ACTION SEQUENCES

A Thesis
Presented to
the Faculty of the College of Computer Studies
De La Salle University Manila

In Partial Fulfillment
of the Requirements for the Degree of
Master of Science in Computer Science

by

DEJA, Jordan Aiko P.

Rafael A. CABREDO, PhD
Adviser

August 11, 2015

**Abstract**

With the growth of technologies and the line between the traditional and virtual classrooms blurring, learning can take place basically anywhere. Self-Initiated Learning Scenarios are environments that enable students to learn on their own without the supervision of a teacher. Self-regulated learners are students who can greatly-benefit from these environments and as such, in this research their activities are tracked to be able to generate a model for positive learning habits. With the use of an annotation tool called Sidekick, these self-regulated learners undergo a process referred to as self-reflection. This is a phase in the learning process that enables students to reflect and improve on their learning habits.

Twenty five (25) undergraduate computing students participated in the study immersing themselves in these scenarios. These students were assessed and categorized into three levels of self-regulation namely low, moderate and high self-regulation.

A model is created based on their interaction data following a machine learning task. A general model covering three types was developed but only performed with a 48% accuracy leading to the need to develop a class-specific model covering these types of learners. The class models specific for students of low, moderate and high regulation obtained accuracy scores of 63.3%, 55.5% and 50.8% respectively; outperforming significantly better than that of the general model.

Their interaction data and models, along with their transitions were used to generate a set of rules called policies, employing a profit-sharing algorithm. Three sets of policies were generated depending on the type of self-regulated learner the student falls into. These sets of policies agree with the specific class precision and recall values from the models, thereby creating a set of rules for activities that when followed would improve or maintain the motivation of students regardless of self-regulation type that they fall to.

Keywords: SideKick, self-regulated learners, self-initiated learning, annotation, activity recognition.

# Table of Contents

# List of Figures

# List of Tables

# 1 Research Description

## 1.1 Overview of the Current State of Technology

Learning is commonly-associated with the knowledge transfer process involving guidance from a teacher. Any scenario that possesses these factors can be referred to as learning scenarios. Learning can be considered a part of personal development, of educating or training. Learning can sometimes be goal-oriented and be even aided by motivation as observed (Klein et al., 2006). It is not an easy task and sometimes requires merit in the form of rewards or reinforcements to not only let it take place, but to sustain it as well. This motivation powers the desire to gain knowledge (Adcock et al., 2006). There are instances where a learner already enjoys the topic at hand which enables him to continue on learning even when on your own; we can call this Tangential learning (Portnow, 2008). On certain cases, a student continues learning for other several reasons which does not necessarily imply that they enjoy the topic. At times there is a need for someone to continue learning in order to fulfill a specific learning task. Motivation enters the equation by fueling the desires of the student to sustain the learning process even without explicitly enjoying it. A person continues to learn not because he enjoys the topic but because even though he does not necessarily enjoy it yet, he is motivated in learning. Furthermore, liking a topic as discussed in tangential learning can be considered a subclass of motivated learning (Baldassarre et al., 2013).To define, motivation refers to the internal state that activates behavior thus affecting how a student behaves with regards to a certain subject matter (Ormrod, 2013). In effect, a motivated learner possesses enhanced cognitive processing abilities, a more direct behavior and an increased initiation and the persistence of these initated activities. These have been well-defined by the Self-Determination Theory by Deci et al. (1991) that describes both intrinsic and extrinsic motivation.

As people grow older they are also expected to know how to learn by themselves (P. Inventado et al., 2012), possibly even on their own. An even bigger challenge is being able to sustain one's own pace of learning (Pintrich, 1999). In a self-initiated learning scenario, student learning takes place even if they are unsupervised by a teacher or an agent and this gives them complete control over their learning. Formally, it is defined as a scenario characterized by the absence of either a human or an automated teacher. As a result, such scenarios require students to manage their learning apart from accomplishing their actual learning goals (P. Inventado et al., 2012). Learners who immerse themselves in self-initiated learning scenarios are referred to Self-Regulated Learners (SRL) (Schunk & Zimmerman, 2012).

Several research have been performed on students to analyze their behaviours (P. S. Inventado et al., 2011), ensuring they have motivation (Si & Wang, 2001), analyzing their transitions (D'Mello & Graesser, 2012) (Baker et al., 2007) and even their over-all general affect (Rodrigo et al., 2007) (D'Mello et al., 2008) (Craig et al., 2004). However, there are limited studies that have been able to model the behaviour of students contained in such scenarios - and such studies employed the use of external observers called coders. There are several studies who have employed educational data mining techniques in analyzing signals such as keystrokes and mouse gestures. However, in these activities, students are either supervised by an agent or by a teacher which does not make their setup fit to be classified as a self-initiated scenario.

Self-reflection covers many phases, which may include self-judgment and self-reaction. Self-judgment is an activity, which involves self-evaluation and causal attribution. A form of self-judgment, called self-evaluation involves comparisons of self-observed performances against a certain previously imposed standard. Feelings of self-satisfaction and positive affect on one's performance will be taken into account and included as part of the annotation data (Zimmerman, 2002). A noted increase in the level of self-satisfaction enhances motivation and a decrease induces any affect that may undermine further efforts to learn, including the act of self-reflection (Schunk, 1989). It has not been discovered if these strategies can be used to improve the habits of regulated learners and to make non regulated learners step up on their pace.

SideKick (P. S. Inventado et al., 2011) is an annotation tool. It is a software tool used to investigate what keeps students motivated in self-initiated learning scenarios. The self- reflection process of Sidekick involves (1) the actual users keying in the tasks at hand that they have done within a certain time-frame, (2) the affect which they may have encountered during that highlighted task and (3) his response on how related were his activities into the success of his current learning session. A learning session in SideKick takes around 2 hours, with 1 hour dedicated to learning activities and the remaining time used for the self-reflection where learners self-annotate. The complexity and tediousness of the work may have an impact to the over-all emotional state of the students when using SideKick. Yet even as such, it cannot still be fully relied on as such processes might be inaccurate, possibly due to the lack of experiences on certain emotions (Ekman, 2007). Such annotation processes can be improved to avoid the prevalence of such issues and increased labor affects levels of emotions, stress, even job-satisfaction and well-being (Pugliesi, 1999). Thus, prolonged hours beyond the designated working time changes the total achievement of self-regulated learners (Deci et al., 1991) if imposed of additional annotation hours worth of effort.

## 1.2 Research Objectives

This section describes the main objective that this research aims to accomplish. The following questions are expected to be addressed:

### 1.2.1 General Objective

What type of model can be created which best describes the patterns and habits of self-regulated learners?

This research aims to determine whether a class-specific or general model can be built that shall hold true for self-regulated learners, best describing their activities in self- initiated learning scenarios. Following this approach, a model in a form of rules and associations can be derived which can be further used to possibly train or improve non regulated learners.

### 1.2.2 Specific Objectives

Specifically, the study aims to answer the following questions:

1. What activities do students undertake in self-initiated learning scenarios?

   It is important to gather and study the data on students immersed in self-initiated learning scenarios. There are several types of tasks that they do when in such environments. These in general can be grouped in activity types and patterns that can be categorized depending on certain definitions such as how can these lead to a successful learning session.

2. What representation should be used for modeling activities of self-regulated learners?

   Self-regulated learners control their activities when performing a learning task on their own. As such, a decision process takes place in scheduling, managing what to learn, how to learn and how long to learn. It is important to employ a methodology that will allow the collection of data. This leads towards determining the features that will represent elements and sequences that defining self-regulation in itself..

3. What metrics can be used to compare the performance between a user-specific model and general model for self-regulated learners.

With the data represented and labelled; features extracted and sequences identified, it is important to validate the model produced if it is applicable for specific types of self-regulated learners or for the general self-regulated learner.

## 1.3 Scope and Limitations of the Research

This research involves a data gathering stage and a data modelling stage. The first stage involves the collection of data from self-regulated learners with the use of the Self- Reflection tool Sidekick. The second stage involves the processing and modelling of data into a set of rules that define student activities and/or sequence. The data collection was held at the University of Santo Tomas. The test subjects involved undergraduate computing students who work on relevant computing tasks such as programming, research writing among many others. Additionally, these students have been assessed to belong in varying schemes of being Self-Regulated. The data collected will undergo pre-processing, feature extraction and model building. The findings are used to create generalizations on the activities of self-regulated learners when deployed in their independent learning scenarios.

Classification and Time-based algorithms were used such as Multi-layered Perceptrons, Hidden Markov Models. Models will be created on the activities performed by the students in self-initiated learning scenarios, the policy value in determining which of their activities are relevant and another model for the transitions in their activities.

The annotations in terms of activities for both phases of the study would be limited within a set of schemes and classifications namely Information Search, View Information Source, Writing Notes, Seeking Help from Peers, Knowledge Application, and Off-task (P. Inventado et al., 2012). Though the affect and the intentions will also be annotated, data involving these two annotation categories are included as part of the model for the first phase but were included in second phase. The raw data would involve the titles of foreground applications, number of mouse clicks, keystrokes, transitions and existing policy data. These raw data would be labeled using the annotations of activity groups earlier mentioned.

## 1.4 Significance of the Research

The research provides new opportunities for learning theorists, educational practitioners, distance learning advocates and computer scientists to validate and use self-initiated learning scenarios as a test-bed in evaluating self-regulated learners. Additionally, the model can be put to the test in evaluating, and observing self-regulated learners in seeing which learning habits can be learned from and possibly be used for the development, training and improvement of non-regulated learners.

The output of this research is a model that determines which activities of self-regulated learners enable them to remain regulated and motivated. Given this, we could discover a set of traits, habits and sequences that can be used to improve existing self-regulated learners and to develop regulation among those who are assessed to be not self-regulated.

The study can also be extended in industries who aim to monitor the activities of their workers. The existing model can be used as a basis to build another model that predicts activities of workers and professionals such as developers and analysts. This way, we can discover and compare if the traits of self-regulated learners are or can be carried over when they become working professionals - a set of personalities with a different set and range of motivational factors. These can help managers and employers determine how productive and motivated their employees are in order to ensure quality results in the workplace. Similarly the same tool and model can be used to monitor the learning schemes of individual and home school-based learners. Without the need of supervision, learners can develop their own learning strategies .

## 1.5 Research Methodologies

This research involved the following stages namely: (1) planning, (2) documentation, (3) review of related literature, (4) data collection, (5) modeling with machine learning, and (6) testing and evaluation. Each of the stages will be discussed in the succeeding paragraphs.

### 1.5.1 Planning and Topic Analysis

This phase involves the selection of the topic and how it was selected by the researcher. Several sessions of consultation and meetings with the adviser took place, and also included collaboration with Dr. Paul Salvador Inventado, the developer of SideKick.

### 1.5.2 Documentation

Documentation is done throughout all stages of the research. The researcher keeps track of his progress through research journals, which are then processed and converted into the findings placed in this manuscript.

### 1.5.3 Review of Related Literature and Concept Formulation

The researcher has already conducted review of related literature where at most 50% of the synthesized studies are used in the study. Studies on machine learning techniques in behaviour detection and prediction were also included, gathered, reviewed and taken into consideration. Studies involving self-supported learning environments and academic emotions were also considered in lieu of supporting the total background of the study.

### 1.5.4 Data Collection

One data collection scheme was implemented. The collection involved gathering the data needed to be able to design the model that will be used as basis in improving the annotation schemes for SideKick. The model can be then be used to develop a prototype which is expected to have implemented an automated annotation of activity sequences.

At least 30 student respondents were asked to install the latest version of SideKick in their personal machines. With this they are expected to allow SideKick to run in the background while they perform their academic activities. Prior to running it in the background, the students are expected to configure into the system the specific objectives that they will be expecting to perform. The respondents filled up a Research Consent form (see Appendix). The respondents are expected to spend at least 10 learning sessions using SideKick where a session lasts at around an hour each. For each learning session, extra time will be alloted by the respondents to annotate three things namely (1) their activities as selected from the pre-set choices by SideKick, (2) their affect for that respective activity and (3) their response as to how related the highlighted activity was into accomplishing their specified learning objectives early on. A post session questionnaire was given to the respondents at the end of each learning session to gauge the relevance of their inputs in comparison to what was actually executed by the students while being monitored by SideKick.

In this data collection procedure, the speed of the user's mouse movement, mouse clicks and keystrokes, along with the titles of the active foreground windows was acquired by SideKick and made part of the data set. The screenshots displaying the currect face of the respondent during a certain moment along with the screen capture of the entire active desktop will be submitted on an optional basis but will not be used for the model building stages entirely. These will basically be used for future references should there be observations or clarifications that might arise leading to the discovery of possibly new insights in the study.

The respondents submitted their compiled database files into a repository where the researcher will be compiling all these together. From this point, the data will be pre-processed to fit for classification training.

### 1.5.5 Modelling and Data Analysis with Machine Learning

This phase involves the modeling of the data with the application of machine learning algorithms. A model will be produced, where it will determine which among the activity types in SideKick is corresponded by the available data provided by the keystrokes, titles and mouse movements. These activity types as defined by (P. Inventado, 2014) are:

- Make learning plan
- Modify learning plan
- Read and understand information
- Practice skills
- Search for info
- Apply acquired info
- Use previous knowledge
- Take notes
- Review notes
- Seek help and
- Off task

.

The model that is to be produced should be able to classify from the given feature sets and sequences which would lead to the given activity types mentioned earlier. Additionally, the annotations and labels acquired from the data collection would be used as labels to help the model classify the given activity types.

The approach of the modeling was sequence-based and as such would need a specialized machine learning algorithm (Miyazaki et al., 1994).There are existing modeling techniques for a limited range and combination of modalities. Since SideKick data would involve screen-captures, mouse movement, keystrokes, and titles of foreground applications, there has not yet been one established technique that deals with multiple-modalities, of varying input data types (visual, time-based, numeric all in all) that can do a predictive analysis on a student's learning pattern/behaviour. Depending on the actual inputs from the gathered data, there will be an appropriate machine learning algorithm that can be employed for improved accuracy. Students in self-initiated learning scenarios use SideKick and as such their inputs can be used to model how they generally transition from one type of task to another. The model is expected to display clustering of the different activity types which have been derived from the sequences found in the data set.Feature selection activities was also employed to determine the most appropriate feature set for the model.

### 1.5.6 Testing and Evaluation

A testing procedure was employed. This procedure was in an iterative manner where the model will be fixed until it produces satisfactory and accurate results. This will ensure a good accuracy rating and an acceptable Kappa statistic value. To test the model, ten (10) fold cross validation testing on the entire data set will be used with the model. Should the results show no significant results, further data gathering may be done or an improved strategy in data pre-processing and Feature Selection will be necessary. Another data collection scheme commenced to validate if the model generated with the initial set is consistent with the newly-gathered data.

### 1.5.7 Calendar of Activities

The Gantt chart seen in Table 1.1 presents the different stages and their respective timelines for this research. Each star represents one week of activity.

Table 1.1: Calendar of Activities

| Activities (2014-15) | Aug | Sep | Oct | Nov | Dec | Jan | Feb | Mar | Apr | May | Jun | Jul |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Planning | * | * | * | | | | | | | | | |
| Documentation | * | * | * | | * | * | * | * | * | * | | * |
| Review of RL | * | * | | | * | * | * | * | | | | |
| Data Collection | | | | | | | | | | | *** | |
| Modeling | | | | * | * | * | | | ** | * | * | ** |
| Testing | | | | | | | | | | * | * | * |

# 2 Review of Related Literature

This section discusses several literatures that have been considered helpful and related to the research topic. The review covers sections, which discusses self-supported learning environments and related activities, the machine learning techniques that are used to model activities involving self-supported learning environments and techniques in annotations.

## 2.1 Related Studies on Self-Regulated Learner Theories

In this study we refer to self-initiated learning environments as an environment where the presence of a supervisor, either a human or an automated teacher is omitted (P. Inventado et al., 2012). With such, students in such learning environments have to manage their time on their own. There are multiple factors that are to be considered when in a Self-Supported Learning Environment. These factors include, (1) Achieving the goals of the session, (2) The amount of time completing a goal, and more importantly (3) the affective states experienced by the students themselves.

Since students are expected to learn by themselves, they have to employ several methodologies that would facilitate a learning process, which they might perceive as something better. It has been understood that the process of teaching taps 3 main concepts namely, structuring, taking responsibility and reflecting (Roscoe & Chi, 2007). Using such concept, students are given their own opportunities to organize their own knowledge, monitor themselves during the process. In a study by Biswas et al. (2010), Students' performance and learning on their own were measured against the performance of a teachable agent. The study involved teaching a 5th Grade Science Topic on River Ecosystems. Students' Activity Sequences were captured and later on interpreted as a corresponding behavior with respect to different conditions.

In a study performed by P. Inventado et al. (2012), a generation of individuals belonging under the category called "Digital Natives" can perform quickly and work more efficiently when faced on their own self-supported learning environments. Similarly, they are also prone to poor learning performance because of curbed undesirable learning behaviors. Digital Natives are considered to be technology-savvy and can carry out many activities when on their own (Roberts et al., 2005). They employ tools where they can easily-switch in between learning activities and non-learning activities such as viewing online tutorial videos and engaging in social network behavior, respectively. Certain studies like as implemented by Kirschner & Karpinski (2010) and Rubinstein et al. (2001) have suggested that there are activities like task-switching between learning and non-learning activities commonly result to poorer performance from such learning outcomes.

The study performed by Legaspi P. Inventado et al. (2012) created a self-supported learning environment, where it has been identified that students who are aware of their own learning behaviors provide a meta-cognitive approach in reflecting if they should adapt to newer habits or not. A similar study conducted by Craig et al. (2004) allowed a rather familiar approach where participants in a self-supported learning environment can input their affective states through the process of a post-test annotation. Students who are able to review how they spent their time learning on their own can get a general idea on the strengths and weaknesses of their own study habits.

Table 2.1 displays the summary of findings on self-supported learning environments.

Table 2.1: Related Studies on Self-Regulated Learner Theories

| Authors | Concept | Contributions |
|---|---|---|
| (P. Inventado et al., 2012) | Elements comprised in a SSLE | A Self-Supported Learning Environment (SSLE) can be described to have the following factors: Achieving the goals of a session, the amount of time in completing the said goal and the affective states experienced by the learner. |
| P. S. Inventado et al. (2011) | Self-Reporting Techniques | Designed a framework and a tool for capturing student activities which allowed students to input annotations based on their affect. |
| Kirschner & Karpinski (2010) & Rubinstein et al. (2001) | Poor Learning Outcomes | They have identified that students with curbed learning habits and constant switching between objective and non objective activities tend to result with poor learning outcomes |
| Roberts et al. (2005) | Learners as Digital Natives | His study has described that individuals aged 8-18 during the Year 2005 range, are classified as Digital Natives, who as identified perform better and more efficient with the use of both online and social tools. |
| Biswas et al. (2010) | Learning with Teaching Agent | Though the presence of a teaching agent voids the definition of a Self-Supported Learning environment, his study has identified that the role of teaching an agent on the part of the student can be linked to several performance indicators as well. |
| Roscoe & Chi (2007) | Description of SSLE | Roscoe has identified that in a SSLE, there are three areas that are involved namely: Structuring activities, Taking Responsibility on your Actions and Reflecting on your performance results. |

## 2.2 Related Studies on Modeling Activities of Self-Regulated Learners

Several machine learning techniques can be used in classifying behaviors. Similarly, these techniques have been used to relate emotions with their possible corresponding learning behavior. In a study conducted by Kinnebrew et al. (2012), Sequential Pattern Mining can be utilized to observe behaviors, activity patterns and derive a conclusion out of it. The study employed a Piecewise Linear Segmentation Algorithm with the same differential sequence pattern mining technique not only to identify but also compare which among the students' behaviors can be classified as productive and unproductive.

The study involved 40 8th grade students whose behaviors where tracked and observed as they instruct a teachable agent. The learning activities, which were conducted by the respondents of the study were classified either as high-performing or low-performing groups. Segments of their activities were categorized also as either productive or unproductive. In this study, the differential sequence pattern mining technique was employed to divide the categories of the students activity segments as either *Hi* or *Lo*. The technique garnered an accuracy rating of 60.3% for the *Hi* group with a Cohen's F value of 0.992.

As described by Sharma & Gedeon (2012) in a survey study, Markov models and Markov chains can be used for time-domain processes where each process is dependent on prior states provided. For a sequence of observations that needs to be used for mining later on, a Hidden Markov Model with a Markov chain can be used. Hidden Markov Models have been mainly used in recognizing stress in forms of speech, voice, and behavioral inputs.

A similar study conducted by Azcarraga et al. (2011) tracked mouse click features such as distance travelled of the mouse pointer, the duration of each click and the number of clicks performed by the user where observed and analyzed. The students in the learning session, though not necessarily on a self-supported learning environment, had their activities sequenced, spliced and analyzed by a certain machine-learning algorithm. A multi-layer perceptron was designed to classify these activities. These actions and sequences were captured and used to find a relationship with academic emotions experienced, further describing the affect that the subjects underwent during the study. The results can be viewed as part of the summary as seen from Table 2.2.

A method in predicting tasks using human movement have been employed by MacKenzie (1992b), Boritz et al. (1991), Epps (1986) among many others. This method analyzes mouse movement and recommends a task for the human user based on the model. These studies combine all techniques of the Shannon Formulation in refining task identification difficulty (MacKenzie & Buxton, 1992), and Fitts' Law (MacKenzie, 1992a) in designing tools for human computer interaction.

Keystrokes, or actions using the keyboard, have been used to ascertain human activities such as user authentication (Gunetti & Picardi, 2005). Models can also be built using keystrokes and their features as attributes. In a study by Gunetti & Picardi (2005), and Noack (2007) the number, length and composition of keystrokes where considered as features for the model.

A notable study performed by Card et al. (1980) applies a keystroke-level analysis on the user's performance when using computer systems. The study proposes a keystroke-level model in predicting the aspect of one's performance. In a self-supported learning environment, one's activities can be tracked on the type of performance it displays while using the computer. Additionally, the study considers the following factors (as directly quoted from the study:

- Time: How long does it take a user to accomplish a given set of tasks using the system
- Errors: How many errors does a user make and how serious are they?
- Learning: How long does it take a novice user to learn how to use the system to do a given set of tasks?
- Functionality: What range of tasks can a user do in practice with the system?
- Recall: How easy is it for the user to recall how to use the system on a task that he has not done for a time?

The prescribed model considers time as one of its features along many others (Card et al., 1980). As such it dissects time into the given equation:

$$T_{task} = T_{acquire} + T_{execute} \tag{1}$$

Where the acquisition time for a task is dependent on the characteristics of the larger task situation in which the task occurs. From the same model, the type of tasks were encoded into operators which we were used as patterns to match the predictions based on the given inputs in contrast with the built model. These operators were modelled based on the same features from the previous equation including newer features as seen below:

$$T_{execute} = T_K + T_P + T_H + T_D + T_M + T_R \tag{2}$$

Where $K$ stands for keystroking, $P$ for pointing, $H$ for homing, $D$ for drawing, $M$ considered as the mental operator and $R$ as the response operator.

Table 2.2 shows the summary of related literature involving these techniques.

Table 2.2: Related Studies on Modeling Activities of Self-Regulated Learners

| Authors | ML Technique | Input/Modality | Findings | Reported Accuracy |
|---|---|---|---|---|
| Kinnebrew et al. (2012) | Piecewise Linear Segmentation Algorithm (PSLA) | Activity Sequences | PLSA is a specialized type of Sequential Pattern Mining developed by Kinnebrew. It utilizes behavior patterns and activity sequences in order to derive a conclusion out of these data. | 68% |
| Sharma & Gedeon (2012) | HMM & Markov Chains | Activity Sequences | Activities and sequences of patterns can be classified as Time-domain processes where prior states are considered dependent. These chains are best used for a sequence of observations be it visual or auditory. | NA |
| Azcarraga et al. (2011) | Multi-layer Perceptron | Mouse Clicks, Brainwave Signals | If combined with features of Brainwaves, prediction accuracy peaked at 92%. Using mouse clicks alone, led to a mere 32-48%. These features were captured with the use of an ITS used by student achievers. | 32-48%, 54-88%, 80% |
| MacKenzie & Buxton (1992); MacKenzie (1992a,b) | Fitts' Law & Shannon Formulation | Mouse Gestures | The technique analyzes mouse movements and clicks such as dragging, selecting. A model was not designed or built. The technique was meant to reduce difficulty for the machine to identify the different mouse activities thus reducing time (10%)consumed to identify such task. | NR |
| (Card et al., 1980) | Keystroke-Level Model | Keystrokes | 1280 respondents given 14 tasks each were used as basis for building the model. It considered. It encoded keystroke patterns into codes which were discussed in the paper. Newer predictions were tested to match the results based on the produced model. | 69% |

# 3 Theoretical Framework

## 3.1 Theories on Self-Regulated Learning and Self-Initiated Learning Scenarios

Students, specifically Digital Natives encounter multiple distractions when immersed in their native learning environments. While these challenges continue to exist, some students maintain a range of motivation to continue learning in a well-controlled, self-decided environment. This is now referred to as Self-Regulated Learning. Formally, it is the generation of thoughts and feelings and actions that are planned and cyclically-adapted for the attainment of a specific goal (Zimmerman, 2002). Several models of Self-Regulation have been devised with Zimmerman (2002) as the most recognized among them. It includes a three-stage cycle categorized as the forethought phase, the performance phase and the reflection phase.

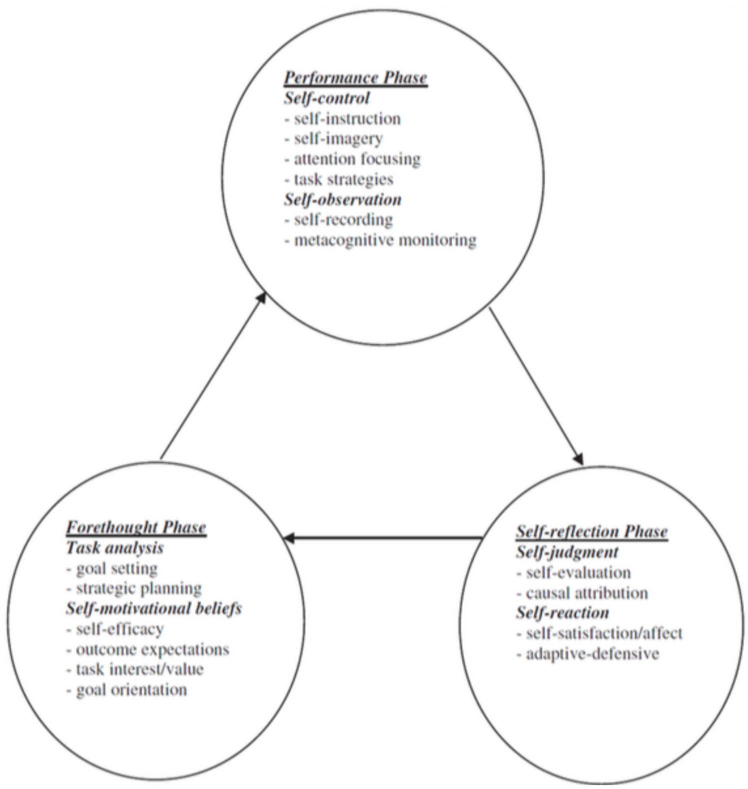


Figure 3.1: Zimmerman's Model on Self Reflection

In focus, the process of self-reflection separates the self-regulated learners from the less self-regulated learners because of its two distinct sub-stages namely self-judgment and self-reaction. These two sub-processes enable a learner to evaluate and assess ones self in order to determine and measure how regulated a person may be. Their previous performance or the performance of others in contrast set against others add attributes that enable one to discover how regulated one can be. Through these stages, less self-regulated learners discover that they have poor management skills and experience difficulty in setting goals for themselves, planning their activities and even monitoring progress while within a certain goal or task.

Existing literature has supported that learners who are more least-likely to be self-regulated encounter difficulty as well in selecting appropriate strategies during tough situations (Garner & Bol, 2011) , and having difficulty as well in incorporating either postive or negative feedback to improve in one's future behaviour(Garner & Bol, 2011). With the help of Zimmerman's Self-regulated learning model, students learn to adapt in their everyday tasks by being able to measure and evaluate ones self while at the same time being able to measure satisfaction, motivation and gauging the affective experience while in a task.

As an improvement (D'Mello & Graesser, 2012) higlighted the importance of determining affect in modeling these dynamics, thus explaining relationships between student's emotions along with the current event that may be happening.

Considering the said figure which focuses on the sequence of student's activities along with the respective affect associate, it is important to note that there are important points indicative that there is a need to make decisions or judgments based on the current event. As such, a student gets to decide if he shall continue or opt out of the said learning activity. This decision making process aids in the definition of a self-regulated learner.

In response to the need of learning more on the online behaviors of self-regulated learners, an environment dedicated for them was devised and created. This environment enable self-regulated learners to learn and plan their learning scenarios while at the same time being able to continously evaluate and reflect on their activities - affecting their overall motivation points. These environments were put into use by (P. Inventado, 2014).

Several studies have been performed by Inventado (P. S. Inventado et al., 2011), (P. Inventado et al., 2012), (P. Inventado, 2014) employed techniques in using student learners in self-initiated learning scenarios. These support self-regulated learning processes as it may allow a self-regulated learner to undergo the different stages of self-reflection. With the help of Sidekick (P. S. Inventado et al., 2011), a tool is utilized to support and enable students analyze and evaluate their learning behavior. This enables them to come up with responsible choices and improved decision calls on matters that lead toward future learning sessions. This as seen, promotes self-regulated learning as the processes are maintained.

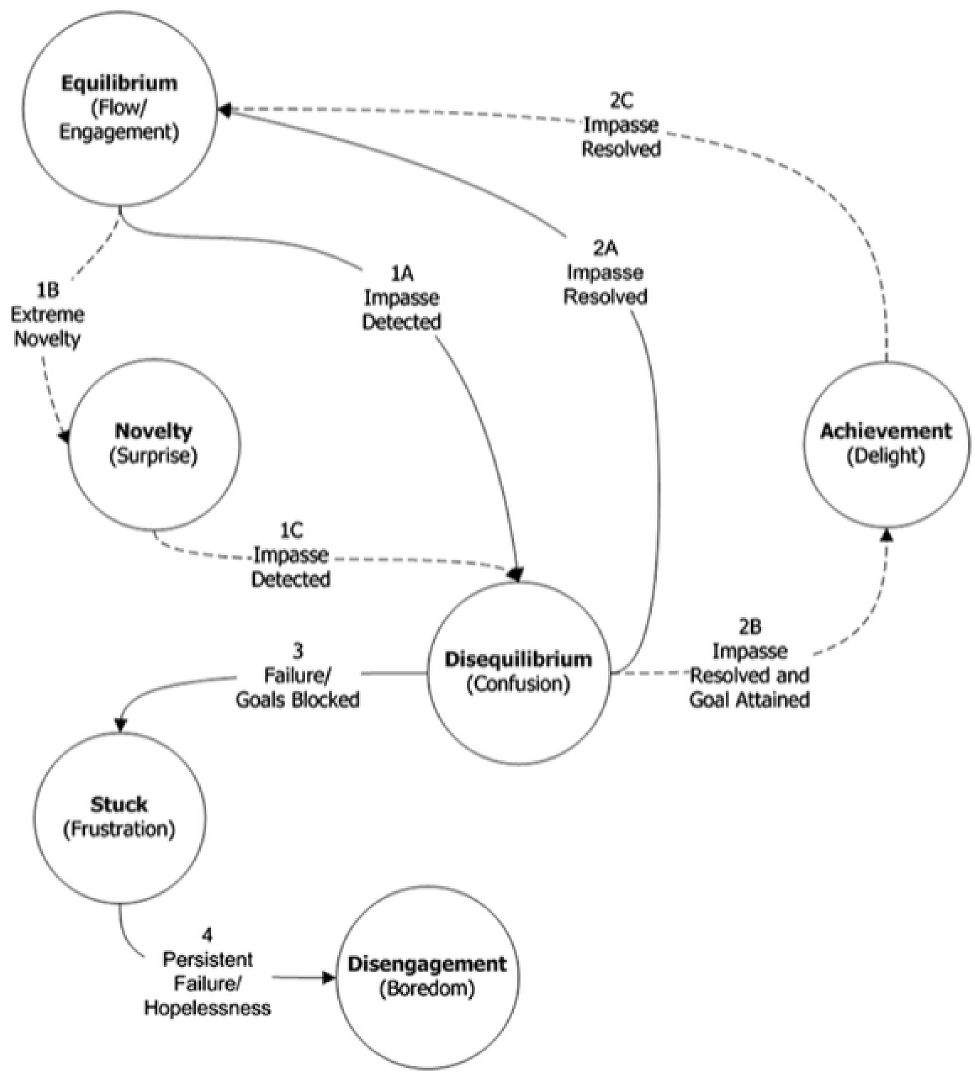


Figure 3.2: D'Mello's and Graesser's model of Affect Dynamics

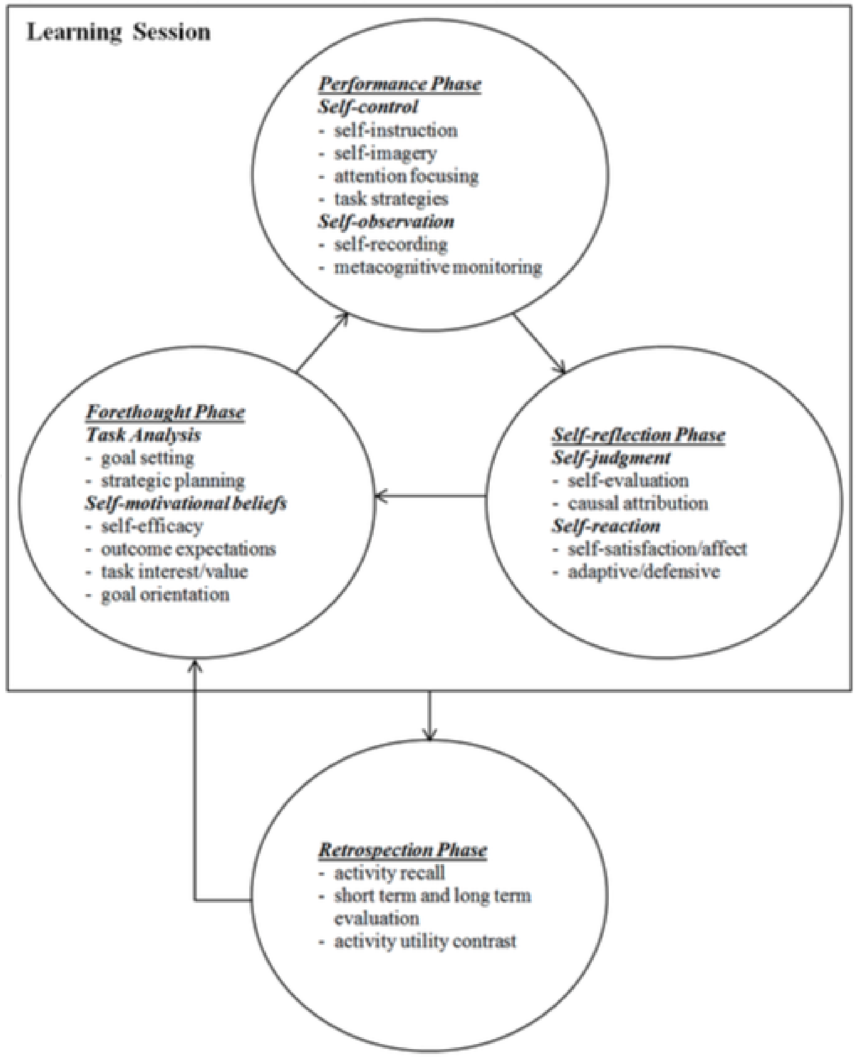


Figure 3.3: Inventado's improvement of the Self-Regulation Model

### 3.2 Machine Learning Techniques in Modeling Self-Regulated Learners

It is essential to use classification algorithms in order to perform prediction tasks. These can be used to model activities and patterns from data such as student gestures and movement. In this study, Multi-Layer perceptron and Hidden Markov models will be used to generate models for student activities, policies and transition data.

#### 3.2.1 Multi-Layer Perceptron

A multi-layer perceptron is a feedforward artificial neural network that maps sets of input data onto a set of appropriate outputs. It has multiple layers, an inner layer referred to as a hidden layer or hidden nodes. It employs non linear activation as opposed to other machine learning classifiers and utilizes a supervised learning technique. This technique is called backpropagation (Rosenblatt, 1961). Learning takes place by changing the weights, where it is called Gradient Descent.

$$\Delta W_{ji} \quad (n) \quad = \quad -\eta \quad \frac{\delta \varepsilon(n)}{\delta v_j(n)} y_i(n) \tag{3}$$

The formula for Gradient Descent is shown above where $\eta$ is the learning rate, $y_i$ is the output of the previous neuron. Multi-layer perceptrons have been widely used for an array of application such as speech recognition, image recognition, machine translation and even activity recognition (Kwapisz et al., 2011), (Mantyjarvi et al., 2001), (Pirttikangas et al., 2006). The MLP can be used to build an initial model to determine if from the given data, a user-specific or a general model will have to be built.

By establishing a standard learning rate $\alpha$, a multi-layer perceptron can be used to model the activities performed by the students since the data can serve as inputs. The changes in the weights accommodate the differences in the movements of the students. With regards to the changes in their activities, its aligned transitions and policies, a separate classifier must be employed.

#### 3.2.2 Hidden Markov Models

Hidden Markov Models (HMM) is a classifier mostly used for a system of processes with hidden observed states such as states in between. A simpler version of HMM referred to as Markov models considers transitions and their probabilities as its input parameters. HMM are used for several areas and have been applied as well on Human Activity Recognition (Brand et al., 1997) (Yamato et al., 1992).

Most HMM employ the Baum-Welch (aka Baldi-Chauvin) algorithm which uses an expectation-maximization algorithm. This way, it is partially-able to derive the maximum likelihood estimates within the HMM given the set of output sequences (Rabiner, 1989).

$$L(\theta; X) \quad = \quad p(X|\theta) \quad = \quad \sum_{z} p(X, Z|\theta) \tag{4}$$

As seen from the formula, X refers to the set of unobserved latent data (which we can refer to also as missing values), Z as the vector of unknown parameters theta and the likelihood function L. The maximum likelihood estimate is produced as an output of the equation and augmenting the classification and estimation done by a markov model.

Since the transitions and patterns of students cannot be quantified or constraint in a single time instance $T_n$, multiple instances should be used and HMM contains a markovian property that can handle such data representations.

## 3.3 Modelling Student Activities

Student activities captured with SideKick can be used to model their tasks, aligned policies and transition information. In certain studies, these have been classified as contextualized action sequences (Burton & Brown, 1979). And a specialized technique has been used to further quantify these sequences, especially for the policy as derived with the help of a profit-sharing algorithm (Arai & Sycara, 2000).

### 3.3.1 Contextualized Action Sequences

Student Activities have been modelled in several studies and areas aimed in being used for many applications. The study of Burton & Brown (1979) covered coaching students in scenarios referred to as informal learning activities. It provided an in-depth view of the philosophy behind such learning environments, the kind of diagnostic and modeling techniques that should be used and the strategies that can be used to infer a student's shortcomings. The study listed tutorial strategies and other feedback techniques that should be given to the students.

Studies having the need to model the activities of students have been an ongoing topic for learning experts. It can be assumed that the paradigm has provided several opportunities and areas for improvements on both the pedagogical strategies of teachers and the personalized learning habits of students. As technology has grown, avenues for learning and self-help have been more available to the point that they have been ubiquitous.

From mobile phones, to sensors, background applications, both students and teachers have found opportunities to improve their learning habits. In a study by Kim et al. (2006), they have modelled the activities of their students with the use of information retrieval and language processing techniques to automate their tasks - minimizing the tasks of both teachers and students involved in online discussions.

Over-all, the goal of studies like these is to model, measure and generalize the change or the growth in student learning (Hill & Rowe, 1998). One way of doing this is thru Contextualized Action Sequences. Kinnebrew & Biswas (2012) has been a pioneer of using this approach where he sometimes refers to it as Differential Sequence Mining (Kinnebrew et al., 2013). It is a technique concerned with finding statistically-relevant patterns among events such as sequences (Mabroukeh & Ezeife, 2010).

### 3.3.2 Profit Sharing Algorithm

Profit Sharing takes as input a sequence of observation-pairs $(O_t, A_t)$ which means performing action $A_t$ when $O_t$ is observed. For this study, P. Inventado (2014) defines $O_t$ as the state of the student being observed. The profit sharing approach works as a model-free reinforcement learning algorithm that handles non-markovian environments. The equation below uncovers motivating activities as an approach in learning domains:

$$W_{n+1} \quad (O_t, A_t) \quad \leftarrow W_n(O_t, A_t) + f(R, T) \tag{5}$$

Given this, the profit sharing approach employs a policy function that is computed as below:

$$f_{n+1}(R, T) = (R - W_n(O_t, A_t))(0.3)^{T-t} \tag{6}$$

Where W represents the weight, $O_t$ is the observation at time t, $A_t$ the action at time t and $f(R, T)$ the reward function relative to the total reward and time (Arai & Sycara, 2000).

# 4 Research Framework

This chapter discusses the steps performed with the aim of modelling the activities of self-regulated learners. The different steps covered data preparation and collection, the usage of sidekick and the modelling activities undertaken from P. Inventado (2014). The figure below shows the over-all framework of this study.

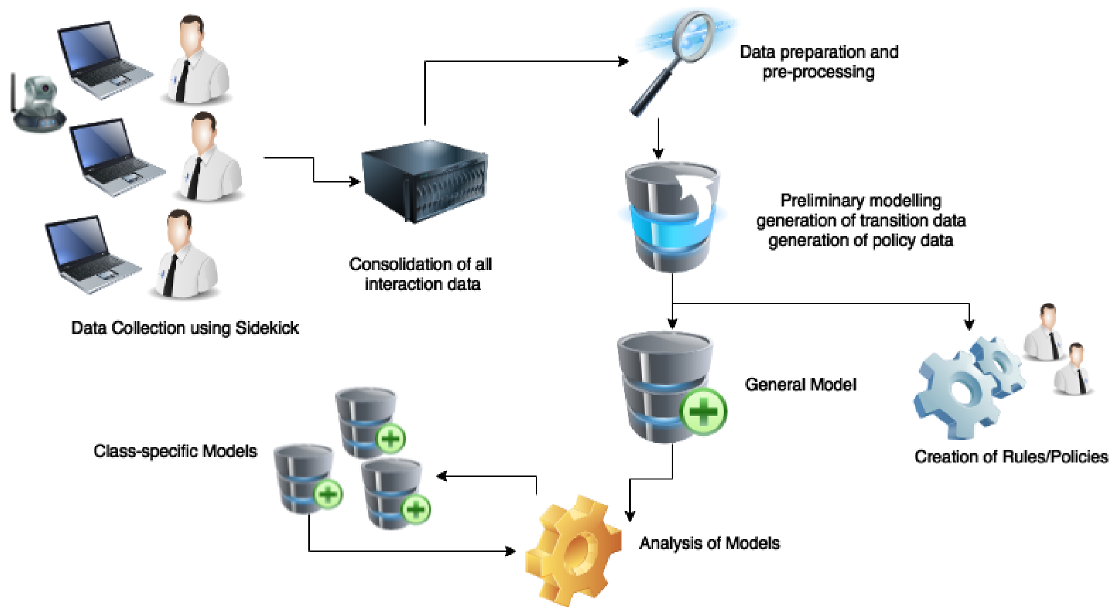


Figure 4.1: Overall Research Framework

## 4.1 Sidekick

The Sidekick learning support tool as developed by P. Inventado (2014) aids the data collection processes in this study. It was developed to monitor the activities of students in environments called self-initated learning scenarios where they get to study on their own and the presence of a 'supervisor' is not critical in the process. According to P. Inventado (2014), the tool enables learners to undergo three stages namely (1) interaction stage, (2) analysis stage and the (3) evaluation stage. All these stages are critical in the self-regulation model as defined by the framework as seen from 3.1.

In the Interaction stage of Sidekick, students use the computer like how they usually do, with the Sidekick application running in the background. Their necessary interaction data are collected by the tool. Such data include mouse clicks, mouse movement, mouse wheel movement, keyboard clicks, webcam face captures and desktop screen captures. Additionally, the title of the background application and its respective process (i.e. Web Browser on chrome.exe) are collected as well as part of the interaction data. In this stage, the students are encouraged to define a learning plan: a usual step found among self-regulated learners especially when learning without the use of computers. These interactions and the definition of the learning plan enables the learners to recall what transpired during the learning session. These subprocesses aid in the succeeding stages of analysis and evaluation stages.

Students are given their own copies of the Sidekick tool in their local machines. They are to create an account, and select their preferences which provide additional information for this study.

These credentials are used by the students to login. The machines have a local copy of the repository and database which contains the interaction data. These files are submitted and are explained further in the Data Preparation subsection of this chapter. On their first login (see Figure 4.3), a questionnaire that measures the autonomy index of the students welcomes the users (see Figure 4.4). This questionnaire was patterned after (Deci et al., 1991) which measures and defines a student's level of being self-regulated. The said questionnaire aims to define the operational definition of what a self-regulated learner is on the context of the profile of the students who are respondents in this study. After answering this questionnaire, before each session, the students are welcomed by another screen (see Figure 4.5) which enables them to learn and plan our their learning sessions. In this module, students can define their learning goals for the session by simply keying-in their inputs on the forms displayed. Students have the option to load their previous learning goal as well (for sessions beyond the first). They also get to configure further the settings and test if the necessary hardware are working. Here, they get to review if their webcam works and if the screen captures work just fine. From this point forward, the students can begin the interaction stage and have their data collected by pressing the "Begin Learning" button. As soon as the learning session begins, the collection module built in sidekick begins the acquiring the interaction data. The data is on a per-second interval.

At the end of each session, the annotation screen shows up (see Figure 4.6) which enables the students to analyse and evaluate their past actions during the recently concluded learning session. The analysis stage as defined by (Deci et al., 1991) encourages students to engage in an activity where they get to recall both short and long term evaluation processes. This aids in the implementation of the Retrospection phase proposed by P. Inventado (2014) in his model as seen from Figure 3.3. The analysis stage proceeds with having the students recall and enter their activities (tasks), affects and the level of contribution they performed. This is done by highlighting portions of the timeline that includes the range and duration of a certain selected performed activity. After highlighting, students identify the task they have performed by clicking one of the categories defined (see Appendix A.3).

P. Inventado (2014) identifies 11 pre-defined task categories that self-regulated learners perform when studying on their own. These pre-defined activities were defined from research that investigates student's learning tasks (Laurillard, 2013) with the use of a specific learning system (Kinnebrew et al., 2012)(Azevedo et al., 2011). Below are the descriptions of each 11 category in detail (excerpt from (P. Inventado, 2014)):

- Make learning plan - involves identifying, ordering and strategizing what tasks will be performed and how they will be performed to achieve the learning goal; this is usually done at the start of the learning task or whenever a learning plan has been completed
- Modify learning plan - involves reviewing the learning plan for tracking progress or identifying what to do next and the removing or replacing previously planned learning activities which are no longer applicable; this may also include adding new activities that will help achieve the learning goal discovered in the course of implementing the learning plan
- Read and understand information - involves reading and understanding information to solve the current problem; although information search involves reading information, this category is differentiated by the intent of the activity wherein the goal is the understanding and learning new information; different mediums may be used to read information such as printed materials, computers or mobile devices
- Practice skills - involves activities that hone the mastery of a skill the stu- dent already knows which may come in the form of test-taking, answering sample problems and the like
- Search for info - involves finding and filtering information needed to solve the current problem; usually involves the use of a web-based search engine but may also refer to finding information from books or other medium
- Apply acquired info - involves the use of information acquired from searching or reading information to solve the current problem (e.g., summarize a paper so it can be included in the review of related literature section of a student's thesis)
- Use previous knowledge - involves the application of previous knowledge to solve a problem (e.g., create and use a computer program to make it easier to solve a complex

equation instead of doing it by hand)

- Take notes - involves storing important information acquired while doing an ac- tivity; different mediums may be used to store information such as a notebook, a computer software or a mobile application
- Review notes - involves the retrieval and use of previously stored notes using the medium they were stored in; may refer to notes taken outside of the software such as notes taken during a lecture
- Seek help - involves communicating with other people to get help with a prob- lem that the student has difficulty solving; students can communicate in many ways such as face-to-face verbal communication, sending/reading emails and send- ing/reading instant messages through a computer or mobile device
- Off task - involves any activity, which may or not be entertainment related, that is not related to achieving the learning goal for the session.

.
It is important to note that Off-task activities are tracked and are considered vital since the students are unsupervised while they are using sidekick. This ensures that the data is as naturalistic as possible. Coincidentally, certain researches support as well that off-task activities are not always primary distractions to students (P. S. Inventado et al., 2011). During certain points, it has been observed that off-task activities may help alleviate stress thus possibly affecting motivation in the long run (Sabourin et al., 2011).

This selection is aided by the viewing and displaying the captured images of the face from the webcam and the screen captures of the desktop in the synced time. After selecting the task performed based on the pre-defined activities, learners then get to pick which type of academic emotion (affect) they think they have experienced during the said task. This is where the captures of the webcam would be most useful as learners, during annotation can use as reference these facial captures in determining which affect they have experienced. Students have the option of identifying and informing the sidekick tool that they may have experienced more than one academic emotion while performing a single task. These inputs on affect are stored and added into the Process Transition Database (PTD).

Lastly, the students are asked to make a third annotation which helps them reflect on their past activities and tasks. This enables them to give a hindsight review of what has been done and allows them to gauge how much of the task has actually contributed to the learning task at hand. This enables the reflection stage. Students select from a 4-point scale (1 the lowest and 4 the highest) on how much did that specific task has contributed to the overall learning task. This cycle is repeated until the timeline has been

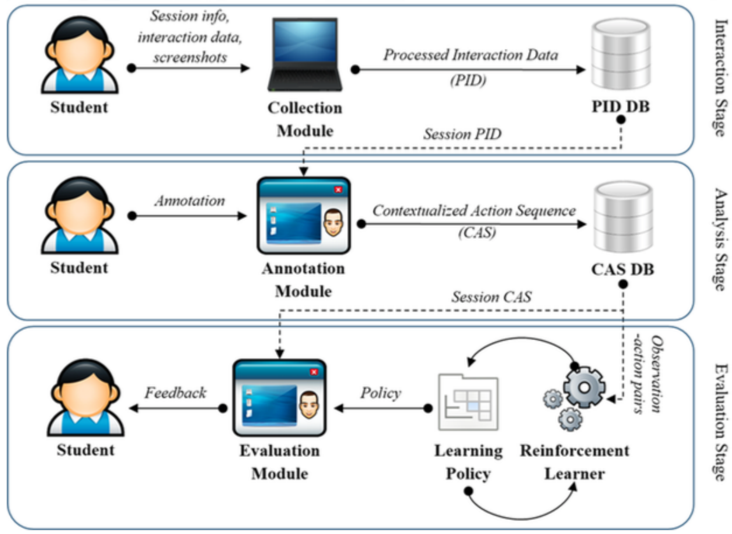


Figure 4.2: Framework of Interaction Stage of Sidekick as derived from (P. Inventado et al., 2012)

fully-annotated and all the interaction data have been synced in the Process Interaction Database (PID). With the help of the evaluation stage, the students are able to contrast two or more activities given a single situation. This encourages them further to reflect on factors that may make an activity preferred or less preferred depending on the learning task at hand. Inside Sidekick, a reinforcement algorithm is employed which deals with the identification of rules for selecting the best action in a given state (Sutton & Barto, 1998), that ensures motivation in the long term (P. Inventado, 2014).

After this stage, students are asked to answer a post-session questionnaire where their evaluation is concretized by their open-ended answers to the questions. This process enables the students and their evaluations to be stored as reminders which can be used to help students in suggesting changes in their behaviors in succeeding learning sessions. P. Inventado (2014) believes that the continuous reflection and re-evaluation processes as observed by the three stages of sidekick enables students to find better ways to become more motivated to learn.

## 4.2 Data Collection and Preparation

The methodology is designed to be in an uncontrolled environment were student-respondents install a local version of Sidekick in their machines. The respondents were explicitly-


Figure 4.3: Login screen of Sidekick

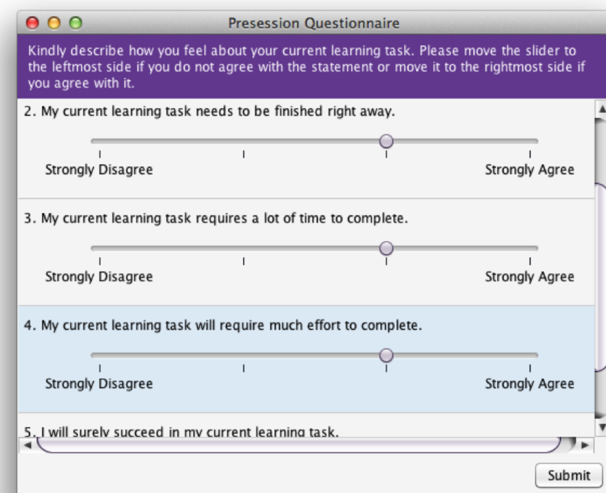


Figure 4.4: Pre-Session Questionnaire that measures tendencies of Self-Regulation

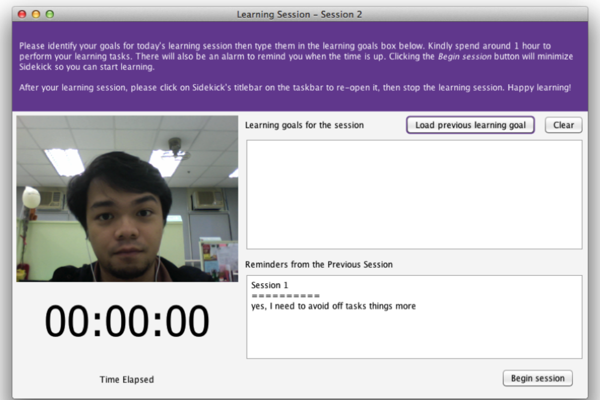


Figure 4.5: Learning Plan Planning Module before each session

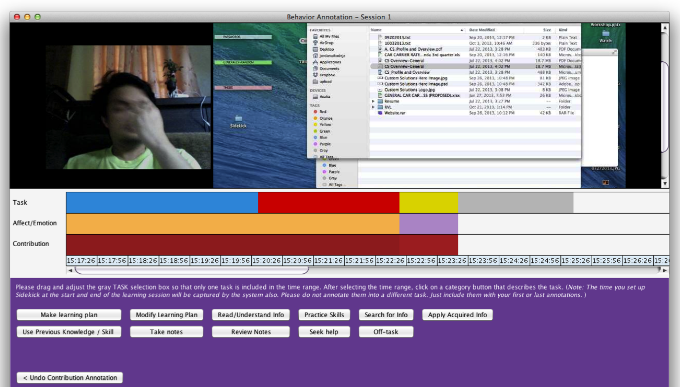


Figure 4.6: Annotation Screen of Sidekick after a learning session

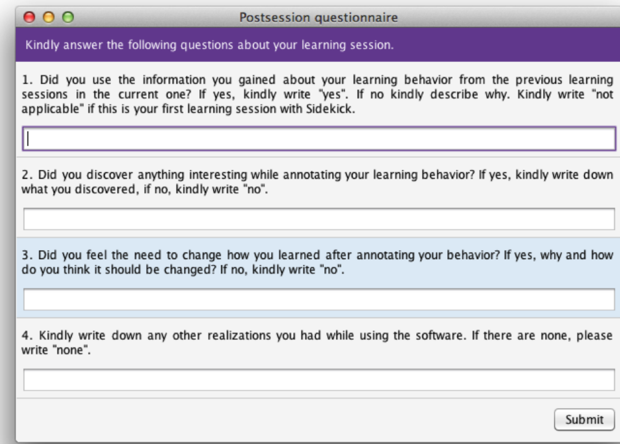


Figure 4.7: Post-Session Questionnaire in Sidekick after a learning session

instructed that they do data collection in an environment where there is no agent or supervisor that can control the learning process. The data collection methodologies was implemented in an Object-Oriented Programming class at the University of Santo Tomas, Manila, Philippines. The students, as an incentive were given bonus points credited to them in exchange of their service and participation in the data collection phases of the study.

The class of 39 students attended a one-hour presentation and demonstration about Sidekick where they were oriented, instructed and given a demonstration on how to use the software. The methodology allows students to use sidekick whenever it is possible especially when they are doing tasks related to the course. 25 students volunteered to participate and were instructed to complete 10 learning sessions.

The students were informed and were registered into the data collection phases with the use of the Consent to Participate in Research Form (see Appendix **??** ). Their rights, private data and other logged information were guaranteed to be preserved and protected by the study and by De La Salle University.

After a two to three week period of data collection, the students submitted their local copy of the database which contained the repository and the individual PID, PTD and POD's in their machines. These were collated and organized based on the needed models to be produced in the study. These repositories were accessible with the use of a Hypersonic 2 (h2.jar) that served as the database management system of the said repositories.

Without the credentials and login information, the data would just appear jibberish to anyone unauthorized to access the data.

The data per student were collated and compiled into organized data sets which represent the Interaction Data, Policy Data and Transition Data. Additionally, the interaction data was also exported per session per student in order to be able to acquire a finer-grained analysis of the interactions later on. The said datasets were prepared in .csv (comma-separated values) format so that they can be imported and fed into the Data Modelling stages of this study.

Additional parsing and cleansing were done to the data to uniform commonly-used applications (WebBrowser: chrome.exe, firefox.exe, iexplore.exe) and eliminate special characters that might hinder the performance of a machine learning classifier.

## 4.3 Data Modeling

The individual activities of each student respondent were identified a historical learning behaviors. As such, they can be categorized as Contextualized Action Sequences. These sequences can run on a model-free environment without compromising data quality and enable multiple domains to converge without the need for a Markovian property (Arai & Sycara, 2000). As considered, this approach was selected with the understanding that the data was acquired in an uncontrolled environment and despite the self-regulatedness of each respondent, it was still uncertain to deal with human behavior in a deterministic approach.

This is where the Profit Sharing Algorithm comes in with the treatment and the modelling of the activities in contrast with the motivation levels of the students, alongside their transitions in consideration. An observation-action pair ( $O_t, A_t$ ) is received as an input of sequences where $O_t$ is the observed from a performing action $A_t$. According to (P. S. Inventado et al., 2011), $O_t$ refers to the state of a student as described by the activity it is in. The activities $A_t$ are categorized into the following rules:

- Short - activities performed for less than five (5) minutes
- Medium - activities performed between five (5) to ten (10) minutes

- Long - activities performed beyond ten (10) minutes

In return, these observation pairs are given a function which returns a value specified by reward $R$. These were derived from Arai & Sycara (2000) 's ARCS model. With this, a weight which measures the motivation level of the students is returned as seen in the equations below.

$$W_{n+1} \leftarrow W_n(O_t, A_t) + f(^T_n, t) \tag{7}$$

$$f(R_t, t) = R(\frac{1}{L})^{T-t} \tag{8}$$

$$\forall t = 1, 2, 3 \cdots, T. \quad L\sum_{j=0}^{t} f(R_j) < f(R, t) \tag{9}$$

$f(^T_n, t)$ is the variable used for a credit returned from an assignment function where $t$ is the rule's position of time with respect to the current episode $T$. Following this approach, the resulting weight values may tend to be updated more than once since there are multiple instances of $t$ with respect to $T$. From here we are able to identify the set of all rules and their weights as the policy itself as contained in the PPD. According to Arai & Sycara (2000), the policy is considered rational and a guaranteed convergance to a solution following the credit assignment function as well satisfies the rationality theorem as seen in the equation above. The variable $L$ represents the number of possible actions in a certain state configuration.

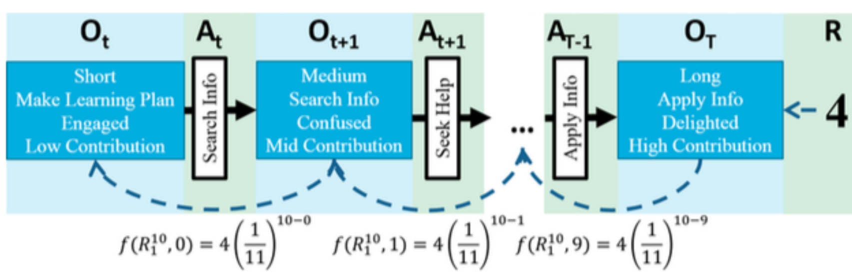


Figure 4.8: Observation-Action Pair Updating Cycle as illustrated by P. Inventado (2014)

To provide an improved illustration of the profit-sharing algorithm, we refer to Figure 4.8 as illustrated by P. Inventado (2014) in his study. The process iterates the computation of the profit sharing algorithm starting from an initial state. In here, the student feels Engaged as the affect while he makes a learning plan for the current session. Given this example it can be seen that the student feels that the task of creating a learning plan might not be contributing much to his current learning task. Following a given short time, the student decides to proceed to performing a different action depicted by $A_t$ (searching for information as a new task). The parameters are then updated as the observation and activity pairs changes, thus with the help of the profit-sharing algorithm adjusting the computed resulting weight until a final weight or policy is arrived at. All these values are relative to the ARCS scale initially defined by Keller (1999). These values particularly mean:

- Attention - engagement of the student to the said activity
- Relevance - relation of the activity of the student's goals and intentions
- Confidence - level of achievement by the student
- Satisfaction - how the said activity applies to the real-life applications

It is important to note that with this function, the values of the weights in the policy may change over time. In consideration of this, the latest value of the policy can be used as a reference point to compare the learning session the student is in and this can be used to measure how his motivation levels have changed over the series of activities. All these policies and weights were computed per student respondent and generalized into another dataset. These formed the Process Policy Database (PPD) as earlier mentioned.

On a different note, all the interaction data from the PID were compiled as earlier mentioned. The interaction data set was fed into a machine learning classifier specially-designed for contextualized action sequences. This MLP was patterned after Ware (2015)'s classifier which was further modified to fit this study. Specifically, the typical multi layer perceptron employs the sigmoid function as categorized by the equation below:

$$\gamma(v_i) = tanh(v_i) \quad and \quad \gamma(v_i) = (1 + e^{-v_i})^{-}1, \tag{10}$$

Where the former describes a hyperbolic tangent while the latter describes a logistic function ranging from 0 to 1, which is similar in shape (Rosenblatt, 1961). The specialized MLP network as created by Ware (2015) employs the same function but allows non-nominal and string-based data such as the labels of sequences and tasks to be part of the attributes in the data set. The standard values for learning rate, the number of hidden nodes and momentum were set with the use of RapidMiner. The experiment, training and testing were deployed on a process which consists of a 10-fold cross validation with 500ms as the minimum training time per iteration.

Additionally, several machine learning tasks were employed to provide a benchmark on the results and performance of the model. These included the Naive Bayes and decision tree classifiers on all groups of the data sets.

### 4.4 Data Analysis

The transition and policy data of the students were divided based on the three different classes: low, moderate and high self-regulation. These classes were based from the answers of the respondents to the modified version of Deci et al. (1991)'s Self-Regulated Questionnaire. It contained a 7-point likert scale and from this categorized three levels of self regulation. The number of the respondents who will form the samples for each sub class will be based on their scores in the said questionnaire. The range of scope of each class of self regulation will be based on the computed standard deviation and median values.

$$\sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N-1}} \tag{11}$$

$$\bar{x} = \frac{1}{2}n + 1 \tag{12}$$

The computed median shall be the benchmark for the moderate self regulation cluster and it shall be distributed in $n$ instances where each jump is based on the value of $\sigma$. All values below the first instance shall all fall under the category of low self regulation while all values above the first instance from the median shall be categorized as of high self regulation.

When the sample have been extracted based on the values corresponding to the likert scale equivalent, the policy and interaction data will be divided based on these samples.

With the policy values, the transition data will be modelled based on the instructions defined in Section 4.3. The combined record instances in the transition will be processed following the formulas aligned with the Profit Sharing Algorithm. From here, a standard sequence that produces a motivation (weight) value will be generated. This sequence would be consists of the parts namely: the initial, midway and final sequence. Specifically, the instances from these states would be extracted and be referred to as the rule.

By tabulating these results and generating the respective line graphs, the growth or change in the motivation levels, from sequences extracted from the 1st, 5th and 10th sessions will be observed. This would give a general overview of the motivation levels of the students belonging to which class. At the same time, an additional insight is given to which types of activities lead towards a more positive growth or change in the motivation levels of that student category. This way, an overview of the best habits and traits or sequences of activities that can be performed to ensure a learning session that contributes to earning or maintaining motivation. This process is iterated until a good learning sequence is identifiable. This is usually determine when upon following the profit sharing algorithm, the motivation values in the form of weights have grown significantly.

With regards to the modelling of the interaction data, the combined data from all three levels are fed into the machine learning task to be able to determine if a general model can be designed. Second, regardless of results and accuracy, class-specific models will be trained. The performance of these models, and the merged specific-classes will be analyzed in contrast to each other to determine the best model that can be used for these self-regulated learners.

When the models have been trained and evaluated, data analysis shall be performed. The performance of the model will be captured considering the accuracy, root mean squared error, precision and recall and kappa statistic. These statistical values along with the values in its confusion matrix will be reviewed to see if they reach within acceptable values (Michalski et al., 2013). The analysis will first be performed on a general model which contains all the interaction data covering the three classes of self-regulated behavior.

Class-specific models which were clustered based on the scores of self regulatedness by the student respondents will be developed as well. These models will undergo the same analysis that was treated with the general model.

# 5 Results and Discussion

## 5.1 Student Profile

The experiment was conducted in an introductory programming class composed of 39 students. The class was handled by the same instructor. From the 39 students, 25 students participated and voluntarily-submitted their data. They were aged between 18 to 20 years old, 20 were males and 5 were females.

The volunteers who participated had varying levels of motivation based on their answers in the self-regulation questionnaire used. A scale of points between 1 to 7 were patterned after Deci et al. (1991)'s questionnaire which determined one's level of being self regulated with 1 as the lowest and 7 being the highest. The average score in the questionnaire among the participants was 5.3 ($\sigma = 1.1, \bar{x} = 5.1$) indicating that there is a diverse mix of intrinsically and extrinsically motivated students within a normal distribution.

Table 5.1: Categorization of Self-Regulated Learners

| Score in Likert Scale | Equivalent | No. of Respondents |
|---|---|---|
| 1-3.9 | Low | 5 |
| 4-5.9 | Moderate | 12 |
| 6-7.0 | High | 8 |

Additionally, a bipolar classification was performed on the students. Instead of the applied 3 categories, the students were divided into two classes of learners with the median 5.1 as the basis. As a result, 13 students were classified into the lower half of the self regulation scale and the remaining 12 were classified into the upper half of the said scale. This was employed to provide additional insight if the model would prove significant changes given the difference in the scale and sample size. The corresponding results will be seen in the latter parts of this section.

## 5.2 Activity Recall

In the data collection, one part is to recall their activities by annotating them with the use of Sidekick. Students spent an average of 58.1 minutes in every learning session. On average, they have annotated 4.4 transitions in each session. The most used activity per student is the Practice Skills task which on total consumed 47.8 minutes per student (80%

of a session on average for the entire 10 sessions). Students in general spent the least time on Taking down Notes; such task took a total of 21 minutes per student (35% on average for the entire 10 sessions) (see Figure A.3). This can be directly-linked to the presence of the internet connection, the ubiquity of information and the modernization of the use of slides in contrast to the use notebooks and writing by pen and paper. Additionally, the most noted affect was the Engaged affect which appeared in a total of 500 instances. Below are other details on the interaction data of the respondents.

Table 5.2: Interaction Data Statistics

| 3 Interaction Data | Average | Max | $\sigma$ |
|---|---|---|---|
| mouse click count (per session) | 43 | 796 | 1.3 |
| mouse movement distance (dpi) | 40287 | 819982 | 1350.7 |
| keyboard input count (per session) | 415 | 21250 | 5.7 |
| contribution level (per session) | 3.1 | 4 | 0.8 |
| face captures (in seconds) | 197 | 3149 | 0.5 |

In comparison as seen in Figure 5.2 there is noticeable differences between the average time students spend seeking for help and searching for information. This could mean that the respondents were independent enough to not ask the help of others for instructions (possibly from elders, mentors, or classmates) as compared to the time it took them to search for information on their own. This can add additional value to their being self-regulatedness were extra points to being independent makes them more regulated and possibly motivated.

The results presented in Figure 5.3 show that among the 11 categories of activities, students spent the most time practicing their skills as evidenced similarly in Figure 5.2. However, it is also notable that Off-tasks form a large average time spent by these self-regulated learners. In the latter part of this analysis, it can be understood further why and how Off-task category is prevalent and is related to maintaining motivation among these self-regulated learners. In conclusion, even if the "Practicing Skills" category is the most-used "useful" task, it is difficult to ignore that these students, both intrinsically and extrinsically self-regulated learners spent an ample amount of time doing off-task things while on their own. This can lead us to conclude that the data we have collected can be as naturalistic as possible and that this can testify on the actual self-regulatedness level of each of our respondents.

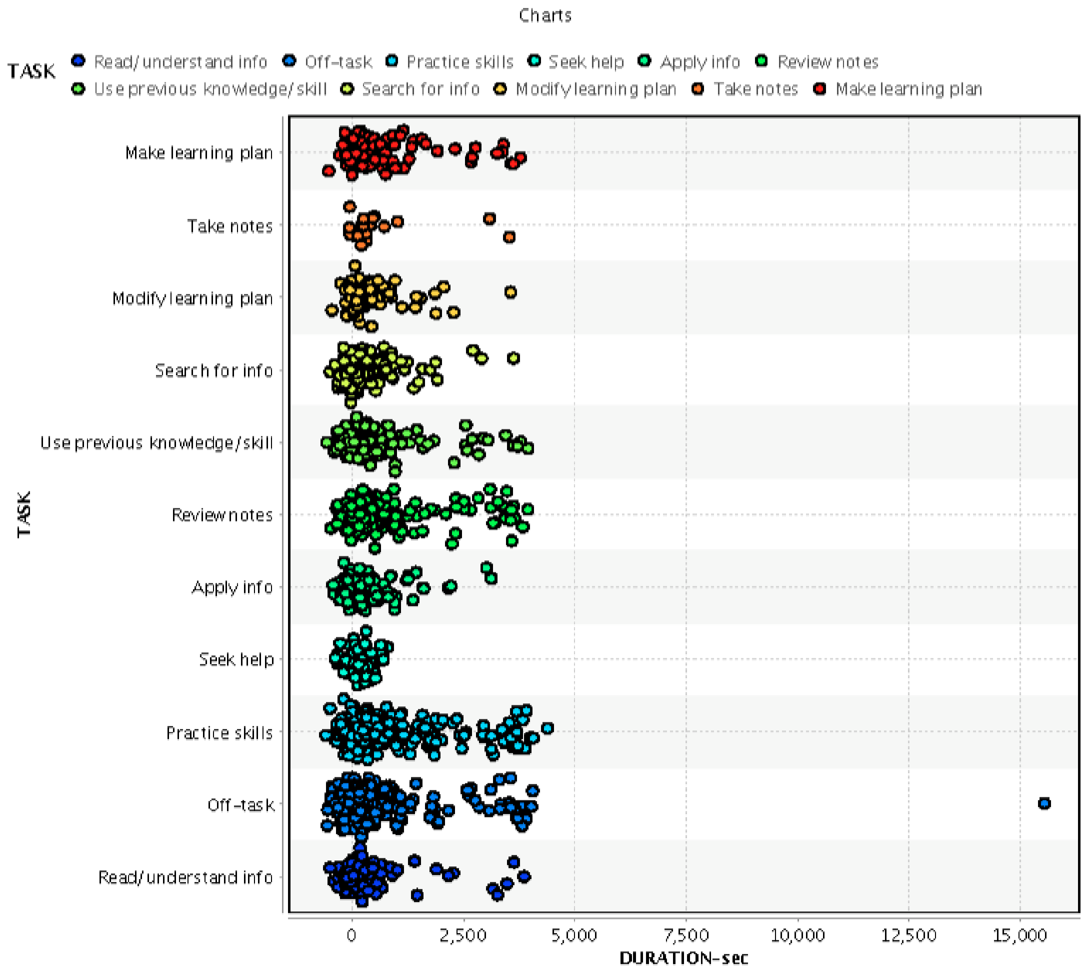


Figure 5.1: Scatterplot of Transitions and their average duration

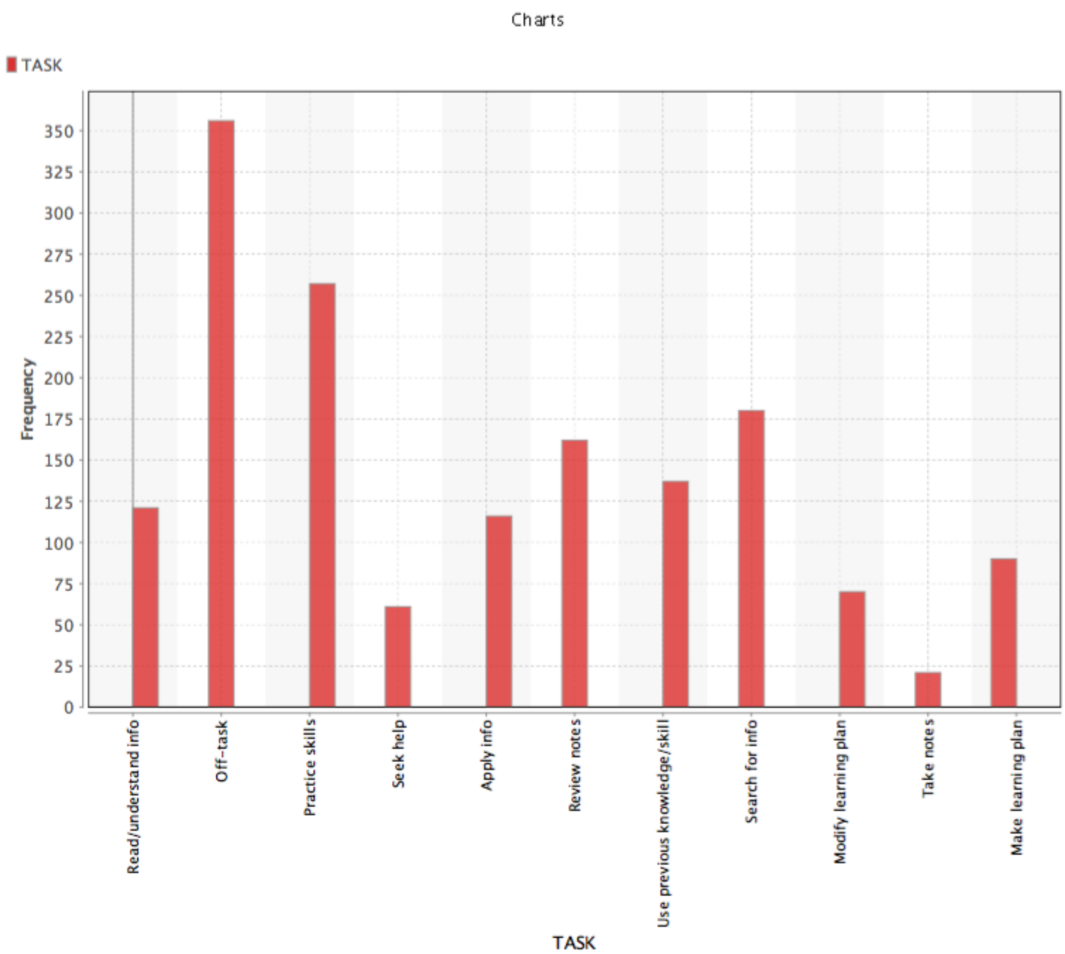


Figure 5.2: Histogram of Transitions and Average Time Spent per student

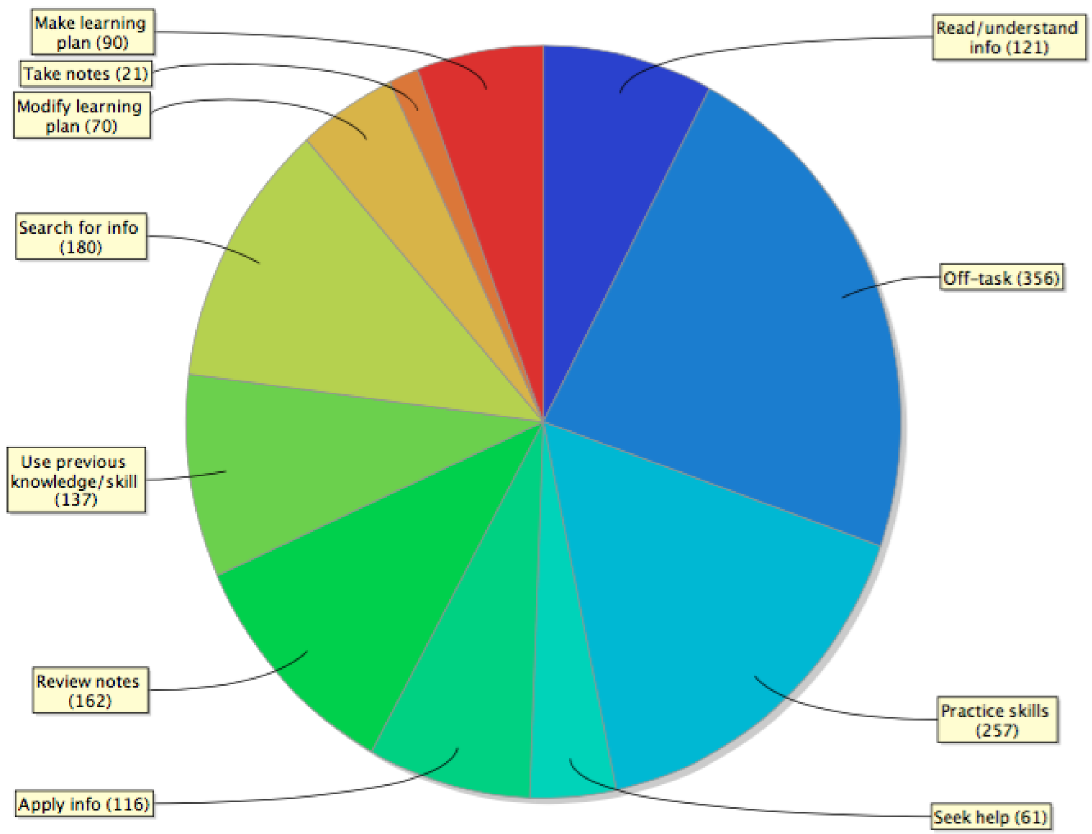


Figure 5.3: Pie chart detailing the task distribution on average per student

## 5.3 Learning Behaviors

Portion of this study is to determine the different activities that students undertake. These were derived by combining the policy and transition data of the interaction of the students while using sidekick. With the use of the data modelling techniques, the following attributes were pulled out to form a behavior model:

- Session Number
- Initial state on that session number
- Final state on that session number
- Type of activity currently being performed
- Affect during the said activity or task
- Weight of the said activity in relation to the motivation of the students

The effectiveness of the reflection phase performed by the students while on each learning session were captured. These were fed into the ARCS model of computing using the Profit-Sharing Algorithm (Keller, 1999) where the ratings were computed at the end of each learning session. These ratings were tabulated and evaluated per student to determine the over-all effect in the self-regulatedness of a particular student. Among the 25 respondents, the highest recorded policy weight is 15.3 ($\overline{x} = 0.5$). To better generalize the learning behaviors observed, 3 types of students were clustered falling under the categories of High Self-Regulation, Moderate Self-Regulation and Low Self-Regulation. These values are seen to have changed as displayed in the succeeding tables:

Table 5.3: Learning Behavior Patterns of Student Class 1 (Low self-regulation)

| Sess No | Source state | State description | Action Performed | Weight |
|---|---|---|---|---|
| 1 | Initial | long-MODIFY-2-Confused | MODIFY | 0.0 |
| 1 | Final | short-MODIFY-2-Bored | MODIFY | 0.0 |
| 5 | Initial | long-MODIFY-2-Frustrated | MODIFY | 0.0 |
| 5 | Final | medium-MODIFY-3-Confused | MODIFY | 0.3 |
| 10 | Initial | long-MODIFY-3-Bored | MODIFY | 3.0 |
| 10 | Midway | long-MAKE-3-Confused | MAKE | 0.3 |
| 10 | Final | long-PRACTICE-3-Confused | PRACTICE | 15.3 |

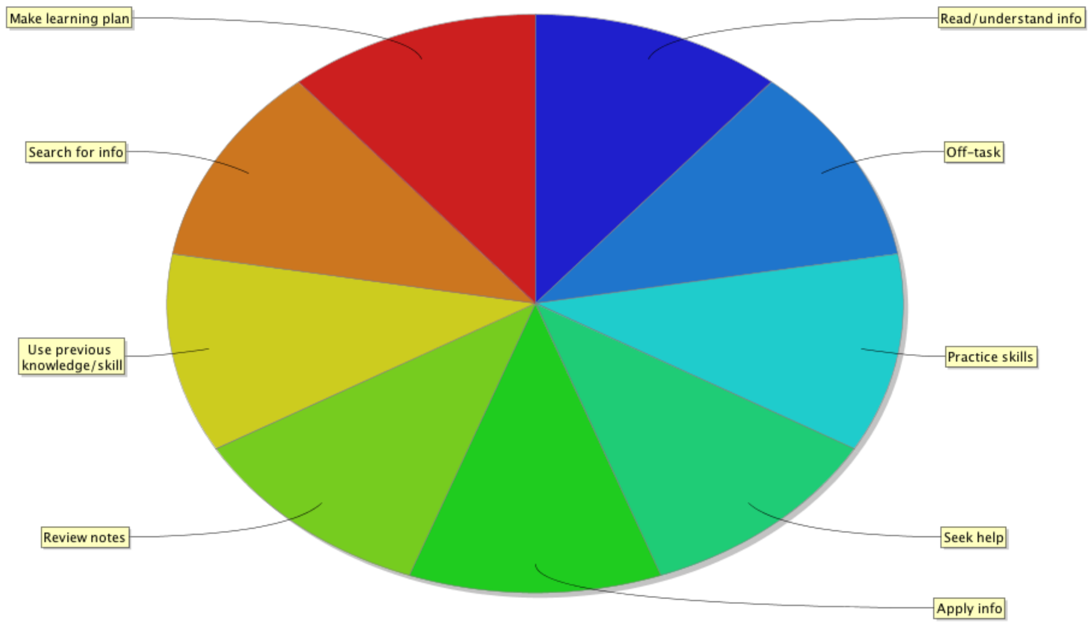


Figure 5.4: Pie chart detailing the task distribution among Class 1 learners

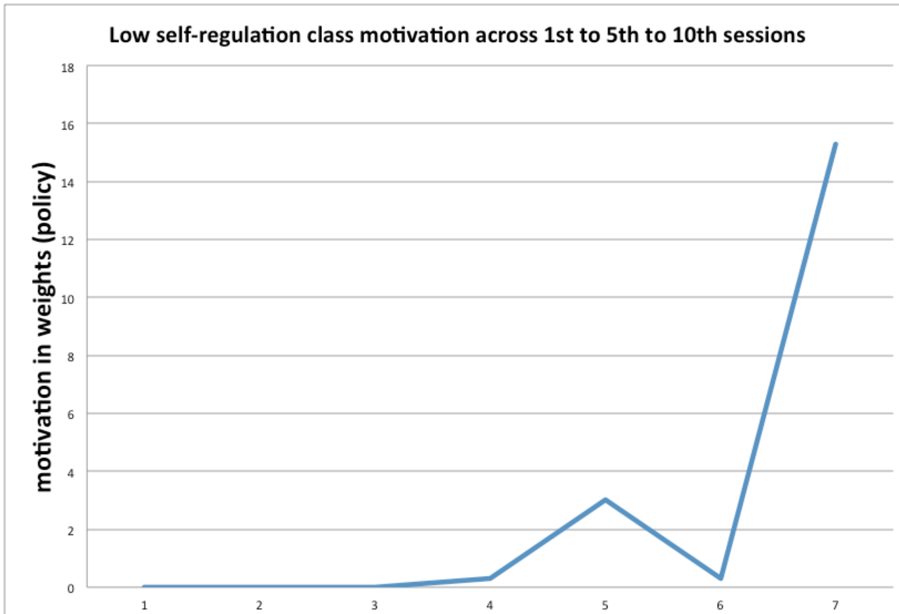


Figure 5.5: Changes in Level of Motivation of Low Self-Regulation Learners

Tables 5.3, 5.4, and 5.5 presents the policies generated with the help of the profit-sharing algorithm featuring their motivation levels from the 1st, the 5th and the 10th sessions of their learnings. The figures 5.5, 5.7, 5.9 also displays these values highlighting the changes or the growth in the motivation levels of each type of student learners. The original data covered multiple instances of the activities performed by the students and their corresponding weights. The values that contained the highest weights and is common among the same type of learners were acquired and used as the value for the state description. The source state describes the state where the description begins. The Action Performed column describes the succeeding activity the student performs from the activity described from the state description column. The weight indicated the level of motivation that the described activity has made the student feel.

These tables show the changes in the learning policy of the students over different sessions. The 1st session as the initial session, the 5th session as the midway breakpoint and the 10th session as an assessment reference point to determine if there has been a change in the values and motivation levels of the learner. Analyzing the review patterns of Student Type 1 (or with the borderline self-regulataion label), it can be observed that their motivation levels begin at a zero point . Since these are borderline regulated students who are not extremely motivated compared to their counterparts, it is safe to assume that the typical activities that self-regulated learners dwell into may be as confusing for these borderline self-regulated learners. Which is why, as seen in Table 5.3, a pattern is seen where it usually takes them a long time modifying their learning plans and this task on most cases gives them the confused affect. By the end of the first session, the time the learners see a need to modify their learning plans further, an indication that they might still be practicing the task of making learning plans to be more regulated further. There is no significant change in motivation level so far. The same learner progresses thru the 5th session, attempts to continuously modify the learning plan is still a noticeable act. The assumption here is that they have finally understood the rationale to have a steady learning plan in order to keep themselves motivation. And as seen that though their affects switch from frustrated and confused, there is a significant jump in their motivation from an absolute zero state (see Figure 5.5) . Progressing thru the 10th session, there is a flux of motivation levels as these learners become more concerned with their learning plans but the motivation has shifted towards completing the learning task at hand and not on focusing on creating the learning plan. The long period spent on creating and modifying their learning plans have earned the students enough motivation to lead towards practicing their skills on long periods of time, with a great jump of motivation to a value of 15.3. What is noticeable about these borderline self-regulated learners is that the confused affect followed by a bored affect provides a boost in their motivation levels. These changes in affect display a self-regulated factor among these students that might indicate that they want to get the job done no matter what (Devillers et al., 2005).

Table 5.4: Learning Behavior Patterns of Student Class 2 (Moderate self-regulation)

| Sess No | Source state | State description | Action Performed | Weight |
|---|---|---|---|---|
| 1 | Initial | short-SEARCH-4-Confused | SEARCH | 3.2E-17 |
| 1 | Final | short-OFF-2-Neutral | PRACTICE | 3.5E-21 |
| 5 | Initial | short-SEARCH-4-Neutral | SEARCH | 0.003 |
| 5 | Final | short-MAKE-3-Bored | REVIEW | 3.5E-05 |
| 10 | Initial | short-OFF-4-Neutral | REVIEW | 3.5E-04 |
| 10 | Midway | short-OFF-3-Neutral | REVIEW | 0.04 |
| 10 | Final | short-REVIEW-4-Engaged | OFF | 0.4 |

An notable observation about Normal self-regulated learners shows (see Table 5.4) that these type of learners may be too difficult to motivate as seen in the low values of weights. Yet, it can be seen that all the activities they have performed are usually short. This leads us to the assumption that they try a multiple combination of activities to ensure that they get the learning task at hand completed. Interestingly, their average affect revolves only between a Confused, then Neutral, Bored then Engaged from the 1st all throughout the 10th session. Across the sessions, their motivation levels have fluctuated but has displayed

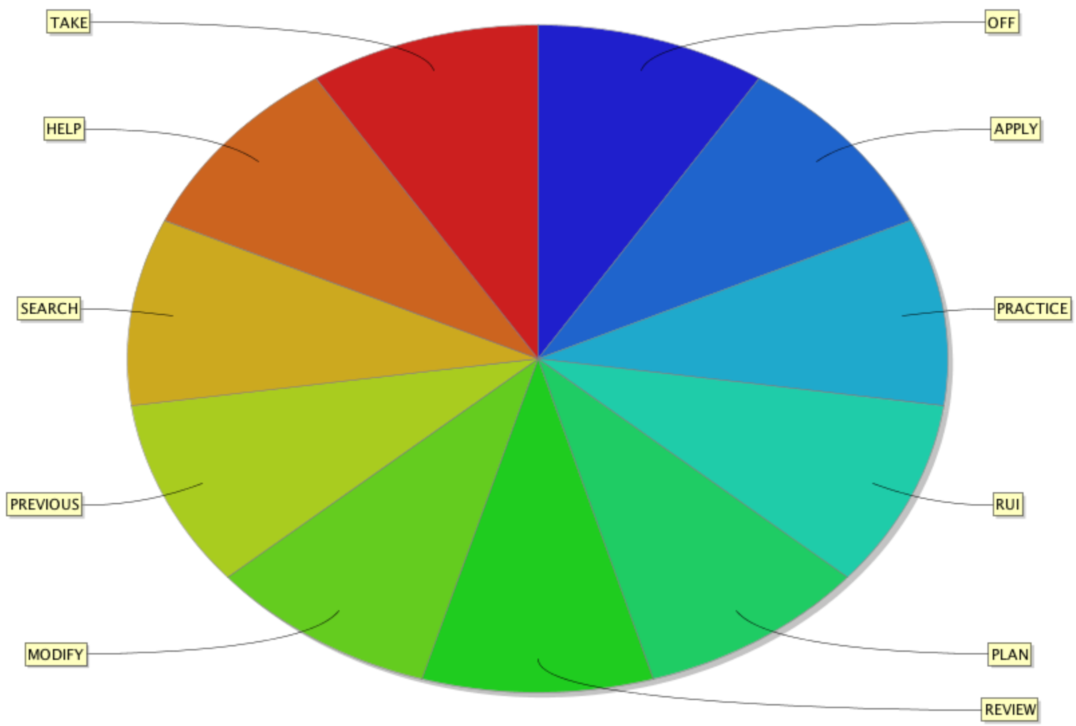


Figure 5.6: Pie chart detailing the task distribution among Class 2 learners

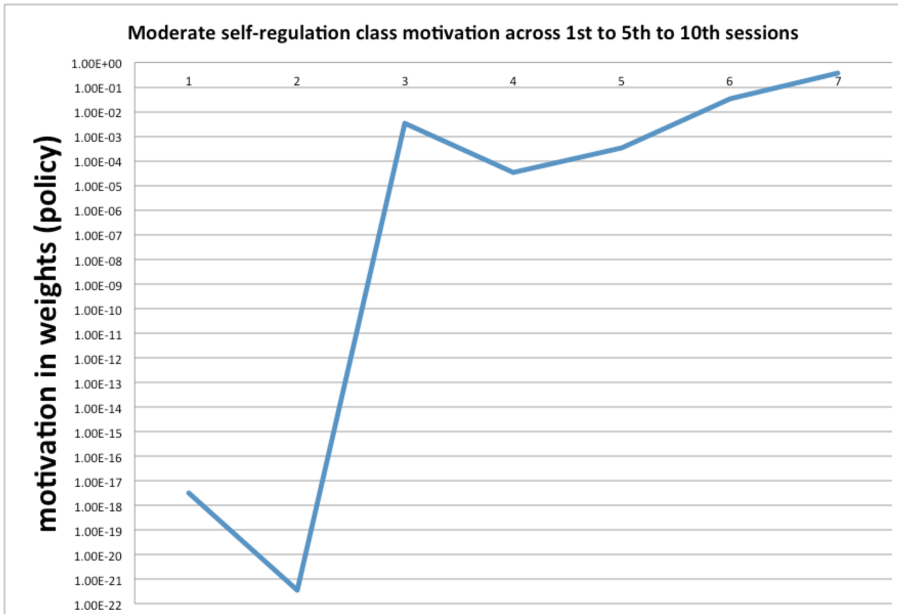


Figure 5.7: Changes in Level of Motivation of Moderate Self-Regulation Learners

a great jump in value when an off-task is performed (see Figure 5.7) . This is a possible indication that normally self-regulated learners consider employing a reward-scheme in fruit of their efforts after a period of intense studying. It can also be noted that they do not spend much time in creating or modifying learning plans, a trait which is possibly innate in them and as such they proceed to moving towards more productive activities such as reviewing, searching for information. Off-tasks are frequent and when an off-task has commenced, growth in their motivation levels are observed. The search for information and review tasks indicate a possible indenpendent learning scenario for these type of learners. As established by Sabourin et al. (2011), off-tasks sometimes keep a learner on-task but these scenarios are still determined on case to case basis.

The third type of learners are those who scored with High Self Regulation from the measurement of self-regulation tendencies questionnaire patterned from Deci et al. (1991). As seen from Figure 5.9, they begin with a high level of motivation (an above zero score of 3.25) on their 1st learning session; an obvious difference from the other two types of learners who began with motivation levels below zero. Interestingly, the length of their activities go through across all three types from short to medium even long tasks. It can be observed with these durations that they are regulated enough to know which tasks should be given greater focus and which should not be. The shift from the 1st session to the 5th

Table 5.5: Learning Behavior Patterns of Student Class 3 (High self-regulation)

| Sess No | Source state | State description | Action Performed | Weight |
|---|---|---|---|---|
| 1 | Initial | short-OFF-4-Bored | OFF | 3.2 |
| 1 | Final | short-MAKE-4-Engaged | OFF | 3.2E-12 |
| 5 | Initial | medium-MODIFY-3-Engaged | OFF | 0.003 |
| 5 | Final | short-MODIFY-4-Engaged | SEARCH | 0.0003 |
| 10 | Initial | short-OFF-3-Neutral | RUI | 0.00002 |
| 10 | Midway | short-SEARCH-4-Engaged | MODIFY | 4.000003 |
| 10 | Final | long-RUI-4-Engaged | SEARCH | 3.5 |

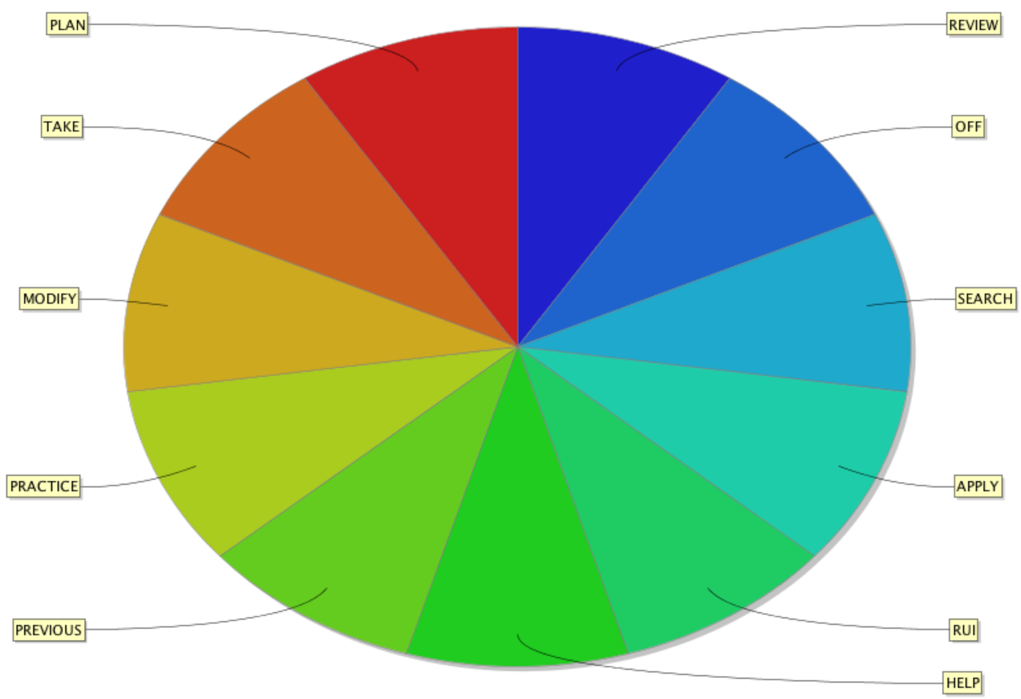


Figure 5.8: Pie chart detailing the task distribution among Class 3 learners

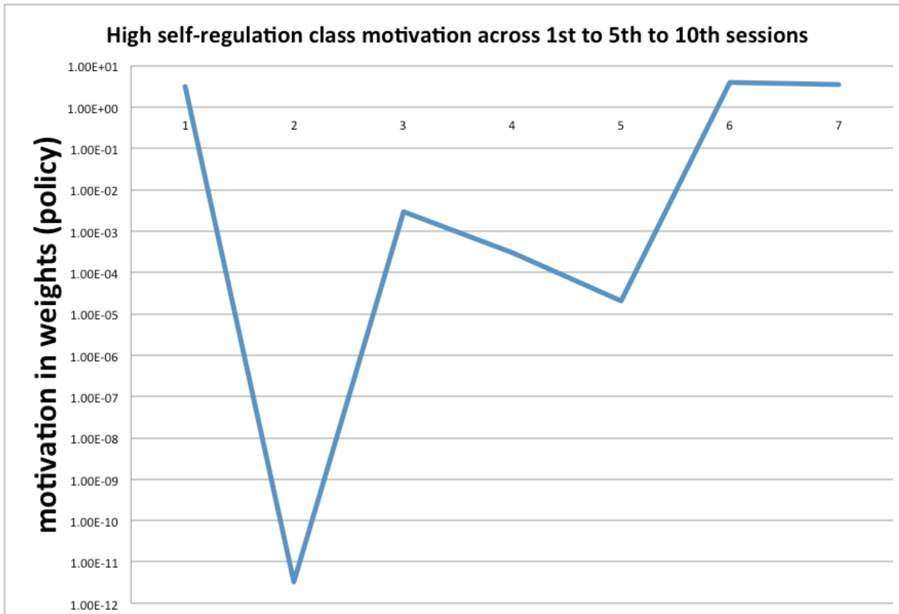


Figure 5.9: Changes in Level of Motivation of High Self-Regulation Learners

session displayed that the learning plan created was not too effective as seen with the drop of motivation and upon modification of it gave a rise to the said value. The number of off-tasks are frequent which indicate a similar approach done by normally self-regulated learners who wish to reward themselves from time to time. Given the 10th session data, a transition from reading additional information boosts not only the motivation level but also the affect (giving the transition between neutral to being engaged).

With the motivation levels between the highly self-regulated learners and normally self-regulated learners, additional data would be necessary to establish if the former establishes a motivation level and intends to keep it a plateau while the latter trying to significantly increase these levels from time to time. One notable difference between these three types of learners can be seen on namely: (1) their initial motivation level when beginning a learning session, (2) the growth and spikes in their motivation along the different learning sessions and (3) the threshold value to where these motivation values limit themselves with. Borderline self-regulated learners begin with a zero value that creates a big jump to a score of 15.3 (the highest recorded motivation weight in this study) while both already self-regulated learners coming from types 2 and 3 have below-zero values of motivation are certain points, yet manage to maintain a momentum in the values that range between not-exceeding values. Type 2 and 3 learners are keen on exploring multiple types of activities as seen in their activity duration and types of tasks performed. P. Inventado (2014) states that exploring other activities can give students a helping boost in getting to discover states and transitions that make them more productive.

## 5.4 Preliminary Analysis of Models

The Interaction data, composed of all the collated interaction data of each respondent was prepared in one data set and is fed with a classifier specially-trained for contextualized sequences. A modification of the multi-layer perceptron patterned after Ware (2015) was created to accommodate contextualized sequences. On the over-all, the perceptron employed additional hidden nodes to accept sequences and numeric labels without the need to use a sigmoid network. The training and validation of the model took around 15.9 hours to complete with the use of RapidMiner (covering almost 705,000 instances, 17 attributes and 1 label). The general model performed with an accuracy of **42.08%** and a kappa statistic value of **0.3** (see Table 5.9).

With the combined interactions of all types of self-regulated learners, the generalized model performed below average than expected. The borderline kappa statistic of 0.3 indicates that even it is at an acceptable range (Michalski et al., 2013), there is a possibility that the data was not fit to work entirely with each other because of its sources. Figure 5.10 displays the individual precision and recall of each class label in the interaction data. Figure A.2 displays the improved neural network model generated with the help of RapidMiner while Figure A.3 displays the spread out values in the form of a scatterplot detailing the derived values from the confusion matrix.

Table 5.6: Feature Extraction on Interaction Data

| Evaluator | Search Method | Selected Attributes |
|---|---|---|
| CfsSubsetEval | Best First | SessionNum, Username |
| CfsSubsetEval | Greedy | SessionNum, Username |
| InfoGainAttributeEval | Ranking | (1,5,4,2,18,9,7,15,13,12,14,8,17,6,10,16,11 : 17) |
| WrapperSubsetEval | Best First | none |

In Figure A.3 the scatterplot of the confusion matrix values from the Interaction data can be seen. It is observed that there is a noticeable scatter among the values which depicts the low accuracy, precision and recall rate generalized from the model. Recall, that the interaction data used to generate this model contains all types of self-regulated learners coming from varying levels of regulation, in one dataset.

| | true RUI | true OFF | true PRACTICE | true HELP | true APPLY | true REVIEW | true PREVIOUS | true SEARCH | true MODIFY | true TAKE | true PLAN | class precision |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| pred. RUI | 8253 | 639 | 2323 | 296 | 331 | 2568 | 2808 | 979 | 246 | 264 | 803 | 42.30% |
| pred. OFF | 1348 | 71841 | 4295 | 1870 | 867 | 5468 | 4114 | 1585 | 885 | 2955 | 3628 | 72.67% |
| pred. PRACTICE | 10208 | 27014 | 102655 | 2929 | 11918 | 37274 | 21119 | 8581 | 9949 | 4807 | 26325 | 39.07% |
| pred. HELP | 8 | 71 | 0 | 40 | 0 | 0 | 4 | 7 | 0 | 0 | 33 | 24.54% |
| pred. APPLY | 1674 | 2738 | 8192 | 938 | 9566 | 2082 | 3345 | 1917 | 2328 | 11 | 1360 | 28.01% |
| pred. REVIEW | 8938 | 9481 | 27989 | 582 | 4483 | 43556 | 7256 | 3526 | 2685 | 1037 | 5926 | 37.72% |
| pred. PREVIOUS | 5454 | 13221 | 15564 | 1179 | 3340 | 16319 | 39746 | 7890 | 5058 | 816 | 11103 | 33.21% |
| pred. SEARCH | 31 | 344 | 0 | 342 | 0 | 354 | 23 | 5486 | 1 | 0 | 222 | 80.64% |
| pred. MODIFY | 606 | 1143 | 3334 | 274 | 1306 | 1692 | 1190 | 536 | 4303 | 23 | 235 | 29.39% |
| pred. TAKE | 0 | 1272 | 1521 | 0 | 0 | 176 | 33 | 11 | 3 | 1346 | 538 | 27.47% |
| pred. PLAN | 527 | 1666 | 6303 | 244 | 164 | 1013 | 3764 | 1065 | 3659 | 0 | 9985 | 35.17% |
| class recall | 22.28% | 55.51% | 59.62% | 0.46% | 29.92% | 39.42% | 47.66% | 17.37% | 14.78% | 11.95% | 16.60% | |

Figure 5.10: Confusion Matrix of Interaction Data

Prior to performing the machine learning task, several feature extraction tasks were employed in order to prepare the dataset. The most notable feature extraction task that was found to be usable was the InfoGainAttributeEval which employs the ranker feature extractor. (see Appendix A.2). The most notable attributes that somehow defined which type of task, along with the intermediary information of the BGApplication were the following:

- 1.101 USERNAME
- 0.29 CONTRIBUTION
- 0.26 AFFECT
- 0.2 SESSIONNUM
- 0.07 HASFACE
- 0.03 KEYINPUTCOUNT
- 0.01 MOUSEMOVEMENTDISTANCE

Ignoring the username which was included in the total dataset before feature extraction, the attributes Contribution and Affect helped define the type of task that was being performed. All other interaction data such as the session no., the hasface attribute, key input count and mouse movement distance were believe to be enough by the ranker attribute evaluator for training the model. The other attributes were still fed into the classifier for completeness.

Training and evaluating a separate model containing these exclusive attributes will have yet to be performed. There is a possibility that the accuracy and kappa statistic might improve but this has to be discovered.

This is why the said data sets were fed into two more classifiers namely the Naive-Bayes classifier and Decision tree classifier. Each of these data sets were fed to be able to provide additional insight with regards to the performance of the model.

Table 5.7: Interaction Data Models Performance using other Classifiers

| Classifier | Percentage Accuracy | Kappa Statistic |
|---|---|---|
| Decision Trees | 31.4% | 0.09 |
| Naive-Bayes | 14.9% | 0.007 |

## 5.5 Analysis of Class-Specific Models

With the low performance of the model containing all three types of self regulated learners, it would be appropriate to create smaller submodels for each class. Following the same modelling techniques as defined in this research's framework (see Section 4.3) but with a smaller population (see values in Table 5.1) because of the identified subclasses, different models and their performances have been designed and evaluated as seen in the following subsections.

### 5.5.1 Model for Low Self-Regulation Class

The interaction data of all respondents whose score fell below the median mark in the self-regulated questionnaire defined by Deci et al. (1991) were included in this sub model. The class model consists of eight (8) respondents whose interaction data from all 10 sessions have been consolidated and fed into the classifier. The same label and set of attributes were used, with a fewer number of instances. See Table 5.9 for the specific results and Figure 5.11 for the details describing the performance of the model. The interaction model for the Low self regulation class perfomed with an accuracy of 63.3% and kappa statistic value of 0.5.

In comparison to the results of the over all interaction data fed into the same classifier (as seen from 5.9 ), the model containing the interaction data of low self-regulated class learners alone perform significantly better than the over all model. The 63% accuracy, beating the 42% accuracy along with their respective kappa values of 0.5 and 0.2 indicate

| | true RUI | true OFF | true PRAC | true SEEK | true APPLY | true REVIEW | true USE | true SEARCH | true MLP | class precision |
|---|---|---|---|---|---|---|---|---|---|---|
| pred. RUI | 10 | 91 | 41 | 11 | 2 | 0 | 0 | 0 | 0 | 6.45% |
| pred. OFF | 88 | 18912 | 1957 | 173 | 4 | 580 | 492 | 6 | 1845 | 78.61% |
| pred. PRAC | 243 | 2947 | 13311 | 168 | 19 | 383 | 249 | 0 | 2909 | 65.80% |
| pred. SEEK | 0 | 4 | 323 | 565 | 243 | 5 | 2 | 0 | 1 | 49.43% |
| pred. APPLY | 0 | 5 | 36 | 57 | 1954 | 35 | 9 | 2 | 2 | 93.05% |
| pred. REVIEW | 0 | 956 | 368 | 69 | 6 | 5972 | 337 | 1 | 1668 | 63.69% |
| pred. USE | 0 | 128 | 3 | 0 | 273 | 1127 | 4294 | 109 | 960 | 62.29% |
| pred. SEARCH | 0 | 2 | 3 | 0 | 0 | 2 | 0 | 1485 | 0 | 99.53% |
| pred. MLP | 1 | 5745 | 8356 | 163 | 27 | 6768 | 2711 | 0 | 27334 | 53.49% |
| class recall | 2.92% | 65.69% | 54.56% | 46.85% | 77.29% | 40.16% | 53.05% | 92.64% | 78.73% | |

Figure 5.11: Confusion Matrix of Interaction Data on Low Self-Regulated Class Learners

that there is great disparity in the values when combined with other classes of self regulated learners.

Referring to the confusion matrix of the said interaction data, it can be noted the precision and recall values of the specific labels. In terms of class precision, a large majority of labels performed beyond 60% which can be referred to as a positive indicator for the model. Similarly, the class recall values highlighted that the model for learners with low self regulation have a gold-standard accuracy of activities that involve Searching for information, Doing Off-task activities and making their learning plans. These values are consistent with the rules and best behaviors observed in the policy data derived from Table 5.3.

### 5.5.2 Model for Moderate Self-Regulation Class

Student respondents whose self-regulation questionnaire scores did not exceed $\pm 1.1$ from the identified median $\bar{x}$. There were 12 respondents who fell in this category and this class took the largest fraction of the interaction data.

The interaction model for the Moderate self regulation class perfomed with an accuracy of 55.9% and kappa statistic value of 0.46. With the largest fraction from the interaction data, the model performed quite fair as but significantly less as compared with the model of the low self regulation class learners (see Table 5.9). Even so, this model outperforms the model for the high self regulation class learners ( see Table 5.9 ) both in terms of accuracy and kappa statistic. Considering the range and the larger size of the model ( $n = 12$ ), it

| | true RUI | true OFF | true PRACTICE | true HELP | true APPLY | true REVIEW | true PREVIOUS | true SEARCH | true MODIFY | true TAKE | true PLAN | class precision |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| pred. RUI | 49675 | 1769 | 4994 | 1133 | 346 | 5635 | 479 | 5142 | 696 | 689 | 17 | 70.39% |
| pred. OFF | 134 | 1812 | 2367 | 138 | 570 | 225 | 168 | 315 | 288 | 105 | 0 | 29.60% |
| pred. PRACTICE | 5314 | 7119 | 58351 | 5483 | 5101 | 8018 | 366 | 9576 | 3462 | 339 | 4 | 56.58% |
| pred. HELP | 252 | 1043 | 1643 | 10404 | 47 | 2271 | 0 | 616 | 355 | 66 | 0 | 62.31% |
| pred. APPLY | 1138 | 150 | 1446 | 66 | 4689 | 23 | 2306 | 888 | 223 | 67 | 0 | 42.64% |
| pred. REVIEW | 2031 | 1080 | 7098 | 2011 | 1525 | 20923 | 0 | 1725 | 1032 | 333 | 100 | 55.27% |
| pred. PREVIOUS | 1088 | 517 | 291 | 42 | 910 | 9 | 5038 | 621 | 240 | 125 | 0 | 56.73% |
| pred. SEARCH | 8413 | 602 | 10708 | 924 | 3646 | 6757 | 1878 | 28431 | 4698 | 626 | 2 | 42.63% |
| pred. MODIFY | 56 | 0 | 0 | 39 | 111 | 0 | 0 | 0 | 541 | 0 | 0 | 72.42% |
| pred. TAKE | 314 | 0 | 1 | 52 | 140 | 21 | 0 | 14 | 0 | 718 | 0 | 56.98% |
| pred. PLAN | 62 | 0 | 1864 | 0 | 0 | 516 | 0 | 0 | 35 | 0 | 3499 | 58.55% |
| class recall | 72.54% | 12.86% | 65.74% | 51.27% | 27.45% | 47.13% | 49.22% | 60.07% | 4.68% | 23.40% | 96.60% | |

Figure 5.12: Confusion Matrix of Interaction Data on Moderate Self-Regulated Class Learners

can act as a submodel from the general model trained earlier only having better scores in terms of accuracy and kappa statistic.

It is interesting to note that with the specific precision and recall values, the class of tasks which were identified in the learning behaviors thru the policy do not necessarily agree with each other. Tasks of Reading and understanding information and Modifying learning plan were tasks that was easier for this model to identify. From the given recall values, the task of Making a Learning Plan achieved a good score of 90% which can be an indicator that moderate self regulated learners differ from low self-regulated learners with their assumed mastery of this basic task. It provides a insight about the standard set of keystrokes, mouse activities that enabled the model to excel in classifying this specific task.

### 5.5.3 Model for High Self-Regulation Class

The interaction data of all respondents whose score went above the median mark in the self-regulated questionnaire defined by Deci et al. (1991) were included in this sub model. The class model consists of eight (12) respondents whose interaction data from all 10 sessions have been consolidated and fed into the same classifier. The same label and set of attributes were used, with a different number of instances. See Table 5.9 for the specific results and Figure 5.13 for the details describing the performance of the model.

The interaction model for the High self regulation class perfomed with an accuracy of 50.7% and kappa statistic value of 0.46. In comparison to the results of the over all interaction data fed into the same classifier (as seen from 5.9 ), the model containing the

| | true RUI | true OFF | true PRACTICE | true HELP | true APPLY | true REVIEW | true PREVIOUS | true SEARCH | true MODIFY | true TAKE | true PLAN | class precision |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| pred. RUI | 32076 | 3705 | 1847 | 671 | 2649 | 838 | 1435 | 13718 | 1261 | 1723 | 652 | 52.95% |
| pred. OFF | 1015 | 17291 | 1086 | 907 | 1368 | 241 | 1639 | 1630 | 852 | 0 | 43 | 66.32% |
| pred. PRACTICE | 60 | 709 | 5541 | 128 | 608 | 39 | 974 | 97 | 26 | 0 | 243 | 65.77% |
| pred. HELP | 1793 | 1249 | 1179 | 6612 | 1506 | 755 | 2390 | 1239 | 2969 | 23 | 85 | 33.39% |
| pred. APPLY | 2049 | 226 | 552 | 925 | 3691 | 587 | 2023 | 315 | 700 | 802 | 915 | 28.87% |
| pred. REVIEW | 34 | 13 | 106 | 4 | 64 | 573 | 100 | 0 | 0 | 0 | 0 | 64.09% |
| pred. PREVIOUS | 4759 | 2153 | 3561 | 772 | 2255 | 363 | 15417 | 1063 | 1720 | 1617 | 1186 | 44.22% |
| pred. SEARCH | 6582 | 2904 | 2646 | 1245 | 1957 | 454 | 1936 | 35921 | 710 | 0 | 3784 | 61.78% |
| pred. MODIFY | 2756 | 3780 | 1755 | 4079 | 2261 | 570 | 1934 | 3803 | 10407 | 204 | 338 | 32.64% |
| pred. TAKE | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 191 | 3268 | 0 | 94.48% |
| pred. PLAN | 108 | 133 | 137 | 12 | 54 | 0 | 132 | 1229 | 46 | 0 | 1108 | 37.45% |
| class recall | 62.61% | 53.76% | 30.10% | 43.06% | 22.49% | 12.96% | 55.10% | 60.87% | 55.12% | 42.79% | 13.26% | |

Figure 5.13: Confusion Matrix of Interaction Data on High Self-Regulated Class Learners

interaction data of low self-regulated class learners alone still performed better than the over all model. The 50.7% accuracy barely getting past the half mark beats the 42% accuracy along with their respective kappa values of 0.4 against the statistic value of 0.2. This can again indicate that there is great disparity in the values when combined with other classes of self regulated learners.

The precision of the Take Notes label was notably high possibly indicating a standard behavior among student learners with high self regulation that this activity might be a common practice. With respect to their online activities and gestures such as mouse movement and keystroke count, students from this said class observe the same set of behavior when taking down notes. It was rather easier for the classifier to learn this classification in contrast to the other classifiers. Similarly, the tasks "Read and understand information" and "Seek for help" achieved acceptable passing marks in terms of class recall. Again, in connection with the learning behaviors derived from the computed policy values as seen from Table 5.5, these two tasks were seen to be most leading towards a greater motivation value. While the rest of the class precision and recall values fell at around the range of 50% mark, the confusion matrix indicated that there was not one activity that can easily be confused of to be another type of activity.

### 5.5.4 Modelling with Two Classes of Self Regulation

With the two-class separation move as earlier explained in Section 5.1, the combined interaction data would be divided with a simpler approach providing a two way overview of how the data is spread out. Grouping the students on the upper and lower half of the self regulation scale created models with different performance scores as seen below:

Table 5.8: Two-way Interaction Data Models

| Category | Classifier | Percentage Accuracy | Kappa Statistic |
|---|---|---|---|
| Low Self Regulation | Decision Trees | 79.01% | 0.7 |
| Low Self Regulation | Naive-Bayes | 28.03% | 0.1 |
| Low Self Regulation | MLP | 51.4% | 0.4 |
| High Self Regulation | Decision Trees | 29.8% | 0.07 |
| High Self Regulation | Naive-Bayes | 16.6% | 0.04 |
| High Self Regulation | MLP | 46.7% | 0.4 |

## 5.6 Combined Analysis Results

With the given values and performance scores of all General and Class specific models, a greater insight on their activities, tasks and the activity features can be derived. The Class specific models were combined using average values and can the values can be seen on Table 5.9.

Table 5.9: Interaction Data Models for both General and Class-Specific Types

| Model Type | Percentage Accuracy | Kappa Statistic |
|---|---|---|
| General | 42.8% | 0.30 |
| Low Self-Regulation Class | 63.3% | 0.55 |
| Moderate Self-Regulation Class | 55.9% | 0.46 |
| High Self-Regulation Class | 50.7% | 0.43 |

The average accuracy and kappa statistic of all the class specific models, when averaged, still outperform the general model which contained the superset of all interaction data. The model for the Low self-regulation learners performed the highest ($accuracy = 63.3\%$) among all the other class models (see Table **??**). The performance of the Moderate and High self-regulation class models performed with decreasing accuracy scores. This could mean establish that the higher the motivation levels are, the more confusing it is for the classifier to correctly identify the different activities these students undertake.

However this assumption will have to be yet further investigated as other factors such as the correctness and the exactness of the self-annotation performed might indeed have an impact as well to the total delivery of the data. While the accuracy scores of both the Moderate and High self-regulation learners performed better than the General model, all these scores are still below an acceptable mark (60%).

# 6 Conclusion and Future Work

This chapter covers the conclusions that could be made from the results of the experiment and a section on the future direction for this research.

## 6.1 Conclusion

This research began with the objective at hand on discovering whether a class-specific or a general model can be built on the activities of self-regulated learners. To determine whether this objective was achieved, three specific objectives were defined.

Students were assessed using the Measurement of Self-Regulation Tendencies Questionnaire which uses a 7-point liker scale following the instrument designed by Deci et al. (1991). Upon assessment, students were categorized into three sub groups specifically Type 1 or Low self-regulation learners, Type 2 or Moderate self-regulation learners and Type 3 High self-regulation learners depending on their scores from the said research instrument. All these student classes were immersed in a self-initiated learning scenario to capture their interaction data. This was done to not only validate their self-regulatedness, but also to measure their motivation levels with the use of the profit sharing algorithm (Keller, 1999) (Arai & Sycara, 2000) and to determine the different types of tasks and activities they dwell on while learning on their own.

Student-learners performed self-reflection by annotating their activities as categorized by (Laurillard, 2013), their affect and the level of contribution they think the learning session task has done to their overall goal of learning. The effect of his retrospection phase following the frameworks of P. Inventado (2014) and Zimmerman (2002) enabled the three different types of self-regulated learners to have their own clustered learning policy. These were measured by their motivation ratings and a difference in their activities can be seen from tables 5.3, 5.4 and 5.5. Each category was able to produce a total 7 rules following the policy and the profit-sharing algorithm. Upon discovery of these learning behavior patterns, the said three categories had observed motivation levels which spiked and fluxed at varying points. Hereby establishing that their motivation levels, their current affect and their productivity can be influenced by the type of activity that they are performing.

Other existing research performed and computed their policies based on simulated data where in this study, these were computed using actual data provided and annotated by the learners themselves. As such it has been observed that the relationships between the student's motivation ratings and the performance of "optimal" actions, when followed enabled more motivated learners across each learning session.

What the learning behavior patterns have showed us is that, the level of self-regulation is directly proportional to the number of activities and combinations of tasks that a learner can perform in a learning session. Several types of activities were performed by a student who is more of a regulated learner as compared to their less-regulated counterparts. A pattern among Class 2 and Class 3 self-regulated learners have also been found that leads to a plateau or a consistent rate of motivation when a set of activities are strictly-followed. These have yet to be further investigated.

Upon creation of a unified data set containing the interactions covering all three types of self-regulated learners, a machine learning task was employed to evaluate and validate its correctness. With the use of a multi layer perceptron that is context-aware, as modified by Ware (2015), the output model performed below average with an accuracy rating of 42% only and having a borderline acceptable (Michalski et al., 2013) kappa statistic of 0.3. This leads us to infer that both the data and the model performed below average because of its diversity, its completeness (since it covered three types of self-regulated learners all in one data set) and its complexity. Intermediary information describing background applications, has been added as an attribute but this was not enough to help improve the performance of the model.

With the performance of the model from the unified data set and the diversity of the rules among the three types of self-regulated learners, it can therefore be assumed that in order for us to model correctly and further the activities of these self-regulated learners, the design, development and use of a class-specific model, instead of a general model would be more appropriate. By dividing the the existing data set into three data sets for each type of learner and by following the recommended features returned by the ranker attribute evaluator (as seen from Table 5.6) perhaps a better performing model for each user type can be generated and validated further.

As such, upon performing clustering on the existing interaction data, more positive results arose. The divided three, based on their level of self regulation where used on a machine learning task and yielded results that were not only average and acceptable but also consistent with rules generated from the policy data. The activities whose class labels fall under which was defined in the classifier where highlighted and very much agree with the activities and sequences identified to be notably motivation-inducing as dictated by the policies. These activities had notably high precision and recall scores which provided a good indicator for these models.

The methodologies and techniques employed in the study were able to generate and identify the learning patterns of self-regulated learners covering three specific classes (borderline, normal and high self-regulation). These can be seen in the identified rules based from the computed policy (weight/motivation points). The data has also provided that a unified general model can be created covering these types of learners and instead requires the modelling of possibly distinct but class-specific model.

## 6.2 Future Work

The future of this research can be directed into several directions, depending on which area of the results would have to be expanded.

The identified policies and rules in this study can be further modelled into temporal likelihood transitions (D'Mello, 2012) which can enable us to determine the changes in the learning behavior should we go beyond 10 sessions. This likelihood function is defined by the equation below:

$$L(A_i, A_{i+1}) = \frac{P(A_{i+1}|A_i) - P(A_{i+1})}{1 - P(A_{Ai+1})} \tag{13}$$

As such, these learning behaviors and the succeeding changes in these activities can be most likely predicted. By following these recommendations, the said activities and learning behavior patterns of these self-regulated learners can be modeled into temporal transition likelihood graphs which can enable us to understand further the changes in their actions from the 10th session and beyond.

With the activities and learning behavior patterns identified, it is also essential that this data can be used to determine, standardize and model the actual length and quantifiable definition of an activity sequence, answering the question "*How long exactly is a sequence on average?* ". These can enable time-based researchers and activity recognition researchers to create a model of activities that hopefully-aim to be deterministic and predictable in nature.

One aspect that has never been tackled in this research is incorporating the concept of feedback which can be used to train non-regulated learners and potentially improve type 1 or low self-regulation learners. With the incorporation of the policy and providing a feedback facility module that interacts with a student beyond the learning session along with further assessment techniques that can be used to determine if the a student has an improved level of self-regulation.

The tool Sidekick can be improved where with the identification of the learning behavior patterns and with the creation of user-specific models, the annotation can be semi-automated to the point that the effort induced when self-annotation would be less. These and all without compromising the self-reflection phase that students undergo when they annotate their data.

Furthermore, the same approach and tool can be used to evaluate the productivity levels of employees and rank and file workers who use the computer. By adjusting a few parameters and using a different instrument to measure self-regulation among adults, the study can be imported to help managers and policy makers determine which activity patterns produce the most motivated and productive employees.

# References


Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron*, *50*(3), 507–517.

Arai, S., & Sycara, K. (2000). Effective learning approach for planning and scheduling in multi-agent domain. In *Proceedings of the 6th international conference on simulation of adaptive behavior* (pp. 507–516).

Azcarraga, J., Ibañez, J. F., Lim, I. R., Lumanas Jr, N., Trogo, R., & Suarez, M. T. (2011). Predicting academic emotion based on brainwaves signals and mouse click behavior. *Chiang Mai, Thailand: Asia-Pacific Society for Computers in Education*.

Azevedo, R., Cromley, J. G., Moos, D. C., Greene, J. A., & Winters, F. I. (2011). Adaptive content and process scaffolding: A key to facilitating students' self-regulated learning with hypermedia. *Psychological Testing and Assessment Modeling*, *53*, 106–140.

Baker, R. S. d, Rodrigo, M. M. T., & Xolocotzin, U. E. (2007). The dynamics of affective transitions in simulation problem-solving environments. In *Affective computing and intelligent interaction* (pp. 666–677). Springer.

Baldassarre, G., Mannella, F., Fiore, V. G., Redgrave, P., Gurney, K., & Mirolli, M. (2013). Intrinsically motivated action–outcome learning and goal-based action recall: a system-level bio-constrained computational model. *Neural Networks*, *41*, 168–187.

Biswas, G., Jeong, H., Kinnebrew, J. S., Sulcer, B., & ROSCOE, R. (2010). Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology Enhanced Learning*, *5*(02), 123–152.

Boritz, J., Booth, K. S., & Cowan, W. B. (1991). Fitts's law studies of directional mouse movement. In *Graphics interface'91* (pp. 216–223).

Brand, M., Oliver, N., & Pentland, A. (1997). Coupled hidden markov models for complex action recognition. In *Computer vision and pattern recognition, 1997. proceedings., 1997 ieee computer society conference on* (pp. 994–999).

Burton, R. R., & Brown, J. S. (1979). An investigation of computer coaching for informal learning activities. *International Journal of Man-Machine Studies*, *11*(1), 5–24.

Card, S. K., Moran, T. P., & Newell, A. (1980). The keystroke-level model for user performance time with interactive systems. *Communications of the ACM*, *23*(7), 396–410.

Craig, S., Graesser, A., Sullins, J., & Gholson, B. (2004). Affect and learning: an exploratory look into the role of affect in learning with autotutor. *Journal of Educational Media*, *29*(3), 241–250.

Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational psychologist*, *26*(3-4), 325–346.

Devillers, L., Vidrascu, L., & Lamel, L. (2005). Challenges in real-life emotion annotation and machine learning based detection. *Neural Networks*, *18*(4), 407–422.

D'Mello, S. (2012). Monitoring affective trajectories during complex learning. In *Encyclopedia of the sciences of learning* (pp. 2325–2328). Springer.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157.

D'Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H., et al. (2008). Autotutor detects and responds to learners affective and cognitive states. In *Workshop on emotional and cognitive issues at the international conference on intelligent tutoring systems.*

Ekman, P. (2007). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life.* Macmillan.

Epps, B. W. (1986). Comparison of six cursor control devices based on fitts' law models. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 30, pp. 327–331).

Garner, J., & Bol, L. (2011). The challenges of e-learning initiatives in supporting students with self-regulated learning and executive function difficulties. In *Annual meeting of the international congress for school effectiveness and improvement, january* (pp. 4–8).

Gunetti, D., & Picardi, C. (2005). Keystroke analysis of free text. *ACM Transactions on Information and System Security (TISSEC)*, *8*(3), 312–347.

Hill, P. W., & Rowe, K. J. (1998). Modelling student progress in studies of educational effectiveness∗. *School effectiveness and school improvement*, *9*(3), 310–333.

Inventado, P. (2014). *Helping students manage learning behaviour in self-initiated learning scenarios.* Unpublished doctoral dissertation, Osaka University.

Inventado, P., Legaspi, R., Cabredo, R., & Numao, M. (2012). Student learning behavior in an unsupervised learning environment. In *Proceedings of the 20th international conference on computers in education* (pp. 730–737).

Inventado, P. S., Legaspi, R., Suarez, M., & Numao, M. (2011). Investigating transitions in affect and activities for online learning interventions.

Keller, J. M. (1999). Using the arcs motivational process in computer-based instruction and distance education. *New directions for teaching and learning*, *1999*(78), 37–47.

Kim, J., Shaw, E., Feng, D., Beal, C., & Hovy, E. (2006). Modeling and assessing student activities in on-line discussions. In *Proc. of the aaai workshop on educational data mining.*

Kinnebrew, J. S., & Biswas, G. (2012). Identifying learning behaviors by contextualizing differential sequence mining with action features and performance evolution. *International Educational Data Mining Society*.

Kinnebrew, J. S., Loretz, K. M., & Biswas, G. (2012). A contextualized, differential sequence mining method to derive students' learning behavior patterns. *Journal of Educational Data Mining*.

Kinnebrew, J. S., Loretz, K. M., & Biswas, G. (2013). A contextualized, differential sequence mining method to derive students' learning behavior patterns. *JEDM-Journal of Educational Data Mining*, *5*(1), 190–219.

Kirschner, P. A., & Karpinski, A. C. (2010). Facebook® and academic performance. *Computers in human behavior*, *26*(6), 1237–1245.

Klein, H. J., Noe, R. A., & Wang, C. (2006). Motivation to learn and course outcomes: The impact of delivery mode, learning goal orientation, and perceived barriers and enablers. *Personnel Psychology*, *59*(3), 665–702.

Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, *12*(2), 74–82.

Laurillard, D. (2013). *Rethinking university teaching: A conversational framework for the effective use of learning technologies.* Routledge.

Mabroukeh, N. R., & Ezeife, C. I. (2010). A taxonomy of sequential pattern mining algorithms. *ACM Computing Surveys (CSUR)*, *43*(1), 3.

MacKenzie, I. S. (1992a). Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction*, *7*(1), 91–139.

MacKenzie, I. S. (1992b). Movement time prediction in human-computer interfaces. In *Proceedings of graphics interface* (Vol. 92, pp. 140–150).

MacKenzie, I. S., & Buxton, W. (1992). Extending fitts' law to two-dimensional tasks. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 219–226).

Mantyjarvi, J., Himberg, J., & Seppanen, T. (2001). Recognizing human motion with multiple acceleration sensors. In *Systems, man, and cybernetics, 2001 ieee international conference on* (Vol. 2, pp. 747–752).

Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (2013). *Machine learning: An artificial intelligence approach.* Springer Science & Business Media.

Miyazaki, K., Yamamura, M., & Kobayashi, S. (1994). On the rationality of profit sharing in reinforcement learning.

Noack, A. (2007). Timing analysis of keystrokes and timing attacks on ssh revisited.

Ormrod, J. E. (2013). *Educational psychology: Developing learners.* Pearson Higher Ed.

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International journal of educational research*, *31*(6), 459–470.

Pirttikangas, S., Fujinami, K., & Nakajima, T. (2006). Feature selection and activity recognition from wearable sensors. In *Ubiquitous computing systems* (pp. 516–527). Springer.

Portnow, J. (2008). The power of tangential learning. *Edge Online*.

Pugliesi, K. (1999). The consequences of emotional labor: Effects on work stress, job satisfaction, and well-being. *Motivation and Emotion*, *23*(2), 125–154.

Rabiner, L. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286.

Roberts, G. R., et al. (2005). Technology and learning expectations of the net generation. *Educating the net generation*, 3–1.

Rodrigo, M. M. T., Baker, R. S., Lagud, M. C., Lim, S. A. L., Macapanpan, A. F., Pascua, S., et al. (2007). Affect and usage choices in simulation problem solving environments. *FRONTIERS IN ARTIFICIAL INTELLIGENCE AND APPLICATIONS*, *158*, 145.

Roscoe, R. D., & Chi, M. T. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, *77*(4), 534–574.

Rosenblatt, F. (1961). *Principles of neurodynamics. perceptrons and the theory of brain mechanisms* (Tech. Rep.). DTIC Document.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 763.

Sabourin, J., Rowe, J. P., Mott, B. W., & Lester, J. C. (2011). When off-task is on-task: The affective role of off-task behavior in narrative-centered learning environments. In *Artificial intelligence in education* (pp. 534–536).

Schunk, D. H. (1989). Social cognitive theory and self-regulated learning. In *Self-regulated learning and academic achievement* (pp. 83–110). Springer.

Schunk, D. H., & Zimmerman, B. J. (2012). *Motivation and self-regulated learning: Theory, research, and applications*. Routledge.

Sharma, N., & Gedeon, T. (2012). Objective measures, sensors and computational techniques for stress recognition and classification: A survey. *Computer methods and programs in biomedicine*.

Si, J., & Wang, Y.-T. (2001). Online learning control by association and reinforcement. *Neural Networks, IEEE Transactions on*, *12*(2), 264–276.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1) (No. 1). MIT press Cambridge.

Ware, M. (2015). *Class multilayerperceptroncs*. Available from `http://weka.sourceforge.net/doc.packages/multilayerPerceptronCS/weka/classifiers/functions/MultilayerPerceptronCS.html`

Yamato, J., Ohya, J., & Ishii, K. (1992). Recognizing human action in time-sequential images using hidden markov model. In *Computer vision and pattern recognition, 1992. proceedings cvpr'92., 1992 ieee computer society conference on* (pp. 379–385).

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice*, *41*(2), 64–70.

# A Appendices and Glossary of Terms and Acronyms

## A.1 Overall Policy Distribution

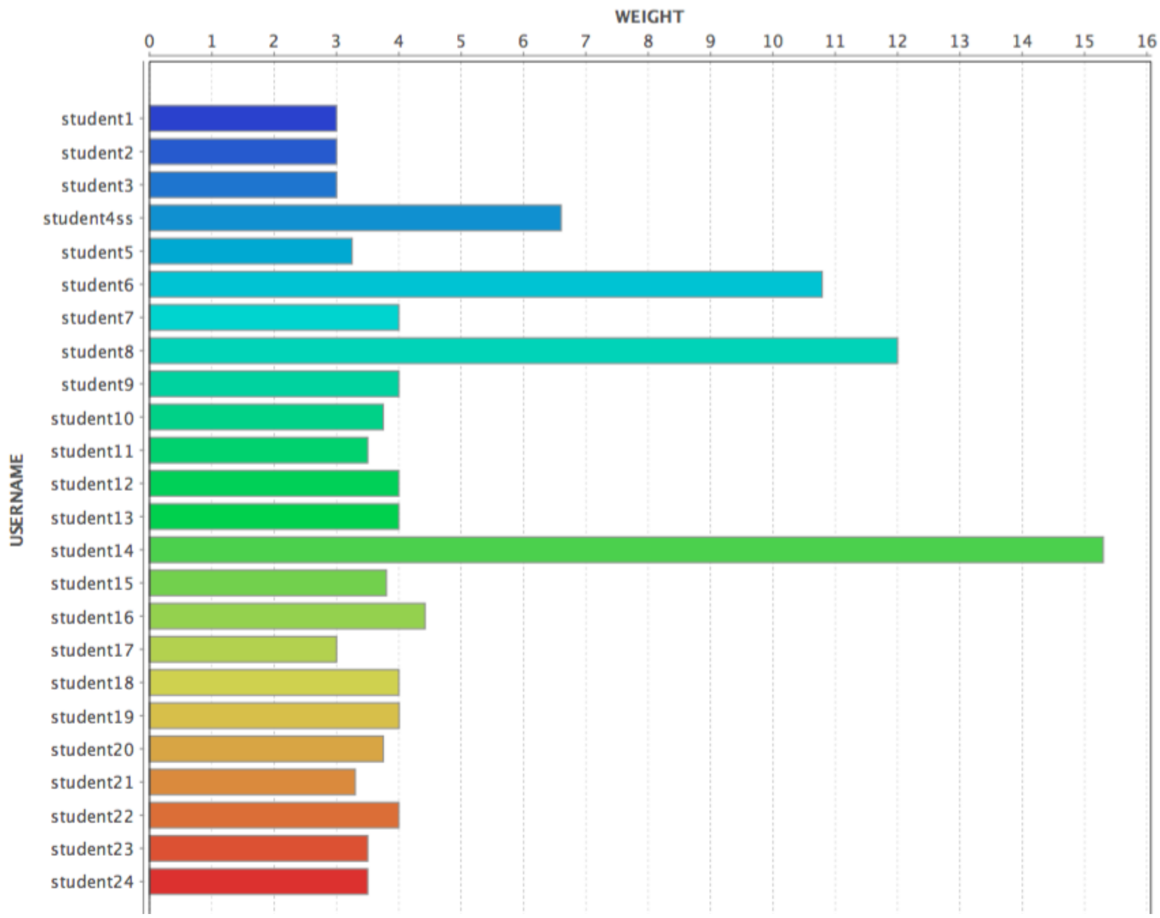


Figure A.1: Policy Distribution for all learners across their weight values

## A.2 Ranked attributes as seen from Table 5.6

Ranked attributes:
1.101965 1 USERNAME
0.298796 5 CONTRIBUTION
0.261403 4 AFFECT
0.221343 2 SESSIONNUM
0.079696 18 HASFACE
0.033172 9 KEYINPUTCOUNT
0.015509 7 MOUSEMOVEMENTDISTANCE
0.009664 15 MOUSEMOVEMENTRIGHTDISTANCE
0.009593 13 MOUSEMOVEMENTDOWNDISTANCE
0.009576 12 MOUSEMOVEMENTUPDISTANCE
0.008664 14 MOUSEMOVEMENTLEFTDISTANCE
0.004832 8 MOUSEWHEELDISTANCE
0.004194 17 MOUSEWHEELDOWNDISTANCE
0.002808 6 MOUSECLICKCOUNT
0.002741 10 LEFTMOUSECLICKCOUNT
0.001181 16 MOUSEWHEELUPDISTANCE
0.000233 11 RIGHTMOUSECLICKCOUNT

## A.3 Labels of Task Acronyms

- Make learning plan (PLAN)
- Modify learning plan (MODIFY)
- Read and understand information (RUI)
- Practice skills (PRACTICE)
- Search for info (SEARCH)
- Apply acquired info (APPLY)
- Use previous knowledge (USE)
- Take notes (TAKE)
- Review notes (REVIEW)
- Seek help (SEEK) and
- Off task (OFF)

.

## A.4 Improved Neural Network Model

### A.4.1 Neural Network Model Diagram

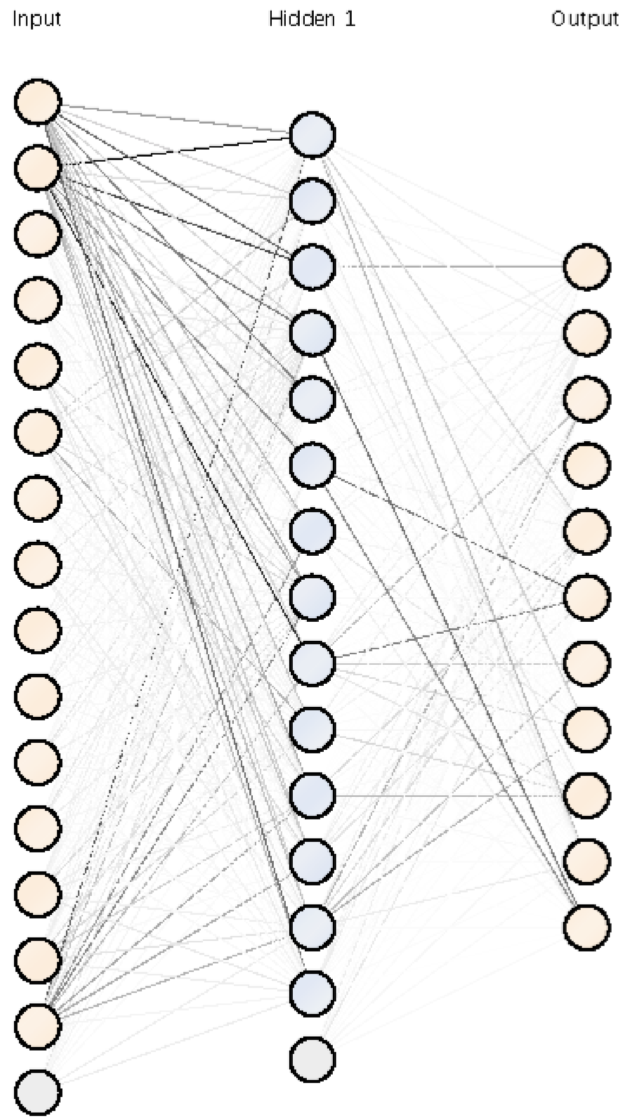


Figure A.2: Improved Neural Network Diagram

Above is the Neural Network Model diagram for the General Model. The combined interaction data have been used. The input nodes represent the specific features and attributes in the data set (see Appendix A.2). The darker the line indicates a connection between the nodes that require a heavy weight. The specific values of each node and their weights can be seen in Appendix A.4.3. The output nodes are the different task categories as seen from Appendix A.3.

### A.4.2 Confusion Matrix Scatterplot

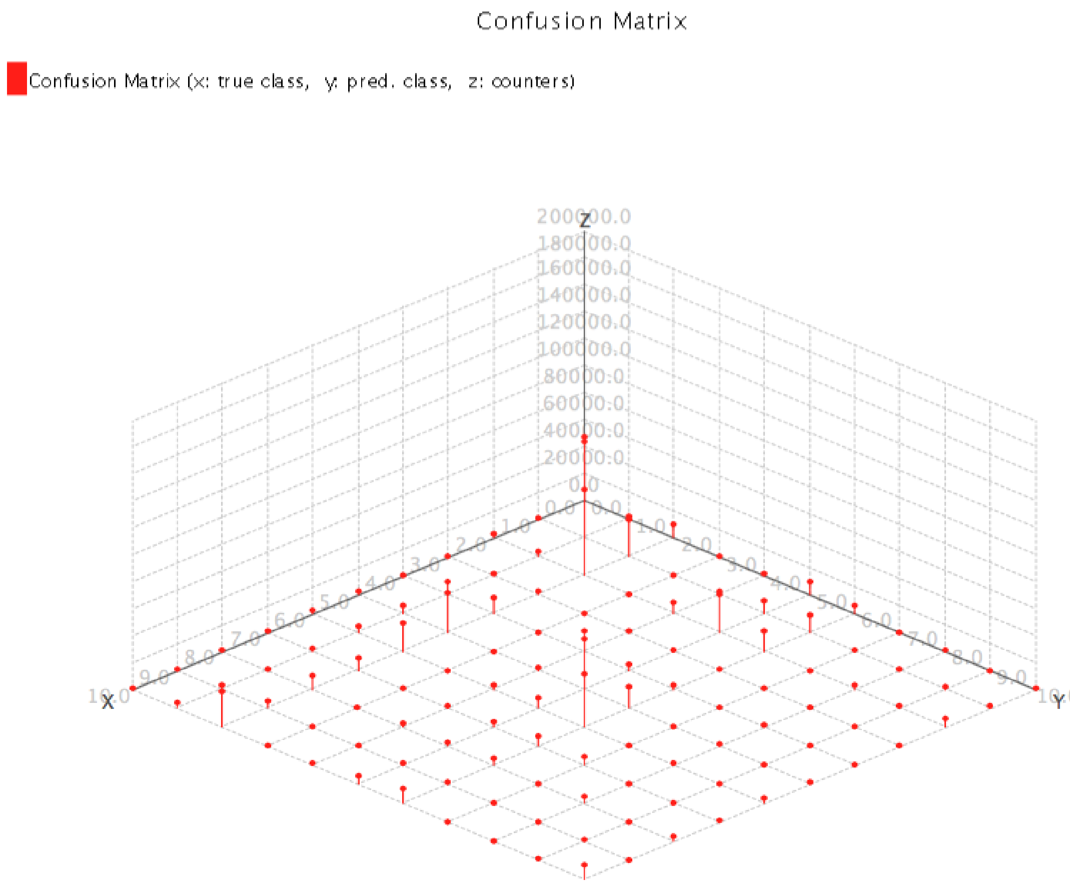


Figure A.3: Scatterplot of Confusion Matrix Values from the Interaction Data

### A.4.3 Details of Nodes and Weights

ImprovedNeuralNet Hidden 1 ========

Node 1 (Sigmoid) ——————- SESSIONNUM: -50.113 CONTRIBUTION: -105.616 MOUSECLICKCOUNT: -4.297 MOUSEMOVEMENTDISTANCE: -5.986 MOUSEWHEELDISTANCE: 0.435 KEYINPUTCOUNT: -2.269 LEFTMOUSECLICKCOUNT: -3.841 RIGHTMOUSECLICKCOUNT: -4.773 MOUSEMOVEMENTUPDISTANCE: -7.793 MOUSEMOVEMENTDOWNDISTANCE: -7.175 MOUSEMOVEMENTLEFTDISTANCE: -3.729 MOUSEMOVEMENTRIGHTDISTANCE: -4.799 MOUSEWHEELUPDISTANCE: -6.471 MOUSEWHEELDOWNDISTANCE: 6.693 HASFACE: -78.282 Bias: 4.124

Node 2 (Sigmoid) ——————- SESSIONNUM: -23.631 CONTRIBUTION: -18.022 MOUSECLICKCOUNT: 3.193 MOUSEMOVEMENTDISTANCE: -2.164 MOUSEWHEELDISTANCE: -4.987 KEYINPUTCOUNT: 24.380 LEFTMOUSECLICKCOUNT: 3.119 RIGHTMOUSECLICKCOUNT: 2.269 MOUSEMOVEMENTUPDISTANCE: -1.120 MOUSEMOVEMENTDOWNDISTANCE: -0.704 MOUSEMOVEMENTLEFTDISTANCE: -2.742 MOUSEMOVEMENTRIGHTDISTANCE: -2.665 MOUSEWHEELUPDISTANCE: -8.117 MOUSEWHEELDOWNDISTANCE: -2.140 HASFACE: 23.995 Bias: -2.034

Node 3 (Sigmoid) ——————- SESSIONNUM: 74.690 CONTRIBUTION: -86.324 MOUSECLICKCOUNT: 1.123 MOUSEMOVEMENTDISTANCE: 2.410 MOUSEWHEELDISTANCE: 1.394 KEYINPUTCOUNT: 15.319 LEFTMOUSECLICKCOUNT: 1.676 RIGHTMOUSECLICKCOUNT: -0.609 MOUSEMOVEMENTUPDISTANCE: 4.660 MOUSEMOVEMENTDOWNDISTANCE: 4.359 MOUSEMOVEMENTLEFTDISTANCE: 1.026 MOUSEMOVEMENTRIGHTDISTANCE: 0.970 MOUSEWHEELUPDISTANCE: -5.287 MOUSEWHEELDOWNDISTANCE: 7.301 HASFACE: -12.576 Bias: -0.148

Node 4 (Sigmoid) ——————- SESSIONNUM: 21.391 CONTRIBUTION: 66.774 MOUSECLICKCOUNT: 9.720 MOUSEMOVEMENTDISTANCE: 11.350 MOUSEWHEELDISTANCE: -7.287 KEYINPUTCOUNT: -4.589 LEFTMOUSECLICKCOUNT: 11.694 RIGHTMOUSECLICKCOUNT: 14.097 MOUSEMOVEMENTUPDISTANCE: 9.623 MOUSEMOVEMENTDOWNDISTANCE: 10.241 MOUSEMOVEMENTLEFTDISTANCE: 12.489 MOUSEMOVEMENTRIGHTDISTANCE: 12.274 MOUSEWHEELUPDISTANCE: -5.615 MOUSEWHEELDOWNDISTANCE: -8.683 HASFACE: 31.540 Bias: -14.244

Node 5 (Sigmoid) ——————- SESSIONNUM: -61.733 CONTRIBUTION: 26.218 MOUSECLICKCOUNT: 6.173 MOUSEMOVEMENTDISTANCE: 7.101 MOUSEWHEELDISTANCE: 3.476 KEYINPUTCOUNT: -6.088 LEFTMOUSECLICKCOUNT: 6.172 RIGHTMOUSECLICKCOUNT: -3.345 MOUSEMOVEMENTUPDISTANCE: 4.898 MOUSEMOVEMENTDOWNDISTANCE: 5.514 MOUSEMOVEMENTLEFTDISTANCE: 7.707 MOUSEMOVEMENTRIGHT-

DISTANCE: 8.156 MOUSEWHEELUPDISTANCE: 1.314 MOUSEWHEELDOWNDISTANCE: 5.369 HASFACE: -0.405 Bias: 3.686

Node 6 (Sigmoid) ————- SESSIONNUM: -9.315 CONTRIBUTION: -56.905 MOUSECLICKCOUNT: -4.170 MOUSEMOVEMENTDISTANCE: 0.730 MOUSEWHEELDISTANCE: 1.587 KEYINPUTCOUNT: 0.320 LEFTMOUSECLICKCOUNT: -4.299 RIGHTMOUSECLICKCOUNT: -3.828 MOUSEMOVEMENTUPDISTANCE: 1.295 MOUSEMOVEMENTDOWNDISTANCE: 1.152 MOUSEMOVEMENTLEFTDISTANCE: 0.326 MOUSEMOVEMENTRIGHTDISTANCE: 0.331 MOUSEWHEELUPDISTANCE: 2.282 MOUSEWHEELDOWNDISTANCE: 0.959 HASFACE: -11.679 Bias: 4.074

Node 7 (Sigmoid) ————- SESSIONNUM: -33.848 CONTRIBUTION: -12.966 MOUSECLICKCOUNT: -2.168 MOUSEMOVEMENTDISTANCE: 1.449 MOUSEWHEELDISTANCE: 1.861 KEYINPUTCOUNT: 7.941 LEFTMOUSECLICKCOUNT: -2.261 RIGHTMOUSECLICKCOUNT: 2.617 MOUSEMOVEMENTUPDISTANCE: 3.228 MOUSEMOVEMENTDOWNDISTANCE: 1.911 MOUSEMOVEMENTLEFTDISTANCE: 0.900 MOUSEMOVEMENTRIGHTDISTANCE: 1.264 MOUSEWHEELUPDISTANCE: 6.712 MOUSEWHEELDOWNDISTANCE: -2.328 HASFACE: -52.316 Bias: -2.776

Node 8 (Sigmoid) ————- SESSIONNUM: 37.208 CONTRIBUTION: 50.756 MOUSECLICKCOUNT: 4.965 MOUSEMOVEMENTDISTANCE: 8.122 MOUSEWHEELDISTANCE: 17.749 KEYINPUTCOUNT: -19.038 LEFTMOUSECLICKCOUNT: 4.791 RIGHTMOUSECLICKCOUNT: 6.628 MOUSEMOVEMENTUPDISTANCE: 6.947 MOUSEMOVEMENTDOWNDISTANCE: 7.864 MOUSEMOVEMENTLEFTDISTANCE: 8.236 MOUSEMOVEMENTRIGHTDISTANCE: 8.962 MOUSEWHEELUPDISTANCE: 17.933 MOUSEWHEELDOWNDISTANCE: 17.590 HASFACE: 59.866 Bias: -6.374

Node 9 (Sigmoid) ————- SESSIONNUM: 20.638 CONTRIBUTION: -93.936 MOUSECLICKCOUNT: 1.878 MOUSEMOVEMENTDISTANCE: 1.765 MOUSEWHEELDISTANCE: 0.044 KEYINPUTCOUNT: 1.108 LEFTMOUSECLICKCOUNT: 1.953 RIGHTMOUSECLICKCOUNT: 2.154 MOUSEMOVEMENTUPDISTANCE: 1.948 MOUSEMOVEMENTDOWNDISTANCE: 1.691 MOUSEMOVEMENTLEFTDISTANCE: 1.491 MOUSEMOVEMENTRIGHTDISTANCE: 1.905 MOUSEWHEELUPDISTANCE: -0.242 MOUSEWHEELDOWNDISTANCE: 0.295 HASFACE: -34.830 Bias: -2.205

Node 10 (Sigmoid) ———— SESSIONNUM: 29.461 CONTRIBUTION: -3.942 MOUSECLICKCOUNT: -5.705 MOUSEMOVEMENTDISTANCE: -4.789 MOUSEWHEELDISTANCE: 6.366 KEYINPUTCOUNT: 37.496 LEFTMOUSECLICKCOUNT: -5.843 RIGHTMOUSECLICKCOUNT: -4.732 MOUSEMOVEMENTUPDISTANCE: -4.713 MOUSEMOVEMENTDOWNDISTANCE: -4.017 MOUSEMOVEMENTLEFTDISTANCE: -4.293 MOUSEMOVEMENTRIGHTDISTANCE: -4.972 MOUSEWHEELUPDISTANCE: -8.378 MOUSEWHEELDOWNDISTANCE: 19.537 HASFACE: -0.206 Bias: 3.368

Node 11 (Sigmoid) ———— SESSIONNUM: 42.693 CONTRIBUTION: -12.813 MOUSECLICKCOUNT: -3.298 MOUSEMOVEMENTDISTANCE: 4.053 MOUSEWHEELDISTANCE: 17.822 KEYINPUTCOUNT: -5.754 LEFTMOUSECLICKCOUNT: -2.592 RIGHTMOUSECLICKCOUNT: -3.625 MOUSEMOVEMENTUPDISTANCE: 4.648 MOUSEMOVEMENTDOWNDISTANCE: 4.584 MOUSEMOVEMENTLEFTDISTANCE: 4.972 MOUSEMOVEMENTRIGHTDISTANCE: 3.507 MOUSEWHEELUPDISTANCE: 9.195 MOUSEWHEELDOWNDISTANCE: 25.519 HASFACE: 44.286 Bias: 2.475

Node 12 (Sigmoid) ———— SESSIONNUM: -27.002 CONTRIBUTION: 28.562 MOUSECLICKCOUNT: 7.165 MOUSEMOVEMENTDISTANCE: 2.745 MOUSEWHEELDISTANCE: -1.321 KEYINPUTCOUNT: 17.091 LEFTMOUSECLICKCOUNT: 7.938 RIGHTMOUSECLICKCOUNT: -0.993 MOUSEMOVEMENTUPDISTANCE: 3.662 MOUSEMOVEMENTDOWNDISTANCE: 3.515 MOUSEMOVEMENTLEFTDISTANCE: 2.203 MOUSEMOVEMENTRIGHTDISTANCE: 2.775 MOUSEWHEELUPDISTANCE: 1.249 MOUSEWHEELDOWNDISTANCE: -3.516 HASFACE: -41.650 Bias: 0.101

Node 13 (Sigmoid) ———— SESSIONNUM: 24.136 CONTRIBUTION: 33.184 MOUSECLICKCOUNT: 2.529 MOUSEMOVEMENTDISTANCE: 1.225 MOUSEWHEELDISTANCE: 11.785 KEYINPUTCOUNT: -3.921 LEFTMOUSECLICKCOUNT: 2.451 RIGHTMOUSECLICKCOUNT: 8.785 MOUSEMOVEMENTUPDISTANCE: -0.398 MOUSEMOVEMENTDOWNDISTANCE: -0.423 MOUSEMOVEMENTLEFTDISTANCE: 3.782 MOUSEMOVEMENTRIGHTDISTANCE: 2.900 MOUSEWHEELUPDISTANCE: 8.507 MOUSEWHEELDOWNDISTANCE: 14.624 HASFACE: -38.520 Bias: -9.014

Node 14 (Sigmoid) ———— SESSIONNUM: -69.574 CONTRIBUTION: 38.242 MOUSECLICKCOUNT: 0.115 MOUSEMOVEMENTDISTANCE: 3.560 MOUSEWHEELDISTANCE: 10.413 KEYINPUTCOUNT: 3.557 LEFTMOUSECLICKCOUNT: 1.664 RIGHTMOUSECLICKCOUNT: 6.816 MOUSEMOVEMENTUPDISTANCE: 4.040 MOUSEMOVEMENTDOWNDISTANCE: 3.381 MOUSEMOVEMENTLEFTDISTANCE: 3.699 MOUSEMOVEMENTRIGHTDISTANCE: 3.625 MOUSEWHEELUPDISTANCE: 11.918 MOUSEWHEELDOWNDISTANCE: 9.132 HASFACE: -5.906 Bias: -10.009

### A.4.4 Details of the Neural Network Output

Output ======

Class 'RUI' (Sigmoid) ———————— Node 1: -5.683 Node 2: -2.454 Node 3: -41.596 Node 4: 0.869 Node 5: -6.473 Node 6: 7.135 Node 7: 4.594 Node 8: -1.933 Node 9: -6.531 Node 10: -1.085 Node 11: -0.447 Node 12: -1.623 Node 13: 0.382 Node 14: 6.897 Threshold: -1.634

Class 'OFF' (Sigmoid) ———————— Node 1: 7.403 Node 2: 7.058 Node 3: -1.033 Node 4: -2.319 Node 5: -6.749 Node 6: -4.989 Node 7: -5.347 Node 8: -1.540 Node 9: 7.125 Node 10: 8.449 Node 11: 0.324 Node 12: 6.007 Node 13: -8.363 Node 14: 7.126 Threshold: -8.792

Class 'PRACTICE' (Sigmoid) ———————— Node 1: 0.949 Node 2: -1.145 Node 3: -0.821 Node 4: 0.230 Node 5: 1.171 Node 6: -1.070 Node 7: 1.286 Node 8: 0.136 Node 9: -33.234 Node 10: 0.288 Node 11: 0.221 Node 12: 8.546 Node 13: -39.069 Node 14: -9.756 Threshold: -1.154

Class 'HELP' (Sigmoid) ————————- Node 1: -0.162 Node 2: -1.367 Node 3: 0.973 Node 4: 0.881 Node 5: 0.294 Node 6: 1.450 Node 7: 0.122 Node 8: -1.064 Node 9: -0.530 Node 10: -2.447 Node 11: 1.404 Node 12: -1.889 Node 13: -2.376 Node 14: 1.451 Threshold: -4.285

Class 'APPLY' (Sigmoid) ———————— Node 1: -22.177 Node 2: 7.244 Node 3: 4.140 Node 4: 2.717 Node 5: -1.087 Node 6: -6.310 Node 7: 0.833 Node 8: 2.074 Node 9: -1.921 Node 10: 5.695 Node 11: -18.907 Node 12: 9.596 Node 13: 18.219 Node 14: 2.259 Threshold: -11.921

Class 'REVIEW' (Sigmoid) ———————— Node 1: -0.416 Node 2: -0.280 Node 3: 1.788 Node 4: 2.591 Node 5: -1.359 Node 6: -67.754 Node 7: 5.639 Node 8: -3.243 Node 9: 61.477 Node 10: -1.156 Node 11: 3.216 Node 12: -14.418 Node 13: 1.456 Node 14: -1.455 Threshold: -1.810

Class 'PREVIOUS' (Sigmoid) ———————— Node 1: -1.128 Node 2: 1.055 Node 3: 0.447 Node 4: -1.780 Node 5: -2.232 Node 6: -1.257 Node 7: 0.730 Node 8: 2.915 Node 9: -28.513 Node 10: -2.121 Node 11: -1.135 Node 12: -0.972 Node 13: -32.695 Node 14: 1.960 Threshold: -1.510

Class 'SEARCH' (Sigmoid) ———————— Node 1: -24.774 Node 2: -0.356 Node 3: -1.774 Node 4: 1.284 Node 5: -1.360 Node 6: 3.555 Node 7: -10.976 Node 8: -0.843 Node 9: 19.315 Node 10: -4.565 Node 11: 3.142 Node 12: 0.553 Node 13: 40.620 Node 14: 0.549

Threshold: -2.426

Class 'MODIFY' (Sigmoid) —————————— Node 1: -0.750 Node 2: -0.356 Node 3: -11.607 Node 4: 0.475 Node 5: 2.558 Node 6: 2.496 Node 7: -1.135 Node 8: 0.524 Node 9: -14.025 Node 10: -20.032 Node 11: -38.072 Node 12: -2.055 Node 13: -1.311 Node 14: -0.662 Threshold: -4.033

Class 'TAKE' (Sigmoid) —————————- Node 1: -37.175 Node 2: -8.831 Node 3: -5.906 Node 4: -3.135 Node 5: -2.079 Node 6: -1.528 Node 7: -0.683 Node 8: 0.895 Node 9: -0.860 Node 10: 2.097 Node 11: -3.541 Node 12: -2.030 Node 13: -9.831 Node 14: -2.079 Threshold: -3.215

Class 'PLAN' (Sigmoid) —————————- Node 1: 1.300 Node 2: 1.689 Node 3: 0.513 Node 4: -69.985 Node 5: 1.420 Node 6: -58.856 Node 7: -7.798 Node 8: 0.563 Node 9: -2.631 Node 10: 0.942 Node 11: 0.076 Node 12: -0.282 Node 13: -1.931 Node 14: -0.439 Threshold: -3.322

# B Forms used

## B.1 Measure of Self Determination and Self Regulation

This questionnaire aims to measure with the use of a 7 point likert scale the autonomy index of a self regulated learner. It was patterned after Deci et al. (1991)'s framework on Self-Determination. The questions are as follows:

- I will participate actively in the courses I am enrolled in because I feel like its a good way to improve my understanding of the materials presented.
- I will participate actively in the courses I am enrolled in because others would think badly of me if I did not.
- I will participate actively in the courses I am enrolled in because I would feel good of myself if I did well.
- I will participate actively in the courses I am enrolled in because a solid understanding of concepts taught in class are important to my intellectual growth.
- I am likely to follow my teacher's suggestion for studying because I would get a bad grade if I did not do what he or she suggests.
- I am likely to follow my teacher's suggestion because I am worried that I am going to perform well.
- I am likely to follow my teacher's suggestion for studying because its easier to follow his/her suggestions than came up with my own study strategies.
- I am likely to follow my teacher's suggestion because he or she seems to have insight about how best to learn.
- The reason that I will work to expand my knowledge in these courses is because it is interesting to learn more about them.
- The reason that I will work to expand my knowledge in these courses is because it is a challenge to really understand how to solve problems I encounter in these courses.
- The reason that I will work to expand my knowledge in these courses is because a good grade in these courses will look positive on my record.
- The reason that I will work to expand my knowledge in these courses is because I want others to see that I am intelligent.

## B.2 Consent to be a Research Subject

Modelling Activities of Self-Regulated Learners as Contextualized Action Sequences

### B.2.1 Purpose and Background

You are being asked to volunteer for a research study. This research serves as a part of a Master's Thesis under the College of Computer Studies, De La Salle University. The purpose of this study is to identify the patterns of self-regulated learners for modeling purposes. The researcher in charge is Jordan Aiko Deja. Your participation in this research will enable you to collect data in an environment you seem fit for your personal learning and does not require a specific laboratory or a complex setup. This will involve approximately 30 volunteers.

### B.2.2 Procedures

If you agree to participate in this research, the following will happen:

- You will be asked to sit down on the sofa and wear a sensor to measure your heart rate or blood volume pulse for a maximum of 10 minutes for at least 3 different sessions.
- We will record the data collection session using a video camera. The camera is positioned in front of the subject.
- During the first 5 minutes, you are requested not to move while heart rate is being measured. The sensor to be used is very sensitive to movement.
- After the first 5 minutes, you may do activities in your seat provided that it will not obstruct the view of the camera and will not disrupt the measurements of the sensor.
- To ensure consistency of our measurement, you are requested to perform the same experiment at least 3 times during the same week.

### B.2.3 Duration

Participation in the study will require at least ten sessions, spending one hour for each session, with a total of at least ten hours.

### B.2.4 Risks and Discomfort

No special or additional sensor/requirement will be needed for the successful implementation of this research. Thus, no known risks or discomfort will be induced into the participants.

### B.2.5 Cost to the Subject

There are no costs for participating in this study.

### B.2.6 Payment to the Participants

In return for your time, effort and travel expenses, simple snacks and a token will be given after the data collection sessions.

### B.2.7 Confidentiality

Data we collect in this research are classified into 2:

- Personal Data means data that allows someone to identify or contact you, including, for example, your name, address, telephone number, email address, as well as any other non-public information about you that is associated with or linked to any of the foregoing data.
- Anonymous Data means data that is not associated with or linked to your personal data; Anonymous data does not, by itself, permit the identification of individual persons.

DLSU CCS ensures that: all research data are used only in appropriate and ethical ways; research participants are protected; and personal data safeguarded. The information provided by you will be used for research purposes. It will not be used in any manner which would allow identification of your individual responses. Anonymised research data will be archived by DLSU in order to make them available to other researchers in line with current data sharing practices. Should other researchers wish to use your data, they need to comply with an end user license, in which they agree to certain conditions on the use of the data, i.e. not using data for commercial purposes, not identifying any potentially identifiable individuals and not sharing data with a 3rd party.

### B.2.8 Consent Form

| **Please tick the appropriate boxes** | **Yes** | **No** |
|---|---|---|
| I have read and understood the project information sheet provided | | |
| I have been given the opportunity to ask questions about the project. | | |
| I agree to take part in the project. Taking part in the project will include being interviewed and recorded (video). | | |
| I understand that my taking part is voluntary. I can withdraw from the study at any time and I do not have to give any reasons for why I no longer want to take part. | | |
| **Use of the information I provide for this project only.** | | |
| I understand my personal details such as phone number and address will not be revealed to people outside the project. | | |
| I understand that my words may be quoted in publications, reports, web pages, and other research outputs. | | |
| I would like my real name used in the above | | |
| **Use of the information I provide beyond this project** | | |
| I agree for the data I provide to be archived by DLSU and by the researcher | | |
| I understand that other genuine researchers will have access to this data only if they agree to preserve the confidentiality of the information as requested in this form. | | |
| I understand that other genuine researchers may use my words in publications, reports, web pages, and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form. | | |
| **So we use the information your provide legally** | | |
| I agree to assign the copyright I hold in any materials related to this project to Jordan Aiko P. Deja | | |

Name of participant: ____________________ Signature: ___________ Date: ________

Name of researcher: **Jordan Aiko P. Deja** Signature: ___________ Date: ________

For further information you may contact Jordan Deja, jordan.deja@dlsu.edu.ph, De La Salle University, 2401, Taft Avenue, 1004 Philippines.

## C Resource Persons

Dr. Rafael Cabredo
Faculty Member
College of Computer Studies
De La Salle University
rafael.cabredo@dlsu.edu.ph

Dr. Judith Azcarraga
Faculty Member
College of Computer Studies
De La Salle University
jay.azcarraga@dlsu.edu.ph

Dr. Remedios Bulos
Faculty Member
College of Computer Studies
De La Salle University
remedios.bulos@dlsu.edu.ph

Dr. Merlin Teodosia Suarez
Dean
College of Computer Studies
De La Salle University
merlin.suarez@delasalle.ph

Dr. Paul Salvador Inventado
Post-Doctoral Fellow
Carnegie-Mellon University
pinventado@cmu.edu

Mr. Juan Lorenzo Hagad
Graduate Student - Humanware Innovation Program
Graduate School of Information Science & Technology
Osaka University
hagad@ai.sanken.osaka-u.ac.jp

# D Personal Vitae

Mr. Jordan Aiko Deja
#12 Wayfield St Laguna BelAir 1,
Sta. Rosa City, Laguna
+63 9177190653; 9062426745
jordan.deja@dlsu.edu.ph