Emotions identification utilizing periodic handwriting on mobile surfaces

Viktors Zagorskis¹, Atis Kapenieks², Aleksandrs Gorbunovs³
¹, ², ³ Riga Technical University, Riga, Latvia

ABSTRACT

The purpose of this study is to between the learners’ emotional characteristics and styles in touch screen environments. We propose a method to identify the variety of learners' boredom in the learning process utilising handwriting data. The novelty of the method is to avoid explicit polling of learners how do they feel. We use the recurring acquiring of personal handwriting data utilising the computational power of both a mobile device and cloud-based resources. Also, we use machine learning-based sentiment detection in the research. We smoothly inject periodic handwriting tests convolving them with learning objects to study the correlation between learners’ emotions dynamic, they demonstrate, and the ability to focus and think critically. With the help of machine-learning methods and new communication protocols, we can step up the student-centric mobile-based education process by taking advantage of the latest achievements in a big data analysis and cloud computing. Also, we clarify the conceptual model for the testbed used in the experiment. The findings may likely impact the future personalized e-learning systems.

Keyword:
Boredom detection
Mobile learning
Handwriting detection
Machine-learning
Online learning

1. Introduction

It is evident that, without a clear rationale, innovative strategies might disrupt long-lasting education standards with incredible speed. Overall, mobile devices with built-in touchscreen interfaces opened a new era of human-machine interactions [1]. In the early days of mobile-based learning, Howard Gardner [2] argued that there is a challenge to build and maintain an educational system that assumes that everyone can learn the same materials in the same way. Nowadays, mobile touchscreen devices provide interacting in a user-friendly fashion: (1) adaptive content on the screen, (2) fast learners response to content by ticking, clicking or dragging, (3) small device footprint, and (4) online internet access in most e-learning environments.

Today, learning systems allow adjustment of content outlook by users helping to nail learner’s attention to the subject [3]. On the e-learning market, most systems are designed to render learning content in a screen-adaptable fashion addressing learners efficiently. Mobile learning does increase the chances of success by providing multiple paths for learning [4].

The latest trends on the e-learning education market introduce personalized content as an intelligent machine response to learners' needs helping reducing response time. In this case, learners receive feedback almost without delay helping to keep learners’ attention on a learning object. In many designs [5], the computational power of mobile devices is enough to solve some fundamental machine learning problems without interaction with high-performance computing systems. For instance, application to recognize handwritten digits can run as a local process on an average performance mobile device. The challenge is to utilize the computational power of mobile devices in new directions.
With the aim of that, we propose the method to identify a variety of students’ emotional states during the coaching process based on handwriting data produced by learners using personal touchscreen devices.

The research organized in the following way: In Section 2 - the reflection of related theories, methods, and approaches. In Section 3, at first, we designate more related methods. Then, we through two research questions related to the implicit detection of learners’ boredom. Next, we specify the environment for the experiment, describe organizational settings, resources, and data flow. Finally, we describe a simple algorithm for implicit boredom detection. In Section 4, that is the most important in the research; we describe the experimental results. Section 5 concludes the paper.

2. Background and related work

2.1. Personal learning environments (PLEs)

Mobile Personal learning environments may utilize combined technologies both local and cloud-based that support a holistic learning process [6] with the aim to collect data and share resources among learners [7].

PLEs constitute technology-enhanced learning (TEL) settings [8]. As Rahimi states [9], TEL provides an environment for (1) interactivity, (2) communication and (3) collaboration with a learner-centric focus on informal learning.

Besides, learning experiences are ongoing and personally motivated [9] regardless of technology. A formalised mobile technology implementation aimed to enhance teaching and learning [10]. Motivation is the main factor to utilise and promote mobile technology.

2.2. Motivation and boredom

In the education process, learners start to tackle the content being engaged just for curiosity. Engagement is a motivational factor. Here, we follow the learners’ motivation model consisting of two essential components [11]: (1) internal motivation based on internal information (e.g., feelings or thoughts), and (2) external motivation based on external information (e.g., environmental stimuli).

Overall, after some time interval of activity and paid attention, learners can get boring. Such switching is a typical human reaction not only to the learning environment and it is content, but also to any environment in general [12].

Sometimes, a negative emotional state may lead to an active search for stimulation. If learners get bored from activity in everyday life, they often turn to their mobile phones to seek that stimulation [13]. Moreover, vice versa, mobile device usage stimulates returning to boring state again. That is, boring states appear periodically.

Although, boredom is not trivial. In modern science, Eastwood [12] propose that boredom can be defined concerning attention. Eastwood states: “boredom is the aversive state that occurs when we (1) are not able to successfully engage attention with internal or external information required for participating in satisfying activity, (2) focused on the fact that we are not able to engage attention and participate in satisfying activity, and (3) attribute the cause of our aversive state to the environment.”

Another model, recently proposed by Goetz and Frenzel [14] includes four types of boredom differentiated based on degrees of attraction (valence) and arousal: (1) indifferent, (2) calibrating, (3) searching, and (4) reactant boredom. This model based on strong empirical evidence indicating that boredom corresponds to many detrimental experiences of learners and their behaviours.

For the current research, we also accept in the 1980s by M. Apter proposed simplified motivation sequence model based on reversal theory [15] to describe the motivation, emotion and personality. Also, M. Apter’s thoughts were adopted to propose a concept for cognitive energy based Virtual Student model [16]. In the motivational sequence model (see Figure 1.), the excitement state defined as the initial and pleasant one, therefore indicating Virtual Student's addiction to learning. Concluding, we claim that a tendency to begin to learn at the emotional excitement phase correlates with the real learners’ motivation keen to learn.

In reality, by standing anxiety as the next state in the sequence, we consider that human behaviour follows the same path: positive emotions and negative emotions are in the constant cycling process. Therefore, we simplify the motivation sequence model avoiding the direct transition to any other state except categorised as emotionally opposite.
2.3. What to avoid by measuring boredom?

Mapping results of Pattyn et al.’s (2007), we conclude that interaction with learners is explicitly asking them to express their emotional states may lead to data quality degradation in the case of a questionnaire-based method by producing of subjective data and having similar to noise impact on results. Levels of boredom consistently increase in tandem with a decrease of activity or performance [17]. Also, we agree with findings that experimental procedures for emotional and psychological state detection might unintentionally lead to feelings of boredom in participants and thereby influence the data collected [18]. We notice this crucial point to avoid in our experiments.

3. Research questions, methods and tools

In this section, we describe some of the existing methods and tools that used for experiments. As tools, we consider the utility of extensive computational power of modern mobile devices that provide functionality and performance. Also, a comprehensive network protocol HTTP/2 selected for experiment making WEB applications faster, simpler, and, what is curious, more robust at the same time. Such a strategy will allow utilising GOOGLE machine-learning services simplifying experimental test-bed environments. Although, on the preliminary stage, proposed settings will operate only in online mode.

3.1. Recent methods and research questions

Recent studies and methods [13] for boredom detection based on automatically inferring boredom from mobile phone usage. M.Pielot et al. introduce the user-independent machine-learning model of boredom–leveraging features related to (1) recency of communication, (2) usage intensity, (2) time of day, and demographics. As results show, M.Pielot et al.’s experiments infer boredom with an accuracy of up to 82.9%.

Before that, Bixler and D’Mello show that the most popular methods for boredom detection are: (1) facial expressions, (2) speech and voice features, text analysis, and physiological signals from mobile devices [19].

According to Bixler and D’Mello, exist methods how to detect boredom during writing tasks through logging the writers’ keystrokes. Here, keystrokes had comparably low predictive power – roughly 11% above chance – for discriminating engagement-neutral and boredom-neutral states. By adding stable traits of the learners to the model, as Bixler and D’Mello argue, prediction performance could be notably improved.

With the aim of that, we formulate research questions, namely:

(RQ1) - is it possible to utilise PLEs surfaces with the aim to provide implicit measures of boredom state building machine-learning model using keystroke identification?

(RQ2) - does indirect data acquisition method allow get data variability related to the boredom?

3.2. Approach overview

To answer RQ1 and RQ2, we propose a combined approach to identify boredom by learners’ handwriting style analysis in touch screen environments. The first component of the approach is the central - organisational landscape or set of organisational activities (see Figure 2) for data collection and addressing. The second component organises data gathering ready for delivery to relevant involved parts. The third component is computing and networking resources involving local - Learner's mobile device and remote - utilising cloud-based machine-learning computing. The last, fourth component is a set of algorithms used to support the experiment.
During the learning process, learners receive small regular assignments with content that relates to learning objectives. Learners are asked to interact with learning content. After learners apply response in a handwriting form captured handwriting style is partially processed, analysed locally and sent to an application running machine-learning algorithms to identify specific patterns.

Figure 2. The Organisational Landscape environment holds and control learners’ data gathering, distribution, and flow among computational and storage resources.

3.3. Organisational activities

From an instructional point of view, we propose the following activities involving:

- Producing a micro-learning content including learning units for interaction with learners with the aim to initiate a response in the form of handwriting;
- Micro-learning content delivery to learners using online network communication protocol HTTP/2 and utilising the internet browser on a mobile device;
- Sequential acquiring of personal handwriting data;
- Partial data processing employing the computational power on mobile device;
- Data delivery for further analysis, particularly, for application of machine-learning algorithms utilising GOOGLE API;
- Building of the unsupervised machine-learning models;
- Identifying indicators of learner's boredom variability over time;
- Recognising boredom related clusters.

3.4. Data processing

Data Processing component involves WEB application, partially sharing the responsibility for experiment with a set of algorithms and a set of resources (see Figure 2.). Mainly, a JavaScript-based algorithm collects user interaction with mobile device surface identifying touching style (one finger or more) coordinates on HTML CANVAS component. At first, utilising HTML CANVAS API collected DATA along with time-stamp is processed for eventually locally controlled feedback. Next, collected DATA prepared for delivery to cloud-computing resources with the aim to utilise GOOGLE API applying machine-learning methods. Finally, processed data used for the preparation of insights, visualisation, and control of the learning process.

3.5. Computing and networking resources

For experiment organisation, we involve both computational and communications resources. A mobile device with the surface for interaction is suitable for the experiment. Today, in 2018, computational power and networking capability of the average mobile device are enough for the operation in research, and therefore more detailed specification is out of the paper scope. The role of portable learning resource in the experiment is the CLIENT.

Another computational power for the experiment provided by cloud-computing resource organised on the resource APACHE CLOUDSTACK deployed on institutional hardware maintained by Riga Technical University and hosted in Riga, Latvia. Cloud-based support allows us to scale resources easily on demand of amount of memory or the number of processors needed. The role of cloud-based resource in the experiment is the SERVER. Also, the SERVER side provides DATA storage according to DATA Model.
As the more complicated for production, however quite easy used for prototyping and experimentation, we decided to employ the side part - GOOGLE computational resources achievable through GOOGLE API. Here, we accept GOOGLE help to engage its ARTIFICIAL INTELLIGENCE using machine-learning libraries (e.g. using TensorFlow, an open-source Python library developed by the Google Brain labs). The role of GOOGLE resource in the experiment is the AI-ML.

All the resources involved in the experimental setting communicate over the network. Mobile devices use a 4G network (depends on the provider). Further quantitative description of connections among SERVER, CLIENT, and AI-ML stays out of this document scope.

Concluding this subsection, a mobile device (CLIENT) used (1) to communicate with a server side of the experimental setting, (2) to deliver micro-learning content to the learner, (3) to gather handwriting data from device surface, (4) for adding a time-stamp to data, (5) to deliver data to the SERVER side. SERVER communicates with AI-ML to delegate resource consuming tasks to GOOGLE. Also, the SERVER side provides delivery of data queries, insights, visualisation, and control of the learning process.

3.6. Introduction to algorithm

In this section, as well as in the whole paper, we omit a detailed description of a software stack used to implement the algorithm. However, the DATA model will be sketched in the text later helping to interpret results.

To build the algorithm, we follow the best findings and practice in the adaptive mobile learning [20, 21]. Also, we take into consideration some relevant findings starting from initial ideas [22], following through the pitfalls, issues and challenges [23] to recent achievements [24, 25] that researchers figured out in handwriting recognition.

Based on recent research findings and available technology, we propose the algorithm adaptable to the variations of learners' perception and specific cognitive abilities.

We name the algorithm as Simple Boredom Identification algorithm, in other words, Simple Algorithm for Boredom Identification - SABI use variables and data models:

- **Response Time Window (RTW)** - a time window that allows to accept all handwritten symbols entered on the mobile device surface;
- **Threshold Time (ThT)** - an average task completion time, individually recorded for each user;
- **Drawing as a Matrix (DaaM)** - a two-dimensional array, holding by handwriting sketched drawing captured from the mobile device surface after forming process completion;
- **Drawing Start Time (DST)**;
- **Drawing End Time (DET)**;
- **Calibration Timeouts (CalTo)** - an array, holding time out settings for each calibration round;
- **Drawing Animation (DAnim)** - an array, holding drawing process animation over the forming time;
- **Reference Drawing (RefD)** - a reference to file on the SERVER side that holds the screen displayed on the CLIENT device. The data model for RefD includes screen (a browser window) settings gathered from CLIENT device;
- **GOOGLE Transactions Model (GTM)** - a data model helping to reconstruct system in dynamics;
- **SERVER Data Model (SDM)** - a data model used by Data Analytics Module.

In short, SABI algorithm operates as follows: IF the Learner's Response Time grows AND becomes longer than expected from calibration results, THEN learner is classified as bored.

 Calibration of the SABI algorithm starts using the game like approach. Let's describe the Calibration Game.

3.7. Calibration experiment

For the first experiment with system calibration, we create a dynamic WEB content consisting of 10 slides displayed in a sequence on the viewable area on the mobile device surface. Each slide holds visible for some arbitrary time. After that, a slide will be changed to another. For example, if each slide show lasts for one second, during the 10 seconds exercise, in that time interval learners got ten different messages.
Initially, at **first round** a user opens the web browser on the mobile personal learning environment, named CLIENT. After connection to the TELECI site, follows formal registration procedure by entering of the familiar nickname. For simplicity sake, we accept only 4-digit "nick numbers", similar to the PIN code in phone or banking card.

In the first introduction round, we give one symbol representing a single digit number per slide. Since all experiment participants are familiar with elementary numbers, the sequence is displayed with the rate equal to 2 digits per second. The first round is quite easy because the numbers are dynamically presented in a natural growing progression: from 0 to 9.

Therefore, after the first-round participants were introduced to this simple information to remember, we assume that all learners remembered presented numbers and their sequence. That is a crucial assumption for calibration.

In the **second round**, we propose to draw over the given numbers by a finger or by stick depending on the device and learners' habits. Complementing the algorithm, we use the following logic: after the moment when the user touches the surface for the first stroke in a number, the algorithm starts a countdown timer to identify time window for input acceptance. By default, the time window size is one second. In the case of writing, lasting longer than one second, the algorithm just branched to waiting until any pressure to the surface will be released. Now, by the handwriting drawn number is prepared for the transfer to the SERVER.

Along with a (1) final drawing of number, (2) start event timestamp, (3) stop event timestamp, (4) a drawing process animation over time, and (5) reference to displayed learning problem (a number shown in the screen) is sent to SERVER. After data received, SERVER applies some normalisation to the drawing and utilise GOOGLE API to identify the number. Supervised machine-learning algorithm responds with a probability distribution among possible outcomes: 0 to 9, and Not a Number (NAN). According to the algorithm, the SERVER should decide regarding the next instructional steps. Since SERVER already knows the number displayed to the CLIENT, it is quite easy to determine how to respond. After entering all of the ten digits, the second round is finished.

The **third round**. We repeated the first round reversely - displaying numbers from 9 to 0.

The **fourth round**. Again, we proposed to draw over the numbers appearing on the screen reversely. The algorithm applying time window for handwriting numbers is equal to described in round two.

The **fifth round**. Calibration includes the task to redraw numbers in up-going sequence - number by number without time limitation. Although, exist hard timeout limitation value stored in (CalTo) array. The algorithm should respond to each handwritten symbol with a typed one number guessed by ML-AI.

The **sixth round**. Calibration includes the task to redraw numbers in down-going sequence - number by number without time limitation. However, exist hard timeout limitation value stored in (CalTo) array. The algorithm should respond to each handwritten symbol with a typed one number guessed by ML-AI.

For the method used in experiments, we use no more calibrations because of concerning to get users in boredom or anxiety.

As follows from the calibration experiment, the system is calibrated to recognise learner's handwritten numbers. The method used for boredom detection involves threshold time (ThT) monitoring in combination with factors from Drawing Animation (DAnim) array learned on the calibration stage. At this stage, the learning system is ready to switch to a micro-learning mode and deliver some learning content. As usually at the micro-learning approach, after content presentation follows assignment. To identify possible switching to state boredom, we propose to formulate assignments notifying learners to respond to tasks using the handwriting method.

4. **Preliminary results and discussion**

Overall, the experiment organisation and technical implementation aspects described in the previous section. In this section, we focused on the discussion of preliminary results from the first runs.

4.1. **Experiment**

To collect data for evidence on the relationship between handwriting style and boredom, we designed and deployed the software stack with the implementation of the SABI algorithm. We launched the study in October 2018. For widespread distribution, we made the application available on the TELECI project site. Participants of the experiment (n=54) were introduced to the research topic and, after the registration in the SERVER, were asked to follow the instructions on the CLIENT device. Experiment organisers and
participants came to the agreement that during algorithm testing users will be honest regarding their feelings. No issues were identified during the SABI algorithm runtime and instructional experiment.

At the instructional level, we injected additional explicit dialogue into user interface about the emotional feelings of a user on the CLIENT side. The specific question was simple: "Do You feel boredom?" On the bottom of the browser window were thrown two buttons: "Yes", "No". By tapping on the button, users were able approved their choice.

Therefore, under the assumption of trust to the data entered by users, subjective data from participants were collected and stored. The subjective data will be used on later stages to validate the SABI algorithm.

4.2. Preliminary results

To illustrate preliminary results, we present the data collected in the experiment. We show the results of the first users' experience. We observe by the SABI algorithm recorded data for each user participating in the calibration experiment (n=54). At first, we show a histogram depicting the average completion time - (ThT) for the calibration task, round two, for a single digit to draw. The Symbol to be drawn is "1" (see Figure 3.). The histogram has 50 bins. On vertical axis - the frequency of occurring of threshold time. The statistics are as follows: mean value is 1.1924, median 1.0957, standard deviation 0.4383.

![Figure 3. Histogram of the average completion time for the calibration task. Round two. On vertical axis - the frequency of occurring of threshold time. Users (n=54) draw the symbol “1”. Mean value 1.1924, median 1.0957, standard deviation 0.4383.](image)

As the next, we present the histogram of how users draw the number “5” (see Figure 4.). Here, we figure out that users spend more time to accomplish the task - mean value is 1.5696, median 1.457, standard deviation 0.6679.

![Figure 4. Histogram of the average completion time for the calibration task. Round two. On vertical axis - the frequency of occurring of threshold time. Users (n=54) draw the symbol “5”. Mean value 1.5696, median 1.457, standard deviation 0.6679.](image)
Comparing histograms, we also find that users spend more time drawing number “5” rather than number “1”. Such observation can be accepted due to more strokes to be drawn. Also, in the case with number “5” drawing process is more widely distributed over the time.

In Figure 5, we show a variation of (ThT) data over the experiment time running the real micro-learning course. Users were asked to interact applying to interact using handwriting on regular time intervals after watching the animated JavaScript content lasting for 10 seconds. On the plot, on vertical axis resides by the SABI algorithm identified threshold time spent for the task completion. The ThT value is computed taking into the account three factors: (1) a user reaction time on a problem appearance identification, (2) cognitive reaction time on the problem, and (3) expected completion time using handwriting response. Factors 2 and 3 belong to metadata factors group because depend on individual skills in the problem domain.

Figure 5. The real micro-learning course. Six handwriting tasks (h1 to h6) delivered for completion. The variation of average task completion time (ThT) data over the experiment time. Each point represents individual user delay data of interaction with learning content at specified a time moment.

In Figure 5, also, we depict six periodic handwriting assignments - tasks: h1, h2, h3, h4, and h6. Each job has its colour on the plot. Each dot on the figure represents a single user response. Following the data in the plot, we conclude that the distribution of ThT values has specific skewness towards the higher values identifying some lazy (or already boredom) users residing quite far from the main group. We claim the possible reason for such as boredom cluster detected is trivial arithmetic’s problems were asked to solve in mind and answer quickly. In the opposite case, the answer delay is subject to study. Task h4, which is the last in the experiment, identifies larger variance of ThT value in comparing to the variation for tasks h1 and h4. Also, the first task (h1), on average, is accomplished in a little bit but longer time interval. We explain this phenomenon as a warm-up factor before users identify the micro-learning workflow.

At this stage, we conclude preliminarily that the research question (RQ1) is answered. For the answer on (RQ2), we need to formulate an appropriate hypothesis and apply inferential statistical analysis to explore more from data gathered by the SABI algorithm.

4.3. Future work

Concluding the section, we formulate the hypothesis that the decrease of the time interval between periodic handwriting tests leads to the early detection of boredom. To test the hypothesis, we consider using more samples of data from both subjective users’ feelings and records gathered by the SABI algorithm. However, an additional report is the topic for the next paper.

As recent findings [25] show, further research directions to develop the SABI algorithm to experiment with online handwriting recognition are perspective to be used in e-learning ecosystems.

5. Conclusion

Current paper reflects some core principles used to organise experiment with hidden or implicit method to identify boredom state of users consuming mobile devices. We conclude that:
Utility of HTTP/2 protocol in online mode and technology based on Progressive Web Apps approach may likely impact the future personalised online learning systems;
Overall, GOOGLE AI & Machine Learning Products are easy to use due to their mapping to most popular programming languages (C#, GO, JAVA, PYTHON, RUBY, NODE.JS) because they are directed to apply in everyday practice for instructors with intermediate programming skills;
Delivery of a learning content with handwriting based interaction and students' feedback is a subject for next research;
Mobile device surface usage for emotional state variability detection can also be further studied.

As in early age of mobile learning, it is still dynamic, individual and comprehensive. It provides restrained by teams or self-paced learning experiences from many sources, enabling learners to select a favoured course, content's format, education method, or learning content provider. By artificial intelligence supervised learning options guarantee more efficient learning outcomes expressed, as usual, in gained knowledge and developed skills carried by solid emotional background.

Acknowledgement
This research has been supported by a grant from the European Regional Development Fund (ERDF/ERA) project “Technology Enhanced Learning E-ecosystem with Stochastic Interdependences - TELECI”, Project No.1.1.1.1/16/A/154.

References