Learning Pulse: Using Wearable Biosensors and Learning Analytics to Investigate and Predict Learning Success in Self-regulated Learning

Daniele Di Mitri, Maastricht University, d.dimitri@student.maastrichtuniversity.nl
Maren Scheffel, Open University of the Netherlands, Maren.Scheffel@ou.nl
Hendrik Drachsler, Open University of the Netherlands, Hendrik.Drachsler@ou.nl
Dirk Börner, Open University of the Netherlands, Dirk.Boerner@ou.nl
Stefaan Ternier, Open University of the Netherlands, Stefaan.Ternier@ou.nl
Marcus Specht, Open University of the Netherlands, Marcus.Specht@ou.nl

Abstract: The Learning Pulse study aims to explore whether physiological data such as heart rate and step count correlate with learning activity data and whether they are good predictors for learning success during self-regulated learning. To verify this hypothesis an experiment was set up involving eight doctoral students at the Open University of the Netherlands. Through wearable sensors, heart rate and step count were constantly monitored and learning activity data were collected. All data were stored in a Learning Record Store in xAPI format. Additionally, with an Activity Rating Tool, the participants rated their learning and working experience by indicating the perceived levels of productivity, stress, challenge and abilities along with the type of activity. These human annotated labels can be used for supervising machine learning algorithms to discriminate the successful learning moments from the unsuccessful ones and eventually discover the attributes that most influence the learning process.

Keywords: Learning Analytics, Biosensors, Affective Computing, Wearable Enhanced Learning

Introduction

This paper presents the development of Learning Pulse, a study designed and conducted within the Technology Enhanced Learning Innovations (TELI) department of the Welten Institute, a research centre at the Open University of the Netherlands. Learning Pulse, funded via the Learning Analytics Community Exchange project[8], is a research initiative of the Learning Analytics and the New Learning Experience thematic working groups of TELI in cooperation with the Department of Data Science and Knowledge Engineering (DKE) of Maastricht University. The study took place from September to December 2015 with the idea to combine wearable technologies with learning activity data in order to analyse and empirically infer the learning patterns of an individual by means of machine learning, data mining and information visualisation techniques. The approach used in Learning Pulse is an example of learning analytics tailored to bridge physical with digital learning spaces (CrossLAK theme 2) and of combining data from varied heterogeneous data sources (CrossLAK theme 4) by means of the new xAPI data standard.

Rationale

Learning Pulse aims at modelling the endeavours of an individual learner in the context of self-regulated learning or cognitive work. Thus, there are three main assets that constitute the study: (1) the employment of biosensors to collect physiological data, (2) the use of regular subjective activity reporting and, (3) the use of predictive and learning platform independent learning analytics. The context and the three research assets are hereby described.

Self-regulated Learning

Self-regulated learning is the process whereby learners set goals for their learning and monitor, regulate, and control their cognition, motivation, and behaviour, guided and constrained by their goals and the contextual features of the environment (Pintrich & Zusho, 2007). The first important assumption that holds for a self-regulated learner is the strong engagement with the learning activity and the desire of improving the learning

[8] http://www.laceproject.eu

performance (Butler & Winne, 1995). Learning Pulse builds on this natural desire and aims at developing a model to support this disposition. The second assumption is that each individual learns differently and has his/her own goals, cognition and motivation (Ryan & Deci, 2000). The predictive models, which will be described later, will therefore be specific and valid only to one specific learner.

Biosensors for Learning

Biosensors are getting increasingly available to the general public: embedded in wearable technologies, biosensors are more and more being used in industries like healthcare, fitness, and sports (Swan, 2012). Multi-sensor approaches, combined with cardiovascular activity, are also a growing trend in the industry (Schneider et al., 2015). Such involuntary responses are easier and cheaper to measure but more difficult to interpret, being the result of a complex system of stimuli (Pijeira-Díaz et al., 2016). The role of the physiological footprints over psychological states has been subject of research for several decades and has already offered interesting insights. Boucsein & Backs (2000) for example relate significant change in physiological responses to common physical and mental activities. Among all physiological responses heart rate is accounted to be the most recurrent and thus most predictive one. In related research there is, however, little focus on the role that biosensors have in enhancing learning (Schneider et al., 2015). Learning Pulse aims to address this challenge, researching for meaningful patterns in physiological responses in self-regulated learning.

Predictive Learning Analytics

The process of exploiting learning data with the aim of understanding and thus optimising the learning practice is usually referred to as learning analytics (LA). Consisting of several different disciplines including learning science, software engineering, statistics, data mining and information visualisation, LA is a modern and powerful tool for sense-making of educational data (Siemens & Baker, 2012). In particular, the capacity to make predictions on learning outcomes makes learning analytics highly valuable for all stakeholders in education (ECAR-ANALYTICS Working Group, 2015). A common drawback on the application of LA is to limit the scope of the learner's activity only to one specific virtual learning environment (VLE) or learning management system (LMS). However, as Suthers & Rosen (2011) point out, learning is often distributed across multiple media, websites and networked environments; the learning activity traces may be fragmented and not match analytic needs. Learning Pulse aims to address this issue by employing platform-independent learning analytics: instead of looking at a particular application or environment, it logs the use of all software in use during the learning activity.

Method

The overarching research question is given below, followed by a possible follow-up question if question 1 is answered positively. This second research question seeks to understand if, by leveraging biosensor data, by scoring and predicting learning success and by constantly feeding back these predictions to the learner, the learning and the cognitive work performance will eventually increase.

1. *Are physiological responses like heart rate and step count, when associated with learners' activity data, predictive for learning and cognitive working performance?*
2. *Can biofeedback techniques be employed to improve learning and cognitive working performance?*

In Learning Pulse the hypothesis range is thus defined by the degree of success in the learning activity. As first theoretical ground for learning success, the concept of Flow is used. Theorised by the Hungarian psychologist Csikszentmihalyi, the Flow is a mental state of operation that an individual experiences when immersed during a state of energised focus, enjoyment and full involvement in the activity process. Being in the Flow means feeling in complete absorption with the current activity and being fed by intrinsic motivation rather than extrinsic rewards (Csikszentmihalyi, 1997). According to Csikszentmihalyi, the Flow happens whenever there is a balance between the level of difficulty of the task (the Challenge dimension) and the level of preparation of the individual for the given activity (the Abilities dimension). When these two dimensions are maximised, the Flow is likely to manifest.

Experimental Setup and Task Description

The experiment lasted twelve working days and involved eight participants, four males and four females, aged between 25 and 35, all of them doctoral students at the TELI group of the Open University of the Netherlands with backgrounds in different disciplines including computer science, psychology and learning science. Being PhD students, they can be considered both learners and cognitive workers. To carry out their own research, the participants used their personal laptops and were asked to install a preconfigured software tracking tool. All participants were also provided with a wearable fitness tracker and were asked to sign an informed consent form about the use of their personal information for research purposes. During the experiment the participants were asked to continue their research activity as usual and, while doing that, rate their learning activity every hour between 8AM and 7PM, for those hours that they worked. The ratings were collected through a web application developed ad hoc, named Activity Rating Tool. In addition, to get more insights into how stressful moments are reflected in the heart rate changes and self-perceived productivity, the participants were asked to do additional tasks, such as delivering presentations or submitting short abstracts about the topic of their research.

Data Sources

Learning Pulse uses four sources of data as detailed below: biosensor data, user activity data, rating data, and weather data. All the collected data are summarised into an Entity-Relation Model shown in Figure 1. Physiological data were collected using Fitbit Charge HR[9], a wristband that every participant wore throughout the whole experiment. A Fitbit is a commercial wireless tracker that embeds different sensors to track a number of statistics in real-time, including heart rate, steps taken, distance travelled, calories burned, stairs climbed and active minutes throughout the day. The two measurements of interest for Learning Pulse are the heart rate and step count, updated respectively every five seconds and every minute. With such frequency the values of these two variables are stored for every participant from 8AM until 8PM during the 12 days of the experiment. The other biosensor, used however for only one participant, observes two measures: skin conductance, updated up to four times every second, and the noise level, updated with the same frequency. The values of this sensor are directly stored in the Learning Record Store in xAPI format (see below).

The activity data was obtained using RescueTime[10], a time management software meant to be a working efficiency tool. RescueTime can be installed on different platforms and generates personal analytics by logging the applications running on the laptop or mobile device. Every five minutes, RescueTime stores an array containing the applications in use, weighted by their duration in seconds, into a proprietary cloud database. Each application is also given a category.

The users' ratings were collected through the Activity Rating Tool, a web application developed in Python running on Google App Engine server. When a user accesses the app and authenticates into the system, he/she is able to click onto one of the past learning intervals (timeframe) of that current day. To simplify the data collection process, the timeframes to be rated have a fixed length of one hour: they begin and end at full hours (e.g. the first timeframe goes from 8AM to 9AM). To rate the activity each participant is asked first of all to choose, from a closed list, the category of the main activity performed during the selected timeframe: (1) reading, (2) writing (e.g. a paper, or a presentation), (3) meeting (both online, offline), (4) communicating (with email, or chat), or (5) other (e.g. going to lunch, having a break). Then, through a sliding button, a value ranging from 0 to 100 has to be chosen for each of the following questions:

- Productivity: How productive were you?
- Stress: How stressed did you feel?
- Challenge: How challenging was the activity?
- Abilities: How prepared did you feel for the activity?

In order to make the ratings as accurate as possible, at the end of each timeframe, participants were encouraged to rate their activities directly after a timeframe concluded by an email reminder to the personal inbox of each participant. All the ratings were stored in Google Datastore and sent in xAPI format to the Learning Record Store as detailed in the section 3.4.

Weather condition may also have an influence on individual learning performance. For this reason it has been decided to model the weather as an extra feature of the learning process. The web service

[9] https://www.fitbit.com/chargehr
[10] https://www.rescuetime.com

Weather Underground[11] was chosen for providing free historical weather data. Updated every 30 minutes, the weather data consist of four attributes: temperature, pressure, humidity, hourly precipitation.

Data Collection

To support the collection of such heterogeneous types of data, Learning Pulse uses a flexible software architecture. In Figure 2 all the different components are divided in three functional layers: (1) the Application Layer, (2) the Controller Layer, and (3) the Data Layer.

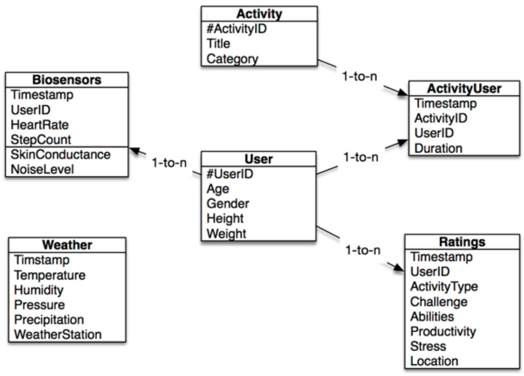

Figure 1. Learning Pulse Entity Relation model

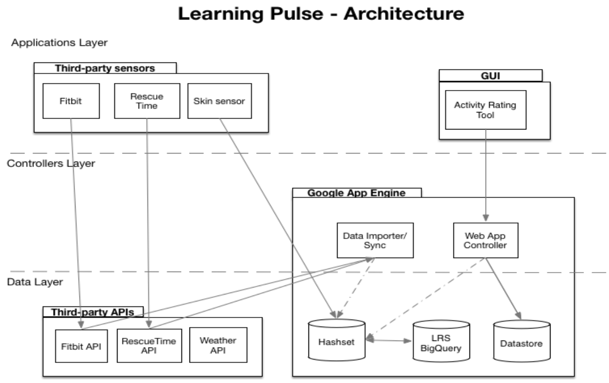

Figure 2. Learning Pulse three-layer architecture

The Application Layer is constituted by the user interfaces, the sensors and the third-party applications which the user directly interacts with. The components of this layer are responsible for collecting the data of the environment and sending them to the Controllers. In this layer fall the Activity Rating Tool, the Fitbit tracker, the skin-conductance sensor and the RescueTime software. The Controller Layer is the core of the software architecture responsible for the processing, manipulation and storing of the data collected. It includes the server-side web application of the Activity Rating Tool, the management of the user accounts, and the data-importing mechanisms to gather the data from the third-party datastores. Part of the Controllers is also the Data Transformer, which prepares the data in the correct representation. The Data Layer is the layer where all the data reside. It includes both the internal databases, i.e. the Datastore and the Learning Record Store, and the third party cloud datastores such as the Fitbit and RescueTime ones.

Data Storing

The standard chosen to store Learning Pulse data is the Experience API (xAPI). The xAPI is an open source API and RESTful web service, with a flexible standard based on learning statements with the format actor-verb-object. The statements, generated in JSON format, are validated by and stored in a Learning Record Store (LRS). The main advantage of xAPI is interoperability: learning data from any system or resource can be captured and eventually queried by third party-authenticated services. In Learning Pulse xAPI statements are opportunely designed: to store for example one heart rate value for the user ARLearn7, the xAPI statement will carry the following meaning "At timestamp 2015-11-24 08:05 ARLearn7 experienced Heart-Rate of value 87". One statement is hence generated for every sensor at any value update. This results in a considerable size of information to be stored. To handle the load of information the Learning Record Store is implemented with Google Big Query Datastore, a non-relational and highly scalable datastore which is able to query massively large datasets in few seconds.

Hypothesis Modelling

A graphical representation of Csikszentmihalyi's model is given in Figure 3. Having sampled, through the Activity Rating Tool, Challenge and Abilities as normalised numerical values the Flow can be calculated as follow:

[11] http://www.wunderground.com

$$F_{ij} = \left(1 - A_{ij}\text{-}C_{ij}\right) * \frac{\left|A_{ij}\text{-}C_{ij}\right|}{2} \quad (1)$$

where F_{ij} is the Flow score for the learner i_{th} at the timeframe j_{th}; A_{ij} and C_{ij} are the values rated by the learner i_{th} at the timeframe j_{th} for, respectively, level of Abilities and Challenge. In the scatter plot in Figure 4 the ratings of one participant are plotted in a two dimensional space and are coloured depending to their value of Flow calculated with formula (1).

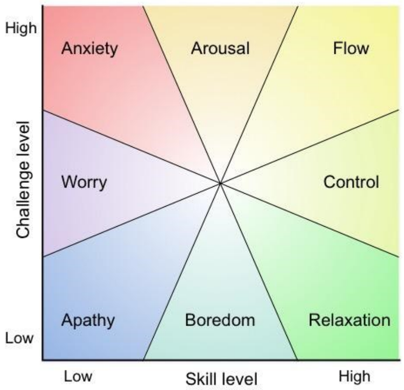

Figure 3. Csikszentmihalyi's Flow model

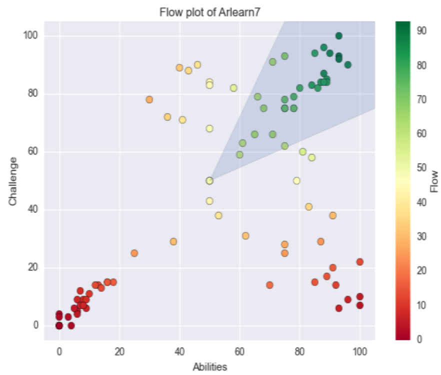

Figure 4. Scatter plot of the ratings on challenges and abilities rated by ARLearn7

To check the validity of the hypothesis, the flow score will be validated by computing its correlations with productivity and stress, in order to check if increasing flow corresponds to increasing productivity and stress. The use of the Flow score enables a representation of the "learning success" of an individual at a particular point of time as a single normalised value. Maximising this value will therefore mean maximising learning success. To further simplify the number of hypothesis, the range of Flow score is divided into three sections: (1) Low success where $0 < F_{ij} \leq 0.33$; (2) Medium success where $0.33 < F_{ij} \leq 0.66$; and (3) High success where $0.66 < F_{ij} \leq 1.00$. The traffic-light classification is popular in the field of predictive learning analytics since it is straight-forward to understand (ECAR-ANALYTICS Working Group, 2015).

Analysis and Further Steps

Once the language of hypothesis is defined, the next step consists of defining the language of learning samples, or in other words, devising a representation of the data convenient for the regression task – i.e. predict the correct Flow class. Given that the data present several one-to-n relations, a suitable representation would be a multiple-time-series in which every data point is a five-minute discretised interval. With such representation each observation can be seen as a stochastic process governed by a set of equations, each of them explaining the previous observations and being of order equal to the number of attributes considered. An expected example of the correlations among the observations can be the following: having a lunch break or a walk is likely to influence the productivity of the next hours. Assuming that there is dependence among the observations restricts the range of possible regression models that can be used. With the choice of regression model other issues need to be addressed. When opting for the five-minute interval representation, those attributes having more than one value every five minute (heart rate above all) need to be represented in a way to avoid information loss. Also the dimensionality can constitute a considerable challenge: events that occurs seldomly, like the use of a particular software application, will turn into very sparse signals presenting few spikes but most of the time zeros.

Finally, once a good performing regression model is trained to discriminate learning success, it can be exploited to make predictions almost in real time. This means for example being able to predict whether the next five, ten minutes or even one hour, is likely to be the time for a successful learning experience. This predictive capability will be explored also in further studies, i.e. *Visual Learning Pulse,* using dashboards to display feedback to the learner.

References

Boucsein, W., & Backs, R. W. (2000). Engineering psychophysiology as a discipline: Historical and theoretical aspects. In R.W. Backs & W. Boucsein (Eds.), *Engineering Psychology: Issues and Applications* (pp. 3–30). Mahwah, NJ: Lawrence Erlbaum Associates.

Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, *65*(3), 245–281. http://doi.org/10.3102/00346543065003245

Csikszentmihalyi, M. (1997). *Finding flow: The psychology of engagement with everyday life*. Basic Books.

ECAR-ANALYTICS Working Group (2015). The Predictive Learning Analytics Revolution: Leveraging Learning Data for Student Success. *ECAR Working Group Paper*, Louisville, CO: ECAR.

Pijeira-Díaz, H. J., Drachsler, H., & Kirschner, P. A. (2016). Investigating collaborative learning success with physiological coupling indices based on electrodermal activity. In *Proceedings of the 6th International Conference on Learning Analytics & Knowledge (LAK'16).* ACM: New York, NY.

Pintrich, P.R., & Zusho, A. (2007). Student motivation and self-regulated learning in the college classroom. In *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 731–810). Springer. http://doi.org/10.1007/1-4020-5742-3_16

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, *55*(1), 68–78. http://doi.org/10.1037/0003-066X.55.1.68

Schneider, J., Börner, D., van Rosmalen, P., & Specht, M. (2015). Augmenting the Senses: A Review on Sensor-Based Learning Support. *Sensors*, *15*(2), 4097–4133. http://doi.org/10.3390/s150204097

Siemens, G., & Baker, R. S. J. (2012). Learning Analytics and Educational Data Mining : Towards Communication and Collaboration. In *Proceedings of the 2nd International Conference on Learning Analytics & Knowledge (LAK'12)* (pp. 252–254). ACM: New York, NY.

Suthers, D., & Rosen, D. (2011). A unified framework for multi-level analysis of distributed learning. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge (LAK'11)* (pp. 64–74). ACM: New York, NY

Swan, M. (2012). Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0. *Journal of Sensor and Actuator Networks*, *1*(3), 217–253. http://doi.org/10.3390/jsan1030217

Acknowledgments

This work was partially funded by the Learning Analytics Community Exchange project (grant no. 619424). We hereby thank the eight participants involved in the experiment and the Department of Data Science and Knowledge Engineering of Maastricht University for the support given.