Machine Learning Approach for Student Engagement Automatic Recognition from Facial Expressions

Vladimir Soloviev

Abstract: Digital revolution can significantly improve the quality of education. There have been already discussions for a long time about the advantages, disadvantages and opportunities for transforming traditional classroom activities. Modern students use smartphones and tablets "from birth", and for the majority of academic subject areas students can often obtain more complete, accurate and up-to-date information from the Internet than from lectures. Is it interesting for students to learn? Are they in time with the professor? Is the presentation clear? How deep are students engaged in learning in the classroom? These issues come to the forefront in the era of digital education. However, it was almost unrealistic to control the level of student engagement until recently: for example, only in the Moscow campuses of the Financial University the classes are held daily from 8.30 to 22.00 in more than 500 classrooms.

Existing information systems for student engagement automatic recognition are focused on analyzing individual engagement of students and schoolchildren. We propose a system that constantly analyzes the flow of data from video cameras installed in classrooms, uses machine learning models to identify students' faces, recognize their emotions and determine the level of engagement, and then aggregates engagement data on student groups, faculties, courses, etc. on interactive dashboards.

The training dataset consisted of 2,000 faces was used for machine learning model identification with boosted decision trees algorithm (ADABoost). The quality metrics (Accuracy, Precision, Recall, AUC) on a test dataset of 500 students faces were all above 0,81.

The system is developed as an elastically scalable cloud service that automatically collects video streams from cameras installed in classrooms and forms the resulting metrics of the students and groups' engagement in the Microsoft Azure cloud.

Vladimir Soloviev is with Data Analysis, Decision Making, and Financial Technology Department, Financial University under the Government of the Russian Federation, 38 Shcherbakovskaya St., Moscow 105187, Russia e-mail: VSoloviev@fa.ru

1 Introduction

Modern students use various computational devices "from birth", and on the Internet for the majority of theoretical and practical academic disciplines students can often get more complete, more accurate and more relevant information than from the classes. At the same time, information on the Internet is often delivered more effectively and eye-catching than in the classrooms.

Is it interesting for students to learn? Are students in time with the professor or the pace of presentation is too fast or too slow? Is the presentation clear? How much are students engaged in learning in classroom? These issues come to the forefront in the era of digital education. However, it was almost impossible to control the level of student engagement until recently: for example, only in the Moscow campuses of the Financial University the classes are held daily from 8.30 to 22.00 in more than 500 classrooms.

Methods for measurement and analysis of student engagement in learning have been actively developed since the 1980s, primarily with the aim to find a possibility to decrease the number of expelling students. Surveys conducted at various universities and schools showed, that from 25 to 60% of students are constantly bored in the classroom and distracted from learning (see, for example, [1, 2]).

Management of students' engagement level is relevant nowadays for the traditional classroom teaching, for MOOCs, for educational games, for simulators, for intelligent teaching systems, etc. [3, 4, 5, 6].

The most common methods for student engagement measurement include self-assessment by students themselves; external monitoring using control charts and subsequent rating; automatic measurement using technical means [7]. For example, the most often used method in Russian studies is self-assessment (see, for example, [8]).

Information systems for automatic measurement of student engagement have been used for a long time. A significant part of them is based on analysis of tests execution speed and accuracy [9, 10]. For example, random answers to easy questions or very short lead times could indicate weak engagement.

Another class of popular techniques for automatic measurement of engagement level is based on data processing from various electro- and neurophysiological sensors [11, 12]. These methods could not be implemented at a large-scale, for example, at the level of a whole university, because it is impossible to provide special sensors to every student at the university.

The third class of techniques for automatic recognition of engagement, which includes the system described in this paper, is based on the use of computer vision [13, 14, 15, 16, 17, 18]. Such techniques allow to assess a student's engagement by analyzing the position and inclination of the head, the view direction, pose, different gestures, and so on. The major advantage of such systems is that the engagement level is measured unobtrusively, without diverting students attention to the engagement measurement process itself.

This paper describes the experience of a cloud service development and implementation for monitoring student engagement in the classroom based on intelligent analysis of video streams from cameras placed in the classrooms, and subsequent aggregation of average engagement for groups, courses, areas of training, education, faculty on interactive dashboards.

In this case, the *supposed engagement* is measured, that is the level of student engagement assessed by external experts.

Based on images of students in the classrooms, the system uses machine learning principles to determine whether or not this student is engaged. Initially, a large number of photos of students' faces made by video cameras in classrooms are presented to experts who divide the photos into two classes (engaged and not engaged). Then the classification model is trained on this dataset, labelled by experts, and after training the classification model is used to predict the level of student engagement in the pictures, which neither experts nor the classification model had previously seen.

The engagement recognition service is deployed in the *Microsoft Azure* cloud. The user identification is based on the *Microsoft Azure Active Directory* directory services synchronized with the on-premise university directory services. The identification of students is based on pictures from the campus access control database, and the identification of classes and professors is based on information from the on-premise classes schedule database.

Currently, the system is being piloted in two buildings of the Financial University, with about 60 video cameras in the classrooms connected to it.

In all known systems engagement is measured on the basis of video streams from cameras placed on individual computers. These systems are capable to measure level of engagement of individual students in computer labs or in distance learning systems (including massive open online courses). Unlike that we propose a system automatically measures the engagement level not only for individual students but also for academic groups, faculties, years, and for the whole university.

2   Cloud Solution Architecture

The architecture of the cloud solution is illustrated by Fig. 1. We use video cameras placed under the ceiling of classrooms as the Internet of Things devices connected to the *Microsoft Azure IoT Hub*. Before being sent to *IoT Hub*, video streams are preprocessed locally: individual frames are captured at a specified periodicity.

Data on the classes schedule are taken in reference to the classrooms in which the video cameras are placed from the on-premise classes schedule database. These data for each class include classroom ID, start and end time, set of academic groups' IDs (there usually is one academic group at a seminar or lab, and more then one academic group at a lecture), academic subject area ID, and the professor's ID.
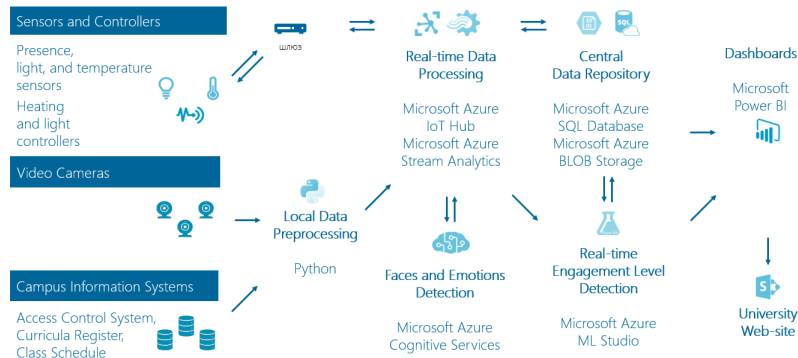
Fig. 1. Engagement Monitoring Service Architecture

When the *IoT Hub* receives an image containing a snapshot of students, it sends it to the *Microsoft Azure Cognitive Services* to recognize students' faces and emotions (the students' pictures in the *Microsoft Azure Cognitive Services* are synchronized with the campus access control database). For each face in the picture the *Microsoft Azure Cognitive Services* return recognized age, gender, Student ID (from the campus information system), head pose, facial landmarks, indicators of lipstick, glasses, mustache, sideburns, beard, recognized emotions (anger, contempt, disgust, fear, happiness, neutral, sadness, surprise), occlusion, etc.

For each face, all the features received from the emotion recognition services, as well as the timestamp of the snapshot, the type of the class (lecture, seminar, computer lab, etc.), academic subject area ID and professor's ID are stored in the *Microsoft Azure SQL Database* while images are stored in the *Microsoft Azure BLOB Storage*.

When a new entry appears in the faces and emotions recognition results table, this entry is automatically submitted to the *Microsoft Azure Machine Learning Studio* web service based on the previously trained classification model. This web service returns the scored probability for classifying the student as engaged, and this probability is stored in the appropriate field in the faces and emotions recognition results table.

We use the *Microsoft Azure Stream Analytics* service for real time event processing.

Another element of the system is the *Microsoft Power BI* service which is used to aggregate the results of engagement level recognition from the *Microsoft Azure SQL Database* to the dashboards published on the university portal.

3   Image Labelling Methodology

We developed a special application for images labelling. This application is published in the *Microsoft Azure* cloud and allows experts to mark each face as engaged on not (Fig. 2).
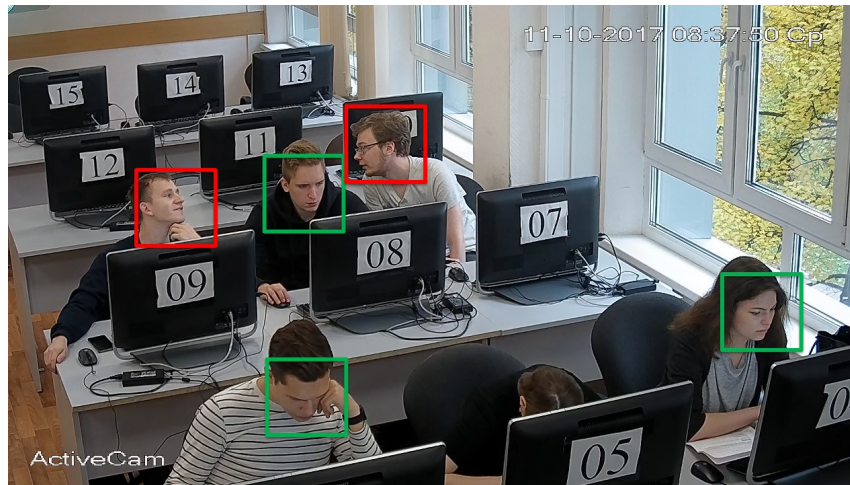
Fig. 2. Engagement Labelling

We asked professors to assess the images of their students. Each professor receives a task to assess engagement level for a certain number of recognized faces. Most of the faces are automatically selected from the images obtained from the cameras during classes taught by this professor (the proportion of such persons is a configurable parameter; we recommend to set its value at the level of 90-95%), and the rest of the faces are taken from the classes of other professors. This is necessary to ensure the adequacy of assessment.

Experts have also the opportunity not to evaluate the engagement for individual picture in cases when the student's engagement evaluation is impossible or unnecessary.

4 Machine Learning Model for Students' Engagement Autodetection

We tried logistic regression, boosted decision trees and random forest models as the predictors of engagement class. The best classification results are obtained using the *ADABoost Two-Class Boosted Decision Tree* model. For this model the following quality metrics were obtained: *Accuracy* = 84.8%, : *Precision* = 0.825, : *Recall* = 81.5%, : *F1 Score* = 82.0%, : *AUC* = 91.2%. These results indicate the legitimacy of using the model to predict engagement level.

Among the factors that have the greatest positive impact on the engagement level, the following features are distinguished (in order of decreasing importance): head pose; recognized age; level of sadness; level of surprise; and also some facial landmarks.

5 Dashboards

The dashboard is placed on a single web page. In the top menu the user chooses the reporting period: semester summary, monthly or daily summary, or individual classes summary.

At any level of the hierarchy the information about engagement is displayed by the universal display unit in the form of a discrete color scale for the selected period.

Initially, the dashboard displays the upper levels available for the particular user (university, faculty, major, or academic subject area).

In the detail mode of the month, the level tree is made in the form of a classic explorer; in front of each level (faculty, major, level of education, year, academic group, student, professor), the information on average engagement for the selected level is displayed by different colors.

A fragment of the dashboard is illustrated by Fig. 3.

◄	Autumn 2017		►
September	October	November	December

Faculty 1

28%	43%	51%	91%

Faculty 2

24%	33%	56%	49%

Faculty 3

45%	39%	41%	53%

Faculty 4

61%	91%	68%	87%

Faculty 5
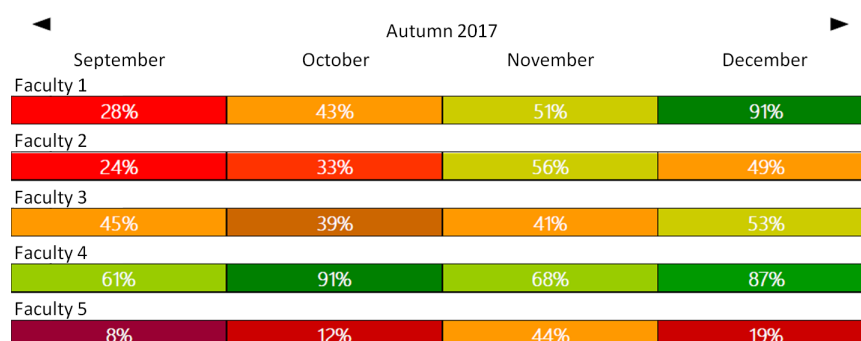
8%	12%	44%	19%

Fig. 3. Engagement Monitoring Service Dashboard

6 Results and Discussion

Most of the tools traditionally used to measure the level of student engagement are too complex to measure the dynamics of the engagement of all the students in the university on an ongoing basis.

Our service developed in the Financial University for monitoring student engagement based on intelligent analysis of video streams from cameras placed in classrooms and subsequent aggregation of averaged data on dashboards is intended for use by the university administration in order to obtain an operational feedback on the dynamics of the average engagement of student groups during the semester, to compare the dynamics of changes in the engagement between faculties, years, groups, etc., and to support the decisions on appropriate corrective actions.

A distinctive feature of the proposed system is that it is built in the form of a cloud service that can be used to monitor the engagement of arbitrarily large groups of students, elastically scaling when the number of students changes. Such a service can be used at the same time by several university or even across the entire education system.

The results of the pilot use of the service demonstrate a sufficient degree of engagement prediction adequacy.

References

[1] Larson, R. and Richards, M. (1991). Boredom in the middle school years: Blaming schools versus blaming students, *American Journal of Education*, 99, 418443.

[2] Dunleavy, J. and Milton, P. (2009). What did you do in school today? Exploring the concept of student engagement and its implications for teaching and learning in Canada, *Canadian Education Association*, 122.

[3] Anderson, J.R. (1982). Acquisition of cognitive skill, *Psychological Review*, 89 (4), 369406.

[4] Mostow, J., Hauptmann, A., Chase, L., and Roth, S. (1993). Towards a reading coach that listens: Automated detection of oral reading errors, In: *Proceedings of the 11th National Conference on Artificial Intelligence (AAAI 1993)*, American Association for Artificial Intelligence, Palo Alto, USA: AAAI Press, 392397.

[5] Koedinger, K.R. and Anderson, J.R. (1997). Intelligent tutoring goes to school in the big city, *International Journal of Artificial Intelligence in Education*, 8, 3043.

[6] VanLehn, K., Lynch, C., Schultz, K., Shapiro, J., Shelby, R. and Taylor, L. (2005). The Andes physics tutoring system: Lessons learned, *International Journal of Artificial Intelligence in Education*, 15 (3), 147204.

[7] Harris, L. (2008). A phenomenographic investigation of professor conceptions of student engagement in learning, *Australian Educational Researcher*, 5 (1), 5779.

[8] Maloshonok, N.G. (2014). Student engagement in learning in Russian Universities, *Higher Education in Russia*, No 1, 3744 (in Russian).

[9] Beck, J. (2005). Engagement tracing: Using response times to model student disengagement, In: *Proceedings of the 2005 Conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology*, Amsterdam, Netherlands: IOS Press, 8895.

[10] Johns, J. and Woolf, B. (2006). A dynamic mixture model to detect student motivation and proficiency, In: *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI 2006), American Association for Artificial Intelligence*, Palo Alto, USA: AAAI Press, 28.

[11] Pope, A., Bogart, E., and Bartolome, D. (1995). Biocybernetic system evaluates indices of operator engagement in automated task, *Biological Psychology*, 40, 187195.

[12] Fairclough, S. and Venables, L. (2006). Prediction of subjective states from psychophysiology: A multivariate approach, *Biological Psychology*, 71, 100110.

[13] Kapoor, A., Picard, R. (2005). Multimodal affect recognition in learning environments, *Proceedings of the 13th Annual ACM International Conference on Multimedia (MULTIMEDIA 2005)*, NY, USA: ACM, 677682.

[14] McDaniel, B., DMello, S., King, B., Chipman, P., Tapp, K. and Graesser, A. (2007). Facial features for affective state detection in learning environments, In: *Proceedings of the 29th Annual Conference of the Cognitive Science Society*, Austin, USA: Cognitive Science Society, 467472.

[15] DMello, S., Craig, S. and Graesser, A. (2009). Multimethod assessment of affective experience and expression during deep learning, *International Journal of Learning Technology*, 4 (3), 165187.

[16] DMello, S. and Graesser, A. (2010). Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features, *User Modeling and User-Adapted Interaction*, 20 (2), 147187.

[17] Grafsgaard, J., Fulton, R., Boyer, K., Wiebe, E., and Lester J. (2012). Multimodal analysis of the implicit affective channel in computer-mediated textual communication, In: *Proceedings of the 14th ACM international conference on Multimodal interaction (ICMI 2012)*, NY, USA: ACM, 145152.

[18] Whitehill, J., Serpell, Z., Lin, Yi-Ch., Foster, A., and Movellan, J.R. (2014). The faces of engagement: Automatic recognition of student engagement from facial expressions, *IEEE Transactions on Affective Computing*, 5 (1), 8698.

[19] How to Detect Faces in Image // *MicrosoftDocs on GitHub*. URL: https:// github.com/MicrosoftDocs/azure-docs/blob/master/ articles/ cognitive-services/Face/Face-API-How-to-Topics/HowtoDetectFacesinImage.md