Impact case study (REF3)

Institution: The Open University
Unit of Assessment: B11 Computer and Informatics
Title of case study: Using predictive learning analytics to improve student retention
Period when the underpinning research was undertaken: 2013-2019
Details of staff conducting the underpinning research from the submitting unit:

Name(s):	Role(s) (e.g. job title):	Period(s) employed by submitting HEI:
Dr Martin Hlosta	Research Fellow	2014 - present
Prof Zdenek Zdrahal	Professor of Knowledge Engineering/ Emeritus Professor	1991 – 2019 & 2019 - present

Period when the claimed impact occurred: 2015-2020
Is this case study continued from a case study submitted in 2014? N

1. Summary of the impact

Professor Zdenek Zdrahal and **Dr Martin Hlosta's** pioneering research in Predictive Learning Analytics has developed new tools which successfully improve higher education student retention, by providing an early warning when students are at risk of failing or abandoning their studies. These tools have reduced from 37% to 19%, the average drop-out rate for first-year students at the Faculty of Mechanical Engineering of the Czech Technical University in Prague, thus saving 422 students from abandoning their studies and preventing the faculty from losing close to GBP1 million in government funding. **Zdrahal** and **Hlosta's** work has also improved professional practice, student outcomes and retention rates at The Open University.

2. Underpinning research

Students dropping out of higher education is a critical problem for the sector, especially for distance education institutions. If universities fail to identify and support those students at risk of abandoning their studies, they risk wasting talent. As tuition fees and funding associated with student numbers are vitally important to universities, reducing drop-out rates is also a financial imperative. Since 2011, Professor Zdenek **Zdrahal** and Dr Martin **Hlosta** have investigated how to use students' data to discover their learning patterns and predict their likelihood of experiencing difficulties. This work on Predictive Learning Analytics offers institutions an early warning system to develop support strategies which deliver timely, effective and meaningful support to students.

In a 2015 article, **Zdrahal** and **Hlosta** explored how machine learning techniques can predict students' success or failure **[O1]**. The researchers collected a sample of Open University students' tutor-marked assignments, weekly interactions through the virtual learning environments and their demographic information. They then used the Minimum Redundancy Maximum Relevance algorithm to select most predictive features and generated multiple models for prediction of students' weekly performance. Finally, **Zdrahal** and **Hlosta** applied the Majority Voting Algorithm to the results of this modelling to create a final forecast of expected student performance in the next assignment. Evaluation of these predictions demonstrated that the approach could identify at-risk students early in the module. It also indicated that the predictive model's performance increases as it collects more data. This approach to predict short-term milestones is distinct from other proposals in the literature. In particular, it allows tutors to monitor student learning continuously and set the most appropriate short-term goals.

Concurrently with this analysis of Open University student performance, the researchers explored how to transfer their proven predictive framework from distance learning to classroom-based university education at Czech Technical University in Prague. **Zdrahal** and **Hlosta** used student data from previous years to build a probabilistic model to predict, for each week, the probability of student success based on the number of credits the student earned to date. Their 2016 formal evaluation of this StudentAnalyse system confirmed its predictive ability **[O2]**. In particular, this research demonstrated the importance of planning interventions before the Christmas break to prevent students from dropping out in the following semesters.

In a 2017 article, **Zdrahal** and **Hlosta** tackled the problem of predicting student performance in new modules, which lack legacy data, by applying their self-learning model to data about the students' first submission for the module in question **[O3]**. In a subsequent publication the same year, the researchers improved this approach by incorporating a problem-based sampling method, which they also generalised to enable the more detailed prediction of students' achievement of specific learning goals within a set deadline **[O4]**.

In the 2015/2016 academic year, **Zdrahal** and **Hlosta** developed a pilot platform based on their research that Open University tutors could use to monitor students' performance. Their formal evaluation of this OUAnalyse dashboard, published in 2019, evidenced that tutors' increased use of the platform resulted in higher student pass rates **[O5]**. These results confirmed the value of the approach and paved the way for deployment at scale within The Open University.

3. References to the research

O1. Kuzilek, J., **Hlosta, M**., Herrmannova, D., **Zdrahal, Z**., Vaclavek J., & Wolff, A. (2015) OU Analyse: analysing at-risk students at The Open University, In Learning Analytics Review, LAK15-1 pp. 1–16. http://oro.open.ac.uk/42529/

O2. **Zdrahal, Z**., **Hlosta, M**., & Kuzilek, J. (2016) Analysing performance of first-year engineering students. In: Learning Analytics and Knowledge: Data literacy for Learning Analytics Workshop, 26th April 2016, Edinburgh. http://oro.open.ac.uk/58597/

O3. **Hlosta, M**., **Zdrahal, Z**., & Zendulka, J. (2017) Ouroboros: Early identification of at-risk students without models based on legacy data. In: LAK17 - Seventh International Learning Analytics & Knowledge Conference, 13-17 Mar 2017, Vancouver, BC, Canada, pp. 6–15. https://doi.org/10.1145/3027385.3027449

O4. **Hlosta, M**., **Zdrahal, Z**., & Zendulka, J. (2018) Are we meeting a deadline? Classification goal achievement in time in the presence of imbalanced data. Knowledge-Based Systems, 160, 278-295. DOI:10.1016/j.knosys.2018.07.021. https://doi.org/10.1016/j.knosys.2018.07.021

O5. Herodotou, C., Rienties, B., Boroowa, A., **Zdrahal, Z**., & **Hlosta, M**. (2019) A large-scale implementation of Predictive Learning Analytics in Higher Education: The teachers' role and perspective, In: Educational Technology Research and Development, 67(5) pp. 1273–1306. https://doi.org/10.1007/s11423-019-09685-0

4. Details of the impact

Zdrahal and **Hlosta** have applied solutions derived from their research **[O1-O5]** on Predictive Learning Analytics in two real-world organisational settings, to improve student retention at the Czech Technical University in Prague and The Open University.

Improving student retention at the Czech Technical University

In the 2015/16 academic year, the Faculty of Mechanical Engineering at the Czech Technical University in Prague (FME-CTU) invited the researchers to deploy the StudentAnalyse system **[O1, O2]** to address its historically high first-year drop-out rate average of 37% **[C1]**. After using the system to identify key student milestones and at-risk students and applying relevant intervention strategies based on the system's data, FME-CTU tutors have reduced this failure rate to an average of 19% between 2015/16 and 2018/19. As a result, 422 additional students have advanced to second and third-year study **[C1]**. In a November 2020 letter, the Vice-Dean for Education at FME-CTU states that the system's success in the early identification of at-risk students had enhanced the "*general wellbeing prospects for affected students*" and the institution **[C1]**. Each year, the Czech Ministry of Education grants funding to the Czech Technical University on the basis of the overall number of students. In his letter, the Vice-Dean explains that the improved student retention delivered by StudentAnalyse for the four academic years, from 2015/16 to 2018/19, has resulted in a total additional financial income of CZK27,900,000 **[C1]**. Converted at the average exchange rate during these years, this corresponds to approximately GBP950,000.

StudentAnalyse also provided additional benefits to FME-CTU during its response to the coronavirus pandemic in 2020. The Vice-Dean noted that the faculty was able to deploy learning strategies proven successful by the system since 2015/16, "*to move to fully online teaching by redesigning the exam strategy for the whole undergraduate programme*" **[C1]**. As a result, they explained, all of the faculty's 2,251 students continued to receive "*all essential educational functions with the maximum possible care for the health and safety of both students and staff*" **[C1]**.

Improving professional practice, student outcomes and retention rates at The Open University

Improving student retention is a critical strategic priority for The Open University (OU). As a distance teaching institution, the university has historically suffered from higher-than-average drop-out rates. In the five academic years to 2015/16, an average of 30% of students abandoned their studies in the first year, 20% more than the average among UK higher education institutions **[C2]**. Since 2015, **Zdrahal** and **Hlosta** have piloted predictive learning analytics solutions to this problem, based on their research **[O1-O4]**. The product of these pilots, the OUAnalyse platform **[O5]**, has had demonstrable success in three critical ways:

1. OUAnalyse usage by tutors is now one of the two significant predictors of whether students will complete and pass a course, alongside previous best score.
2. Teachers who regularly access the student performance platform have better student retention rates. Those who use it at least 41% of the time have average student retention of 56%, compared to 48% for those who do either limited or no use of the platform.
3. Teachers had better student outcomes in the years they used OUAnalyse than in previous years **[C2]**.

In a November 2020 letter, The Open University's Director for Students explained how, in April 2019, this evidence "*led the university internal review panel to recommend using Predictive Learning Analytics in all faculties*" **[C2]**. They also described how, as a result of this decision, from November 2020 "*all of the OU's 3,581 tutors in the undergraduate modules have been given training materials on how to use OUAnalyse directly in their primary teaching portal - TutorHome. As such, Learning Analytics and monitoring students' progress using OUAnalyse has become part of teaching for 118,963 students*" **[C2]**.

Interviews with tutors also reveal that using OUAnalyse complements existing teaching practices. In feedback collected by email in December 2019, Open University tutors wrote that the system is "*the best method*" for monitoring student performance they have worked with and provides an efficient way "*to work out if there may be potential problems*" **[C3**, p.2**]**. Others explained that OUAnalyse "*pushed*" them to monitor and contact struggling students more quickly than they usually would. One tutor, who explained they had a particularly challenging group, said that without OUAnalyse, they "*would have lost around four or five of them, but all of them made it 'til the end and passed*" **[C3**, p.2 and p.3**]**.

In its 2016 report, From Bricks to Clicks, which includes an entire case study about the use of OUAnalyse at The Open University, the UK House of Lords' Higher Education Commission pointed out that "*apart from the OU the Commission does not believe that any UK institution has made significant headway in this area*" **[C4**, p.53**]**. Finally, in September 2020, the DataIQ Awards recognised OUAnalyse by awarding the prize for 'Best use of data by a not-for-profit organisation' **[C5]**.

5. Sources to corroborate the impact

C1. Letter from the Dean of the Faculty of Mechanical Engineering, Czech Technical University in Prague, evidencing both i) the increase in student retention resulting from the well-targeted interventions driven by the predictions calculated by StudentAnalyse, and ii) the positive role the system played during the coronavirus pandemic. 20th November 2020.

C2. Comprehensive testimonial from the OU's Director for Students and Chair of the Digital Strategy Steering Group, explaining the key role played by OUAnalyse in improving student outcomes and retention rates at The Open University. 30th November 2020.

C3. Tutor feedback report evidencing the impact of OUAnalyse on their professional practice. 15th December 2020.

C4. Report from The Higher Education Commission, From Bricks to Clicks. 26th January 2016. Link: https://www.policyconnect.org.uk/hec/research/report-bricks-clicks-potential-data-and-analytics-higher-education

C5. DataIQ 2020 Awards - Best use of data by a not-for-profit organisation: Open University OUAnalyse. 30th September 2020. Link: https://www.dataiq.co.uk/articles/articles/dataiq-awards-2020---best-use-of-data-by-a-not-for-profit-organisation-open-university-ouanalyse