Open University: predicting students performance

Mariia Markovska, Rahul Puniani, Tedo Gogoladze

Background 🎓

The Open University is a public non-profit higher education institution based in the UK. The Open University delivers knowledge exclusively via distance learning in over 100 countries worldwide.

Demographics and performance

- **Young** (<35)
- **Adults** (35–55)
- **Seniors** (55+)

Majority of the Open University students are young people with education of A level (43%) or lower (40%).

A major issue is prevailing of Withdraws and Fails over Passes and Distinctions

Interaction with VLE*

The most engaging pages are quizzes, homepage, main content, forums and glossary. Majority of activities are not engaging (a user clicks certain page on average less than 2 times).

Average number of clicks (per course page per capita) is the highest for Distinction and Pass. Number of unengaging pages is high (a user clicks certain page on average only 3 times).

We are predicting:
1) Passing vs Not Passing binary class label
2) Completing vs Withdrawing of the course binary class label

For both of the classification prediction we used Random Forest model.

Variables included: course module and semester, demographics of a student, number of previous attempts of the course, studied credits, learning environment interaction score, average submission delay of a student, registration date of a student on the course.

Conclusions and recommendations

Main issue of the Open University are: prevalence of negative final results over positive, poor interaction with VLE, delays in submissions.

Our recommendation is to use prediction model to identify those students who are likely to fail or withdraw and provide them with customized and more supportive interface based on their learning needs.

Data

Data consists of 7 interconnected tables which contain variables such as: student demographic information, students scores, students interaction with VLE (sum of the clicks, registration/withdrawal dates etc.).

* VLE – virtual learning environment

---

<table>
<thead>
<tr>
<th>Performance of Open Education Level</th>
<th>Students (Students)</th>
</tr>
</thead>
</table>
| Distinction                        | 100%
| Pass                               | 81%
| Fail                               | 19%

<table>
<thead>
<tr>
<th>Performance of Open Education Level</th>
<th>Students (Students)</th>
</tr>
</thead>
</table>
| Complete                           | 79%
| Withdraw                           | 21%