KU LEUVEN

NATIONAL RESEARCH UNIVERSITY

Measuring students' activity in MOOCs using extensions of the Rasch model

Dmitry Abbakumov [1, 2], Piet Desmet [1], & Wim Van den Noortgate [1]

[1] KU Leuven (University of Leuven), Leuven, Belgium

[2] National Research University Higher School of Economics, Moscow, Russia

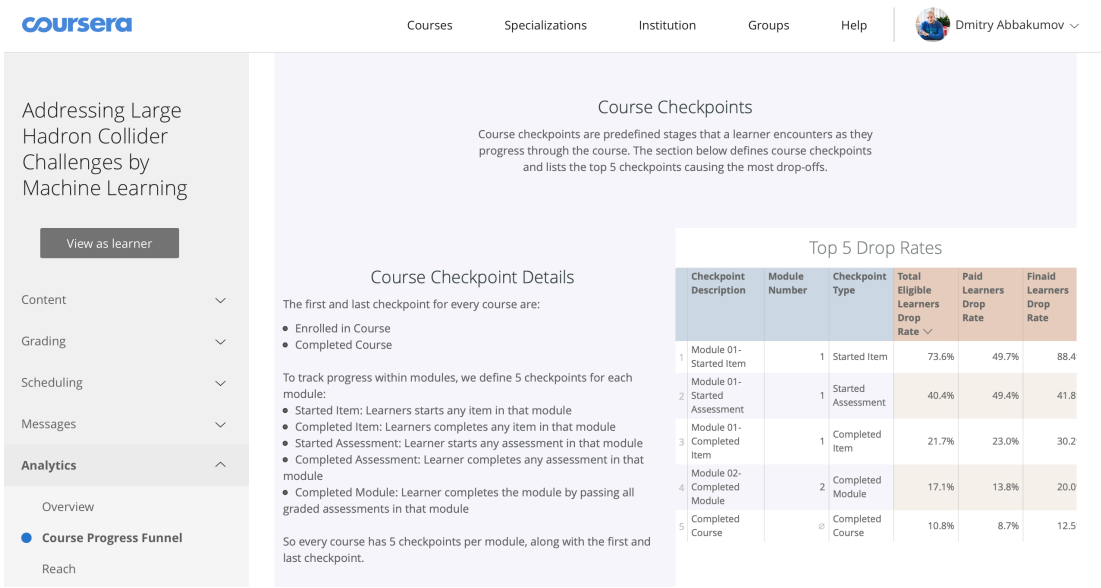
A Massive Open Online Course is a large scale web-based course developed by a university, solely or in cooperation with industrial partners, in which anyone with internet access can participate.

A MOOC consists of video lectures, reading assignments, assessments, and forums.

Now 800 universities offer more than 9,400 MOOCs.

Coursera: 40 million students and 3,000 courses.

Students' activity in MOOCs is typically described through the proportions (videos viewed, assessments taken). These measures are simple and intuitive. However, by aggregating the information per student, *we lose information on how a student interacts with individual units of content*.



Framework

The Rasch model $Logit(\pi_{ij}|\theta_j) = \ln(\pi_{ij}/1-\pi_{ij}) = \boldsymbol{\theta_j} - \boldsymbol{\delta_i}$ *and* $Y_{ij} \sim Bernoulli(\pi_{ij})$

Y_{ij} is the interaction of student *j* with a unit of content *i*, where 1 means that unit *i* was completed by student *j*, and 0 means that unit *i* was not completed by student *j* during the MOOC

$\boldsymbol{\theta_j}$ is student's activeness

$\boldsymbol{\delta_i}$ is the effect of a unit of content on the probability of interaction

The reformulation presented by Van den Noortgate, De Boeck, & Meulders (2003)

$Logit(\pi_{ij}) = \boldsymbol{b_0} + \boldsymbol{u_{1j}} + \boldsymbol{u_{2i}}$ *and* $Y_{ij} \sim Bernoulli(\pi_{ij})$

where $u_{1j} \sim N(0, \sigma^2_{u1})$ and $u_{2i} \sim N(0, \sigma^2_{u2})$

Very flexible for making extensions.

The dimensionality check did not reveal substantial evidence against unidimensionality (p = .37).

Extension

Extension $Logit(\pi_{ij}) = b_0 + b_{10} * \boldsymbol{type_i} + (b_{20} + b_{2j}) * \boldsymbol{week_{ij}} + u_{1j} + u_{2i}$

$\boldsymbol{type_i}$ is a dummy variable, which distinguishes between two major types of content in a MOOC (where 0 stands for video lecture and 1 for reading assignment)

$\boldsymbol{week_{ij}}$ is 0, 1, 2, …, n-1 and means the first, the second, the third, …, or n-th week of a MOOC at which student *j* interacts with unit of content *i* respectively

b_{2j} and u_{1j} are assumed to follow the bivariate normal distribution $N(\mathbf{0}, \boldsymbol{\Sigma_j})$

Thus, $\theta_j = u_{1j} + (b_{20} + b_{2j}) * week_{ij}$

$\delta_i = (b_0 + b_{10} * type_i + u_{2i})$

Illustration: Data

MOOC: Introduction to Neuroeconomics: How the Brain Makes Decisions
Language: English
Length: 11 weeks
N = 11,826
Female (48%)
Africa (3.4%), Asia (21.5%), Europe (35.5%), N.America (28.2%), Oceania (2.5%), and S.America (8.9%)
Age: 18-24 (18.2%), 25-34 (49.5%), 35-44 (18.7%), 45-54 (7.7%), other age groups (5.9%)
We used the data from five weeks (modules) from the course
27 videos and 10 reading assignments
437,568 observations

Illustration: Results

			Rasch Model	Extension		Extension with Type*Week Interaction	
Fixed	Intercept		-1.55 (0.16) ***	1.67 (0.15) ***		1.41 (0.19) ***	
	Type			-0.90 (0.17) ***		-0.30 (0.28)	
	Week			-4.32 (0.08) ***		-4.18 (0.10) ***	
	Type*Week					-0.48 (0.14) ***	
Random	Intercept	Student	3.43	3.11	*Corr.*	3.11	*Corr.*
	Week	Student		3.51	*.27*	3.52	*.27*
	Intercept	Content	2.29	0.54		0.52	
AIC			245,952	206,819		206,811	

Illustration: Results

			Rasch Model	Extension		Extension with Type*Week Interaction	
Fixed	Intercept		-1.55 (0.16) ***	1.67 (0.15) ***		1.41 (0.19) ***	
	Type			-0.90 (0.17) ***		-0.30 (0.28)	
	Week			-4.32 (0.08) ***		-4.18 (0.10) ***	
	Type*Week					-0.48 (0.14) ***	
Random	Intercept	Student	3.43	3.11	*Corr.*	3.11	*Corr.*
	Week	Student		3.51	*.27*	3.52	*.27*
	Intercept	Content	2.29	0.54		0.52	
AIC			245,952	206,819		206,811	

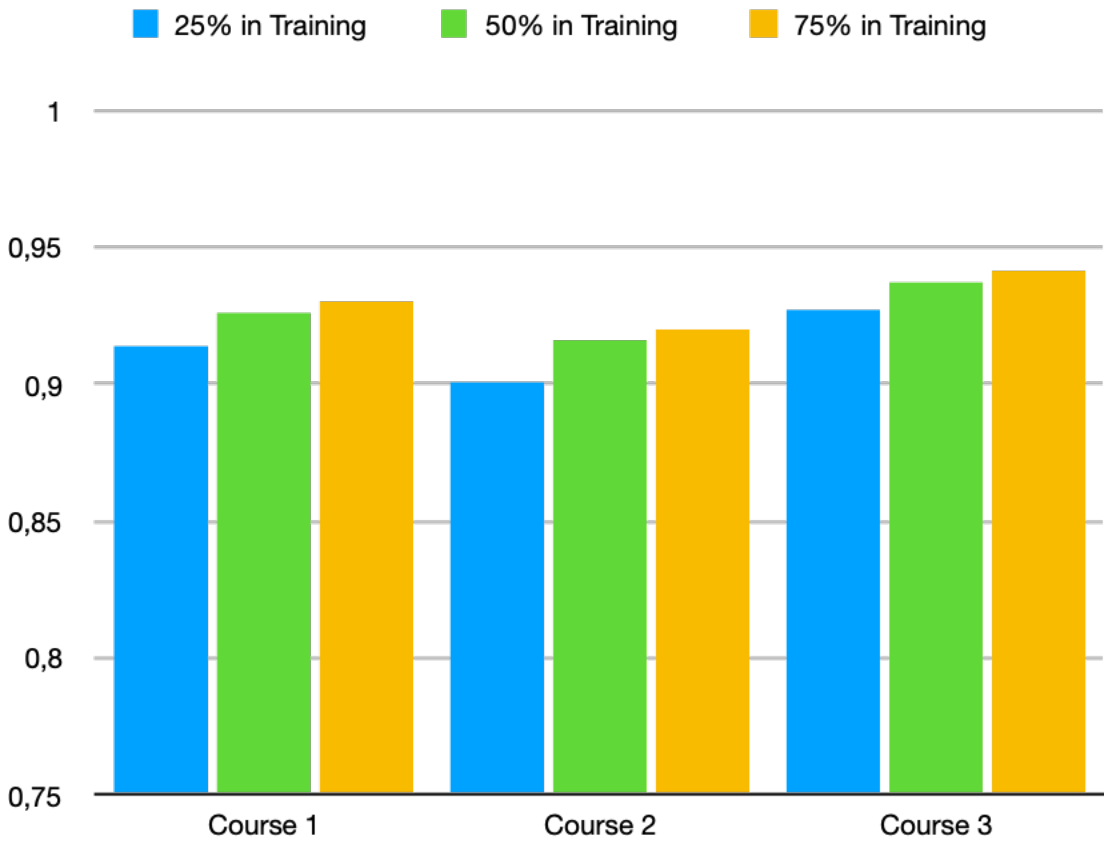
Illustration: Results

			Rasch Model	Extension		Extension with Type*Week Interaction	
Fixed	Intercept		-1.55 (0.16) ***	1.67 (0.15) ***		1.41 (0.19) ***	
	Type			-0.90 (0.17) ***		-0.30 (0.28)	
	Week			-4.32 (0.08) ***		-4.18 (0.10) ***	
	Type*Week					-0.48 (0.14) ***	
Random	Intercept	Student	3.43	3.11	*Corr.*	3.11	*Corr.*
	Week	Student		3.51	*.27*	3.52	*.27*
	Intercept	Content	2.29	0.54		0.52	
AIC			245,952	206,819		206,811	

Illustration: Results

			Rasch Model	Extension		Extension with Type*Week Interaction	
Fixed	Intercept		-1.55 (0.16) ***	1.67 (0.15) ***		1.41 (0.19) ***	
	Type			-0.90 (0.17) ***			
	Week			-4.32 (0.08) ***		-4.18 (0.10) ***	
	Type*Week					-0.48 (0.14) ***	
Random	Intercept	Student	3.43	3.11	*Corr.*	3.11	*Corr.*
	Week	Student		3.51	*.27*		*.27*
	Intercept	Content	2.29	0.54		0.52	
AIC			245,952	206,819		206,811	

Cross-Validation



Key result:

- better insight on students' activity in a MOOC

Limitation:

- binary dependent variable

Practice:

- application in a recommender tool

Thank you!