Detecting Cheaters in MOOCs
Using
IRT & Learning Analytics

Giora Alexandron  Sunbok Lee  Zhongzhou Chen  David E. Pritchard
Massachusetts Institute of Technology

PALE’16, Halifax
The problem: Detecting cheaters

• Cheating in MOOCs:
  • **Copying Using Multiple Accounts**:  
    • 10% of certificatees in 8.MReVx (Alexandron et al. 2015, Ruipérez-Valiente et al. 2016)  
    • 1.3% of certificates in 69 MITx courses (Northcutt et al. 2016)

• Limitations of the two methods:
  • Rely on IP* (both teams work to remove this limitation)  
  • Detect a specific cheating method

• Goal: A general algorithm that does not assume a specific method

• Process:
  1. Tag known cheaters with previous algorithm (move from unsupervised to supervised)  
  2. Features engineering: EDM + Item Response Theory (IRT) to build behavioral profiles  
  3. Feature Selection  
  4. Classification (generalizing from the known cheaters to the unknown ones)
Classification model

• Data:
  • The 2014 run of *8.MReVx Introductory Physics MOOC*
  • Known cheaters: 10% of certificate earners on > 1% of correct answers

• Feature engineering & selection:
  • Fraction of videos watched (rationale: cheaters spend less time on resources)
  • Fraction of correct answers in < 30s (based on *Palazzo et al., 2010*)
  • Mean time for submitting a correct answer
  • Student’s IRT ability parameter (cheating worth .5 σ!)
  • Standard error for student’s IRT ability estimate:
  • Guttman error* (*Meijer, 1994*)

* The number of item pairs in which an easier item is answered incorrectly and a more difficult item is answered correctly.
Results

Distribution (certificate earners, N=502)

Cheaters/non-cheaters

Performance of the Classifier

AUC = 0.826 (3-fold cv X 500 runs)
Lessons learned and next steps

• ‘Multiple account’ cheaters can be classified with good accuracy using ‘post-factum’ behavioral/IRT analytics

• Generalization:
  • False positives – students using other forms of cheating?
    • ~30% are suspected to be students using team-work
    • Other 70% -- misclassification? Other forms of cheating? => FUTURE WORK!
Acknowledgments
RELATE Group, MIT

giora@mit.edu