

# Recommendation of New Questions in Online Student Communities

**Jakub Mačina**  
Supervisor: Ivan Srba

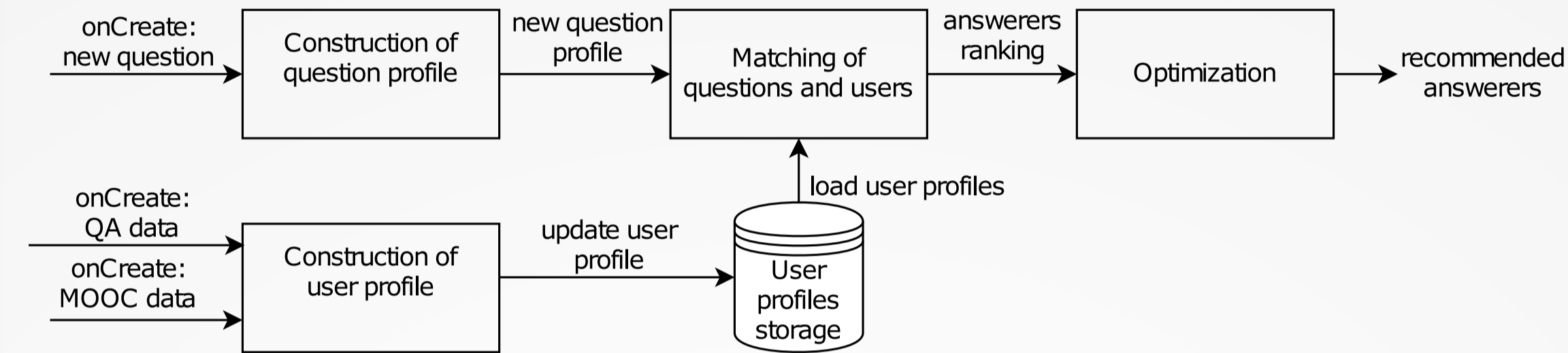
## Motivation

- Students' performance in Massive Open Online Courses (MOOCs) is enhanced by high quality discussion forums or recently emerging educational Community Question Answering (CQA) systems.
- Problems of discussion tools in MOOCs:
  - A small number of students answer questions asked by their peers.
  - An increasing proportion of unanswered questions (up to 50%).
  - Dropout rates for courses can be as high as 94%.

## Contributions

- While existing methods primarily focus on askers' needs, we take an answerer-oriented approach by considering not only students' expertise, but also willingness to answer a question.
- Incorporating additional non-QA data from the course to involve more students in question answering and reduce the burden on individual users.
- Online experiment as an A/B test of the proposed method with more than 4600 MOOC students.

## Educational question routing framework

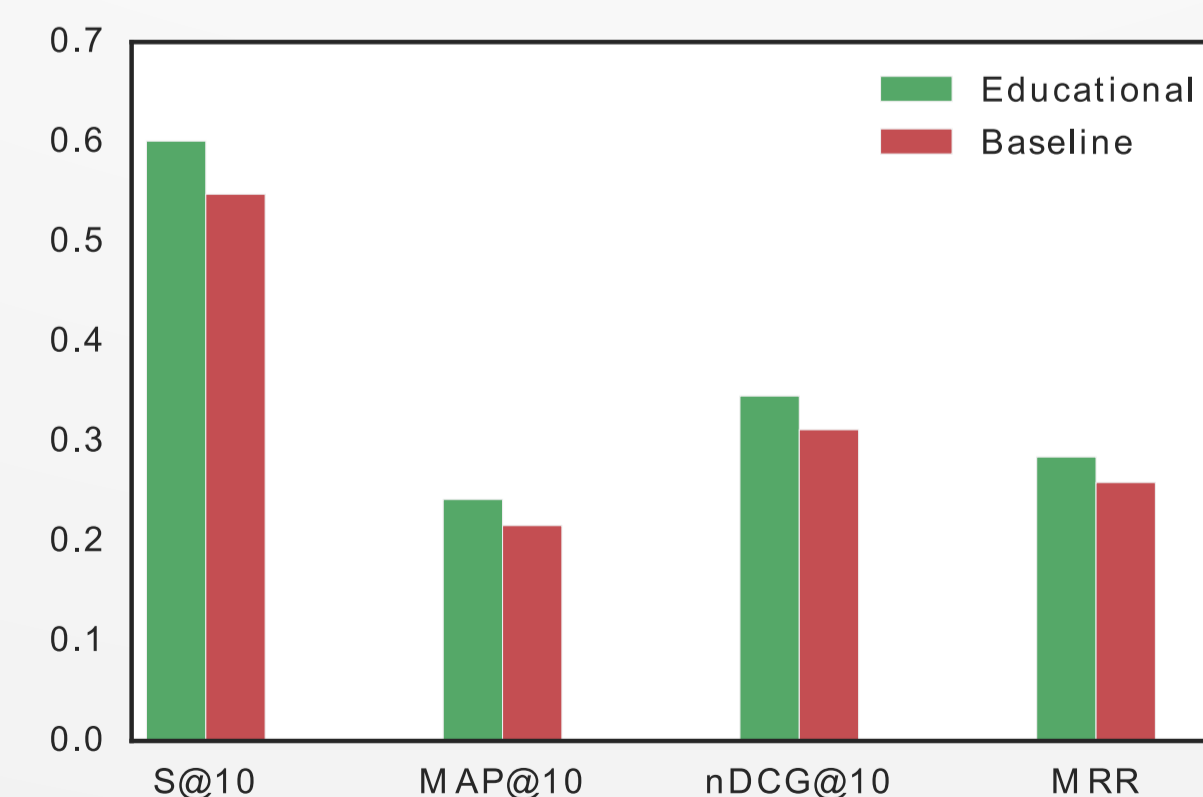


## Matching of questions and users

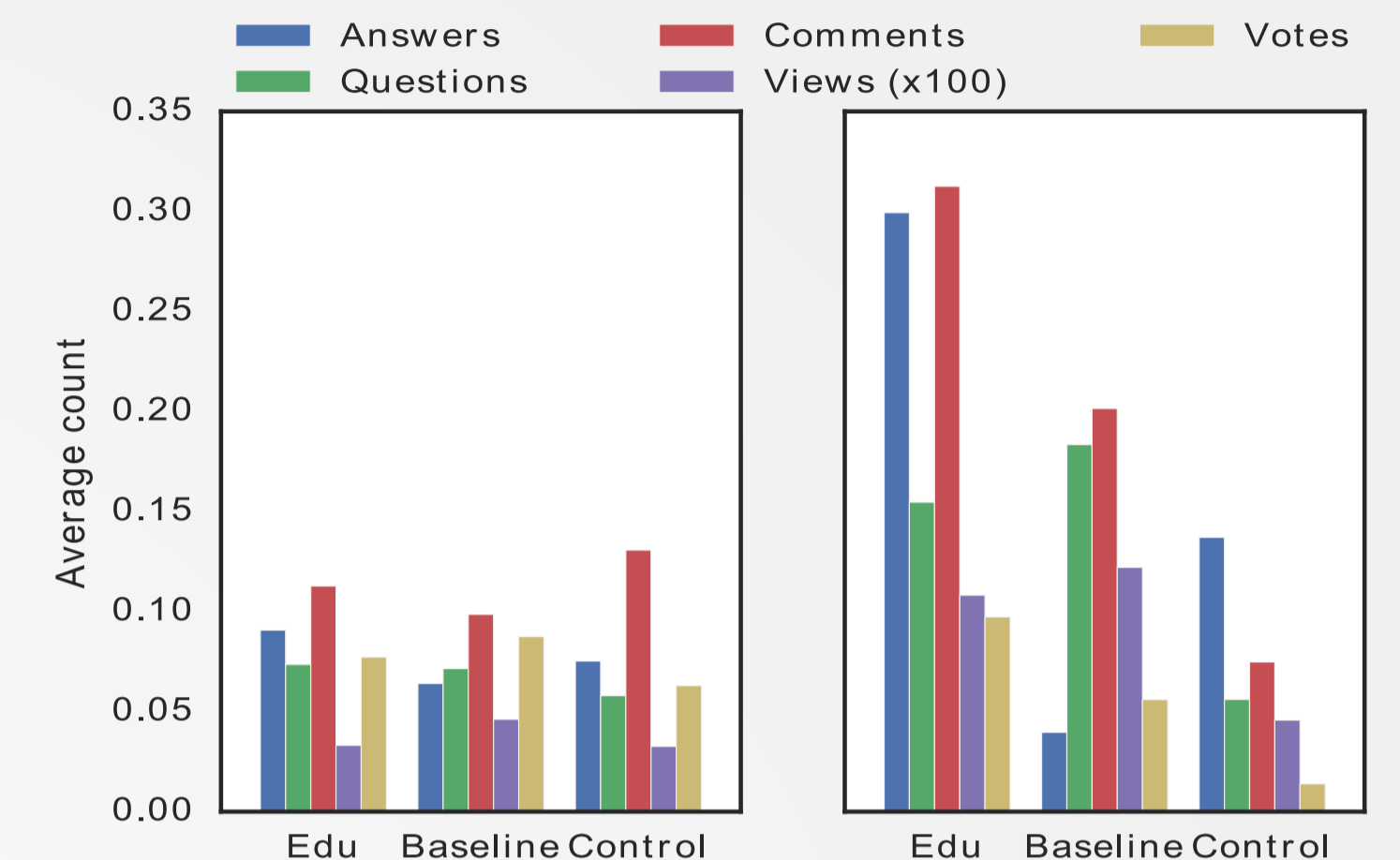
- Ensemble classifier
  - Predicting whether a user has sufficient expertise to answer a new question.
  - Predicting user willingness to answer a new question.
- Combination:

$$P(y = 1) = P(\text{expertise} = 1) * P(\text{willingness} = 1)$$

## Offline experiment



## Online experiment



Metric	Quantity
Students started the course	4618
Users participating in CQA (with any question view)	1098 (24%)
Users contributing in CQA	377 (8%)
Questions	361
Questions with answer	299 (83%)
Answers	386
Comments	476
Teachers evaluations of answers	24

## Accuracy of question recommendation

	Educational	Baseline	Statistical significance
CTR	23.04%	19.05%	$\chi^2(1, N=1320) = 6.65, p < 0.01$
Success rate	17.42%	11.36%	$\chi^2(1, N=132) = 1.97, p = 0.16$

## Number of users contributing to CQA system

	Educational	Baseline	Control
Before experiment	60 (4.60%)	70 (5.52%)	76 (5.81%)
During experiment	39 (12.91%)	25 (9.03%)	29 (8.95%)

## Conclusion

- Higher accuracy of answerers prediction
- Higher interest of users in the routed questions and engaged more students, who in addition provided more contributions
- Lower dropout rate of active contributors in CQA and the lower instructors workload