# Learning Analytics Challenges: Trade-offs, Methodology, Scalability

Radek Pelánek


LAK 2020

# Context

- LAK 2020 theme: Shaping the future of the field
- LAK 2019 keynote – Ryan Baker's challenges
  - six challenges, focus on clear specification/goals

my argument: we should focus on hard-to-grasp, ill-defined problems

# Netflix Prize

- closely related are: recommender systems
- Netflix Prize: 1 million dollars, well-structured task
- impulse for research, lot of attention
- limited practical impact

# Challenges

- trade-offs
- methodology
- scalability

# Trade-offs

- mastery learning thresholds: over-practice vs under-practice
- engagement vs learning
- hints: support for learning vs risk of gaming
- interests of students vs researchers
- model accuracy vs implementation simplicity

# Trade-offs

- hard to perform research studies – evaluation is difficult
- but practically very important

research directions:

- visualization of trade-offs
- optimization with multiple criteria

# Methodology

Baker's challenges and typical current research:

- briefly described data
- results for a specific performance metric (e.g., AUC)

methodological details matter:

- biases in data
- choice of metric (AUC / RMSE / MAE / ...)
- details of metric computation (averaging)
- train-test set data division

Note: Deep knowledge tracing paper

# Methodology

challenges:

- clarification of methodological issues
- replication, reproducation
- “what works when”

# Scalability

- computational scalability: using techniques on real life traffic / data
- development scalability: developing systems under real life constraints

# My Setting

- `umimeto.org`
- adaptive practice for Czech students (K–12)
- mathematics, Czech, English, programming, ...
- 2 computer scientists + 6 content creators (few hours a week)
- $\sim$ 10 000 students daily

# Development Scalability

- developing and managing content (tens of thousands of items)
- "debugging perspective" – identifying most important bugs
- student models: taking into account implementation simplicity, number of parameters

very practical but still can be research-based

# Conclusions

- well-structured challenges (Netflix Prixe, Baker's challenges)
  - clear focus on a specific problem
  - short-term progress
- ill-structured challenges
  - unclear progress
  - long-term progress