

Exploration of the Robustness and Generalizability of the Additive Factors Model

Tomáš Effenberger, **Radek Pelánek**, Jaroslav Čechák



LAK 2020

Student modeling

TODO fig umuai paper

Additive Factors Model

- family of “logistic models”
- Q-matrix
- used in many studies in last 10 years – see paper for overview

Q-matrix

	−	+	×	()
$10 + 3 \times 2$	0	1	1	0
$(7 - 4) \times 3$	1	0	1	1
$2 + (3 + 5)$	0	1	0	1
$8 - (6 + 2)$	1	1	0	1
$5 - 2 \times 6$	1	0	1	0

Additive Factors Model

$$P(Y_{ij}|\alpha, \beta, \gamma) = \sigma \left(\alpha_i + \sum_{k=1}^K \beta_k q_{jk} + \sum_{k=1}^K \gamma_k q_{jk} t_{ik} \right)$$

- i is student index, j is item index,
- Y_{ij} is the binary response of a student i on a item j ,
- $\sigma(x) = 1/(1 + e^{-x})$ is the standard logistic function,
- K is the number of skills, J is the number of items,
- Q is the $J \times K$ binary matrix – q_{jk} is the indicator that item j uses skill k ,
- α_i is the proficiency (prior skill) of a student i ,
- β_k is the easiness of skill k ,
- γ_k is the learning rate for skill k ,
- t_{ik} is the number of times student i has practiced skill k (opportunity count).

Additive Factors Model

TODO illustration log. function...

AFM: Simplifying Assumptions

- learning is linear (on the logit scale)
- effect of practice not related to observed performance
- observed outcomes are binary (ignoring response time, common wrong answers)
- Q-matrix is binary
- compensatory model of skills
- ignores difficulty of items
- ignores biases in data

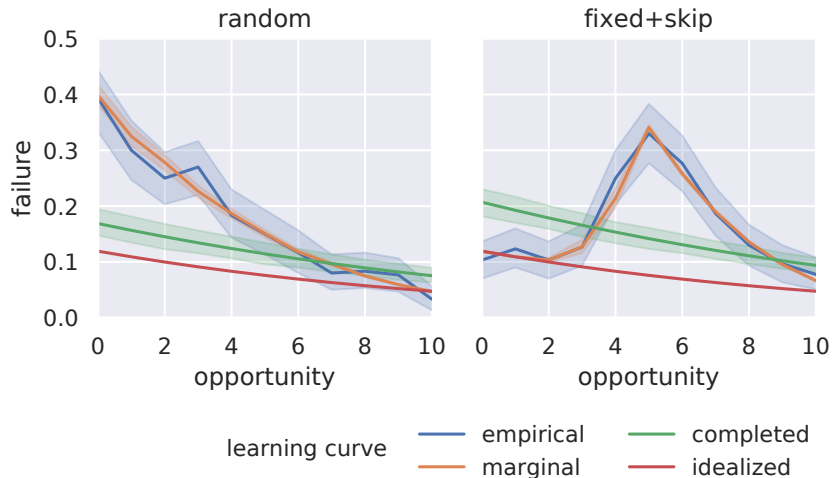
Learning Curves

TODO illustration simple

Types of Learning Curves

Type	Attempt	Opportunity	Success
<i>empirical</i>	observed	observed	observed
<i>marginal</i>	observed	observed	predicted
<i>completed</i>	observed	simulated	predicted
<i>idealized</i>	simulated	simulated	predicted

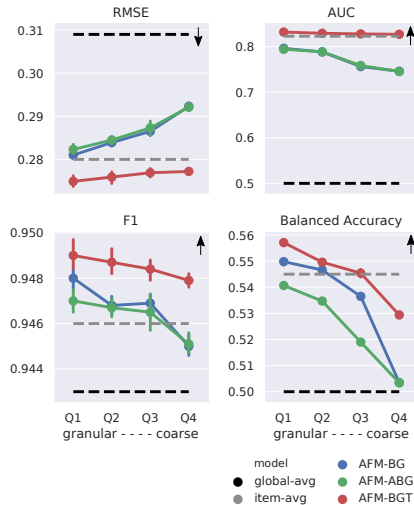
Learning Curves



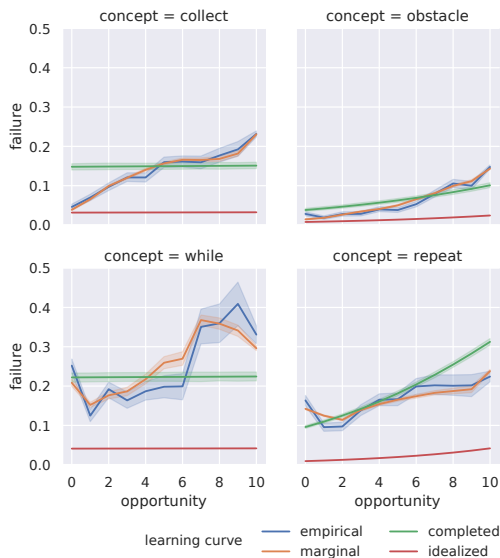
Case Study: Programming

- RoboMission...

Results: Model Comparison



Results: Learning Curves



Conclusions

- studies using AFM: more caution necessary
- AFM has many simplifying assumptions, not satisfied in practice
- possibly misleading conclusions
- basic precaution: comparison with “item average” predictor