

Lab session 6

Machine Learning for Behavioral Data (CS-421)

March 31, 2021

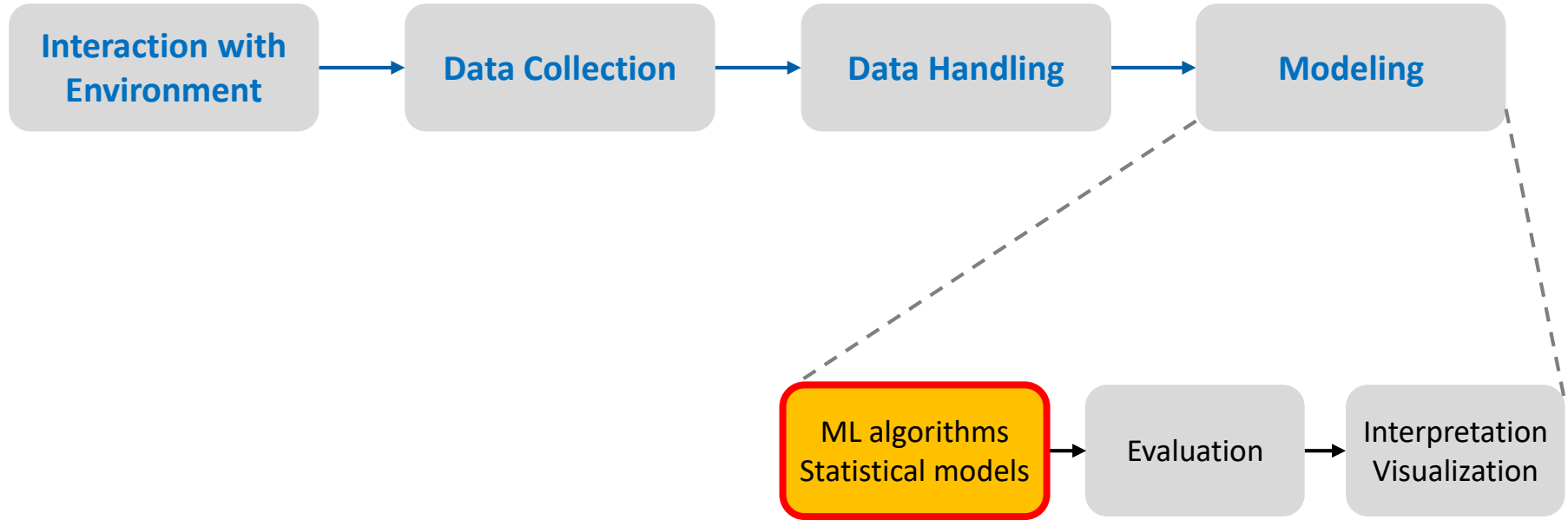
Today

- **08:15-08:50** Tutorial on Additive Factors Model
 - **08:50-09:00** SHORT BREAK
 - **09:00-09:30** Tutorial on Performance Factor Analysis
 - **09:30-10:00** Homeworks and Project
 - Next steps for the project and questions time
 - Solution HW3, introduction HW5 and questions time
-

Where we are

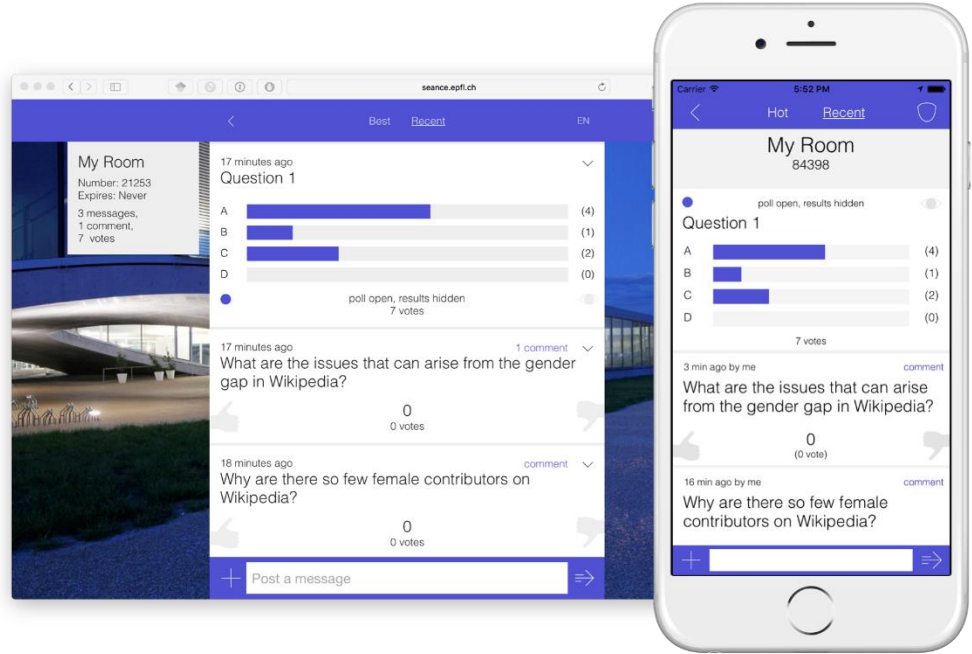
Week	Lecture	Lab Sessions	Project
1	Introduction	Tutorial	
2	Data Handling	Tutorial + Homework	
3	Regression & Classification	Tutorial + Homework	
4	Model Selection & Evaluation	Tutorial + Homework	Presentation of data sets and research questions
5	Latent Variable Models	Tutorial + Homework	M1: Preferences on team members and data sets
6	LatVarMod +Unsupervised Learning	Tutorial + Homework + PO	
7	Spring Break	Spring Break	Spring Break

ML for Behavioral Data: Modeling



SpeakUp

- **Android / iOS:**
<http://speakup.info/>
- **Web App:**
<https://web.speakup.info/>
- **Room number:**



AFM

SpeakUp: How do you feel about Additive Factors Models (AFM)?

A: I have never heard of AFM.

B: I am **not confident at all** about using AFM.

C: I am **slightly confident** about using AFM.

D: I am **fairly confident** about using AFM.

E: I am **very confident** about using AFM.

PFA

SpeakUp: How do you feel about Performance Factors Analysis (PFA)?

A: I have never heard of PFA.

B: I am **not confident at all** about using PFA.

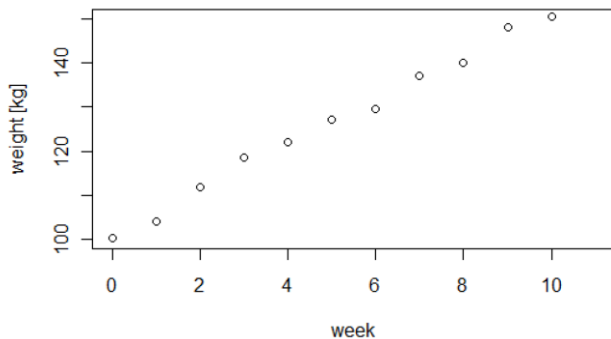
C: I am **slightly confident** about using PFA.

D: I am **fairly confident** about using PFA.

E: I am **very confident** about using PFA.

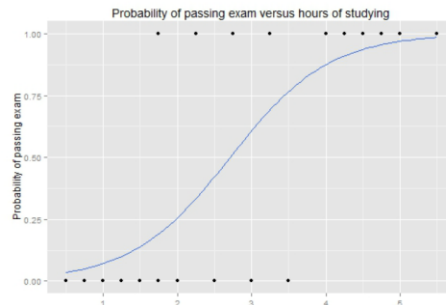
Generalized Linear Models revisited

- Example 1: strength gain by weight training



$$y_n = \beta_0 + \beta_1 x_{n,1}$$

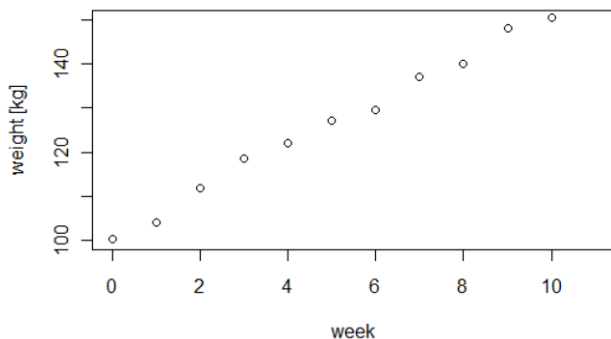
- Example 2: probability of passing exam depending on hours studied



$$\log\left(\frac{y_n}{1 - y_n}\right) = \beta_0 + \beta_1 x_{n,1}$$

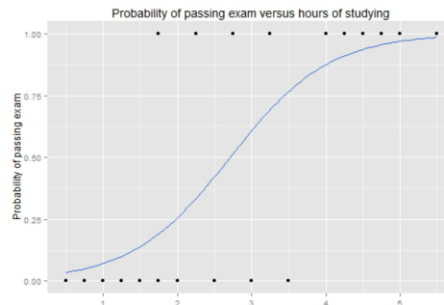
Generalized Linear Models revisited

- Example 1: strength gain by weight training



$$y_n = \beta_0 + \beta_1 x_{n,1}$$

- Example 2: probability of passing exam depending on hours studied

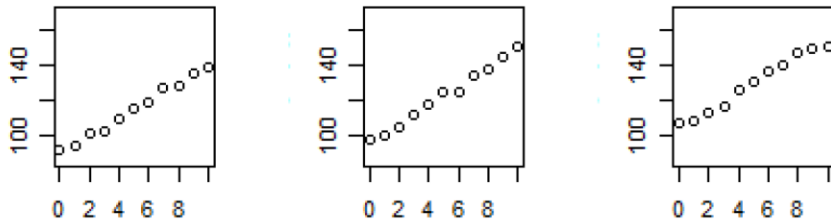


$$\log\left(\frac{y_n}{1 - y_n}\right) = \beta_0 + \beta_1 x_{n,1}$$

“Fixed” Effects

Generalized Linear Mixed Effects Model

- Example 1: strength gain by weight training
 - Each person has individual starting strength



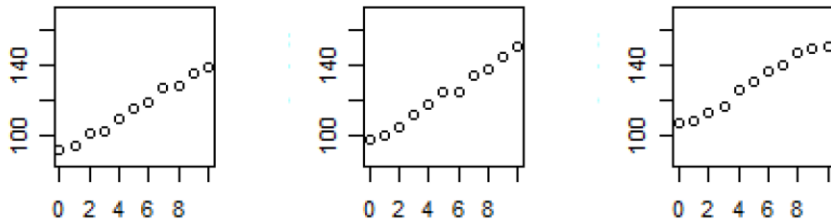
$$y_n = \beta_0 + u_n + \beta_1 x_{n,1} \quad u_n \sim N(0, \sigma_u^2)$$

“Fixed” Effects

“Random” Effect

Generalized Linear Mixed Effects Model

- Example 1: strength gain by weight training
 - Each person has individual starting strength



$$y_n = \beta_0 + u_n + \beta_1 x_{n,1} \quad u_n \sim N(0, \sigma_u^2)$$

“Fixed” Effects

+

“Random” Effect

=

“Mixed” Effects

Fitting the parameters:

- Fixed effects only: linear least squares
- Mixed effects: maximum likelihood estimation

Item Response Theory (IRT)

- **Goal:** explain the relationship between **latent traits** (unobservable characteristic or attribute) and their **manifestations** (i.e. observed outcomes, responses or performance)
-

Item Response Theory (IRT)

- **Goal:** explain the relationship between **latent traits** (unobservable characteristic or attribute) and their **manifestations** (i.e. observed outcomes, responses or performance)



**Binary answer to
an item (task)**

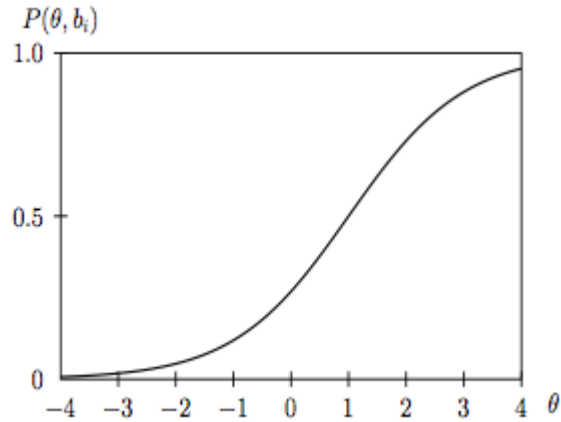


Ability

Rasch Model

$$\log\left(\frac{p_{i,n}}{1 - p_{i,n}}\right) = \theta_n - b_i$$

Probability that student n will solve item i correctly.



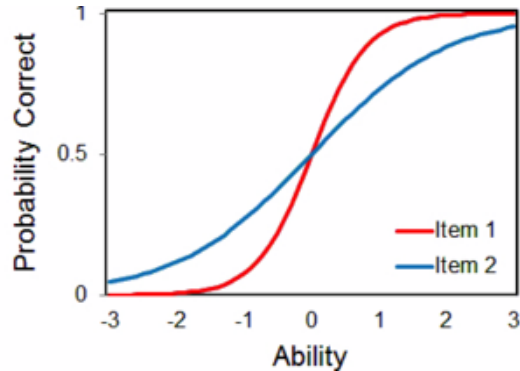
$$\theta_n \sim N(0, \sigma_\theta^2)$$

θ_n : student ability
 b_i : difficulty of item i

Rasch Model

$$\log\left(\frac{p_{i,n}}{1 - p_{i,n}}\right) = \theta_n - b_i$$

Probability that student n will solve item i correctly.



θ_n : student ability
 b_i : difficulty of item i

Additive Factors Model (AFM)

$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

Probability that student n will solve task i correctly.

Additive Factors Model (AFM)

$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$

Additive Factors Model (AFM)

$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$

Student proficiency:

$$\theta_n \sim N(0, \sigma_\theta^2)$$

Additive Factors Model (AFM)

$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$

Student proficiency:

$$\theta_n \sim N(0, \sigma_\theta^2)$$

$q_{ik} = 1$, if item i uses skill k

Additive Factors Model (AFM)

$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$

Student proficiency:
 $\theta_n \sim N(0, \sigma_\theta^2)$

Difficulty of
skill k

$q_{ik} = 1$, if item i uses skill k

Additive Factors Model (AFM)

$$p_{n,i} = \frac{1}{1 + e^{-\pi_{n,i}}}$$

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$

Student proficiency:
 $\theta_n \sim N(0, \sigma_\theta^2)$

$q_{ik} = 1$, if item i uses skill k

Difficulty of
skill k

Learning rate
at skill k

Number of practice
opportunities
student n had at
skill k

AFM - Assumptions

- Students may initially know more or less
 - Students learn at the same rate
 - Some skills are more likely to initially be known
 - Some skills are easier to learn than others
 - Students learn with each practice opportunity
 - Each item belongs to one or more skills
-

Performance Factors Analysis (PFA)

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} \cdot (\beta_k + \gamma_k \cdot T_{n,k})$$

Performance Factors Analysis (PFA)

$$\pi_{n,i} = \theta_n + \sum_k q_{i,k} + (\beta_k + \gamma_k \cdot s_{n,k} + \rho_k \cdot f_{n,k})$$

Number of prior
successes student
 n had at skill k

Number of prior
failures student n
had at skill k

PFA - Assumptions

- Students may initially know more or less
 - Students learn at the same rate
 - Some skills are more likely to initially be known
 - Some skills are easier to learn than others
 - Students learning rate differs for correct and wrong practice opportunities
 - Each item belongs to one or more skills
-

Tutorial 6

Tutorial 6

- **Task:** Pull Tutorial 6 from GitHub. We will then alternate walk-through and independent work phases based on your speed. Use SpeakUp to tell us when you're ready!
- **Virtual environments:**
 - <https://janakiev.com/blog/jupyter-virtual-envs/>
 - Create virtual environment: `python -m venv myenv`
 - Activate virtual environment: `source myenv/bin/activate`
 - add to Jupyter (**deactivate virtual environment first**)
`python -m ipykernel install --user --name=myenv`

Homeworks and Project
