Understanding Student Procrastination via Mixture ModelsUnderstanding Student Procrastination via Mixture Models

by Jihyun Park, Renzhe Yu, Fernando Rodriguez, Rachel Baker, Padhraic Smyth and Mark WarschauerJihyun Park, Renzhe Yu

presented @ EDM 1818

Roadmap

1

Why
should we care?

2

What
are the contributions?

3

How
do they do it?

1. Motivation

MOTIVATION

Students are complex human beings

- Generalized models have obvious limits regarding per-student performance.

We can gain by looking at groups of similar students

- But, the task of clustering can be intricate and definitely isn't trivial and could benefit from data-driven and automated tools.

By understanding in more details the **different behavioral patterns**, we also get closer to an accurate probabilistic modeling of students, and subsequently, of accurately producing some **synthetic test data** 🏆

2. Contributions

CONTRIBUTIONS

Using **click-stream data** from an online learning platform, the authors answer the following research questions in their work:

- We can cluster students that **procrastinate from the non-procrastinators** using **mixture models**.
- Furthermore, we can compute their **consistency** regarding the (non-)procrastination and **TM** score.
- Per cluster, the students obtain significantly **different learning outcomes**, motivating the usefulness of inferring such information.

CONTRIBUTIONS

Using **click-stream data** from an online learning platform, the authors answer the following research questions in their work:

- We can cluster students that **procrastinate from the non-procrastinators** using **mixture models**.
- Furthermore, we can compute their **consistency** regarding the (non-)procrastination and **TM** score.
- Per cluster, the students obtain significantly **different learning outcomes**, motivating the usefulness of inferring such information.

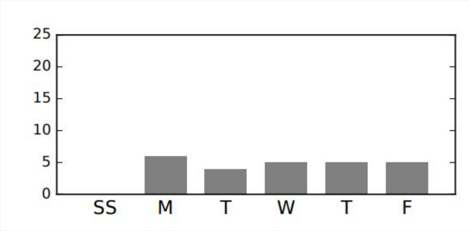
1) The authors show that using **mixture models is a great way to cluster students**. Their technique has the advantage of producing **highly interpretable outcomes**, allowing to extract even more information about the students and clusters in a **post-processing** step.

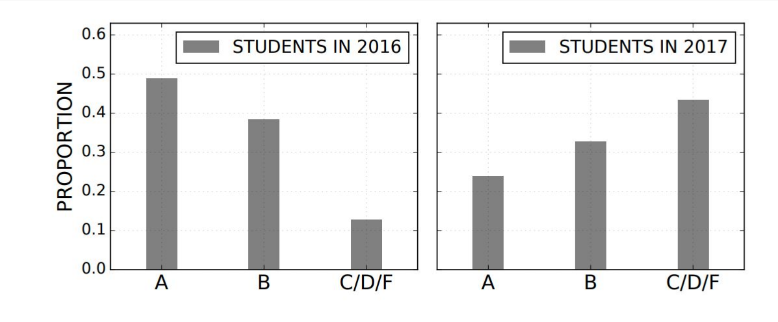
2) The authors **apply** their technique to capture information about **procrastination** of students using **real-life click data**, to group them into similar groups with great accuracy, and finally **extract and analyse** the learning patterns.

3. Going in depth

DATA SOURCES

Aggregated daily activities + grades





Data is collected in 2016 and 2017. It is quite different (instructor, type of data, ...):

- No comparison should be made between results
- But if it works on both, results are stronger.

METHODS

Task:

Group the students according to their learning patterns.

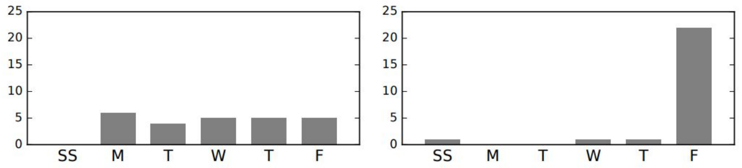


Figure 2: *Aggregated daily task counts* **across weeks (y_i) for the two students shown in Figure 1.**

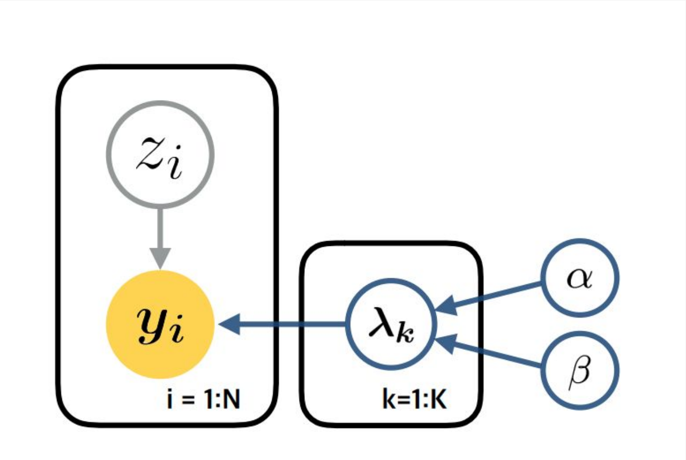
METHODS

<u>Mixture Model:</u>

y_i are the aggregated daily activities.

z_i are the marginal mixing weights.

λ_i prior of the cluster.



K components

METHODS

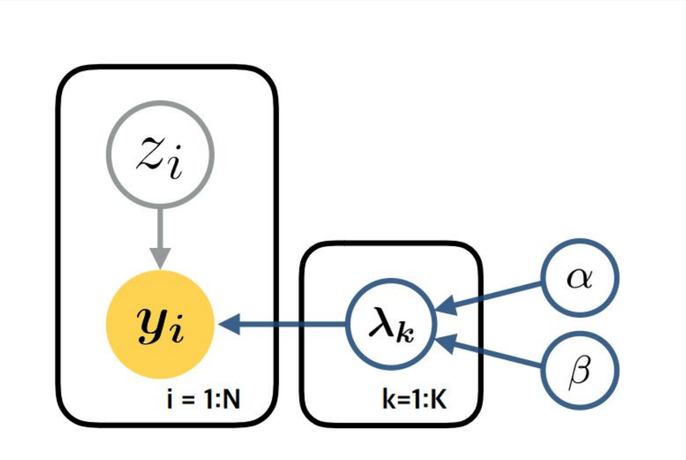
Mixture Model:

y_i are the aggregated daily activities.

z_i are the marginal mixing weights.

$\boxed{\lambda_i \text{ prior of the cluster.}}$

Through EM



K components

METHODS

<u>Mixture Model:</u>

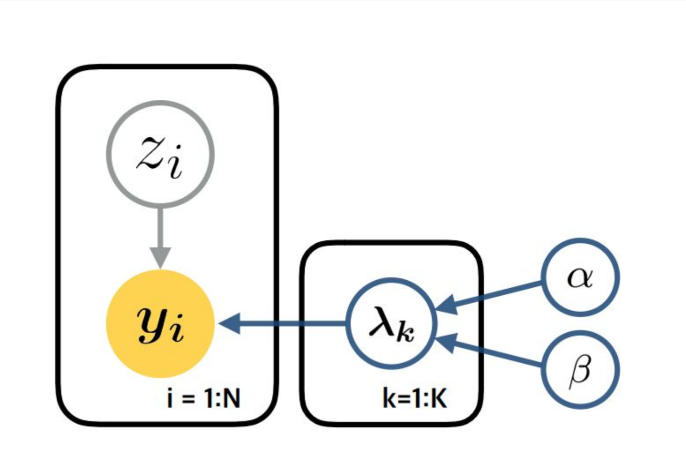
y_i are the aggregated daily activities.

z_i are the marginal mixing weights.

λ_i prior of the cluster.

Through EM

Likelihood: $p(y \mid \lambda)$

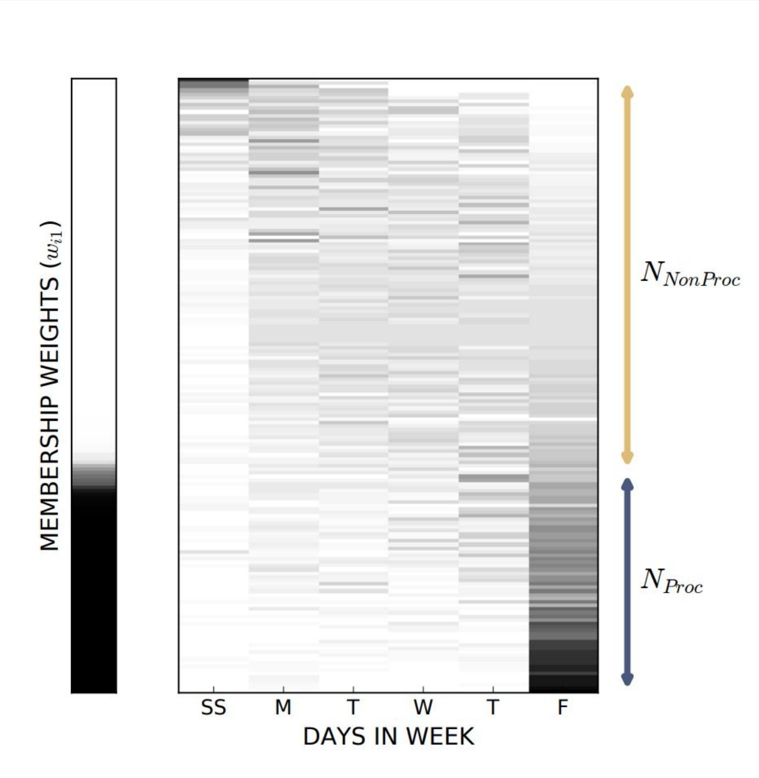


K components

Membership weights: w_{ik}, i.e., the probability that a student is a (non-)procrastinator .

$$w_{ik} = p(z_i = k | \boldsymbol{y_i}, \boldsymbol{\lambda}, \alpha, \beta)$$

RESULTS



Aggregated daily task counts shown along with the membership weights.

RESULTS

<u>Association between Behaviors and Grades:</u>

Non-procrastinators tend to get significantly more A grades than the procrastinators, whereas the procrastinators get more C, D, and F's.

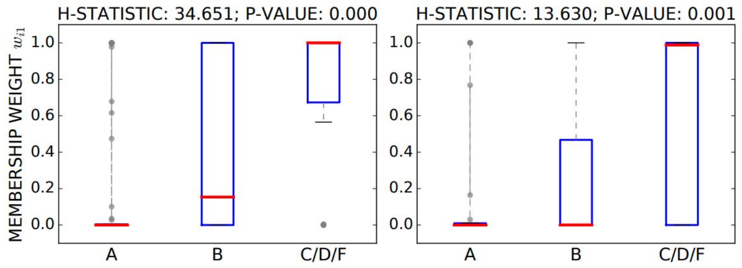


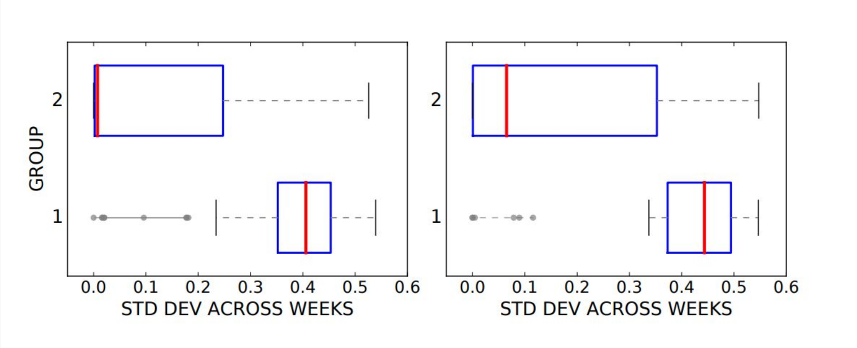
Figure 8: Distribution of w_{i1} in different grade groups of class in 2016 (left) and in 2017 (right). H-statistic comes from a Kruskal-Wallis test. (w_{i1}: membership weight on the procrastinating group)

RESULTS

Regularity:

$$SD_i = \left(\frac{1}{M-1} \sum_{j=1}^{M} \left(w_{ij1} - \overline{w}_{i\cdot 1} \right)^2 \right)^{1/2}$$

By definition, a higher value for SD signifies more volatile behavioral patterns.
Students with better grades in general have lower levels of SD, hence are more regular learners.



RESULTS

TM Score:

$$TM_i = (1 - 2w_{i1})(1 - SD_i)$$

RESULTS

Both negatively correlated ✅

TM Score:

$$TM_i = (1 \boxed{- 2w_{i1}})(1 \boxed{- SD_i})$$

RESULTS

Both negatively correlated ✅

TM Score:

$$TM_i = (1 \boxed{- 2w_{i1}})(1 \boxed{- SD_i})$$

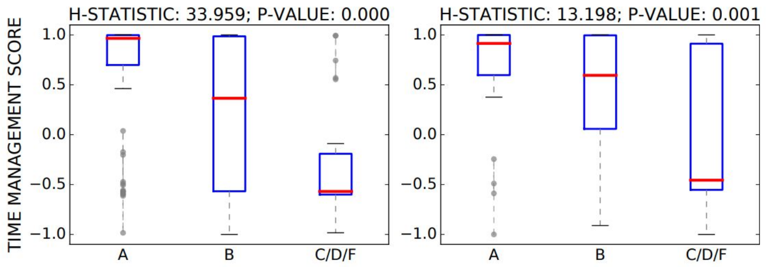


Figure 14: Distribution of Time Management Score (TM_i) in different grade groups of class in 2016 (left) and in 2017 (right). H-statistic comes from a Kruskal-Wallis test.

RESULTS

Relationship with Student Background:

Although it was not used, demographic data was available.
Interestingly, while **TM** is a strong predictor of course outcomes, it is not strongly and significantly related to students' demographics or prior academic achievement.

These results suggest that, as a whole, procrastination behaviors seem to be more of an inherent characteristic.

There is hope for everyone.

CONCLUSION

We have seen a **data-driven methodology to cluster students using mixture models**.

- ★ Grouping students eventually helps **towards their learning goals**, but also to understand the **learning patterns**.
- ★ Regarding the task of clustering according the **procrastination** behaviors, MM shows great accuracy.
- ★ By **post-processing** the outputs of the mixture models, the authors derive some scoring metrics such as the **regularity** of (non-)procrastination and the **time management score**.
- ★ It can be **extended** to other behavioral pattern, but, the number of clusters has to be **explicitly given**.