Learning Analytics: an investigation on the influence of school quality in overcoming social inequalities

A non-parametric analysis on INVALSI data

Final Presentation

3 February 2021

POLITECNICO
MILANO 1863

INVALSI

The Data

We based our analysis on the results of **28145 students** from **658 schools** of Italy, observing **305 variables** for each record in the dataset:

INVALSI

S.Y. 2013/14 Data

- Primary School
- Middle School
- High School
  - 1st year results (S.Y. 2011/12) (Grade 6)
    - Italian
    - Maths
  - 3rd year students (Grade 8)
    - Italian
    - Maths

\+ School management questionnaire

Data were also considering **Socio-Economical** status of each students, in both **aggregated** and **disaggregated** flavours:

- Parents' education
- City/Countryside
- Family status
- Family size

⟷ **ESCS** ECONOMIC SOCIAL CULTURAL STATUS

School Level Aggregation

After the good results obtained at a student level, we considered them under a **school level perspective**, as the **literature** seems to **lack** an **in-depth analyisis** at that level

Student-level data

- **Categorical** features → We selected the **most meaningful classes** and computed the school-wise proportion (e.g. *prop_laurea_madre*).
- **Numerical** features → We selected **a set of aggregating functions** (mean, standard deviation, skew, quantiles) and aggregated on school level using those (e.g. *mean_ESCS*).

School management Questionnaire

- **Multiple choice** (often in **«ordinal» fashion**) and binary answers

+

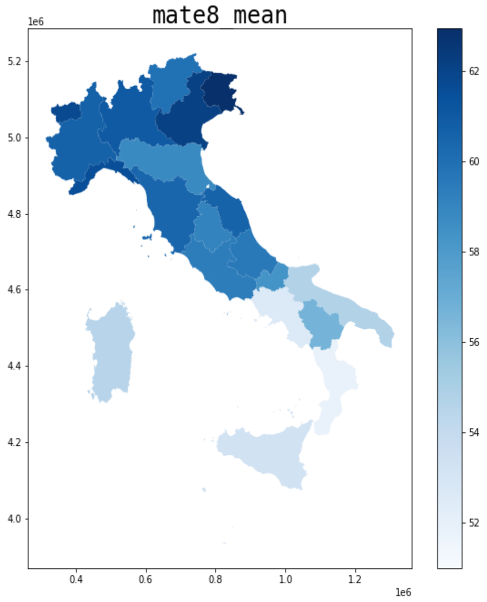
- Presence of **sub-questions**

→ We encoded the answers with **ordered integers** and kept the **average** of the subquestions as feature (e.g. *opinione_invalsi*)

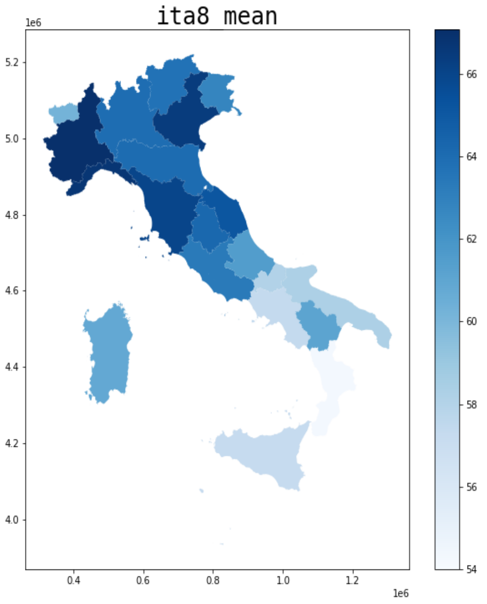
North – Center/South Differences

The **North/South differences** found at student level have a clear reflection also by looking at **schools**:

Differences in absolute **outcomes** and relative improvements.



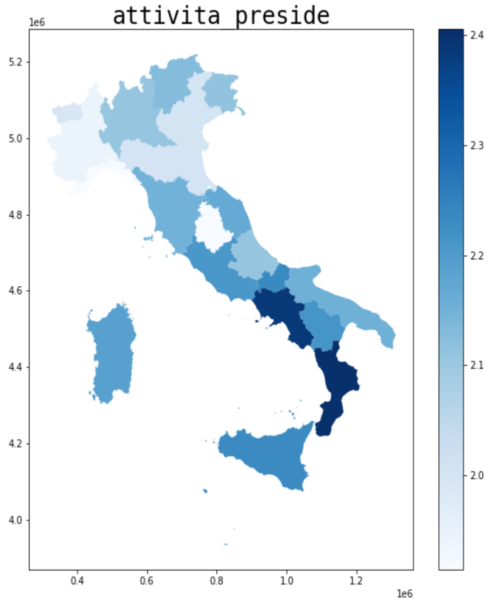
P-value: 0



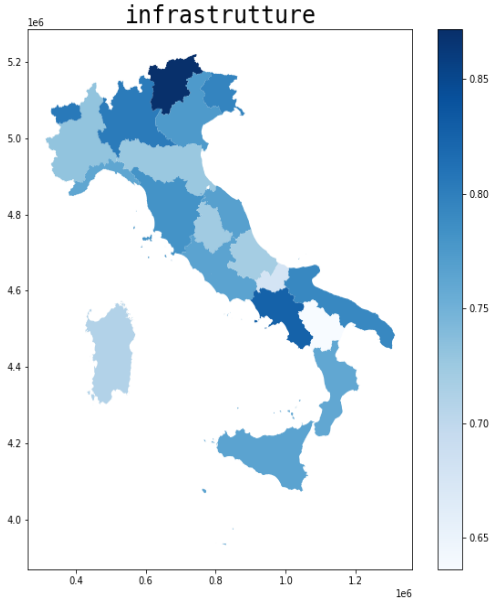
P-value: 0

Better absolute outcomes in the North for Maths and Italian.

Differences in **indices** and **factors**.



P-value: 0

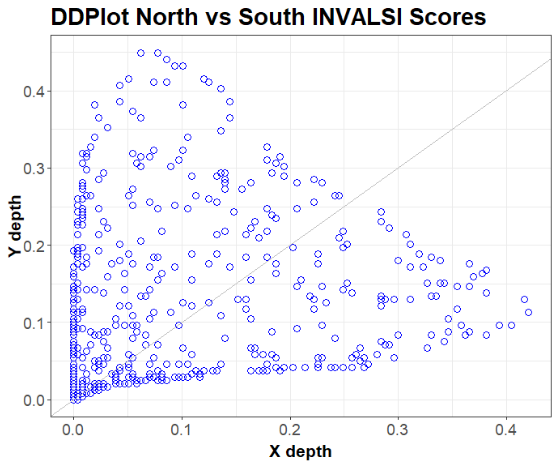


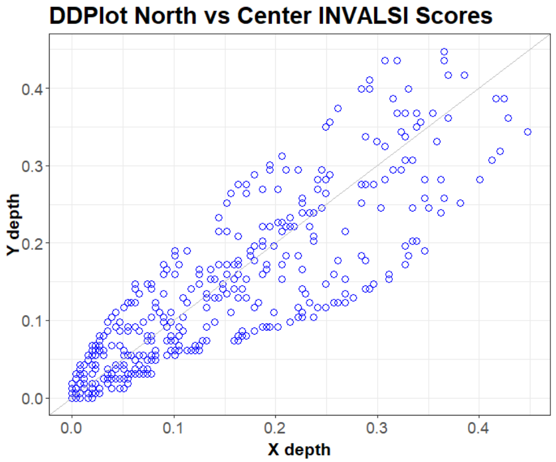
P-value: 0.0047

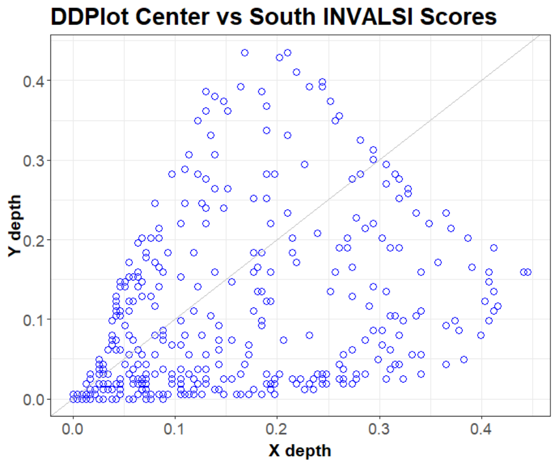
More principals' self-reported activity in the south and better reported infrastructure in the North.

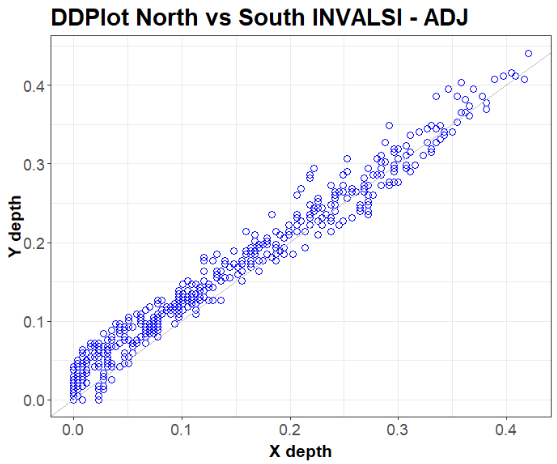
North – Center/South Differences: A DDPlots Investigation

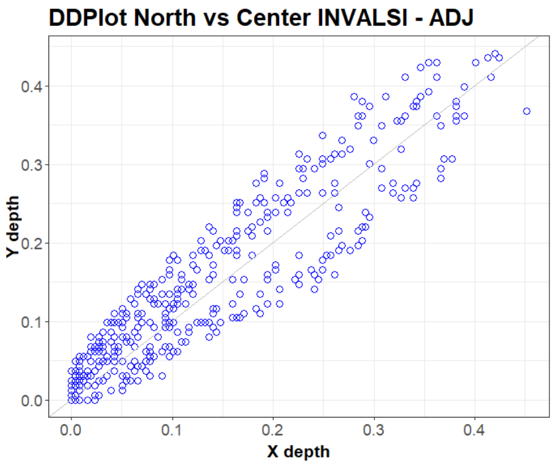
After the **Location-Scale adjustments** the **distributions** are **comparable**, suggesting a **difference** only **affecting the parameters** of the underlying distribution and **not the distribution itself**.

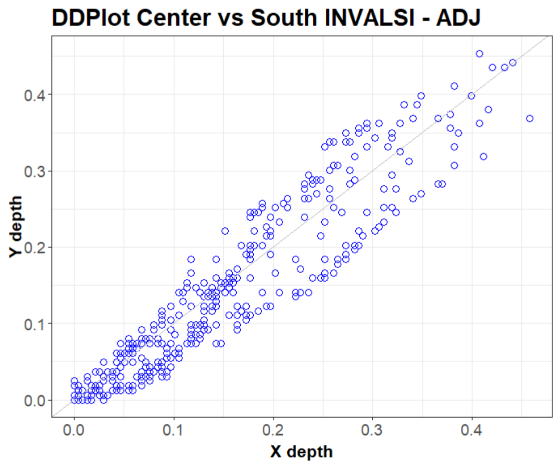












Connecting Managerial Practices to Outcomes

We focus on finding possible connections between **managerial practices** in schools (*indices* and *debiased answers*) and **school outcomes**.

- Bootstrap Reverse Percentile Intervals for Slope Estimation
- Univariate Permutational Tests
- Permutational MANOVA
- Mann-Whitney Test

Further grouping and interpretation

260 Questions → **29** Leads → **7** Results → **4** Insights

Stress testing

Actionable Insights to School Management

We have synthesized our findings in 4 insights:

Negative link between the tendency of responsibilization of teachers and Maths; **positive link** between the tendency of setting objectives for everyone and Maths.

Strong leadership and judgement

Relationship with parents

Parents' involvement **lowers** the **variance** of scores and we see a possible **positive** direct **link** between parents' pressure and absolute scores.

Usage of INVALSI

Positive association, twice at a regional level, between Maths and Invalsi usage/consideration.

Gender differences

For both Italian and Maths, a male principal relates to a **higher variance** in the scores.

School performances driving factors

In order to find out what differentiates **high** and **low impact schools**, we fitted a **mixed effects model** on the INVALSI outcome at 3rd year taking into account the schools as factors (discarding schools with the **lowest correction coefficient**):

$$score_{third\ year} \sim score_{first\ year} + school.effect\ + school.effect * score_{first\ year}$$

We decided to focus on Maths performances only, where we achieved an adjusted R^2 of **0.5655** w.r.t. **0.4505** of the model **without the school factor**. This allowed us to divide the schools into 2 groups:

'TOP' SCHOOLS
Schools in the top 20% of effects, having a significant p-value (<0.01) in the regression

Higher minimal ESCS among students (0.0376)
Higher percentage of educated fathers (0.0278)
More variance in ESCS among students (0.0420)

'FLOP' SCHOOLS
Schools in the worst 20% of effects, having a significant p-value (<0.01) in the regression

More discussion of INVALSI scores with teachers (0.0274)

Discussing the results with teachers is **not effective**: School Inspectorate should promote **more substantial actions** in schools showing poor performances.

Predicting the Schools at Risk

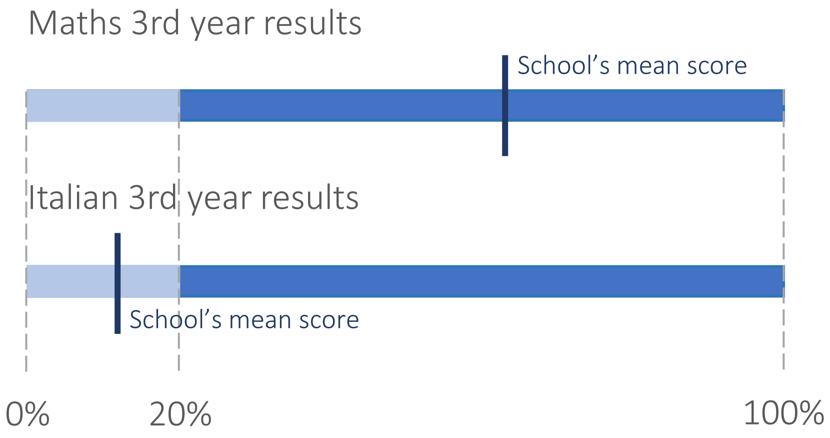
A model able to **predict poor performing schools** could be a **powerful tool** in the hands of **School Inspectorate**

INVALSI Scores at 1st year
School's data
School's management questionnaire data

→ **MODEL** (Binary classification) →

School at Risk (SaR) binary classification

We defined **Schools at Risk** as the ones getting results **below the 20th percentile** at 3rd year, separately in Maths and Italian



Example:
School at 65% percentile in Maths results and at 13% percentile in Italian results at third year:

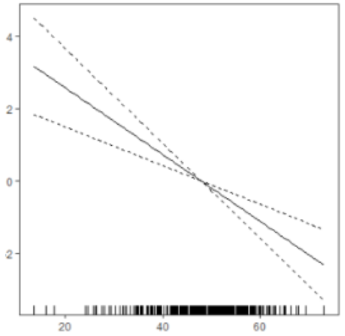
SaR-Maths = 0, SaR-Italian = 1

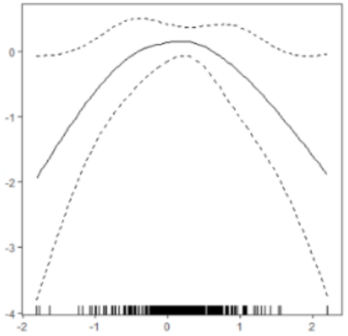
We considered **Generalized Additive Models** and **Explainable Boosting Machines** in order to achieve this goal.

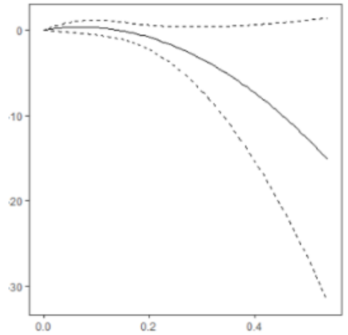
Predicting the Schools at Risk: GAM approach

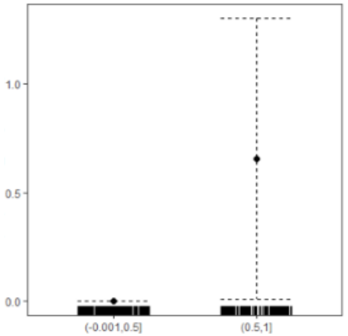
We first considered **Logistic GAMs**, for which we selected relevant features via **4-fold stratified cross-validation** performances.

Maths

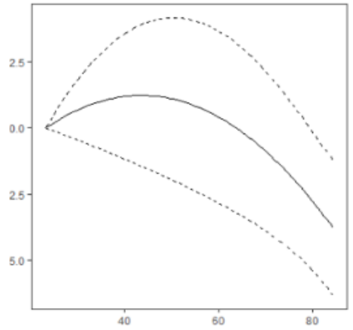

Mean Maths – 1st year

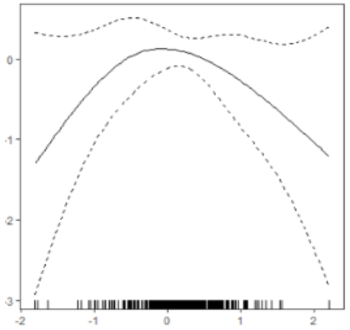

Mean school ESCS

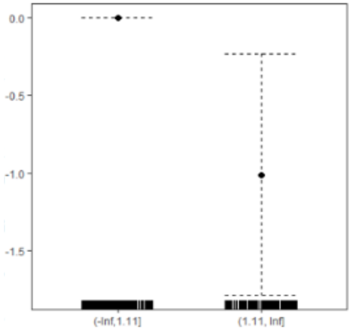

% Mothers with degree

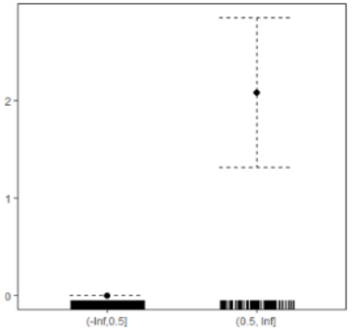

Discussion of scores with teachers

Italian


Mean Italian – 1st year


Mean school ESCS


Mean parents' involvement

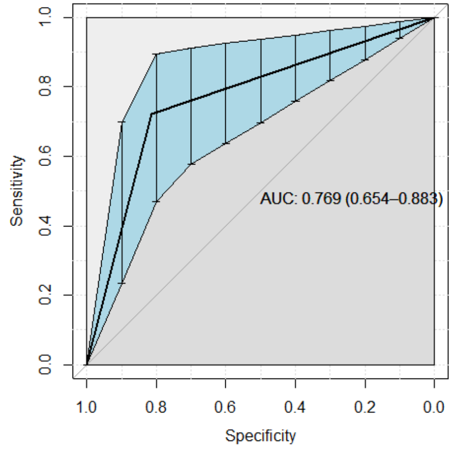

School location in the South

The models achieve **33.4%** of explained deviance in Maths and **31.5%** in Italian

Predicting the Schools at Risk: GAM approach performances

We achieved **good performances** in 4-fold stratified cross-validation:

Maths



Confusion Matrix – Maths

	Predicted No SaR	Predicted SaR
True No SaR	257	64
True SaR	19	52

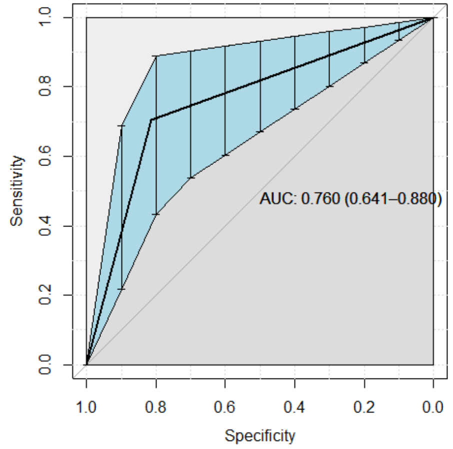
Accuracy: 0.7883 (0.0226)

AUC: 0.7670 (0.0054)

Sensitivity: 0.7328 (0.0212)

Recall@20th: 0.5240 (0.0194)

Italian



Confusion Matrix – Italian

	Predicted No SaR	Predicted SaR
True No SaR	255	71
True SaR	20	50

Accuracy: 0.7703 (0.0199)

AUC: 0.7483 (0.0389)

Sensitivity: 0.7148 (0.1183)

Recall@20th: 0.5742 (0.0324)

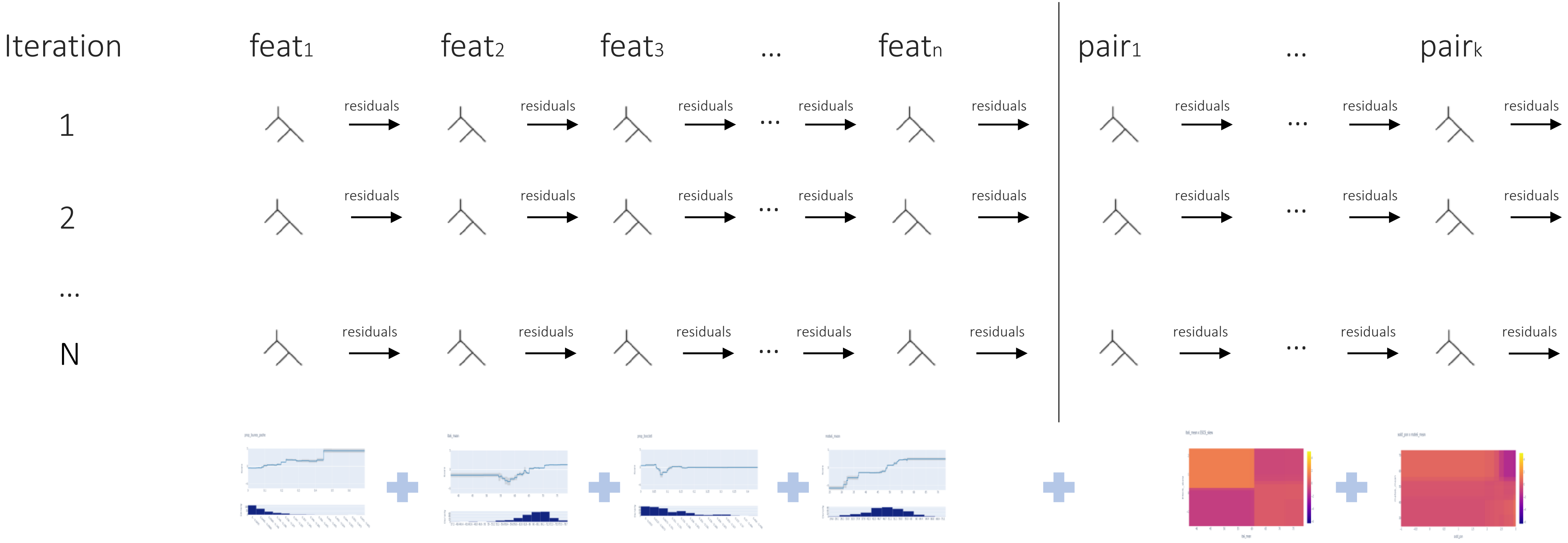
Nevertheless, we were **not satisfied yet**…

Explainable Boosting Machines

Our *desiderata* for the final model included the following features:

- It should spot some **nonlinear relationships and interactions** between variables
- It **shouldn't overfit** too easily
- It should keep a **high level of interpretability**.

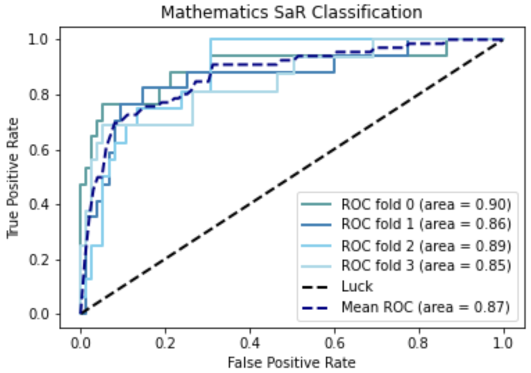
We came up with **Explainable Boosting Machines**:



Predicting the Schools at Risk: EBM approach performances

We achieved **very good performances** in **4-fold stratified cross-validation**:

Maths



Confusion Matrix – Maths

	Predicted No SaR	Predicted SaR
True No SaR	248	52
True SaR	15	51

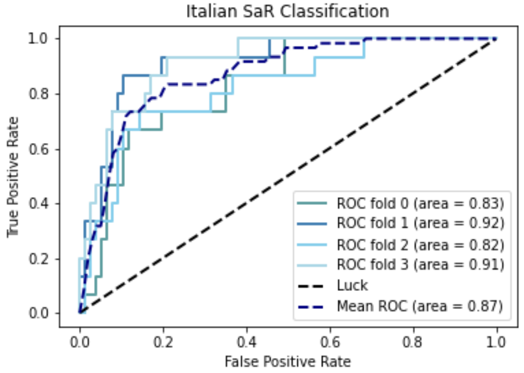
Accuracy: 0.8166 (0.0732)

AUC: 0.8742 (0.0019)

Sensitivity: 0.773 (0.0236)

Recall@20th: 0.6389 (0.0621)

Italian



Confusion Matrix – Italian

	Predicted No SaR	Predicted SaR
True No SaR	263	43
True SaR	16	44

Accuracy: 0.8553 (0.0189)

AUC: 0.8701 (0.0441)

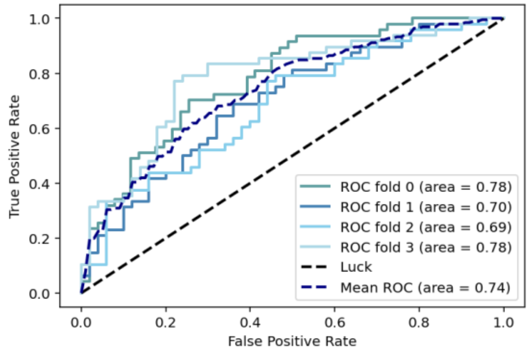
Sensitivity: 0.7028 (0.1977)

Recall@20th: 0.5694 (0.0461)

Predicting the Improvement: EBM approach performances

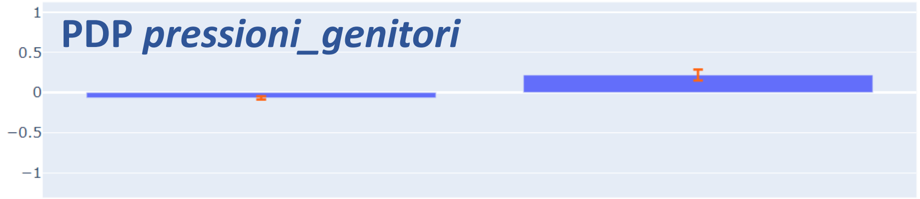
Using **4-fold stratified cross-validation**, we classified **schools showing an improvement** (> national median) in INVALSI scores:

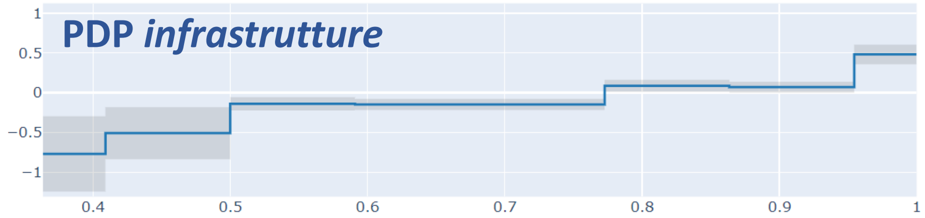
Maths



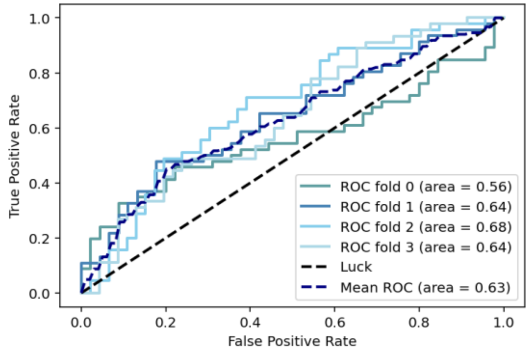
AUC: 0.74 (0.0454)
Recall@20th: 0.8026 (0.0436)

- Most important features are infrastructures and family background information.
- Overall good performance, strong dependence on starting level.



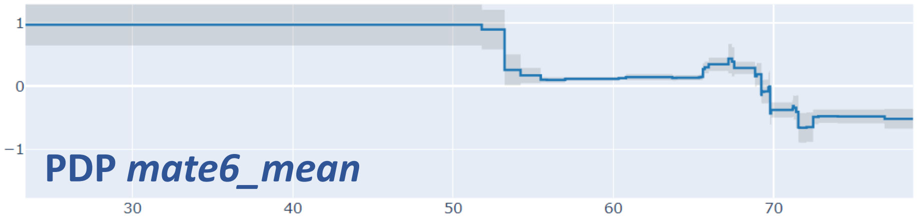


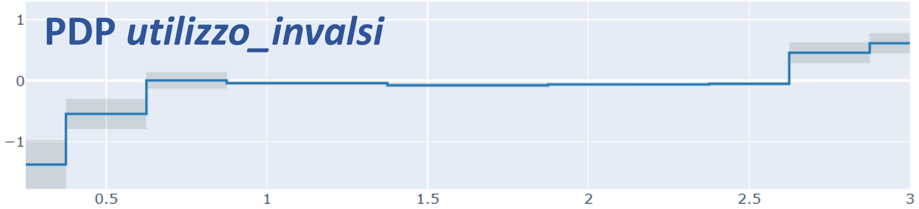
Italian



AUC: 0.63 (0.0408)
Recall@20th: 0.7222 (0.0393)

- Difficult task
- Geographical Area and management's usage of INVALSI scores are the most relevant features
- Strong dependence on starting level, also on its skew.





Conclusions & Further Advancements

Summing up, we can summarize our findings over these 3 axes:

1. Pointed out **structural regional differences,** not only from an **outcome perspective**, but also from a **factors / indices point of view.**

2. Uncovered **influences of managerial practices on school performance**, ending with a selection of **4 strategic suggestions.**

3. Successfully implemented GAM/GA2M prediction models to obtain an **accurate estimate of the future performance of schools**.

Further advancements to our work could be the following:

- Use suitable models for nonparametric **causal inference**.

- Re-analyse indices and more management-related questions considering the **user de-biasing procedure**.

- Replicate the analysis using DEA (**Data Envelopment Analysis**).

- Interesting to test the quality of the analysis using data from **INVALSI tests of the following years** and possibly including the **time-dependency** of the phenomenon

Thank You!

Yin Lou, Rich Caruana, Giles Hooker, Johannes Gehrke, **Accurate Intelligible Models with Pairwise Interactions** - KDD'13, August 11-14, 2013, Chicago, Illinois, USA, 2013

Harsha Nori, Samuel Jenkins, Paul Koch, Rich Caruana, **Interpretml: A unified framework for machine learning interpretability** - arXiv preprint arXiv:1909.09223, 2019

Chun-Hao Chang, Sarah Tan, Ben Lengerich, Anna Goldenberg, Rich Caruana, **How Interpretable and Trustworthy are GAMs?**, arXiv preprint arXiv:2006.06466, 2020

Chiara Masci, Kristof De Witte, Tommaso Agasisti, **The influence of school size, principal characteristics and school management practices on educational performance: An efficiency analysis of Italian students attending middle schools** - Socio-Economic Planning Sciences, Volume 61, Pages 52-69, 2018

Fritz Schiltz, Chiara Masci, Tommaso Agasisti, Daniel Horn. **Using regression tree ensembles to model interaction effects: a graphical approach** - Applied Economics, 50:58, 6341-6354, DOI: 10.1080/00036846.2018.1489520, 2018

Chiara Masci, Anna Maria Paganoni, Francesca Ieva, **Semiparametric mixed effects models for unsupervised classification of Italian schools** - Journal of the Royal Statistical Society: Series A, 2019