

AcolheEdu: Predicting Psychosocial Vulnerability in the School Context with Histogram-Based Gradient Boosting Trees

Beatriz Brum¹, Iasmin Dembinski², Cristhiano Vasconcellos¹, Carlos Santos¹

¹ Federal Institute Farroupilha (IFFar)
Alegrete – RS – Brasil

²Federal University of Pampa (UNIPAMPA)
Alegrete – RS – Brasil

beatriz.06059@aluno.iffar.edu.br
iasmingsilva.aluno@unipampa.edu.br
cristhiano.vasconcellos@iffarroupilha.edu.br
carlos.santos@iffarroupilha.edu.br

Abstract. *AcolheEdu proposes a data-driven screening approach to support schools in the early identification of students potentially vulnerable to psychological distress. As a proof of concept, it uses public, anonymized PeNSE 2019 microdata and a calibrated Machine Learning model to estimate a risk score from self-reported psychosocial factors. The calibrated HistGradientBoosting achieved ROC-AUC 0.86 and PR-AUC 0.77 in stratified cross-validation, and ROC-AUC 0.859 on a holdout test set. With an operational threshold of 0.30, recall reached 80.1% with 61.7% precision. The product flow is currently a navigable Figma prototype; REST API integration (FastAPI) is planned for implementation. These results indicate feasibility for high-sensitivity, data-driven school mental health screening.*

1. Introduction

Mental health in the school environment has increasingly been recognized as a relevant challenge in Brazil. Data from the National Survey of School Health indicate a high prevalence of feelings of sadness and loneliness among adolescents (PeNSE 2019) [IBGE 2021]. At the institutional level, pedagogical teams and psycho-pedagogical support units face human-resource constraints alongside high demand for services, which hinders the early identification of students experiencing psychological distress. The lack of systematic, data-driven screening instruments means that interventions occur predominantly in a reactive manner, when signs are already more severe. This scenario contributes to increased risks such as school dropout, declining academic performance, and the worsening of psychological distress [Andifes 2019].

In this context, AcolheEdu was conceived as a support solution for psycho-pedagogical and pedagogical support units, based on a high-sensitivity automated screening approach using public data. The Federal Institute Farroupilha (IFFar), in the context of integrated upper secondary education, serves as the initial adoption setting for the solution as a pilot project, with workflow governance assigned to the Support Unit of the Alegrete campus.

The proposal is innovative because it integrates public data, a calibrated predictive model, and an operational school-screening workflow, which has not yet been institutionalized at IFFar. Therefore, this study is characterized as a proof of concept (PoC), aimed at evaluating the technical feasibility of the solution as a preliminary step toward institutional adoption. The main contributions of this work are:

- the design of AcolheEdu as a data-driven school-screening workflow to support psycho-pedagogical and pedagogical teams;
- the evaluation of a calibrated machine learning model using public and anonymized PeNSE 2019 microdata;
- the definition of an operational threshold prioritizing high sensitivity for initial screening;
- the presentation of a navigable prototype for communicating risk scores, risk bands, and recommendations without replacing specialized human assessment.

2. Methodology

The process followed the CRISP-DM methodology [Shearer 2000, Martinez-Plumed et al. 2021], organized into four main stages: (i) data preparation, (ii) modeling and calibration, (iii) operational threshold selection, and (iv) prototyping of the intended usage flow, as detailed below.

We used PeNSE 2019 microdata (42,422 records of upper secondary students). The target variable was operationalized as a self-report-based risk *proxy* defined by frequent/constant sadness. PeNSE does not provide a clinical diagnosis of depression; therefore, we adopted proxies from the mental health module in the following order: (1) frequent sadness (last 30 days)—recommended for screening; and (2) “life is not worth living” (last 30 days). We selected 46 predictors grouped into thematic blocks (sleep, violence/bullying, social support, physical activity, substance use, routines, screens/internet, and school context), excluding variables directly related to mental health to mitigate semantic leakage.

We compared Logistic Regression (an interpretable baseline) and HistGradientBoosting [Xia et al. 2024]. The final model was calibrated via *Platt scaling* [Silva Filho et al. 2023] to improve probability reliability for threshold-based use [Niculescu-Mizil and Caruana 2005]. For initial screening, we prioritized *Recall*, given the severity of false negatives. Multiple thresholds were evaluated, and 0.30 was selected as the operating point balancing high *Recall* with acceptable precision.

The intended product flow involves a mobile application that (i) collects responses to a questionnaire, (ii) queries a REST API (FastAPI), and (iii) returns a *risk score*, risk band, and recommendations to support school management. At this stage, however, the prototype is implemented as a navigable Figma flow (without functional API integration); API implementation and end-to-end integration are part of the system’s evolution roadmap. In this phase, we used exclusively public, already anonymized microdata from the National Survey of School Health (PeNSE 2019), made available by IBGE¹. As a proof of concept (PoC), the goal was to reduce technical and feasibility uncertainties—such as the data pipeline and model, calibration, threshold selection, and presenta-

¹<https://www.ibge.gov.br/estatisticas/sociais/saude/9134-pesquisa-nacional-de-saude-do-escolar.html>

tion of results in accessible language—before advancing to field stages involving institutional stakeholders and formal validation procedures.

As next steps, we plan to conduct requirements elicitation and exploratory validation with school stakeholders (psychologists, student support staff, and management teams) through semi-structured interviews and/or co-design workshops, reporting the number of participants (N) and professional profiles to refine requirements, recommendation wording, and referral workflows. We will also structure an institutional pilot with applicable informed-consent terms (student and/or guardian, depending on local policy and study design) and submit the protocol for ethical review by a Research Ethics Committee (CEP), when applicable, prior to any collection/use of non-public data or any direct/indirect identification of participants.

3. Proposed Solution

The proposed solution integrates a calibrated predictive model into an operational school screening workflow, emphasizing high sensitivity and privacy-by-design principles. The goal is to support pedagogical and psychoeducational teams in prioritizing student support and referrals, reducing reliance on reactive interventions and strengthening the capacity for early identification of potentially vulnerable students.

The target architecture includes a REST API implemented in FastAPI, responsible for hosting the inference pipeline (input validation, preprocessing, application of the calibrated model, and post-processing). This backend exposes a `/predict` endpoint that receives questionnaire responses and returns three main components: (i) a calibrated probability associated with the *risk score*, (ii) a risk-band classification (low/moderate/high), and (iii) actionable, domain-oriented recommendations (e.g., sleep, bullying/violence, social support, routines and screen use), accompanied by institutional guidance and support contacts (for example, CVV 188), enabling the school to adopt measures consistent with its support network and internal protocols.

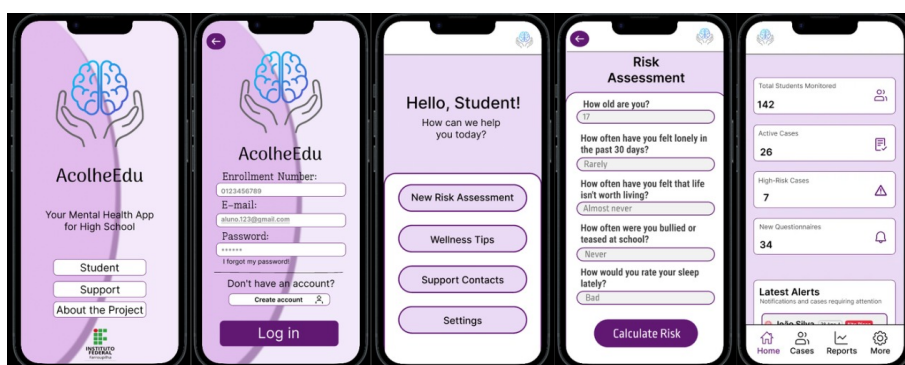


Figure 1. AcolheEdu prototype: home screen, login, student menu, risk assessment questionnaire, and monitoring panel for the Support Unit with indicators and alerts.

On the *front-end*, a mobile application will collect responses to a questionnaire derived from PeNSE items and present the output in an understandable way, avoiding diagnostic language and emphasizing its screening-and-prioritization purpose. At this stage, the interface and navigation flow are available as a navigable prototype in Figma,

while the model is evaluated *offline*; the end-to-end functional integration (app-API) is planned as the next implementation step, including robustness tests and validation of system behavior under different operational thresholds. Figure 1 shows the main prototype screens of the AcolheEdu app.

Regarding compliance and data governance, the application provides for data minimization and anonymization of personal data, as well as traceability mechanisms through auditable *logs*, access control, and assurance of the right to deletion, in compliance with the Brazilian General Data Protection Law (LGPD) [Brasil 2018]. Finally, the system does not provide a clinical diagnosis: its results must be used exclusively to support institutional screening and the management of student support actions, complementing—and not replacing—specialized human assessment.

4. Results and Discussion

Table 1 presents the calibrated model’s performance in two scenarios: (i) stratified 5-fold cross-validation, which estimates the expected average performance on samples similar to the training data, and (ii) a *holdout* test, which simulates generalization to unseen data.

Table 1. Performance of the calibrated model.

Scenario	ROC-AUC	PR-AUC / Precision / Recall
Cross-validation (5-fold)	0.86	PR-AUC = 0.77
Test (<i>holdout</i>)	0.859	Precision = 0.617; Recall = 0.801 (thr = 0.30)

In cross-validation, $\text{ROC-AUC} \approx 0.86$ indicates good overall ranking ability between students at higher and lower risk: in probabilistic terms, there is about an 86% chance that the model assigns a higher *score* to a positive case than to a negative one. Because the problem tends to be imbalanced (positives relatively rare), $\text{PR-AUC} \approx 0.77$ is particularly informative, as it evaluates the trade-off between *precision* and *recall* as the decision threshold varies, better reflecting operational usefulness when the priority is to recover potential cases at the lowest possible cost in false alarms.

On the *holdout* set, $\text{ROC-AUC} \approx 0.859$ remains consistent with cross-validation, suggesting stability and no relevant degradation when moving beyond the training data. By adopting an operating point with a 0.30 threshold, the model achieves $\text{Recall} \approx 0.801$, i.e., it identifies approximately 80% of the students classified as “high risk” by the *proxy*. This high sensitivity is consistent with the logic of initial screening (minimizing false negatives), in which the cost of “not flagging” a potentially vulnerable case is generally higher than the cost of generating an additional alert. In contrast, $\text{Precision} \approx 0.617$ indicates that, among those flagged above the threshold, about 62% belong to the positive group defined by the *proxy*; therefore, a fraction of referrals will require human review, reinforcing the system’s role as support for prioritization rather than as a diagnostic instrument.

Model calibration adds operational value because it transforms outputs into more consistent probabilities, allowing the school to adjust the threshold according to its support capacity (for example, raising the threshold to reduce the volume of flagged cases, or lowering it to maximize sensitivity in a scenario with greater staff availability). In addition, interpretability mechanisms (e.g., SHAP) can support responsible communication

by indicating which psychosocial factors contributed to the *score* at an aggregate level, without attributing causality and without replacing specialized assessment.

To move from technical feasibility to deployment feasibility, a two-stage pilot is proposed: (i) a *closed pilot*, conducted with a multiprofessional team (psychology/guidance/management), to validate questionnaire language, referral workflows, thresholds, and care-oriented messages; and (ii) an *assisted pilot*, with supervised application in selected classes, with criteria and service capacity defined in advance, monitoring flagging rate, response time, and workflow adequacy.

As limitations and operational risks, it is emphasized that the *risk score* is based on secondary, self-reported data and on a *proxy* outcome rather than on clinical diagnosis or institutionally validated records. Therefore, the results should be interpreted exclusively as evidence of technical feasibility for initial screening, not as proof of clinical or institutional effectiveness. Because the data are cross-sectional (PeNSE 2019), changes in the school context may induce *dataset shift*, requiring local calibration and validation before continued use. In addition, self-report biases, under-/over-reporting, and differences across school contexts may affect generalization and generate false positives and/or false negatives. False positives may increase workload and unnecessary concern, whereas false negatives may delay support for vulnerable students. For this reason, the system requires conservative threshold selection, interpretability support, systematic human review, and continuous monitoring of referral volume according to institutional capacity.

Real-world validation is planned as a staged process. First, a closed validation round with the Support Unit will assess the clarity of the questionnaire, the adequacy of risk-band communication, and the feasibility of the referral workflow. Second, semi-structured interviews and/or co-design workshops with psychologists, pedagogical staff, and school managers will be used to refine the recommendations and define operational thresholds according to institutional capacity. Third, an assisted pilot in selected classes will monitor practical indicators such as flagging rate, review workload, response time, and adequacy of referrals. No autonomous decision-making is intended: all alerts generated by AcolheEdu must be reviewed by qualified staff before any institutional action.

5. Conclusion and Future Work

AcolheEdu represents an innovative data-driven school screening proposal, with potential for replication in other institutions. The results indicate technical feasibility for an initial high-sensitivity screening approach in the context of school mental health, using public, anonymized microdata from PeNSE 2019 and a calibrated model (ROC-AUC 0.859 on the *holdout*; with a 0.30 threshold, *Recall* 80.1% and precision 61.7%). At this stage, the product flow is embodied as a navigable prototype in Figma, aiming to validate requirements and governance prior to functional implementation through a REST API/FastAPI. Therefore, AcolheEdu should be understood as an initial PoC that demonstrates technical feasibility, while real-world effectiveness, acceptability, and operational impact remain to be evaluated in institutional pilots.

As future work, we plan to expand the experimental and institutional validation stages. This includes extending the evaluation protocol with additional threshold-dependent and calibration-oriented metrics, such as F1-score, specificity, balanced accuracy, and Brier score; comparing the proposed approach with stronger tabular-learning

models, such as XGBoost, LightGBM, CatBoost, Random Forest, and TabNet; and conducting validation with end users and institutional stakeholders through interviews, co-design workshops, and assisted pilots. These steps will support the refinement of the questionnaire, risk-band communication, referral workflow, operational thresholds, privacy safeguards, consent terms, and CEP submission, when applicable.

Acknowledgements

This study was partially funded by the National Council for Scientific and Technological Development (CNPq), Brazil.

Ethical approval

This study used exclusively the public PeNSE 2019 dataset, which consists of anonymized records with no identifiable information. The study did not collect participant data or access personal data; therefore, it did not require additional informed consent or approval from an ethics committee.

Use of generative AI tools

We used generative AI tools only to assist with minor language revisions during manuscript review. All study design, data collection, data analysis, and scientific interpretation were conceived and carried out by the authors, who assume full responsibility for the content of this work.

References

- Andifes (2019). V pesquisa nacional de perfil socioeconômico e cultural dos(as) graduandos(as) das IFES – 2018. Technical report, Associação Nacional dos Dirigentes das Instituições Federais de Ensino Superior (Andifes), Brasília.
- Brasil (2018). Lei geral de proteção de dados pessoais (LGPD). Lei nº 13.709, de 14 de agosto de 2018.
- IBGE (2021). Pesquisa nacional de saúde do escolar: 2019 (PeNSE 2019). Technical report, Instituto Brasileiro de Geografia e Estatística (IBGE), Rio de Janeiro.
- Martinez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernandez-Orallo, J., Kull, M., Lachiche, N., Ramirez-Quintana, M. J., and Flach, P. (2021). CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. *IEEE Transactions on Knowledge and Data Engineering*, 33(8):3048–3061.
- Niculescu-Mizil, A. and Caruana, R. (2005). Predicting good probabilities with supervised learning. In *Proceedings of the 22nd International Conference on Machine Learning (ICML '05)*, pages 625–632.
- Shearer, C. (2000). The crisp-dm model: the new blueprint for data mining. *Journal of Data Warehousing*, 5(4):13–22.
- Silva Filho, T., Song, H., Perello-Nieto, M., Santos-Rodriguez, R., Kull, M., and Flach, P. (2023). Classifier calibration: a survey on how to assess and improve predicted class probabilities. *Machine Learning*, 112(9):3211–3260.
- Xia, L., Zheng, P., Li, J., Huang, X., and Gao, R. X. (2024). Histogram-based gradient boosting tree: A federated learning approach for collaborative fault diagnosis. *IEEE/ASME Transactions on Mechatronics*, 29(4):2637–2648.