Multi-Modal Education Equity Monitor: A Machine Learning System for Identifying School Infrastructure and Resource Inequities

Project Proposal: Multi-Modal Education Equity Monitor

1. Introduction & Problem Statement

School quality varies widely across districts, often due to unequal funding, aging infrastructure, and socioeconomic disparities. Current evaluation methods—like inspections or self-reported data—are slow and inconsistent, leaving underserved schools unidentified for years.

Goal: Build a system that uses **tabular data and text** to estimate an **Equity Risk Score** for each school. This can help districts and policymakers identify where additional support is needed.

2. Data Sources

- **NCES (National Center for Education Statistics)** datasets: funding, enrollment, test scores, student–teacher ratios, free/reduced lunch rates.
- **U.S. Census/ACS data:** neighborhood income and broadband access.
- **School inspection reports** and district facility audits (PDFs/text).

3. Methods / Techniques / Technologies

- **Tabular Models:** XGBoost/Random Forest to analyze funding levels, performance metrics, and demographics.
- **Text Models:** BERT-based models to extract issues (e.g., maintenance, safety) from inspection reports.
- **Multi-Modal Fusion:** Combine image, tabular, and text outputs into one **Equity Risk Score**.
- The labels in the dataset come entirely from the district-provided ESA/FCA scores, which reflect the educational suitability and facility condition of each campus. Every data modality (text, demographics) is linked to a school via its ID, so it simply inherits that school's ESA/FCA label. The school is labeled, not the individual data sources, and all modalities contribute features toward predicting those known labels.

4. Project Deliverables

1. **Tabular model** predicting resource and performance risk.
2. **Text extraction model** identifying key issues from inspection reports.
3. **Combined multi-modal Equity Risk Score** for each school.

4. **Final report** describing data, methods, experiments, and results.

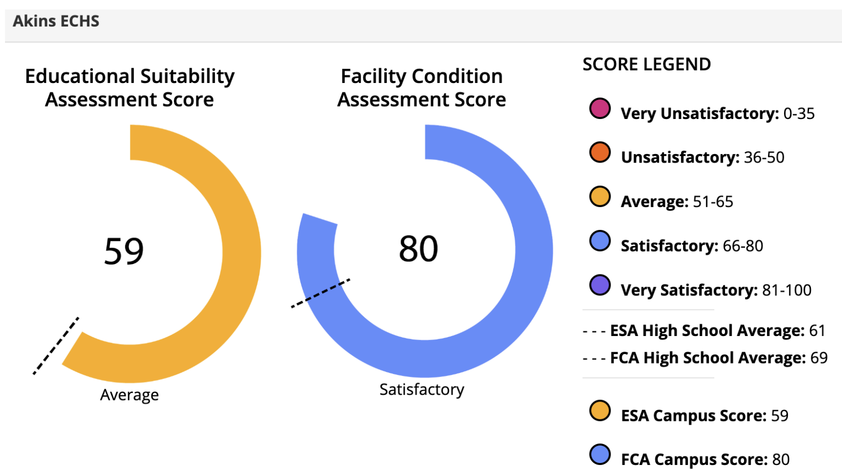
Data Set and Preprocessing Pipeline (with a focus on Austin ISD High Schools)

1. **Tabular model (scikit-learn + XGBoost)**
    a. Input: schools_tab numeric/categorical features (NCES + ACS).
    b. Model:
        i. baseline: RandomForestRegressor
        ii. strong: XGBoostRegressor
    c. Task: predict continuous risk_score (regression).
    d. Metrics: RMSE, MAE, R^2; maybe compare to simple baseline (mean predictor).
        i. AISD shapefile (school locations)
            1. https://www.austinisd.org/planning-asset-management/school-maps-gis
            2. Austin ISD provides their school-location data in standard GIS formats such as:
                a. .shp
                b. .dbf
                c. .shx
                d. .prj
            3. Will need to extract
                a. School names
                b. Addresses
                c. Type (High School)
        ii. NCES Demographics
            e. CSV file with data types
                i. Directory → School name, address, IDs
                ii. Membership → Enrollment + race
                iii. Staff → Teacher FTE
                iv. School Characteristics → Locale, Title I, grade span
                v. Lunch Program Eligibility → Free/reduced lunch counts

SCHOOL_YEAR	FIPST	STATENAME	ST	SCH_NAME	LEA_NAME	STATE_AGENCY_NO	UNION	ST_LEAID	LEAID	ST_SCHID	NCESSCH	SCHID	MSTREET1	MSTREET2	MSTREET3	MCITY	MSTATE
2023-2024	1	ALABAMA	AL	Albertville Middle School	Albertville City	1		AL-101	100005	AL-101-0010	10000500870	100870	600 E Alabama Ave			Albertville	AL
2023-2024	1	ALABAMA	AL	Albertville High School	Albertville City	1		AL-101	100005	AL-101-0020	10000500871	100871	402 E McCord Ave			Albertville	AL
2023-2024	1	ALABAMA	AL	Albertville Intermediate School	Albertville City	1		AL-101	100005	AL-101-0110	10000500879	100879	901 W McKinney Ave			Albertville	AL
2023-2024	1	ALABAMA	AL	Albertville Elementary School	Albertville City	1		AL-101	100005	AL-101-0200	10000500889	100889	145 West End Drive			Albertville	AL
2023-2024	1	ALABAMA	AL	Albertville Kindergarten and PreK	Albertville City	1		AL-101	100005	AL-101-0035	10000501616	101616	257 Country Club Rd			Albertville	AL

2. **Text model (HuggingFace Transformers)**
    a. Using raw text per school and ESA/FCA categorical labels fed into a BERT-based classifier. Performance is evaluated with precision and recall
    b. AISD Facility Assessment PDFs
        i. Extract text for narrative per school- > tokenize using BERT's tokenizer

8 - Academics & Learning (cont.)		1	2	3	4	5	N/A	Comments		
8.6	Is there a Professional Learning Center in the learning neighborhood?	1						Comments	1.00	**1.00**
1	Very unsatisfactory - No dedicated space for Professional Learning Center							None		
2	Unsatisfactory - Too small, poor condition, disconnected									
3	Average - Adequate size, adequate condition, needs some renovations or updates to function well									
4	Satisfactory - Good size, good condition, connected									
5	Very Satisfactory - Generous size, excellent condition, connected									
8.7	Are there adequate places to display the evidence and artifacts of learning (both 2D and 3D)?			3				Comments	1.00	**3.00**
1	Very unsatisfactory - No dedicated space							Adequate display in fine arts building. Some display at main entry in main building.		
2	Unsatisfactory - Too small, non-dedicated, poor condition									
3	Average - Adequate size, adequate condition, needs updates or renovations to function well									
4	Satisfactory - Good size, good condition									
5	Very Satisfactory - Generous size, excellent condition									
N/A	Does not apply									
8.8	Does the building / learning neighborhood provide adequate opportunities for students/teachers to collaborate in varying group sizes?		2					Comments	1.00	**2.00**
1	Very unsatisfactory - No space to support collaboration in various sizes.							Limited to no collaboration areas within the building. Outside courtyard areas have the opportunity to provide additional collaboration/informal meeting space, though it's lacking proper outdoor furniture and shade.		
2	Unsatisfactory - Too small, inflexible, only one option, does not support a variety of group sizes									
3	Average - Adequate space for collaboration, needs updates or renovations to function well.									
4	Satisfactory - Good size, good condition, multiple options to support a variety of group sizes									
5	Very Satisfactory - Generous size, excellent condition, multiple options to support a variety of group sizes									
N/A	Does not apply									
8.9	Does the building provide opportunities for informal interaction?		2					Comments	1.00	**2.00**
1	Very unsatisfactory - No space							Space in certain locations for informal interaction but furniture upgrades required to function well.		
2	Unsatisfactory - Too small, poor condition, layout and furnishings do not support informal interaction well									
3	Average - Adequate size and condition, needs updates or renovations to function well.									
4	Satisfactory - Good size, good condition, layout and furnishings support informal interactions									
5	Very Satisfactory - Generous size, excellent condition, layout and furnishings support informal interactions									



ii.

iii. ESA → How well the educational spaces support learning

iv. FCA → The physical condition of the building

v. Primary labels (or secondary proxy labels) for training both the text and vision models.

1. Each campus has associated ESA and FCA scores (0–100)
    a. discretize into five ordered categories (Very Unsatisfactory, Unsatisfactory, Average, Satisfactory, Very Satisfactory)

using the district's published thresholds.

Score	Label
0–35	Very Unsatisfactory
36–50	Unsatisfactory
51–65	Average
66–80	Satisfactory
81–100	Very Satisfactory

More Followup Papers/Notes/Thoughts/Ideas…

PyTorch, torchvision, HuggingFace Transformers, scikit-learn, and XGBoost/Random Forest.

- Vision model → F1 score on condition categories
- Tabular model → evaluating regression using RMSE / MAE
- Text model → Precision/recall on extracted issues

Fusion model → Overall calibration + correlation with known district risk metrics

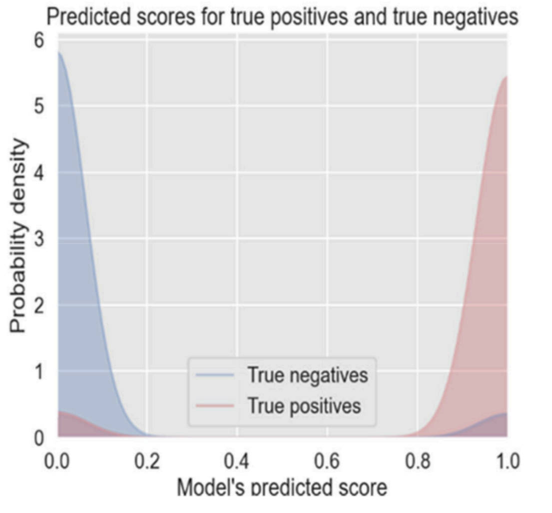
Proxy labels from:

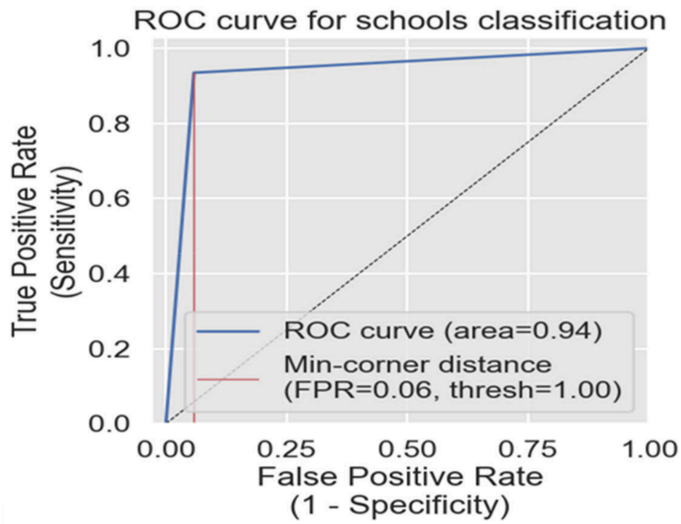
- Facility condition ratings
- Funding percentile rankings
- District quality reports
    - These serve as supervised targets.

Evaluate performance variance across:

- Income brackets
- Urban vs rural schools
- Racial demographics to ensure no systematic disadvantage.

Example from paper





Dataset

- NCES data, and a large collection of inspection reports depending on district availability

  Example from paper

Preprocessing Pipeline:

1. Clean and normalize tabular data
2. Download and filter school images
3. Extract text from PDF reports
4. Align modalities by school ID
5. Train models individually
6. Combine in the fusion layer

*Biggest bottleneck: Collecting and cleaning multi-modal data, especially text reports, which vary by district and format.

https://pmc.ncbi.nlm.nih.gov/articles/PMC10893965

https://www.mdpi.com/2072-4292/14/4/897

https://docs.edtechhub.org/lib/VINQBTJ5

https://www.sciencedirect.com/science/article/pii/S1226798825003186

school_id	NCES: poverty_rate	NCES: teacher_ratio	geolocation: median_income	ESA_class
0	54%	15.9	58,000	2
1	72%	17.8	42,000	1
2	18%	14.1	110,000	4

NCES Demographics →
Geolocation Data → Model → Predict ESA/FCA (label)
Images (Vision) →
Facility Text →

1. geographic data and aerial imagery
2. NCES demographic variables
3. publicly available facility assessment text
4. authoritative campus-level condition labels (ESA/FCA).
5. T**he only ground-truth labels used for supervised learning are the Educational Suitability Assessment (ESA) and Facility Condition Assessment (FCA) scores provided by Austin ISD**.
    a. All other data modalities serve as *input features* describing each school.

1. Label Construction (Ground-Truth Targets)

Each Austin ISD school receives two district-generated condition metrics:

- **ESA Score (0–100):** measures how well learning spaces support educational programs
- **FCA Score (0–100):** measures the physical condition of the facility

These ESA/FCA categories are the only supervised labels used across all models. Every data modality associated with a school inherits its ESA/FCA labels via school_id. We map each school's ESA/FCA score into a class label: ESA_class, FCA_class $\in$ {0, 1, 2, 3, 4}

2. NCES Demographic Data (Tabular Features)

NCES provides school- and district-level demographic variables, including:

- Percent economically disadvantaged
- Racial/ethnic composition
- Enrollment
- Student-teacher ratio
- Percent English Learners
- Percent students with disabilities

These variables are **not labels**. They function as **predictor features** in the tabular model and in the final fusion model.

Preprocessing:

1. Download NCES Non-Fiscal School Universe files.
2. Filter to Texas and match records by NCES School ID.
3. Standardize numeric features using z-score normalization.
4. Join with ESA/FCA labels using school_id.

3. Geolocation and Aerial Imagery (Vision Features)

Each school's latitude/longitude is obtained from the Austin ISD facilities shapefile.

Aerial Imagery Collection:

- Use Google Static Maps API or Mapbox to retrieve overhead imagery centered on each campus.
- Retrieve multiple zoom levels (e.g., 17–20) to capture building footprint and surrounding context.

Preprocessing:

- Resize images to 224×224
- Normalize pixel values using ImageNet statistics
- Apply data augmentation (random crops, flips, color jitter)

All images belonging to a school inherit the school-level ESA/FCA label. These images constitute the vision dataset used to train a ResNet-18/EfficientNet-B0 classifier to predict FCA_class.

4. Facility Assessment Text (Text Features)

Austin ISD publishes facility assessment reports for each campus, including narrative descriptions of:

- Educational suitability findings
- Condition deficiencies (HVAC, roofing, ADA issues, safety issues)
- Overall campus context and modernization needs

Text Preprocessing:

1. Extract text from PDFs using pdfplumber with OCR fallback.
2. Remove headers/footers, normalize whitespace, and clean OCR artifacts.
3. Tokenize using BERT-base (max length 512).
4. Associate each document with the ESA/FCA class of its school.

These processed texts form the input to our BERT-based classifier, which predicts ESA_class and generates 768-dimensional text embeddings for fusion.

5. Unified Multimodal Dataset Structure

  - One shared supervisory signal (ESA/FCA)
  - Multiple input feature types feeding into the model

Example final merged row:

school_id	tabular_features	images	text_embedding	ESA_class	FCA_class
227901017	…35 features…	12 images	768-d vector	2	3