Absenteeism Prediction App

Target Users

Primary Users

The system is designed for non-technical HR coordinators, team leads, and operations managers who need quick estimates of absenteeism hours to plan staffing and support resources. These users typically:

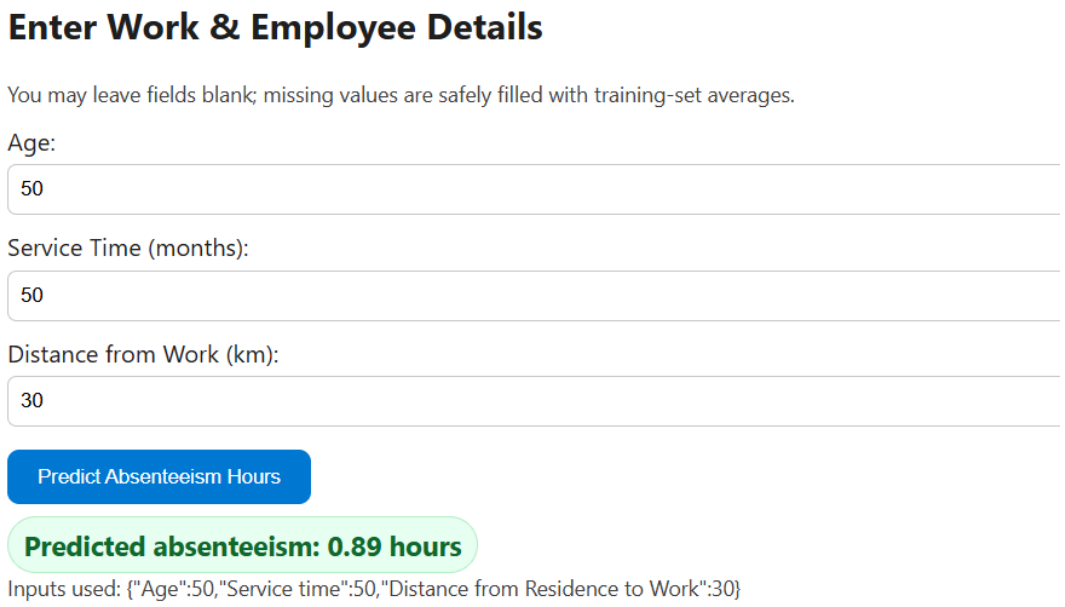
- Have limited ML background and need clarity on what the model can and cannot do
- Require fast, accessible predictions without complex forms
- Need to understand fairness implications to make responsible decisions
- Use predictions for planning purposes, not punitive actions

Secondary Users

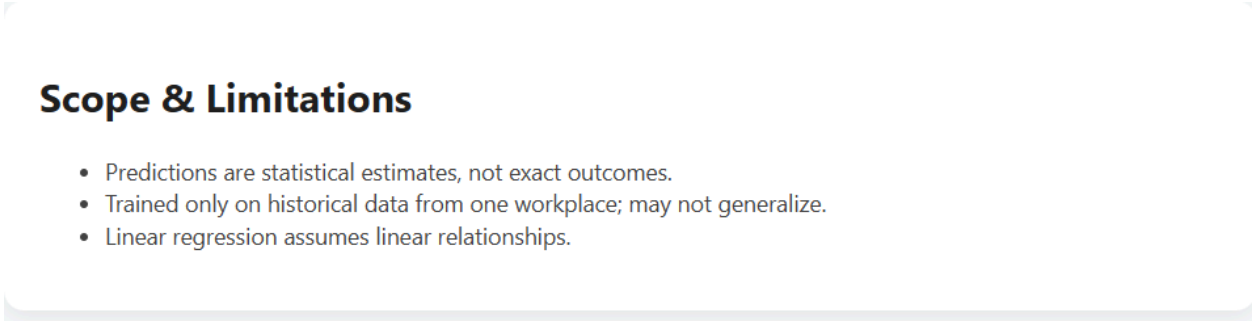
Instructors and students, learning applied ML, fairness, and responsible AI practices can use this as a teaching example that demonstrates transparency, fairness evaluation, and ethical considerations.

User Goals

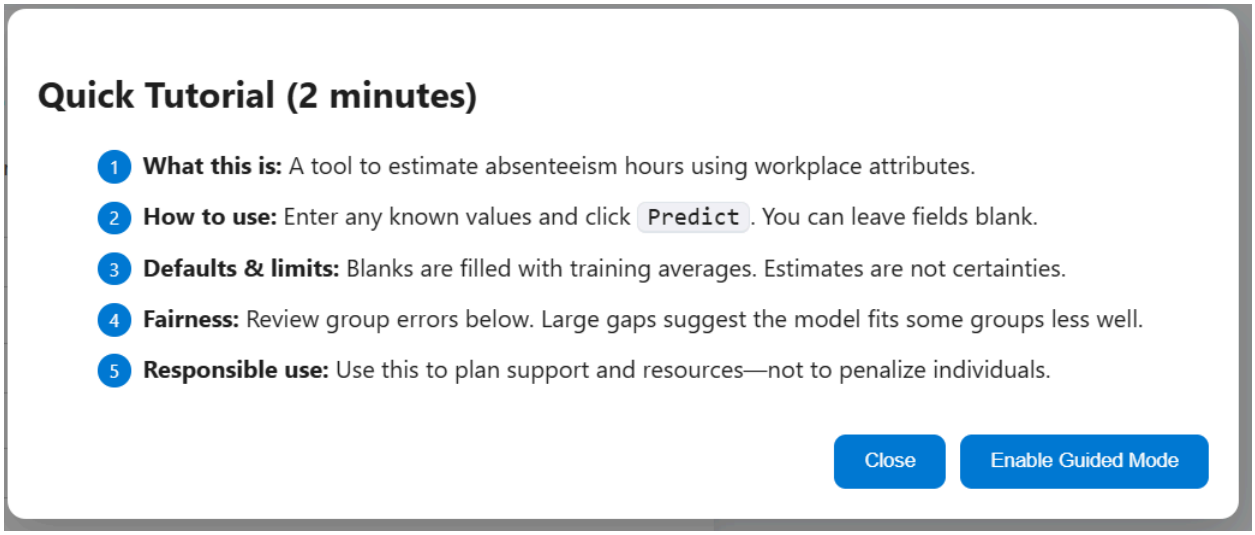
- The interface accepts partial inputs (Age, Service Time, Distance from Work). Users can leave any field blank, and the backend automatically fills missing values with training-set averages.



- Clear input fields with placeholder examples (e.g., "e.g. 35") and guidance text stating "You may leave fields blank; missing values are safely filled with training-set averages.
- A dedicated "Scope & Limitations" section lists: (1) predictions are statistical estimates, not exact outcomes; (2) trained on one workplace's data and may not generalize; (3) linear regression assumes linear relationships. A "How to Use & Interpret Results" section emphasizes using results for planning support, not punitive decisions.



- A "Fairness & Performance Overview" table shows average prediction error by age groups (18-30, 30-40, 40-50, 50-60). The section explicitly labels this as "Fairness (before mitigation)" and displays mitigation status (currently "None applied"). This transparency helps users spot potential disparities and understand that no bias mitigation has been implemented.
- A "Guided Mode" feature provides: A concise 5-step tutorial accessible via "Open tutorial" link explains: what the tool is, how to use it, defaults and limits, fairness, and responsible use.

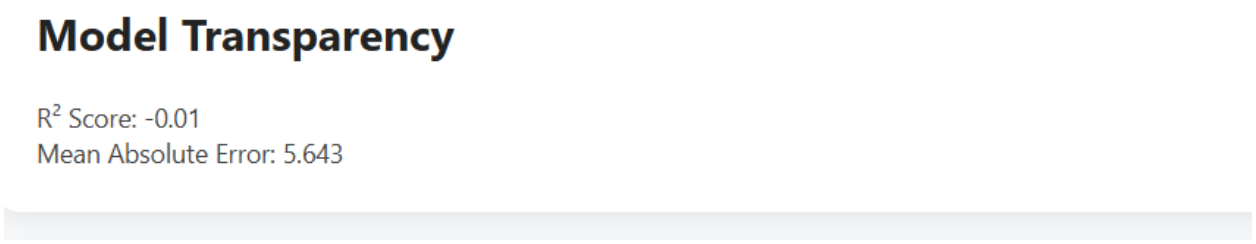


Prediction Output

- A loading spinner during the API call (visual feedback that processing is occurring)
- An animated green badge showing "Predicted absenteeism: X hours" where X counts up from 0 to the final value (draws attention to the result)
- The single numeric estimate (hours) clearly communicates the model's task: predicting a continuous value representing absenteeism time
- Predictions are automatically clipped to non-negative values (hours cannot be negative), which is enforced by the backend and reflects domain constraints

Model Performance

The "Model Transparency" section displays:
- **R^2 Score**: -0.01 (indicating the model explains almost none of the variance, signaling low predictive power)
- **Mean Absolute Error**: 5.643 (average absolute difference between predicted and actual hours)



Fairness

The "Fairness & Performance Overview" section includes:
- A table with columns "Age Group" and "Average Error" showing:
  - (18, 30]: 3.386
  - (30, 40]: 5.801
  - (40, 50]: 5.862
  - (50, 60]: 26.037
- A mitigation status line: What mitigation has been applied.

Fairness & Performance Overview

Fairness (before mitigation)

Age Group	Average Error
(18, 30]	3.386
(30, 40]	5.801
(40, 50]	5.862
(50, 60]	26.037

Mitigation: None (applied: no) — No fairness mitigation applied; values shown are baseline.

3. HAX-Based Evaluation Results

HAX Principles – Quick Evaluation

Principle	Status	Notes
Purpose & Fit	Meets	Shows goal, scope, limitations.
Data & Assumptions	Partial	Defaults shown; dataset source noted.
Performance	Meets	Shows MAE and R^2.
Fairness	Partial	Shows group errors; no mitigation yet.
Interpretability	Partial	Simple model; add more features if needed.
Control	Meets	Optional inputs; safe defaults.