## NAME

xai.py – explainable multi-objective optimization using decision trees

## SYNOPSIS

**./xai.py** [*OPTIONS*] [*ACTION*] [*FILE*]

## DESCRIPTION

**xai.py** is a lightweight, zero-dependency Python framework for multi-objective optimization. Unlike "black box" AI models, it focuses on interpretability by generating decision trees that isolate the "best" outcomes ("heaven") from the rest. It uses a recursive partitioning algorithm ("backpacking" code style) to cluster data, maximizing the separation between high-scoring and low-scoring rows based on weighted Euclidean distance to specific goals.

## DATA FORMAT

The input must be a CSV file. The first row (header) defines column roles using casing and suffixes:

**[A-Z]* (Starts with Uppercase)**
Numeric columns (e.g., "Age", "Weight").

**[a-z]* (Starts with Lowercase)**
Symbolic/Categorical columns (e.g., "job", "color").

**name+ (Ends with +)**
A numeric goal to **maximize** (e.g., "Salary+").

**name- (Ends with -)**
A numeric goal to **minimize** (e.g., "Cost-").

**nameX (Ends with X)**
A column to **ignore** (e.g., "idX").

**?** Denotes a missing value in the data.

## OPTIONS

Configuration flags control the internal algorithms.

**–b** *bins* Set the number of bins for discretizing continuous values. Default is **7**.

**–B** *Budget*
Set the max number of rows to evaluate during optimization. Default is **30**.

**–C** *Check*
Set the number of guesses to check during stochastic search. Default is **5**.

**–d** *file* Set the default data file to load. Default is **data.csv**.

**–h** Print the help message and exit.

**–l** *leaf* Set the minimum size for leaf nodes. The tree stops growing if a node has fewer than *2 * leaf* rows. Default is **2**.

**–s** *seed* Set the random number generator seed. Default is **1**.

## ACTIONS

Action flags trigger specific analysis routines.

**––all** Run all test suites and examples.

**––bins** Show the rankings of bins found in the data.

**−−clone**
Test internal table replication logic.

**−−csv** Test the CSV loader; prints raw parsed rows.

**−−data** Test column generation; prints column statistics.

**−−distx**
Sort/print rows by distance to a specific row (independent vars).

**−−disty**
Sort/print rows by distance to "heaven" (dependent goals).

**−−num** Run unit tests for the **Num** class.

**−−sym** Run unit tests for the **Sym** class.

**−−tree** Run the decision tree optimizer once and display the rule hierarchy.

**−−xais** Run the optimizer 20 times and display stability statistics.

## EXAMPLES

**Generate a decision tree for a custom dataset:**

```
./xai.py −−tree data/auto93.csv
```

**Run stability analysis with a specific seed:**

```
./xai.py −s 12345 −−xais
```

**Debug data parsing:**

```
./xai.py −−csv data.csv
```

## AUTHOR

Written by Tim Menzies (2025).

## LICENSE

MIT License.