

# MLighter

Reference manual

## Record of revisions

[illegible]

## Table of contents

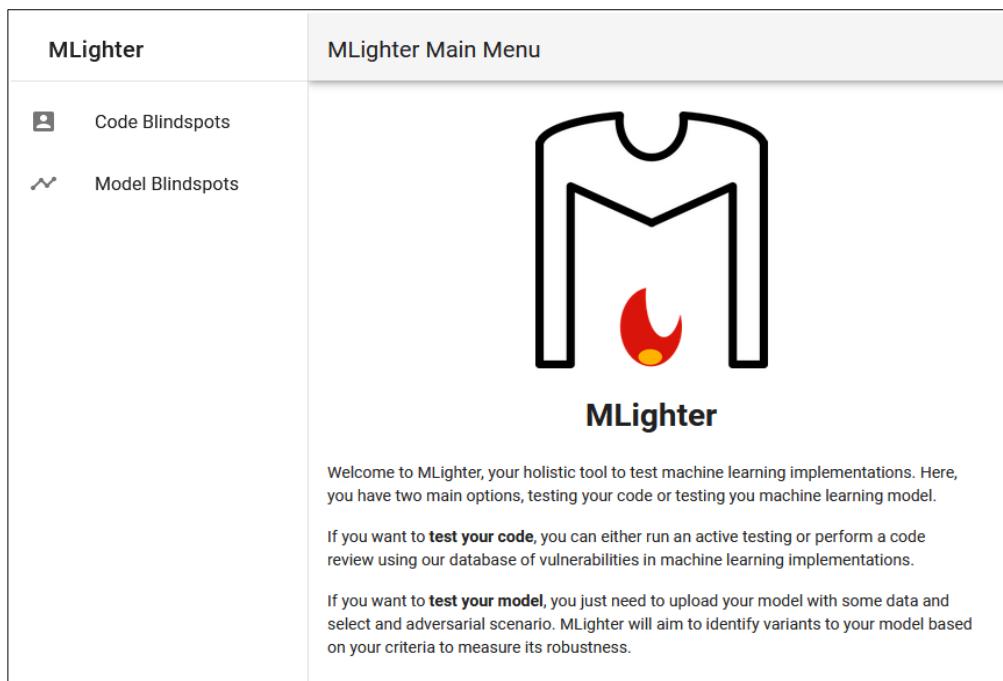
Introduction.....	4
User Interface.....	5
Code Blindspots.....	6
Code testing.....	6
Code Review.....	7
Bug Report.....	8
Model Blindspots.....	9
Dataset.....	9
Model.....	9
Evasion.....	10
Run.....	11
Report.....	12

## Introduction

Welcome to the manual for MLighter, your holistic tool to test machine learning implementations.

There are two different options for testing:

- **Code Blindspots** for testing your code.
- **Model Blindspots** for testing your model.

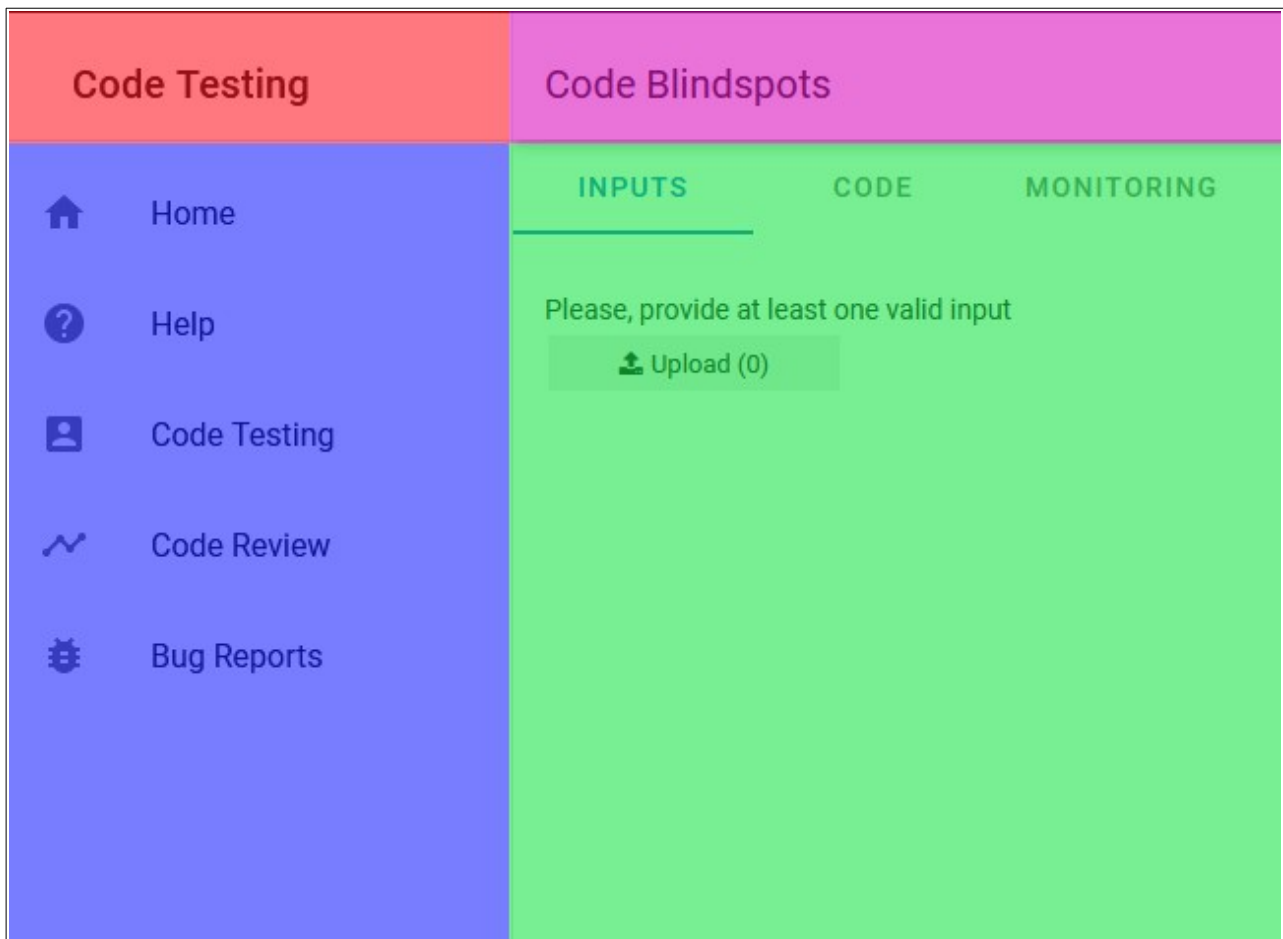


*MLighter main menu*

## User Interface

After connect with Mlighter servers the main window is showed in your browser. There are four main separated areas:

- The **red area** is the tab option title, where the selected tab option is showed
- The **Magenta area** is the main option title where the main option (code or model blindspot) is showed
- The **blue area** is the side bar where the main options are showed
- The **green area** is the tab and working area where the option tabs and upload options are showed.



*Main user interface*

## Code Blindspots

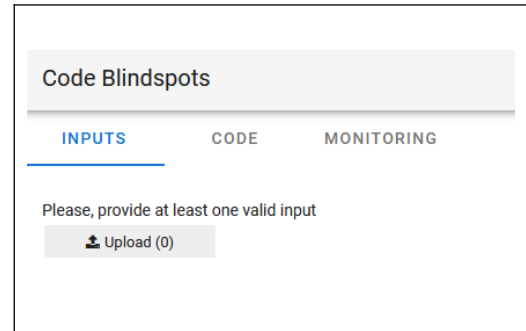
**Code Blindspots** allows you test your code or review your code using our database of vulnerabilities.

### Code testing

In the *Inputs tab* upload your inputs. The input should be **??**

In the *Code tab* upload your code. The code should be in **.CVS** and press start testing

The *monitoring tab*



The screenshot shows the 'Code Blindspots' application interface. At the top, there's a header 'Code Blindspots'. Below it, there are three tabs: 'INPUTS' (which is active and underlined), 'CODE', and 'MONITORING'. Under the 'INPUTS' tab, there is a message 'Please, provide at least one valid input' and a button labeled 'Upload (0)' with a small upload icon.

## Code Review

For reviewing your code, upload your code. The code should be in **.CSV** and press review code.

Upload (0)

REVIEW CODELOAD REPORTS

No data available

Rows per page: 10 - < >

## Bug Report

In the bug report option you can load the bug find in your code or select one from our database in order to create a report for it.

First you choose the bug in *load bug tab* and get the report in the *Bug report tab*.

Code Blindspots

LOAD BUGS

BUG REPORT

Select Bug From Database

None

Select Bug From Tester

None



## Model Blindspots

**Model Blindspots** allows you to test your model against a selected adversarial scenario. MLighter will aim to identify variants to your model based on your criteria to measure its robustness.


### Dataset

Select data set in the side bar menu For testing your code first upload all the datasets that you want to test. The code should be in .CSV

In the *Clean features tab*, you can refine the test removing those classes that are not relevant for the model or you don't want to test.

In the *Select class tab* you choose the feature that you want to predict such as target, etc.

Please provide at least one sample from your dataset (including headers)

 Upload (1)

Unnamed: 0	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	target
0	5.1	3.5	1.4	0.2	0
1	4.9	3	1.4	0.2	0
2	4.7	3.2	1.3	0.2	0

### Model


In the *model tab*, upload the model that you want to test. The models should be compatibles with SkLearn.

In the *pre evaluation tab*, you can run the pre evaluation test and check the results.

**MODEL**    PREEVALUATION

---

Please provide the model you want to evaluate

 Upload (1)

Model Ready

## Evasion

In *Evasion tab*, you can choose the evasion strategies as well as the type of evasion.

Only Random noise works in the Lite version.

This mutation adds a random permutation of the selected features.

Choose between discrete which allows you use integers numbers and continuous which allows floating points numbers.

How many variants you want to generate per provided input.

Noise is the level of perturbation

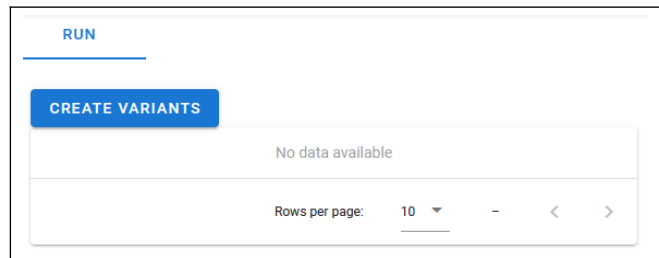
Shift is where you centred your noise.

In the *choose features tab* you choose the features where the perturbation will be applied.

The screenshot displays the 'Evasion' configuration interface. It features two tabs: 'EVASION' (active) and 'CHOOSE FEATURES'. Under the 'EVASION' tab, there are two dropdown menus: 'Evasion Strategy' set to 'Random Noise' and 'Type of Evasion' set to 'Discrete'. Below these, a section titled 'Select the number of variants per input:' contains three sliders: 'Noise' (set to 1), 'Shift' (set to 0), and another unlabeled slider (set to 0).

## Run

The *Run tab* allows you run the test with the previous parameters in order to generate the evasive variants.

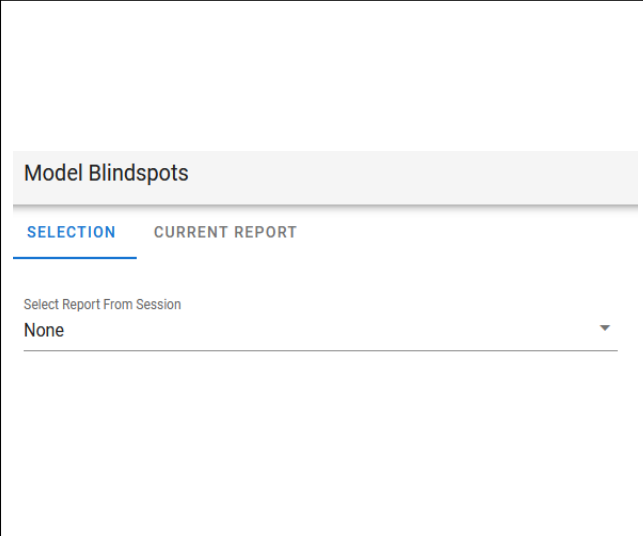


## Report

The Report option allows you get information about model blindspot findings.

In the *Selection tab* you can select between the different running scenarios of the current session.

In the *Current Report tab* you can retrieve the reports of the selected scenario.



The screenshot shows a web interface titled "Model Blindspots". It features two tabs: "SELECTION" (which is active and underlined in blue) and "CURRENT REPORT". Below the tabs, there is a label "Select Report From Session" followed by a dropdown menu. The dropdown menu is currently open, showing the option "None" with a downward-pointing arrow to its right.