

User Guide for Local Usage of MLAI

V 1.0

PREPARED BY: James Anibal, Isaac Withrow, Alec Peltekian, Ian Post, Long Phan, and Daniel Le

We released this version because of server performance issues and hacking, should the grader want to test large directories.

SETUP:

Downloading the files:

Run the command:

```
git clone https://github.com/janibal54/MLAI-EECS393.git
```

Then, run the command:

```
cd MLAI-EECS393
```

Installations:

Docker:

Docker for Windows: <https://docs.docker.com/docker-for-windows/>

Docker for Mac: <https://docs.docker.com/docker-for-mac/install/>

Node.js:

<https://nodejs.org/en/download/>

MongoDB:

This command will import the user database:

```
docker-compose up -d
```

Python Dependencies:

```
pip3 install --no-cache-dir -r requirements.txt
```

```
python3 -m spacy download en_core_web_lg
```

To run MLAI

First, run the command: `python3 app.py`

Next, open a second terminal and run the following commands:

```
cd MLAI-EECS393
cd templates
cd mlai
```

Third, run: `npm install` in the second terminal window

Finally, run: `npm start` in the second terminal window

You can now run MLAI locally without any concerns regarding our cloud server:

Our software product has two main use cases, which are regular users and developers. The sign-up and login process for each is the same.

1. Go to <http://34.70.151.69:3000/> (We host MLAI on the Google Cloud). This will bring you to the default user sign in page (Fig. 1, left). If you already have a user account you can use your credentials to sign in to the application here. If you want to sign up for a user account, there is a button that will take you to a form to do so. Once you have created an account, You will then be taken back to the main login page and will be able to log in using the same credentials to the user login.

2. If you want to sign in as a developer and are on the user sign in page, you must click "Sign in as a developer". This will take you to the developer sign-in page (Fig. 1, right). You then are given the option to sign in as a developer or to create a developer account. This process is similar to the user sign-up/sign in process. Now that we have discussed how to create an account and login, we'll go over how to leverage the functionality and features our software provides.

The image shows two side-by-side login forms. The left form is titled 'Sign in' and features a red icon of a computer monitor. It has input fields for 'username *' and 'Password *', a 'SIGN IN' button, and links for 'Don't have an account? Sign Up' and 'Sign in as a developer'. The right form is titled 'Developer Sign in' and features a red icon of a person. It has input fields for 'devname *' and 'Password *', a 'Sign In' button, and links for 'Create new Developer Account' and 'Sign in as an user'. Both forms have 'EECS 393 Spring 2020' at the bottom.

Figure 1: User and Developer Login Options

User Guide

The image shows the main interface for users. On the left is a navigation menu with the title 'Isaac Withrow' and links for 'Upload New Data', 'Previous Models', 'Account', and 'Settings'. The main area has a header with a blue folder icon and the text 'Click or drag file to this area to upload' and 'Support for a single or bulk upload. Strictly prohibited from uploading company data or other banned files'. Below this is a text prompt: 'Or, if you prefer to upload an entire directory, go below.' followed by an 'Upload Directory' button and a 'Start Upload' button.

Figure 2: Main interface for users

As a user, you will be able to upload either a file or directory of files that you want to run our software on. After uploading the data, you will be prompted to fill out more information about the type of data you are uploading. **MLAI accepts .xlsx, .csv, .jpg, and .png files.**

For your convenience, we included several datasets and the descriptions for each in the folder entitled *SampleData_LOCAL*, which contains a ReadME.

After uploading, you will be prompted for additional details, a time constraint, and a file of labels for your data (if applicable).

Additional Details: This is where you describe your data so that MLAI can choose the algorithms that will be applied. Be as detailed as possible - the more specific you are, the more targeted the algorithm selection will be. We provided some descriptions

Time Constraint: You will have 5 options

- 1 - As fast as possible, a single train/test split, and **no deep learning**
- 2 - Two train test splits, deep learning with low-dimensionality parameters
- 3 - Cross-validated approach, deep learning with moderate-dimensionality parameters
- 4 - Minimal Grid search, cross validation, deep learning with moderate-dimensionality parameters
- 5 - Full grid search, cross validation, optimized deep learning parameters (this takes a LONG time)

Labels:

If your data has labels and you want supervised analysis, you have several options:

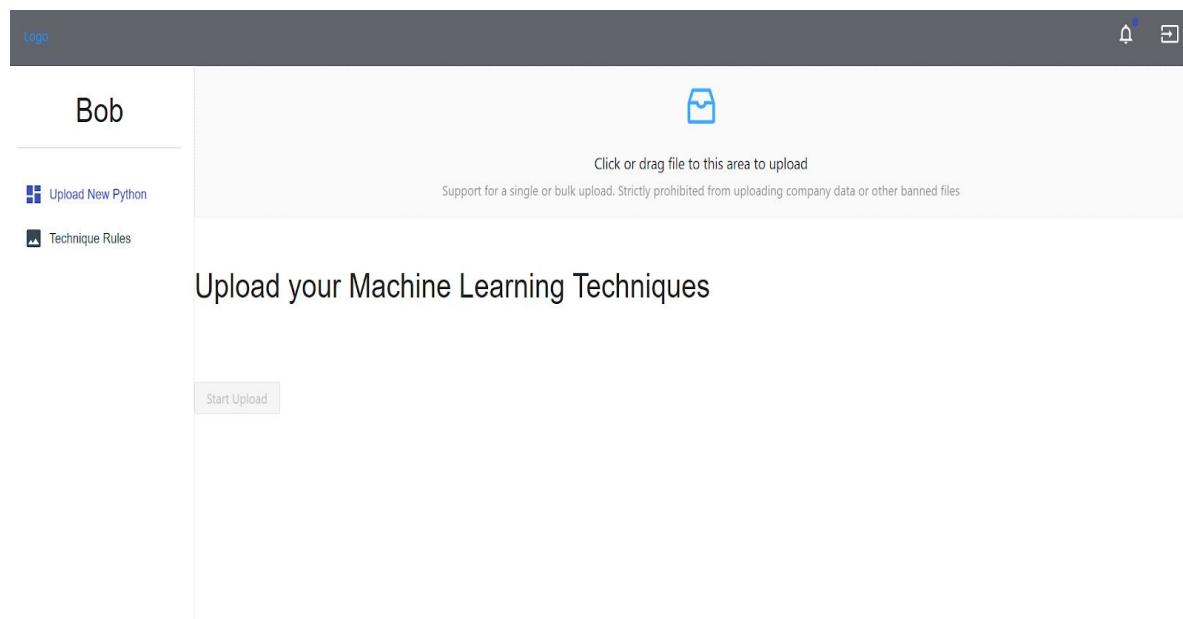
1. If your data is a .csv/.xlsx file, the labels can be in the file itself, in a column named "labels" or "label"
2. If you upload a directory of data files, each corresponding to a sample you want analyzed, you can have a separate file with the labels and the names of the corresponding file (i.e. labels.csv). You can also do this for single files if necessary. In this case, each index of the label file must correspond to an index in the data file.

Blind Test Data: If you want to use prior data to predict new unknown samples, upload your training data, your labels, and the test data. The test data **MUST** be in a file with the word test in the filename (i.e. testfile.csv). **After you click submit and the models are trained and analyzed, you will be brought to a results page that describes the best model for your data along with an accuracy rate (or evaluation metric).**

Additional Notes: We built in some protections to prevent MLAI from returning invalid or misleading results.

- If your data has labels but all the labels are the same value, the analysis type will be unsupervised - otherwise there would be a misleading accuracy value of 100 for predictions
- If your training data has 10 or fewer samples, we return no results, as the model would be extremely overfit and the results would be invalid
- If your description does not provide any usable keywords, MLAI will choose two very common approaches to fit to your data (i.e. SVM)

Developer Guide



1. After login, the developer will be taken to the page shown in figure X and will be asked to upload a python file (ONLY .py files are accepted at this time).
2. Once you've selected the .py files you want, press "Start Upload." Below you see a "details" question, where you will be asked to upload a basic explanation of the projects so we can have information about this approach for our text mining/reinforcement learning algorithms.

3. After you click submit the file and the description will be sent to our server.

support for a single or bulk upload. Strictly prohibited from upload

More Details

* Details:

On the left hand side of the webpage, you can see some “rules” we have to help users make sure their file is correct. This of course will change as people start uploading actual techniques so we can make sure people better understand what we want. If you would like to log out of the developer page, click the button in the top right.