

## Project: PluMA 2.0 Machine Learning Pipeline Refactor

**Students:** Adrian Alpizar

**Mentor:** Dr. Trevor Cickovski, PhD, Florida International University, Bioinformatics Research Group (BioRG)

**Instructor/Faculty:** Dr. Masoud Sadjadi, PhD, Florida International University

### PROBLEM

In order to enable developers to easily implement a new algorithm in their preferred programming language, as well as test and debug it in a larger pipeline alongside stages in other languages that may use different file formats, PluMA aims to make the construction of flexible and lightweight analysis pipelines easier. This is done using PluMA's plugins, of which there are many in its plugin pool and which are implemented in different programming languages for both the CPU and GPU. Through our user interface, plugins may be quickly added, deleted, or changed. They can also be run consecutively to create a pipeline. You may build a pipeline by downloading PluMA and just the plugins you require from the pool since plugins are dynamically loaded.

### SYSTEM DESIGN AND VERIFICATION

We fulfilled every demand our product manager had set:

1. We were able to deconstruct a machine learning project, utilizing different methods to reconstruct and organize ML models into modules.

2. Modules were then annotated and verified for easy execution, something that was under constant testing so it could be versatile and used in possibly any plugin.

3. Modules were then organized into possible plugins, which allowed for interchangeable data and which were developed for easy access and input into any machine learning pipeline.

#### The PluMA Initiative: Plugin-Based Microbiome Analysis

PluMA is an effort to transform microbiome analysis from individual to community-driven.

**Milestone: PluMA Hits 500 Plugins**

CUDA Plugin Expansion with support from

**New Release: PluMA 2.0 Now Available**

**Emphasis: User Interface**

**Key Features:** Improved Compilation, Back End Redesign, Pipeline Visualizer, Beta GUI, Alpha Windows Version, 400+ Plugins)

| Machine Learning |                                                                                                 |          |
|------------------|-------------------------------------------------------------------------------------------------|----------|
| Name             | Short Description                                                                               | Language |
| BLSOM            | BLSOM unsupervised neural net (Nakao et al, 2013) (Generated by PluGen)                         | C++      |
| Cafta            | Deep learning through Convolutional Architecture for Fast Feature Embedding (Fajet et al, 2019) | Python   |
| CONV             | Convolutional Neural Network (Lawrence et al, 1996)                                             | Python   |
| DNN              | Deep Neural Network (Nakamoto, 2017)                                                            | R        |
| NaiveBayes       | Naive Bayes classifier (Webb et al, 2010)                                                       | Python   |
| NeuralNet        | Neural network for virus classification                                                         | R        |
| PoSVM            | Python Support Vector Machine (Hearst, 1998)                                                    | Python   |
| RandomForest     | Random Forest classifier (Liaw and Wiener, 2002)                                                | R        |
| RBM              | Restricted Boltzmann Machine (Smolensky et al, 1986)                                            | R        |
| RNN              | Recurrent Neural Network (Williams et al, 1986)                                                 | Python   |
| SoNMF            | Supervised Non-Negative Matrix Factorization (Cai et al, 2017)                                  | R        |
| SVM              | Support Vector Machine (Hearst, 1998)                                                           | R        |

Visitors since 10/5/2020:

#### Download PluMA 2.0

PluMA GitHub Site

PluMA Userguide

PluMA Plugin Pool

PluMA Pipeline Pool

#### PluMA Add-Ons

GUI

PipeViz

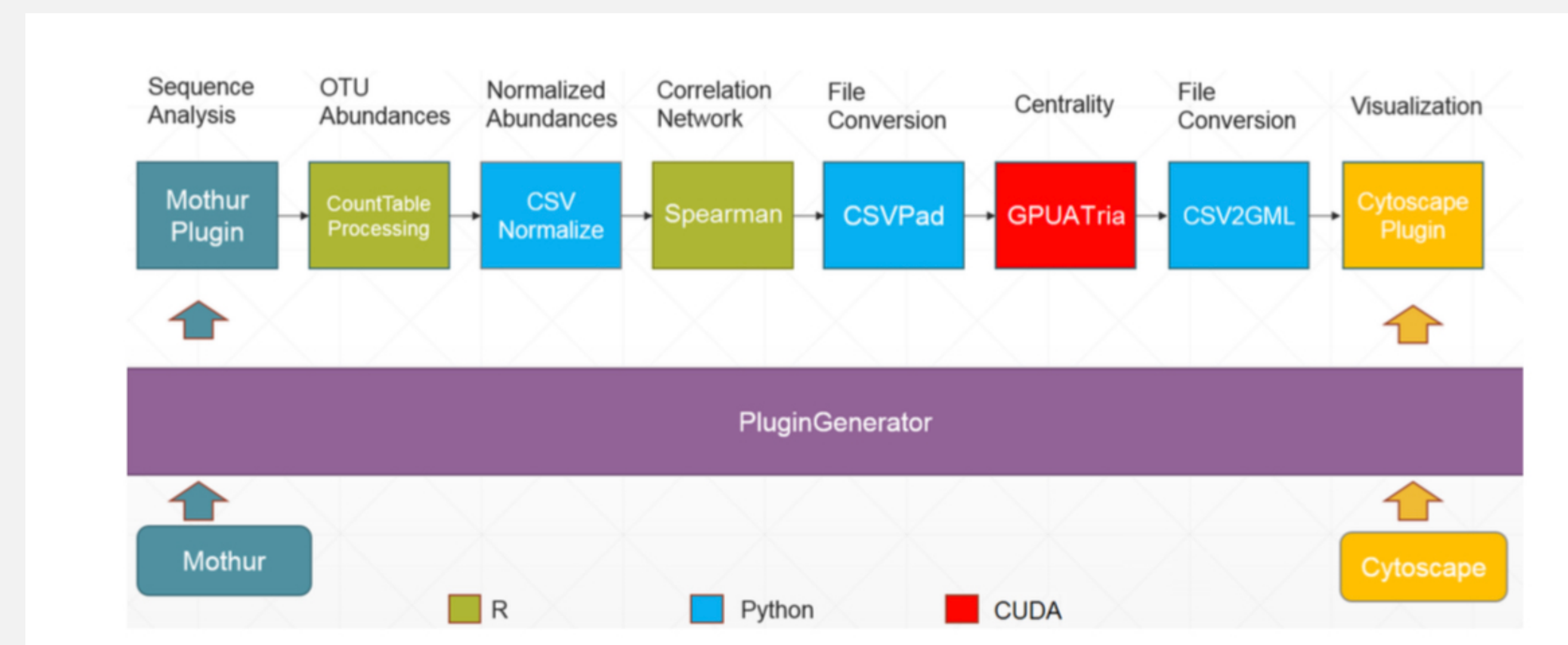
Upstream-Downstream

Windows

### CURRENT SYSTEM

The current system supports only 12 Machine Learning plugins, the goal is to increase the total plugin pool for Machine Learning projects to be tested quickly and efficiently. The current system lacks full pipeline functionality, with the addition of data extraction and importing, data preprocessing plugins, along with plugins for splitting and deploying ML models to test the data, users will have an entire beginning to end project they can deploy and test in minutes.

### OBJECT DESIGN



### SUMMARY

We achieved all the requirements:

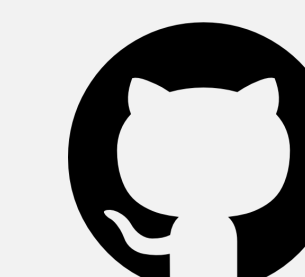
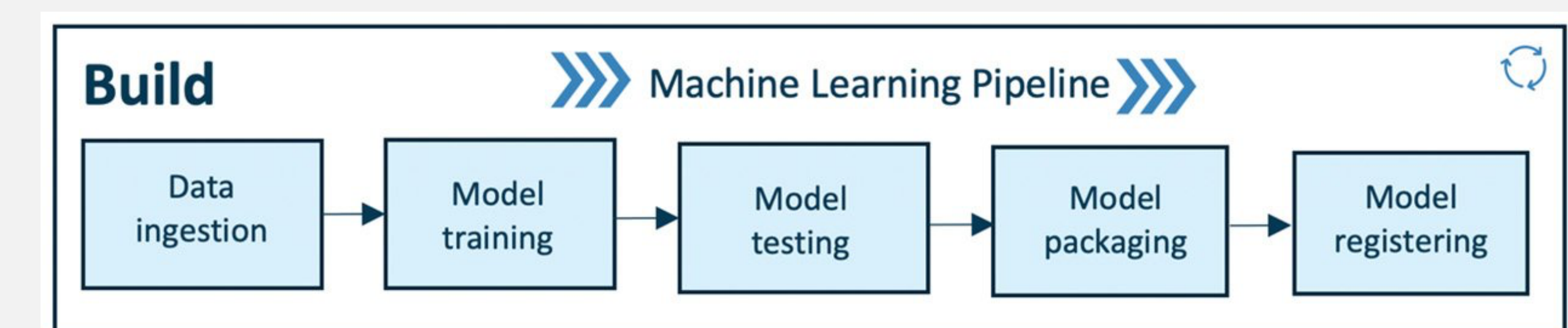
We were able to add around 30 possible plugins, many from different types of ML algorithms. Modularization of different machine learning project deployments, allowed for plugins that allow for quick data extraction, preprocessing, graphing, splitting, and finally testing for different values.

### REQUIREMENTS

As previously stated, the goal of PLUMA is to enable users to create and test unique algorithms as plugins in the languages of their choosing, and then to seamlessly link these plugins with other plugins in an inventive way to create pipelines and work as framework for building a variety of pipelines across several languages.

I was able to work on the modularization of various ML models, of which are now able to become plugins on their won for quick deployment.

### IMPLEMENTATION



<https://github.com/FIUBioRG/PluMA>

### REFERENCES

Principal Investigator: Prof. Giri Narasimhan  
Project Manager: Dr. Trevor Cickovski  
Senior Developer: Mr. Joseph Quinn

**ATTENTION:** Subheadings shown above is a recommended structure for a research/scientific poster. Main elements of a poster are Introduction, Research and Conclusion. Researchers can customize the section headers based on their projects' needs.  
**REMOVE THIS TEXT ONCE YOU START WORKING ON YOUR POSTER**

### ACKNOWLEDGEMENT

The material presented in this poster is based upon the work supported by XXXXX (name of the project sponsor: federal, state or local government, or corporate partners, where applicable). We/I thank XXXX for his/her/their assistance and mentorship that I/we received throughout the senior design project.