

Title: Waffles (machine learning)

URL: [https://en.wikipedia.org/wiki/Waffles_\(machine_learning\)](https://en.wikipedia.org/wiki/Waffles_(machine_learning))

PageID: 32867182

Categories: Category:Data mining and machine learning software, Category:Free data analysis software, Category:Free science software, Category:Software optimization

Source: Wikipedia (CC BY-SA 4.0). Content may require attribution.

Waffles is a collection of command-line tools for performing machine learning operations developed at Brigham Young University . These tools are written in C++ , and are available under the GNU Lesser General Public License .

Description

The Waffles machine learning toolkit contains command-line tools for performing various operations related to machine learning , data mining , and predictive modeling . The primary focus of Waffles is to provide tools that are simple to use in scripted experiments or processes. For example, the supervised learning algorithms included in Waffles are all designed to support multi-dimensional labels, classification and regression , automatically impute missing values, and automatically apply necessary filters to transform the data to a type that the algorithm can support, such that arbitrary learning algorithms can be used with arbitrary data sets. Many other machine learning toolkits provide similar functionality, but require the user to explicitly configure data filters and transformations to make it compatible with a particular learning algorithm. The algorithms provided in Waffles also have the ability to automatically tune their own parameters (with the cost of additional computational overhead).

Because Waffles is designed for script-ability, it deliberately avoids presenting its tools in a graphical environment. It does, however, include a graphical "wizard" tool that guides the user to generate a command that will perform a desired task. This wizard does not actually perform the operation, but requires the user to paste the command that it generates into a command terminal or a script. The idea motivating this design is to prevent the user from becoming "locked in" to a graphical interface.

All of the Waffles tools are implemented as thin wrappers around functionality in a C++ class library. This makes it possible to convert scripted processes into native applications with minimal effort.

Waffles was first released as an open source project in 2005. Since that time, it has been developed at Brigham Young University , with a new version having been released approximately every 6–9 months. Waffles is not an acronym—the toolkit was named after the food for historical reasons.

Advantages

Some of the advantages of Waffles in contrast with other popular open source machine learning toolkits include:

Waffles automatically takes care of many issues related to data format in order to simplify its tools.

Because it is implemented in C++, many of its algorithms are particularly fast. Also, the lack of dependency on any virtual machine makes it easier to deploy in conjunction with other applications.

The functionality included in Waffles is very broad, including algorithms for dimensionality reduction , collaborative filtering , visualization, clustering, supervised learning, optimization, linear algebra, data transformation, image and signal processing, policy learning, and sparse matrix operations.

Disadvantages

Although Waffles provides significant breadth, it lacks the depth of many toolkits that focus on a particular area of machine learning. The Weka (machine learning) toolkit, for example, provides many more classification algorithms than Waffles provides.

Waffles only has a limited graphical interface.

See also

Free and open-source software portal

Weka (machine learning)

RapidMiner (formerly YALE (Yet Another Learning Environment)), a commercial machine learning framework implemented in Java

List of numerical analysis software

References