Machine Learning has been added to Tiki23 as a built-in (but optional) feature, using the Rubix ML library, which provides 40+ supervised and unsupervised learning algorithms.

Introduction

In this basic documentation to machine learning feature in Tiki, you'll discover how to create Machine learning (ML) models, train against data stored in trackers and query models.

What is Machine Learning?

Machine Learning (ML) is a form of Artificial Intelligence (AI) that uses data to train a computer to perform tasks. Unlike traditional programming, in which rules are programmed explicitly, machine learning uses algorithms to build rulesets automatically. At a high level, machine learning is a collection of techniques borrowed from many disciplines including statistics, probability theory, and neuroscience combined with novel ideas for the purpose of gaining insight through data and computation.

Machine Learning is further broken down into subcategories based on how the learners are trained and the tasks they handle.
In this documentation we assume that you already have a basic understanding of the different types of machine learning such as classification and regression. If not, we recommend the section on What is Machine Learning? to start with.

What is Rubix ML?

Rubix ML is a high-level machine learning and deep learning library for the PHP language that includes implementations of several machine learning algorithms, so you can define a model object in a single line or a few lines of code, then use it to fit a set of points or predict a value. It is Open Source and free to use commercially.

The library provides tools for the entire machine learning life cycle from ETL (extracting, transforming, loading, manipulating and summarising data) to training, cross-validation, and production with over 40 supervised and unsupervised learning algorithms. A number of algorithms in the library support Deep Learning including the Multilayer Perceptron classifier and MLP Regressor.

This library has been added in tiki starting with version 23 in order to give Tiki users access to the power of artificial intelligence and machine learning.
How can this be used in Tiki?

It opens up so many possibilities for Tiki users:

- can be used to build machine learning pipelines, train classifiers, and run evaluations without having to write a single line of code.
- To get an idea of the diversity of the use cases, look at the "Tutorials and Example Projects" section here: [https://rubixml.com/](https://rubixml.com/) (Yes, we will now be able to do this in Tiki!)
- Machine Learning will be deployed all over Tiki to leverage both
  - system data (ex. logs of user activity) and
  - content managed by Tiki users, like Trackers and Webmail.

How to get started?

step 01: Enable ML feature in admin features.

![Admin Panel](https://example.com/admin_panel.png)

To continue to the next step you need also to have the Trackers feature enabled.
Check it is enabled on the same admin control panel page.

step 02: Give permissions to admin ML models to relevant users.
step 03: Preparing or Importing a Dataset using Tiki tracker

To import a dataset into tiki, we need to create a tiki tracker with the corresponding data fields then import the csv file.

The tracker will be used as our data set to create a machine learning model. we can even create a fake dataset (tracker) to train a model by populating fields and items manually.

so trackers are define as input sources - one or more tracker fields would be the feature set and item ID will be the label.
Suppose that you went out and asked 4 couples (2 married and 2 divorced) to respond to 3 features - their partner’s communication skills (between 1 and 5), attractiveness (between 1 and 5), and time spent together per week (hours per week). You would structure the data in tracker like in the example below.

step 04: Create a ML model

Go to the "List Models" under the Machine Learning menu to launch the tiki-ml-list page containing the list of models; from there you can click on the add button to launch the new model creation form.

- Input the name for your ML Model.
- Input a description of your ML Model.
- select the tracker that will be the input source dataset. (see step 03)
- Select a predefined ML model template designed to solve a common problem or start with a blank model to create Machine Learning models from scratch.
- select one or more relevant tracker fields that would be the dimensions fields. Depending on the model, MLT works best with one field at the moment.
- Check model creation process - you can choose from basically all Transformers from rubix base and extras packages.
- then specify a learner or estimator (regressor, classifier, etc.) then tweak the parameters and save.

step 05: Train the model.

Go to the "List Models" under the Machine Learning menu to launch the tiki-ml-list page containing the list of models; from there you can click on the train button for the corresponding model.
If data source is not big, you can train via web interface. Otherwise, it is better to use the new console command (ml:train) and do this with a scheduler - e.g. train every night.

```php
php console.php ml:train <mlmId>
```

Something important I see here is that trained models are stored as cached content on disk. Clearing Tiki cache will require to re-train the model.

**step 06: Making Predictions.**

Go to the "List Models" under the Machine Learning menu to launch the tiki-ml-list page containing the list of models; from there you can click on the use button to use the corresponding model.

Use this model by entering a sample information in the form this execute a query against the trained model. This will produce results based on the chosen estimator and show you the most relevant matches.
step 07: Model Evaluation

Go to the "List Models" under the Machine Learning menu to launch the tiki-ml-list page containing the list of models; from there you can click on the test button to evaluate the corresponding model. You can test models at every time against the real data. All exceptions/errors are shown during params tweaking and testing/training.

Others

1. Code added

- Tiki23 will be released in June 2021 and you can use and test it today via a Daily Build. After sufficient testing, the machine learning code could be backported to Tiki22 (ex: in 22.3 or 22.4) given that it's pretty self-contained (if you don't activate the feature, nothing changes)
- Here is some of the code:

1. feedback & Improvement

That's in short. I suppose you will have a lot of feedback, so curious to hear how to continue/improve here. One way to experiment is different source options - more than trackers, e.g. wiki pages, action log items, calendar items, files, etc. Files should probably give us the ability to upload a csv of a sample data to train against.

See also Machine Learning Models
Video about Rubix ML in English

Video about Rubix ML in French

Related links


other page names for this page

Rubix ML
RubixML