Introduction

In this basic documentation to machine learning feature in Tiki, you'll discover how to create Machine leaning (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 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/ (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.

Getting Started

To start using Machine Learning in Tiki, the Machine Learning feature must first be enabled in control panel.
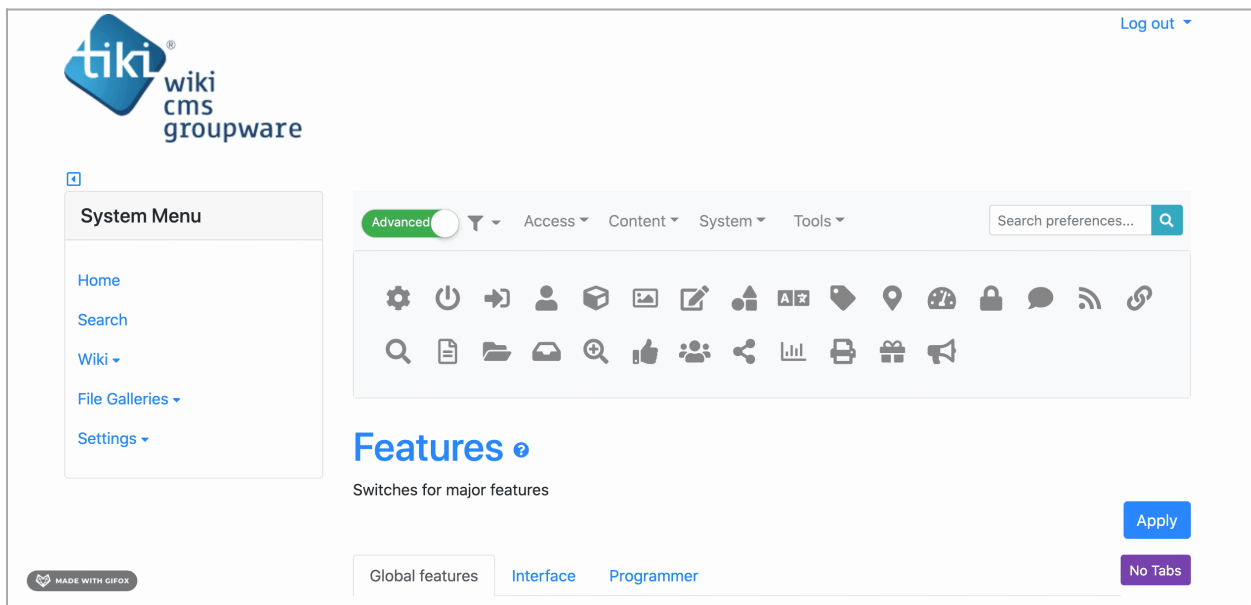
Machine Learning models in Tiki use data from Trackers to learn. This makes it compulsory to have the Trackers feature also enabled. Check it is enabled on the same admin control panel page.

Additionally, you might want only selected users to manage and use the Machine Learning models. In that case, you can give the permissions to the relevant users in the **User Settings** page.
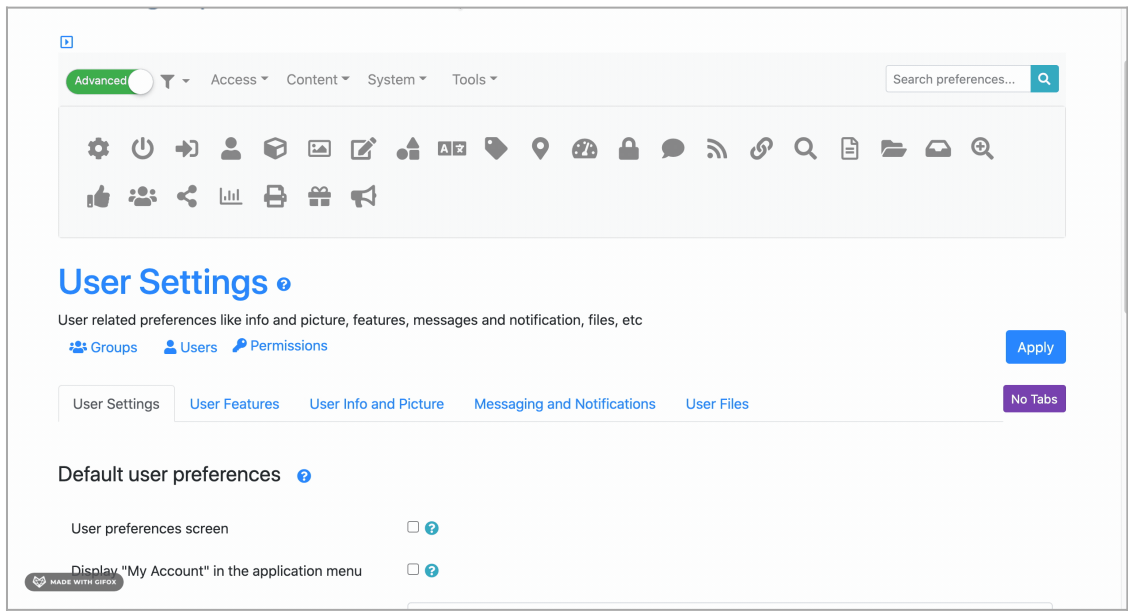
At this point, you are ready to start creating and using Machine Learning models in Tiki with only a few more steps ahead. The next steps being to prepare the training dataset and then creating the Machine Learning model. Model configuration and training steps come last.

A Machine Learning model after being trained can then be used to make predictions as described in Using Machine Learning Models.
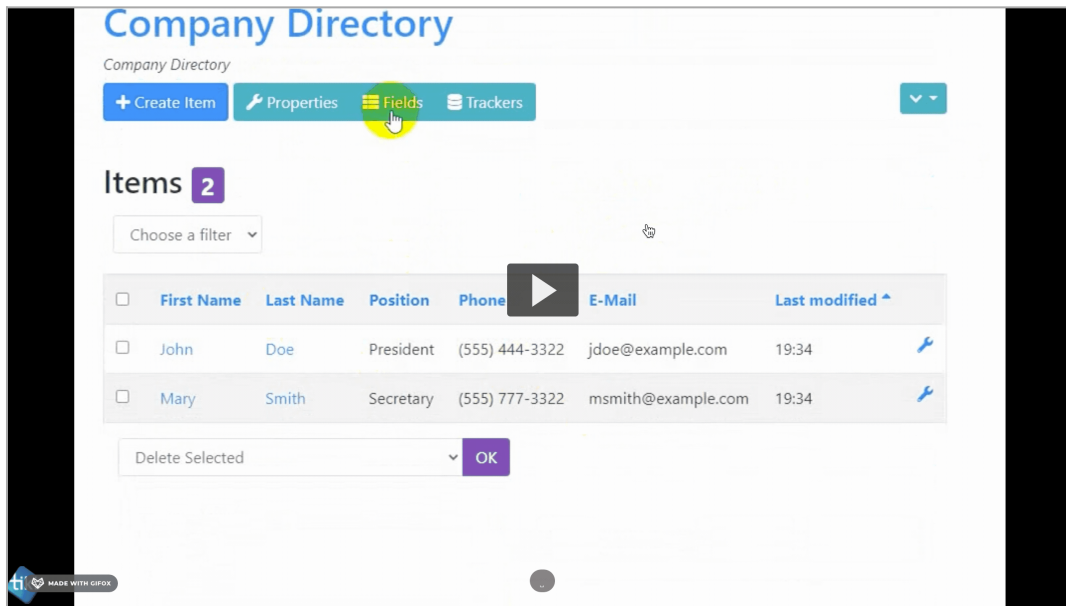
An Example

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 follow the following steps to setup a Machine Learning model in Tiki that will predict whether other couples will get divorced or remain married.

Step 01: Preparing or Importing the Dataset

Get the data you collected by asking the four couples in our example into a 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.

As a side note, we can even create a fake dataset (tracker) to train our 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.


Continuing with our example, structure the data gotten from the couples in a tracker like in the example below.
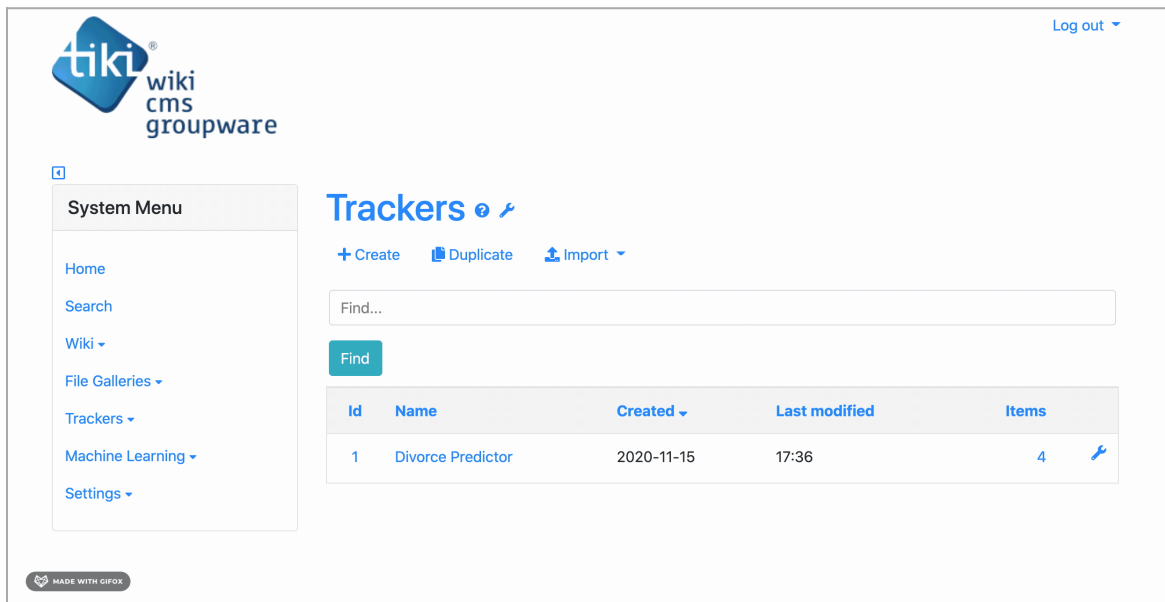
Step 02: Creating the Machine Learning model

Go to the **List Models** page 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 Machine Learning model.
- Input a description of your Machine Learning model.
- Select the tracker that will be the input source dataset. (see Step 01)
- Select a predefined model template designed to solve a common problem or start with a blank model to create Machine Learning models from scratch.
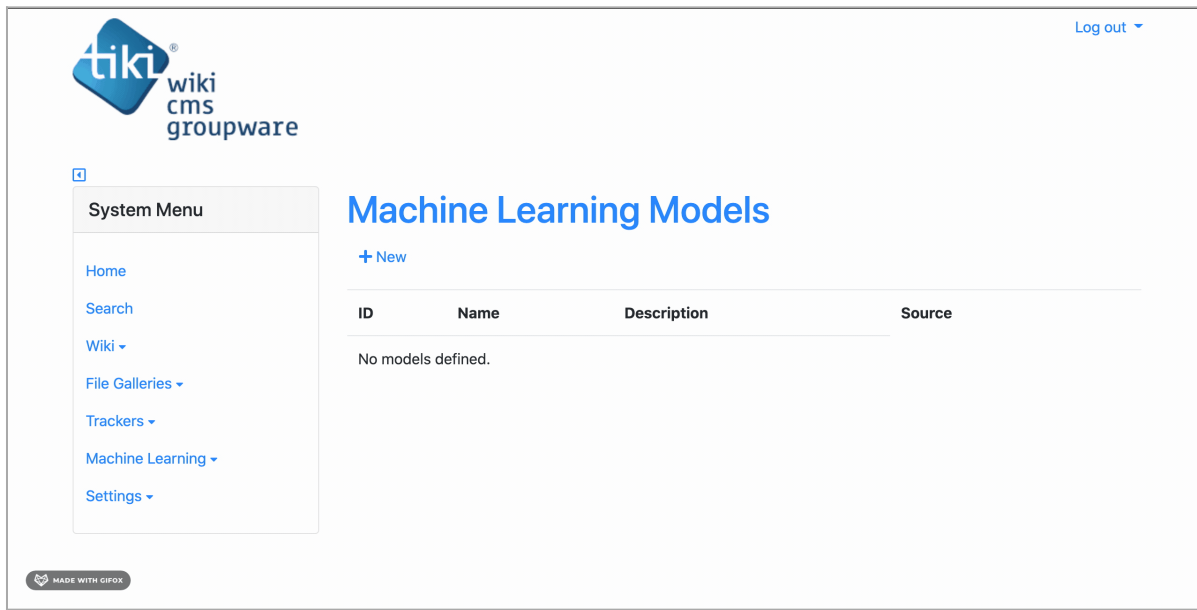
Step 03: Configuring the Machine Learning model

On the Machine Learning model **Edit Page**

- Select one or more relevant tracker fields that would be the dimensions fields. Depending on the model. More Like This (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 04: Training the Machine Learning model

Go to **List Models** page 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.
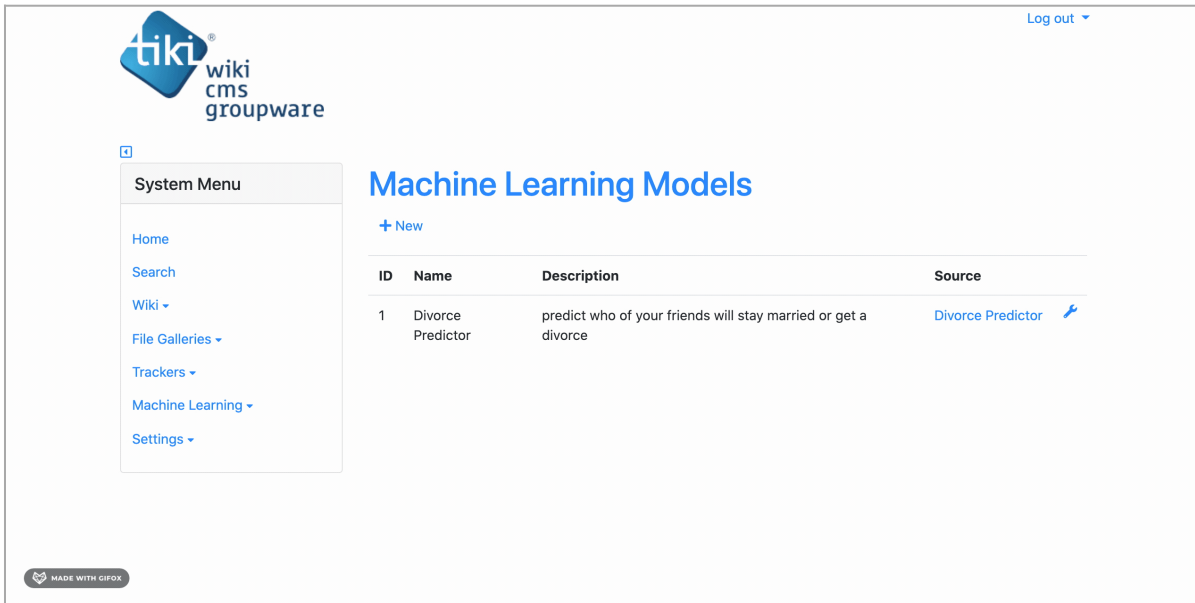
```
php console.php ml:train
```

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 05: Using the Model - 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.

Others

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:
    - https://gitlab.com/tikiwiki/tiki/-/blob/master/lib/core/Services/ML/Controller.php
    - https://gitlab.com/tikiwiki/tiki/-/blob/master/lib/ml/mllib.php

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.

Videos

- About [Rubix ML](#) in English

- About [Rubix ML](#) in French

Related links

- Preparing Machine Learning Dataset
- Creating Machine Learning Models
- Configuring Machine Learning Models
- Training Machine Learning Models
- Using Machine Learning Models
- Machine Learning Models
- Trackers
- Tracker Import Export
- https://www.we-rc.com/blog/2021/04/04/named-entity-recognition-in-php
- https://rubixml.com/

Other page names for this page

Rubix ML
RubixML