Preparing Machine Learning Dataset

Preparing dataset in Tiki involves creating a new Tiki tracker and adding items to it or using an existing tracker and making any necessary modifications to it. This is an important step because using Machine Learning models to make predictions require the models to be trained beforehand on already existing data. Trackers in Tiki are used for this purpose because of the ease in working with data contained in them just within Tiki, eliminating the need for special tools.

Creating Dataset

You create a dataset in Tiki by simply creating a tracker. Do this by going to the **List Trackers** page, click on the **Create** button and fill in the **Create tracker** form like you would for every other tracker.

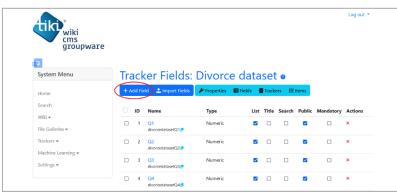
				Log out 👻
wiki				
groupware	2			
System Menu	Tradition			
System Menu	Trackers			
Home	+ Create 🗘 In	nport 🔹 🔚 Manage Tabula	r Formats	
Search	ld Name	Created +	Last modified	Items
Wiki v	No records found	L		
File Galleries 🕶				
Trackers 🔻				
List Trackers				
Manage Tabular Formats				
Machine Learning 👻				

Use the create button on List Trackers page to create a tracker

tiki	Create Tracker ×	Log out *
cms groupware	General	
	Name	
System Menu	Tr Divorce dataset	
Home	+ Description	
Search	Training data for Divorce Predictor model	Items
Wiki 🕶	No	
File Galleries 🕶	Description is wiki-parsed	
Trackers -	Field Prefix	
List Trackers	divorcedataset	
	Short string prepended by default to all fields in this tracker.	
Manage Tabular Formats	Permanent Name	
Machine Learning 👻	divorcedataset	
Settinos =	Required for Advanced Shopping Cart and some other	

Fill in the Create tracker form

Give it a name and a concise description and save; then add fields to the tracker to match your dataset attributes. That is, you add a field for every dataset attribute.



Add a tracker field by clicking on the Add Field button

tiki wiki cms	Add Field	×				Log out *
groupware	Name					
System Menu	Тг туре		0			
Home	+ Text Field	~	rs i≣	Items		
Search	More types available	×	Search	Public	Mandatory	Actions
Wiki ♥ File Galleries ♥	More field types may be enabled from the administration panel.					×
Trackers 🕶	Description					×
Machine Learning - Settings -			0			×
	Description contains wiki syntax	li.				×

Add new field form

The tracker field type you choose will depend on the attribute type. Tracker Field Types has a list of available Tiki Tracker field types. Do not use multi-value field types like Checkbox Tracker Field. Only use single-value field types such as Text Tracker Field. Use Numeric Tracker Field for all continuous attributes.

Note that the field types you choose will directly affect the user experience during model use. For categorical attributes, consider using field types like Dropdown Tracker Field or Radio Tracker Field to ease user input during usage.

See Creating a Tracker for more information on creating trackers. To learn how to add fields to a tracker, see Adding fields to a tracker.

Adding Samples to Dataset

In Tiki, each tracker item represent a sample. You add dataset samples by adding items to the tracker. A simple way to do this is to use the **Create Item** form displayed after clicking on the **Create Item** button in the view tracker page.

tiki																			Lo	og out	•
cms groupward	е																				
System Menu		Divo	orc	e	da	ta	set	t													
Home		ining da		-																	•
Search		+ Creat	e Iten	ソ	P Pro	pertie	s I	E He	lds	81	racke	rs									
Wiki 🕶	It	ems		150																	
File Galleries 🕶		CIII		150																	
Trackers 🔻			Q1	Q2	Q3	Q4	Q5	06	07	Q8	Q9	Q10	Q11	Q12	013	014	015	Q16	017	018	01
Machine Learning -		• •	2	2	4	1	0	0	0	0	0	0	1	0	1	1	0	1	0	0	-
Settings 🕶		•	4	4	4	4	4	0	0	4	4	4	4	3	4	0	4	4			
		• •		2	2	2	1	3	2	1	1	2	3	4	2	3	3	3	3	2	

Click on Create Item button to add tracker item

							Lo	og out *		^
UIKI wiki	Create Item ×									П.
cms groupware	Q1									
	٥									
System Menu	If one of us apologizes when our discussion deteriorates, the discussion ends. (1-5) E.g. 1,									
Train	Q2									
Home	0									
Search	I know we can ignore our differences, even if things get hard sometimes. (1-5) Eg. 5.									
wiki v	Q3									
File Galleries 🕶	0									
Trackers 👻	When we need it, we can take our discussions with my spouse from the									
Machine Learning -	beginning and correct it. (1-5) E.g. 2.	2	Q13	Q14	Q15	Q16	Q17	Q18 Q	n	
Settings •	Q4	0	1	1	0	1	0	0		
4	When I discuss with my spouse, to contact him will eventually work. (1-5)	3	4	0	4	4	4	4		
	E.g. 2.									
· · · · · · · · · · · · · · · · · · ·		4	2	3	3	3	3	3		×
Create tracker item form	L									

Samples can also be collected from user input if the dataset tracker is embedded in a wiki, article or blog page using PluginTracker.

Importing Samples

If dataset already exist in an external source such as a Comma-Separated Values (CSV) file, you can import samples from it to a tracker in Tiki using Tracker Import Export for this. Say you are working on Divorce Prediction project, you can download the project's dataset into your computer, then using Tracker Tabular in Tiki, import the samples into a tracker.

			Log out 💌
wiki			
cms groupware			
2			
System Menu	Tabular Formats		
Home	+ New 🔔 Create Tracker from File	E Trackers	
Search	Name	Tracker	
Wiki 🕶	Divorce dataset TT	Divorce dataset	1
File Galleries 🕶			
Trackers 🔻			
List Trackers			
Manage Tabular Formats			
Machine Learning 🕶			
Settings 🔻			

Tracker Tabular list



Use the Import button to import items into tracker

See Tracker Import Export to learn how to import tracker items using Tracker Tabular.

Related links

- Machine Learning
- Creating Machine Learning Models
- Configuring Machine Learning Models
- Training Machine Learning Models
- Using Machine Learning Models
- Trackers
- Tracker Field Types
- https://github.com/RubixML/Divorce