Models

Handle your models at Giza!

In Giza, a model represents a container for versions of your machine learning model. This design allows you to iterate and improve your ML model by creating new versions for it. Each model can have multiple versions, providing a robust and flexible way to manage the evolution of your ML models. This traceability feature ensures that you have a clear record of the originalonnx model used for transpilation, who performed the transpilation, and the output generated.

Remember, you need to be logged in to use these functionalities!

Create a Model

Creating a new model in Giza is a straightforward process. You only need to provide a name for the model using the--name option. However, you can also add a description of the model using the--description option. This can be helpful for keeping track of different models and their purposes.

Here's how you can do it:
```

Copy

```
giza models create --name my_new_model --description"A New Model" [giza][2023-09-13 13:24:28.223] Creating model ✅ { "id": 1, "name": "my_new_model", "description": "A New Model" }
```

Typically, thetranspile command is used to handle model creation. During this process, the filename is checked for an existing model. If none is found, a new model is automatically created. However, manual creation of a model is also supported. For more information, refer to the transpile documentation ([cairo](#) and[ezkl](#)).

List Models

Giza provides a simple and efficient way to list all the models you have stored on the server. This feature is especially useful when you have multiple models and need to manage them effectively.

To list all your models, you can use thelist command. This command retrieves and displays a list of all models stored in the server. Each model's information is printed in a json format for easy readability and further processing.

Here's how you can do it:
```

Copy

```
giza models list [giza][2023-09-13 13:11:39.403] Listing models ✅ [ { "id": 1, "name": "my_new_model", "description": "A New Model" }, { "id": 2, "name": "Test", "description": "A Model for testing different models" } ]
```

Retrieve Model Information

You can retrieve detailed information about a model stored on the server using its unique model id. This includes its name, description, and other id:
```

Copy

```
giza models get --model-id 1# When we create model for you we output it in the logs so be aware [giza][2023-09-13 13:17:53.594] Retrieving model information ✅ { "id": 1, "name": "my_new_model", "description": "A New Model" }
```

Now we can see that we have a model successfully transpiled! Now if we want we could download it again!

Transpile a model

Note: This is explained extensively in the transpile documentation [Orion Cairo](#) and[EZKL](#)). Transpiling a model in Giza is a

crucial step in the model deployment process as an endpoint. Transpilation is the process of converting your machine learning model into a format that can be executed on Giza. Depending the ZKML framework chosen, this process involves converting the model into a series of Cairo instructions or performing the setup using EZKL.

When you execute the 'transpile' command, it initially checks for the presence of the model on the Giza platform. If no model is found, it automatically generates one and performs the transpilation. Here is an example of the command:

```
Copy giza transpile --framework CAIRO awesome_model.onnx --output-path my_awesome_model
```

It's worth noting that if you already have created a model, you can transpile it by specifying the model ID:

```
Copy giza transpile --model-id 1 --framework CAIRO awesome_model.onnx --output-path my_awesome_model
```

This method proves useful when you intend to create a new version of an existing model.

For more information, refer to the transpile documentation (cairo andezkl).

Download a successfully transpiled model

For this we can use thedownload command available in the CLI which only needs amodel_id to download it again!

```
Copy

    giza models download 1 [giza][2023-08-04 10:33:14.271] Transpilation is ready, downloading! ✓ [giza][2023-08-04 10:33:15.134] Transpilation saved at: cairo_model
```

Let's check the downloaded model:

```
Copy

    tree cairo_model/ cairo_model ├── inference │ ├── Scarb.toml │ └── src │ └── lib.cairo └── initializers ├── node_l1 │ ├── Scarb.toml │ └── src │ └── lib.cairo
```

Last updated17 days ago