



Quick Start

This page helps you quickly create your first **source**, **dataset**, **model**, and **prediction**.

To get started with **BigML.io** you need:

1. Your **username** and your **API key**.
2. A terminal with **curl** or any other command-line tool that implements standard HTTPS methods.
3. Some sample data. You can use:
 - A **csv** file with some data. You can download the "**Iris dataset**" or "**Diabetes dataset**" from our servers.
 - Even easier, you can just use a URL that points to your data. For example, you can use <https://static.bigml.com/csv/iris.csv> or <https://static.bigml.com/csv/diabetes.csv>.
 - Even even easier, you can just send some inline test data.

Jump to:

- [Getting a Toy Data File](#)
- [Authentication](#)
- [Creating a Source](#)
- [Creating a Remote Source](#)
- [Creating an Inline Source](#)
- [Creating a Dataset](#)
- [Creating a Model](#)
- [Creating a Prediction](#)

Getting a Toy Data File

If you do not have any dataset handy, you can download **Fisher's Iris dataset** using the **curl** command below or by just clicking on the link.

```
curl -o iris.csv https://static.bigml.com/csv/iris.csv
```

curl

Authentication

The following snippet will help you set up an environment variable (i.e., **BIGML_AUTH**) to store your **username** and **API key** and avoid typing them again in the rest of examples. See this [section](#) for more details.

Note: Use your own **username** and **API Key**.

bash

```
export BIGML_USERNAME=alfred
export BIGML_API_KEY=79138a622755a2383660347f895444b1eb927730
export BIGML_AUTH="username=$BIGML_USERNAME&api_key=$BIGML_API_KEY"
```

If you have a [private deployment](#), you have to connect to your own API server. Ask to your server administrators to know what is your API server. All the BigML [bindings](#) are able to set the appropriate API server domain using the env var **BIGML_DOMAIN**. Read the documentation of the bindings you're using for further information.

Creating a Source

To create a new **source**, POST the file containing your data to the source base URL.

curl

```
curl "https://bigml.io/andromeda/source?$BIGML_AUTH" -F file=@iris.csv
```

To create more **sources** simply repeat the **curl** command above using another file. Make sure to use the full path if the file is not in your current directory.

Creating a Remote Source

You can also create a source using a valid URL that points to your data or some public data. For example:

curl

```
curl "https://bigml.io/andromeda/source?$BIGML_AUTH" \
-X POST \
-H 'content-type: application/json' \
-d '{"remote": "https://static.bigml.com/csv/iris.csv"}'
```

Creating an Inline Source

You can also create a source using some inline data. For example:

```
curl "https://bigml.io/andromeda/source?${BIGML_AUTH}" \
-X POST \
-H 'content-type: application/json' \
-d '{"data": "a,b,c,d\n1,2,3,4\n5,6,7,8"}'
```

curl

BigML.io will respond with a JSON object containing preliminary information about your new **source**. As with all **BigML.io** resources, the new source will have a **resource** key with a unique **resource/id**. You can use the **source/id** to retrieve the source or to create new datasets.

```
{
  "code": 201,
  "content_type": "application/octet-stream",
  "created": "2012-03-01T05:29:07.217968",
  "file_name": "iris.csv",
  "md5": "d1175c032e1042bec7f974c91e4a65ae",
  "name": "iris.csv",
  "number_of_datasets": 0,
  "number_of_models": 0,
  "number_of_predictions": 0,
  "private": true,
  "resource": "source/4f52824203ce893c0a000053",
  "size": 4608,
  "source_parser": {
```

json

For nicely formatted JSON responses, pipe the **curl** command to

```
| python -m json.tool or to | *jq* "." .
```

You can get [jq](#) [here](#).

Creating a Dataset

To create a **dataset**, POST the **source/id** from the previous step to the **dataset** base URL as follows.

```
curl "https://bigml.io/andromeda/dataset?$BIGML_AUTH" \  
-X POST \  
-H 'content-type: application/json' \  
-d '{"source": "source/4f52824203ce893c0a000053"}'
```

BigML.io will return a **dataset** resource if the request succeeds. **BigML.io** detects types for each field and will begin computing the histograms and summary statistics. In the [Datasets Section](#) you can learn how to customize the parsing rules and other options when converting a datasource to a dataset. Each field in your source is automatically assigned an id that you can later use as a parameter in models and predictions.

```
{  
  "code": 201,  
  "columns": 5,  
  "created": "2012-03-04T02:58:11.910363",  
  "fields": {  
    "000000": {  
      "column_number": 0,  
      "name": "sepal length",  
      "optype": "numeric"  
    },  
    "000001": {  
      "column_number": 1,  
      "name": "sepal width",  
      "optype": "numeric"  
    }  
  }  
}
```

Creating a Model

To create a **model**, POST the **dataset/id** from the previous step to the model base URL. By default **BigML.io** will include all fields as predictors and will treat the last non-text field as the objective. In the [Models Section](#) you will learn how to customize the input fields or the objective field.

```
curl "https://bigml.io/andromeda/model?$BIGML_AUTH" \  
-X POST \
```

```
-H 'content-type: application/json' \  
-d '{"dataset": "dataset/4f52da4303ce896fe3000000"}'
```

If the request is successful a new model will be created. The model will contain a **model/id** that you can use to retrieve it later or to create predictions.

```
{  
  "code": 201,  
  "columns": 5,  
  "created": "2012-03-04T03:46:53.033372",  
  "dataset": "dataset/4f52da4303ce896fe3000000",  
  "dataset_status": true,  
  "holdout": 0.0,  
  "input_fields": [],  
  "max_columns": 5,  
  "max_rows": 150,  
  "name": "iris' dataset model",  
  "number_of_predictions": 0,  
  "objective_fields": [],  
  "private": true,  
}
```

json

Creating a Prediction

To create a prediction, POST the **model/id** and some input data to the **prediction** base URL.

```
curl "https://bigml.io/andromeda/prediction?${BIGML_AUTH}" \  
-X POST \  
-H 'content-type: application/json' \  
-d '{"model": "model/4f52e5ad03ce898798000000", "input_data": {"000000": 5, "000001": 3}]'
```

curl

If the request succeeds, **BigML.io** will return a new **prediction** resource with its own **prediction/id**. You can use this id to retrieve the prediction later on. The predicted value is found in the **prediction** object, keyed by the corresponding **objective field id**.

```
{  
  "code": 201,  
}
```

json

```
"created": "2012-03-04T04:11:10.433996",  
"dataset": "dataset/4f52da4303ce896fe3000000",  
"dataset_status": true,  
"fields": {  
  "000000": {  
    "column_number": 0,  
    "datatype": "double",  
    "name": "sepal length",  
    "optype": "numeric"  
  },  
  "000001": {  
    "column_number": 1,
```

[Documentation](#)

[Tools](#)

[Certifications](#)

Copyright © 2024 BigML, Inc.