



(<https://www.mimacom.com/en/about/#offices-anchor>)

Getting started with Tensorflow and Java-Spring

June 29, 2018

Java (<https://blog.mimacom.com/tag/java/>)

by Enrique Llerena Domínguez (<https://blog.mimacom.com/author/enriquedominguez/>)

Spring (<https://blog.mimacom.com/tag/spring/>)

Machine Learning (<https://blog.mimacom.com/tag/machine-learning/>)

Introduction

One of the goals I set for this year is to explore Machine Learning (ML), so after having done a couple of courses here and there, I decided to do a - rather simple- starting project, where I could deal with some of the basic stages of the ML: Get the data, prepare it, choose a model, train it, evaluate it, export it, and make the predictions available for use. For this first project, I chose:

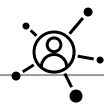
- Tensorflow (<https://www.tensorflow.org/>) as the framework,
- Python (v3.5), with Jupyter Notebooks as the model generating part,
- Java, using Spring Boot as the prediction serving part.

Environment setup

Note: These steps were executed in Windows. The full code for this post can be found at <https://github.com/ellerenad/Getting-started-Tensorflow-Java-Spring>

Training part

- Download Anaconda (<https://www.anaconda.com/download/>).
- Create an environment using:



```
1 | conda create -n tensorflow pip python=3.5
```

In this example of the conda create command, tensorflow is used as the name of the environment. ([https://www.mimacom.com/en/about/#offices-](https://www.mimacom.com/en/about/#offices-anchor)

- Activate your recently created environment

anchor)

```
1 | activate tensorflow
```

- Install Jupyter, Tensorflow, and other required packages:

```
1 | pip install --ignore-installed --upgrade jupyter
2 | pip install --ignore-installed --upgrade tensorflow
3 | pip install --ignore-installed --upgrade scipy
4 | pip install --ignore-installed --upgrade pandas
5 | pip install --ignore-installed --upgrade sklearn
```

- Execute Jupyter Notebook

```
1 | jupyter notebook
```

From now on, to run your notebooks, you just need to open the Anaconda Prompt and execute:

```
1 | activate tensorflow
2 | jupyter notebook
```

Using the jupyter notebook webapp (by default opened at <http://localhost:8888> (<http://localhost:8888>)), create a new notebook, or use the following, found on the Github repository (<https://github.com/ellerenad/Getting-started-Tensorflow-Java-Spring>): `./training/TF_iris_data.ipynb`

Serving part

- Install Java SDK
- Install Maven
- Use the following parent POM and dependencies:

```
1 | <parent>
2 |   <groupId>org.springframework.cloud</groupId>
3 |   <artifactId>spring-cloud-starter-parent</artifactId>
4 |   <version>Camden.SR7</version>
5 | </parent>
6 | <dependencies>
7 |   <dependency>
8 |     <groupId>org.springframework.boot</groupId>
9 |     <artifactId>spring-boot-starter-web</artifactId>
10 |   </dependency>
11 |   <dependency>
12 |     <groupId>org.tensorflow</groupId>
13 |     <artifactId>tensorflow</artifactId>
14 |     <version>1.8.0</version>
15 |   </dependency>
16 |   <dependency>
17 |     <groupId>org.springframework.boot</groupId>
18 |     <artifactId>spring-boot-starter-test</artifactId>
19 |     <scope>test</scope>
20 |   </dependency>
21 |   <dependency>
22 |     <groupId>com.google.code.gson</groupId>
23 |     <artifactId>gson</artifactId>
24 |     <scope>test</scope>
25 |   </dependency>
26 | </dependencies>
```

Note: This POM shows just the required dependencies, and is not in the required final form.

Describing the problem and the data set

We will use the famous Iris Data Set, where different types of Iris flowers are classified based on some of its features, like the length and width of its petal and sepal, resulting into 3 different categories: *Setosa*, *Versicolour*, or *Virginica*. We will use supervised training and a neural network classifier. More information about the data set at the scikit learn website (http://scikit-learn.org/stable/auto_examples/datasets/plot_iris_dataset.html) and Wikipedia (https://en.wikipedia.org/wiki/Iris_flower_data_set).

Describing the -rather basic- architecture

As previously discussed, we have 2 components: The training component, written in Python, and the server component, written in Java. The output of the former is a trained and evaluated model, which is the one of the inputs of the latter. This is possible because we are using the Tensorflow framework on both components.

Describing the training component

In this component  we will perform the following steps:



1. Get the data
2. Prepare the data
3. Partition the data into train and evaluation/test sets
4. Format the data as required by Tensorflow
5. Train the model
6. Evaluate the model
7. Export the model

(<https://www.mimacom.com/en/about/#offices->

anchor)

Now, let's proceed with the code: (Finally!)

Prepare the data and train the model

Import the modules:

```
1 | import tensorflow as tf
2 | import pandas as pd
3 | import numpy as np
4 | from sklearn import datasets
```

Define "constants" for the names of the features:

```
1 | FEATURE_SEPAL_LENGTH = 'SepalLength'
2 | FEATURE_SEPAL_WIDTH = 'Sepalwidth'
3 | FEATURE_PETAL_LENGTH = 'PetalLength'
4 | FEATURE_PETAL_WIDTH = 'Petalwidth'
5 | LABEL = 'label'
```

Get the data:

```
1 | # load the data set
2 | iris = datasets.load_iris()
```

Since the data comes originally with all the examples ordered, we need to shuffle it to get a meaningful test set. To achieve this, we first need to add the target (the label each set of measures correspond to) to the data, and then shuffle it:

```
1 | iris_data_w_target = [];
2 |
3 | # add the target to the data
4 | for i in range(len(iris.data)):
5 |     value = np.append(iris.data[i], iris.target[i])
6 |     iris_data_w_target.append(value)
```

Create a Pandas Data Frame to operate with, and shuffle the data:

```
1 | columns_names = [FEATURE_SEPAL_LENGTH, FEATURE_SEPAL_WIDTH, FEATURE_PETAL_LENGTH, FEATURE_PETAL_WIDTH, LABEL]
2 |
3 | df = pd.DataFrame(data = iris_data_w_target, columns = columns_names )
4 |
5 | # shuffle our data
6 | df = df.sample(frac=1).reset_index(drop=True)
```

Having done the shuffling of the data, we can partition it into training and evaluation/test sets. We will reserve 20% of the original set for evaluation/testing, whilst the model will be trained with the rest 80%:

```
1 | test_len = (len(df) * 20)//100;
2 | training_df = df[:test_len:]
3 | test_df = df[test_len:]
```

After that, we format the data for Tensorflow. So far, we have stored our data in a Pandas Data Frame, which represents a data table, but Tensorflow expects to receive a map, where the keys are the names of the features, and the correspondent values are arrays storing the same data as the columns from our Pandas Data Frame. So, we first declare the columns we will be using, and then we create the map using the Pandas Data Frame with our training data.

```
1 | iris_feature_columns = [
2 |     tf.contrib.layers.real_valued_column(FEATURE_SEPAL_LENGTH, dimension=1, dtype=tf.float32),
3 |     tf.contrib.layers.real_valued_column(FEATURE_SEPAL_WIDTH, dimension=1, dtype=tf.float32),
4 |     tf.contrib.layers.real_valued_column(FEATURE_PETAL_LENGTH, dimension=1, dtype=tf.float32),
5 |     tf.contrib.layers.real_valued_column(FEATURE_PETAL_WIDTH, dimension=1, dtype=tf.float32)
6 | ]
7 |
8 | x = {
9 |     FEATURE_SEPAL_LENGTH : np.array(training_df[FEATURE_SEPAL_LENGTH]),
10 |    FEATURE_SEPAL_WIDTH : np.array(training_df[FEATURE_SEPAL_WIDTH]),
11 |    FEATURE_PETAL_LENGTH : np.array(training_df[FEATURE_PETAL_LENGTH]),
12 |    FEATURE_PETAL_WIDTH : np.array(training_df[FEATURE_PETAL_WIDTH])
13 | }
```

Then, we instantiate the model and train it. We will use a Neural Network Classifier, with 5 nodes and 5 hidden layers, which has an output of 3 different classes.

```

1 classifier = tf.estimator.DNNClassifier(
2     feature_columns = iris_feature_columns,
3     hidden_units = [5, 5],
4     n_classes = 3)
5
6
7 # Define the training inputs
8 train_input_fn = tf.estimator.inputs.numpy_input_fn(
9     x = x,
10    y = np.array(training_df[LABEL]).astype(int),
11    num_epochs = None,
12    shuffle = True)
13
14 # Train model.
15 classifier.train(input_fn = train_input_fn, steps = 4000)

```



(<https://www.mimacom.com/en/about/#offices->

anchor)

Evaluate the model

Once we have the model trained, we proceed to evaluate it using the evaluation/test set we separated earlier:

```

1 x = {
2     FEATURE_SEPAL_LENGTH : np.array(test_df[FEATURE_SEPAL_LENGTH]),
3     FEATURE_SEPAL_WIDTH  : np.array(test_df[FEATURE_SEPAL_WIDTH]),
4     FEATURE_PETAL_LENGTH  : np.array(test_df[FEATURE_PETAL_LENGTH]),
5     FEATURE_PETAL_WIDTH   : np.array(test_df[FEATURE_PETAL_WIDTH])
6 }
7
8 # Define the training inputs
9 test_input_fn = tf.estimator.inputs.numpy_input_fn(
10    x = x,
11    y = np.array(test_df[LABEL]).astype(int),
12    num_epochs = 1,
13    shuffle = False)
14
15 # Evaluate accuracy.
16 accuracy_score = classifier.evaluate(input_fn=test_input_fn)["accuracy"]
17
18 print("Test Accuracy: ", accuracy_score)

```

Output:

```

1 INFO:tensorflow:Saving dict for global step 4000: accuracy = 1.0, average_loss = 0.03394517, global_step = 4000, loss = 1.018355
2 Test Accuracy: 1.0

```

Now, we can also do some more manual testing of our model, if required. To do so, we take some arbitrary records from the original data set, including its respective target, and feed them to the model.

```

1 x = {
2     FEATURE_SEPAL_LENGTH : np.array([5.0, 6.7, 7.4]),
3     FEATURE_SEPAL_WIDTH  : np.array([3.5, 3.1, 2.8]),
4     FEATURE_PETAL_LENGTH  : np.array([1.3, 4.4, 6.1]),
5     FEATURE_PETAL_WIDTH   : np.array([0.3, 1.4, 1.9])
6 }
7
8 expected = np.array([0, 1, 2])
9
10 # Define the training inputs
11 predict_input_fn = tf.estimator.inputs.numpy_input_fn(
12    x = x,
13    num_epochs = 1,
14    shuffle = False)
15
16 predictions = classifier.predict(input_fn = predict_input_fn)
17
18 for pred_dict, expec in zip(predictions, expected):
19     class_id = pred_dict['class_ids'][0]
20     probability = pred_dict['probabilities'][class_id]
21
22     print('\nPrediction is "{}" (certainty {:.1f}%), expected "{}"'.format(class_id, 100 * probability, expec))
23

```

Note: Since these records are hardcoded, it is highly probable (80%, to be precise ;)) they are part of the training data set. To limit overfitting, it is important not to test the model with the data it was trained with. If this testing step is important for you, please consider improving this piece of code.

Output:

```

1 Prediction is "0" (certainty 100.0%), expected "0"
2
3 Prediction is "1" (certainty 98.0%), expected "1"
4
5 Prediction is "2" (certainty 99.8%), expected "2"

```

Export the model

Now that we have evaluated our model, we proceed to export it. To do that, we need to define a function describing the input it will receive, and then call to the `export_savedmodel` method of the classifier itself.

```

1 | def serving_input_receiver_fn():
2 |     serialized_tf_example = tf.placeholder(dtype = tf.string, shape = [None], name = 'input_tensors')
3 |     receiver_tensors = {'predictor_inputs' : serialized_tf_example}
4 |     feature_spec = {FEATURE_SEPAL_LENGTH : tf.FixedLenFeature([25], tf.float32),
5 |                     FEATURE_SEPAL_WIDTH : tf.FixedLenFeature([25], tf.float32),
6 |                     FEATURE_PETAL_LENGTH : tf.FixedLenFeature([25], tf.float32),
7 |                     FEATURE_PETAL_WIDTH : tf.FixedLenFeature([25], tf.float32)}
8 |     features = tf.parse_example(serialized_tf_example, feature_spec)
9 |     return tf.estimator.export.ServingInputReceiver(features, receiver_tensors)
10 |
11 |
12 |
13 | model_dir = classifier.export_savedmodel(export_dir_base = "stored_model",
14 |                                         serving_input_receiver_fn = serving_input_receiver_fn,
15 |                                         as_text = True)
16 | print('Model exported to ' + model_dir.decode())

```



(<https://www.mimacom.com/en/about/#offices->

anchor)

The following output means we have properly exported the model, and it is at `.\stored_model\1530093489`

```

1 | INFO:tensorflow:SavedModel written to: b"stored_model\\temp-b'1530093489'\\saved_model.pbtxt"
2 | Model exported to stored_model\1530093489

```

Having exported our trained model, we are now ready for loading it in Java for the server component :D

Describing the process in java

In this component we will perform the following steps:

1. Publish two GET endpoints to retrieve predictions: The predicted class und a set of probabilities per class.
2. Load the previously saved model.
3. Feed the input to the model and fetch the prediction.
4. Integration testing.
5. Examples of usage

Defining the domain objects

We need two domain objects: The Iris and the possible types of Iris, represented by an enum: IrisType:

```

1 | public class Iris {
2 |
3 |     private float petalLength;
4 |     private float petalwidth;
5 |     private float sepalLength;
6 |     private float sepalwidth;
7 |
8 |     public Iris() {
9 |     }
10 |
11 |     public Iris(float petalLength, float petalwidth, float sepalLength, float sepalwidth) {
12 |         this.petalLength = petalLength;
13 |         this.petalwidth = petalwidth;
14 |         this.sepalLength = sepalLength;
15 |         this.sepalwidth = sepalwidth;
16 |     }
17 |     // ...
18 |     // (Getters and setters omitted)
19 | }

```

```

1 | public enum IrisType {
2 |     SETOSA,
3 |     VERSICOLOUR,
4 |     VIRGINICA
5 | }

```

Exposing the endpoints

Here we expose the two required GET endpoints. Both expect as parameters the features of the Iris:

- petalLength
- petalWidth
- sepalLength
- sepalWidth

The Spring framework will read them from the URL and inject the Iris object in the method.

- The `/iris/classify/class` endpoint returns the predicted class, *Setosa*, *Versicolour*, or *Virginica*
- The `/iris/classify/probabilities` returns the probabilities the input has for each given class to appear

```

1 | @RestController
2 | public class IrisController {
3 |
4 |     @Autowired
5 |     IrisClassifierService irisClassifierService;
6 |
7 |     @GetMapping(value = "/iris/classify/class")
8 |     public IrisType classify(Iris iris) {
9 |         return irisClassifierService.classify(iris);
10 |     }
11 |
12 |     @GetMapping(value = "/iris/classify/probabilities")
13 |     public Map<IrisType, Float> classificationProbabilities(Iris iris) {
14 |         return irisClassifierService.classificationProbabilities(iris);
15 |     }
16 |
17 | }

```



(<https://www.mimacom.com/en/about/#offices->

anchor)

On "Examples of Usage" section we will show examples of CURL get requests.

Loading the Tensorflow model

As you might have noticed, we have a service where the classification logic is encapsulated.

```

1 | public interface IrisClassifierService {
2 |
3 |     /**
4 |      * Method to fetch a classification from the model
5 |      * @param iris the data to classify
6 |      * @return the predicted type
7 |      */
8 |     IrisType classify(Iris iris);
9 |
10 |    /**
11 |     * Method to fetch from the model the probabilities of all the types
12 |     * @param iris the data to classify
13 |     * @return A map relating the type with its predicted probabilities
14 |     */
15 |    Map<IrisType, Float> classificationProbabilities(Iris iris);
16 | }

```

In this implementation of the service, we will use the Tensorflow framework to load the previously trained model and feed the inputs to fetch the outputs. Here we use the `SavedModelBundle.load()` method to load the model, and create a session out of it. Such session will be used later to interact with the model.

```

1 | public class IrisTensorflowClassifierService implements IrisClassifierService {
2 |
3 |     private final Session modelBundleSession;
4 |     private final IrisType[] irisTypes;
5 |
6 |     //...
7 |
8 |     @Autowired
9 |     public IrisTensorflowClassifierService(@Value("${irism1.savedModel.path}") String savedModelPath,
10 |      @Value("${irism1.savedModel.tags}") String savedModelTags) {
11 |         this.modelBundleSession = SavedModelBundle.load(savedModelPath, savedModelTags).session();
12 |         this.irisTypes = IrisType.values();
13 |     }
14 |
15 |     //...
16 |
17 | }

```

Feeding the input to and fetching the prediction from the Tensorflow model

Before we can feed the input to the model, we need to build it: The Tensorflow framework uses Tensors (https://www.tensorflow.org/programmers_guide/tensors) to do this. Here we see how it is built.

```

1 | public class IrisTensorflowClassifierService implements IrisClassifierService {
2 |     //...
3 |     private static Tensor createInputTensor(Iris iris){
4 |         // order of the data on the input: PetalLength, Petalwidth, SepalLength, Sepalwidth
5 |         // (taken from the saved_model, node dnn/input_from_feature_columns/input_layer/concat)
6 |         float[] input = {iris.getPetalLength(), iris.getPetalwidth(), iris.getSepalLength(), iris.getSepalwidth()};
7 |         float[][] data = new float[1][4];
8 |         data[0] = input;
9 |         return Tensor.create(data);
10 |     }
11 |     //...
12 | }

```

Notice the importance of the order of the parameters on the array. This order was obtained from the node `dnn/input_from_feature_columns/input_layer/concat` on the `saved_model.pbtxt` file. In that node, we can see how the name of the parameters match with those described on the `serving_input_receiver_fn` on the export of the model at the (Python) training section, and in this case, the order happens to be alphabetical.

Once we have a standard way to build the input for the model, we proceed to feed them and fetch a prediction. The different kinds of predictions are returned when we *query* for an operation. In this case, we have two fetch operations. We also need to define the input.

```

1 public class IrisTensorflowClassifierService implements IrisClassifierService {
2     //...
3     private final static String FEED_OPERATION = "dnn/input_from_feature_columns/input_layer/concat";
4     private final static String FETCH_OPERATION_PROBABILITIES = "dnn/head/predictions/probabilities";
5     private final static String FETCH_OPERATION_CLASS_ID = "dnn/head/predictions/class_ids";
6     //...
7 }
8

```



(<https://www.mimacom.com/en/about/#offices->

The exact names of the fetch and feed operations were found on the saved_model.pbtxt file

Now, we can fetch the predicted class for the given input:

anchor)

```

1 @Service
2 public class IrisTensorflowClassifierService implements IrisClassifierService {
3     // ...
4     @Override
5     public IrisType classify(Iris iris) {
6         int category = this.fetchClassFromModel(iris);
7         return this.irisTypes[category];
8     }
9
10    private int fetchClassFromModel(Iris iris){
11        Tensor inputTensor = IrisTensorflowClassifierService.createInputTensor(iris);
12
13        Tensor result = this.modelBundleSession.runner()
14            .feed(IrisTensorflowClassifierService.FEED_OPERATION, inputTensor)
15            .fetch(IrisTensorflowClassifierService.FETCH_OPERATION_CLASS_ID)
16            .run().get(0);
17
18        long[] buffer = new long[1];
19        result.copyTo(buffer);
20        return (int)buffer[0];
21    }
22    // ...
23 }

```

And below we see how we fetch the predicted probabilities for each possible class, and build a map to return.

```

1 @Service
2 public class IrisTensorflowClassifierService implements IrisClassifierService {
3     // ...
4
5     @Override
6     public Map<IrisType, Float> classificationProbabilities(Iris iris){
7         Map<IrisType, Float> results = new HashMap<>(irisTypes.length);
8         float[][] vector = this.fetchProbabilitiesFromModel(iris);
9         int resultsCount = vector[0].length;
10        for (int i=0; i < resultsCount; i++){
11            results.put(irisTypes[i],vector[0][i]);
12        }
13        return results;
14    }
15
16    private float[][] fetchProbabilitiesFromModel(Iris iris) {
17        Tensor inputTensor = IrisTensorflowClassifierService.createInputTensor(iris);
18
19        Tensor result = this.modelBundleSession.runner()
20            .feed(IrisTensorflowClassifierService.FEED_OPERATION, inputTensor)
21            .fetch(IrisTensorflowClassifierService.FETCH_OPERATION_PROBABILITIES)
22            .run().get(0);
23
24        float[][] buffer = new float[1][3];
25        result.copyTo(buffer);
26        return buffer;
27    }
28
29    // ...
30 }

```

Notice how the buffer is a matrix, whose second dimension matches the dimension of the expected output.

Performing integration test

Having done the required services, we can do some integration testing. First, the easiest: The `/iris/classify/class` endpoint to get the class given the Iris features. For both endpoints, we feed the same numbers we used on the training (Python) manual testing section, and we expect the endpoint to return the same classes contained in a response with status OK.





```

1 public class IrisControllerTest extends BaseControllerTest {
2     // ...
3     @Test
4     public void classify() throws Exception {
5
6         String urlTemplate = "/iris/classify/class?petalLength=%.1f&petalWidth=%.1f&sepalLength=%.1f&sepalWidth=%.1f";
7
8         // Locale.US to make sure the numbers are with period instead of comma.
9         String urlRequest = String.format(Locale.US, urlTemplate, 1.3f, 0.3f, 5.0f, 3.5f);
10        MvcResult mvcResult = this.mockMvc.perform(get(urlRequest))
11            .andExpect(status().isOk()).andReturn();
12        assertEquals(IrisType.SETOSA.toString(), mvcResult.getResponse().getContentAsString().replace("\\", ""));
13
14        urlRequest = String.format(Locale.US, urlTemplate, 4.4f, 1.4f, 6.7f, 3.1f);
15        mvcResult = this.mockMvc.perform(get(urlRequest))
16            .andExpect(status().isOk()).andReturn();
17        assertEquals(IrisType.VERSICOLOUR.toString(), mvcResult.getResponse().getContentAsString().replace("\\", ""));
18
19        urlRequest = String.format(Locale.US, urlTemplate, 6.1f, 1.9f, 7.4f, 2.8f);
20        mvcResult = this.mockMvc.perform(get(urlRequest))
21            .andExpect(status().isOk()).andReturn();
22        assertEquals(IrisType.VIRGINICA.toString(), mvcResult.getResponse().getContentAsString().replace("\\", ""));
23
24    }
25    // ...
26 }

```

(<https://www.mimacom.com/en/about/#offices-anchor>)

Then, we test the `/iris/classify/probabilities` endpoint, which retrieves the a map of the probabilities for each class. Since the exact number can be slightly different depending on the instance of the model, we will assert the following:

- The class with the highest probabilities is the one we expect.
- The amount of entries is the same as the amount of possible classes.
- All the possible classes have an entry
- All entries have a class and a probability.

Here we see the assertions:

```

1 public class IrisControllerTest extends BaseControllerTest {
2     // ...
3     private void assertProbabilitiesResponse(MockHttpServletResponse mockHttpServletResponse, IrisType expectedType) throws Unsu
4     // Extract the probabilities response
5     Gson gson = new Gson();
6     LinkedTreeMap<String, Float> probabilities;
7     probabilities = (LinkedTreeMap<String, Float>) gson.fromJson(mockHttpServletResponse.getContentAsString(), Map.class);
8     // Assert
9     assertEquals(expectedType.toString(), getPredictedType(probabilities));
10    assertProbabilities(probabilities);
11 }
12
13 private String getPredictedType(LinkedTreeMap<String, Float> probabilities) {
14     // The predicted type is the one with the highest probabilities
15     String predictedType = probabilities.entrySet().stream().max(Map.Entry.comparingByValue()).get().getKey();
16     return predictedType;
17 }
18
19 private void assertProbabilities(LinkedTreeMap<String, Float> probabilities) {
20     // The same amount of entries in the map as the possible values
21     assertEquals(probabilities.size(), IrisType.values().length);
22
23     // All the types have a probability value
24     for(IrisType irisType: IrisType.values()){
25         assertTrue(probabilities.containsKey(irisType.toString()));
26     }
27
28     // All the entries have a value
29     probabilities.entrySet().stream().forEach(probabilityEntry -> {
30         assertTrue(probabilityEntry.getKey() != null);
31         assertTrue(probabilityEntry.getValue() != null);
32     });
33 }
34 // ...
35 }

```

And here we perform the calls and evaluate the results:

```

1 public class IrisControllerTest extends BaseControllerTest {
2     // ...
3     @Test
4     public void classificationProbabilities() throws Exception {
5
6         String urlTemplate = "/iris/classify/probabilities?petalLength=%.1f&petalWidth=%.1f&sepalLength=%.1f&sepalWidth=%.1f";
7
8         // Locale.US to make sure the numbers are with period instead of comma.
9         String urlRequest = String.format(Locale.US, urlTemplate, 1.3f, 0.3f, 5.0f, 3.5f);
10        MvcResult mvcResult = this.mockMvc.perform(get(urlRequest))
11            .andExpect(status().isOk()).andReturn();
12
13        assertProbabilitiesResponse(mvcResult.getResponse(), IrisType.SETOSA);
14
15        urlRequest = String.format(Locale.US, urlTemplate, 4.4f, 1.4f, 6.7f, 3.1f);
16        mvcResult = this.mockMvc.perform(get(urlRequest))
17            .andExpect(status().isOk()).andReturn();
18
19        assertProbabilitiesResponse(mvcResult.getResponse(), IrisType.VERSICOLOUR);
20
21        urlRequest = String.format(Locale.US, urlTemplate, 6.1f, 1.9f, 7.4f, 2.8f);
22        mvcResult = this.mockMvc.perform(get(urlRequest))
23            .andExpect(status().isOk()).andReturn();
24
25        assertProbabilitiesResponse(mvcResult.getResponse(), IrisType.VIRGINICA);
26    }
27    // ...
28 }

```


Packaging the application, executing it, and curl examples with output:

First, create the jar file with maven (on the /serving folder), and then, execute the jar:

```
1 | mvn clean package
2 | java -jar ./target/tensorflowdemo-0.0.1-SNAPSHOT.jar
```

The default port is 7373, but is configurable using the application.yml. Again, we are using the same manual testing values.

Example for /iris/classify/class endpoint, where we expect versicolour:

```
1 | curl -GET "localhost:7373/iris/classify/class?petalLength=4.4&petalWidth=1.4&sepalLength=6.7&sepalWidth=3.1"
```

Output:

```
1 | "VERSICOLOUR"
```

Example for /iris/classify/probabilities endpoint, where we expect setosa to have the highest probabilities:

```
1 | curl -GET "localhost:7373/iris/classify/probabilities?petalLength=1.3&petalWidth=0.3&sepalLength=5.0&sepalWidth=3.5"
```

Output:

```
1 | {"SETOSA":0.999987,"VIRGINICA":3.4298865E-15,"VERSICOLOUR":1.2982294E-5}
```

Conclusions

Using Tensorflow, Java, and Python, we have demonstrated with a simple project how to perform the basic steps to train, evaluate, and export a model, so it can be later used by another application, in this case a Java Spring Boot application, exposing a REST endpoint to fetch predictions.

The full code can be found at <https://github.com/ellerenad/Getting-started-Tensorflow-Java-Spring> (<https://github.com/ellerenad/Getting-started-Tensorflow-Java-Spring>)

Thanks for reading!

About the author: Enrique Llerena Domínguez

Passionate in code and in life. Likes football (both american and the real one ;). Goal oriented. Keep movin', keep movin'!

COMMENTS



Autocomplete with Elasticsearch - Part ...

a year ago • 3 comments

You know searches from Google or DuckDuckGo. As you start typing the ...

Building a REST API with Liferay DXP with ...

a year ago • 2 comments

With the arrival of Liferay DXP and JAX-RS services, we discovered a whole ...

Code Folding for JSON Data

3 months ago • 1 comment

In Pair Programming sessions, we often exchange tips and tricks. ...

Issuing Certificate With BOSH

a year ago • 1 comment

In any on-premise deployment there are usually a bunch of ...

3 Comments mimacom Disqus' Privacy Policy

Recommend Tweet Share Sort by Best ▾

Join the discussion...

LOG IN WITH OR SIGN UP WITH DISQUS

Alejandro Piedra Abarca • 10 months ago

JOIN US

Are you creative and passionate about software development? Do you think unconventionally and act with initiative? Do you want to achieve great things within our team?

[Check out Our Job Offers \(https://www.mimacom.com/en/jobs/\)](https://www.mimacom.com/en/jobs/)