Data Science & Machine Learning

Productionizing H2O

geral@performetric.net

Productionizing H2O with Spring Boot

- **H2O** - Open source machine learning and artificial intelligence platform.
- **Spring Boot** - Open source Java-based framework used to create a micro Service; It is developed by Pivotal Team and is used to built stand-alone and production ready spring applications.

H2O + JAVA

*"H2O allows you to convert the models you have built to either a **Plain Old Java Object (POJO) or a Model ObJect, Optimized (MOJO)**.*

...

H2O-generated MOJO and POJO models are intended to be easily embeddable in any Java environment. "

http://docs.h2o.ai/h2o/latest-stable/h2o-docs/productionizing.html#productionizing-h2o

H2O + JAVA

- **POJO**
  - Some customers encountered issues with large POJOs not compiling (note that POJOs are not supported for source files larger than 1G)
  - Do not score in real-time
  - Needs to be compiled inside the project

- **MOJO**
  - Can then be deployed for scoring in real time
  - MOJOs do not have a size restriction
  - The resulting executable is much smaller and faster than a POJO
  - At large scale, new models are roughly 20-25 times smaller in disk space, 2-3 times faster

http://docs.h2o.ai/h2o/latest-stable/h2o-docs/productionizing.html#productionizing-h2o

Spring Boot

- **Auto Configuration** - Basically Spring Boot auto-configures dependencies just by looking at the class path. It uses sensible defaults and makes a lot of choices on its own that are mostly battle-tested.

- **Starter Modules** - Spring Boot comes with a bunch of starter modules. Each starter module basically supports a feature. This makes it easy to plug and play with new features. You simply include the package in a POM or Gradle file and Spring Boot will try to provide you with that feature.

5 Essential Requirements for a Spring Boot Application in Production

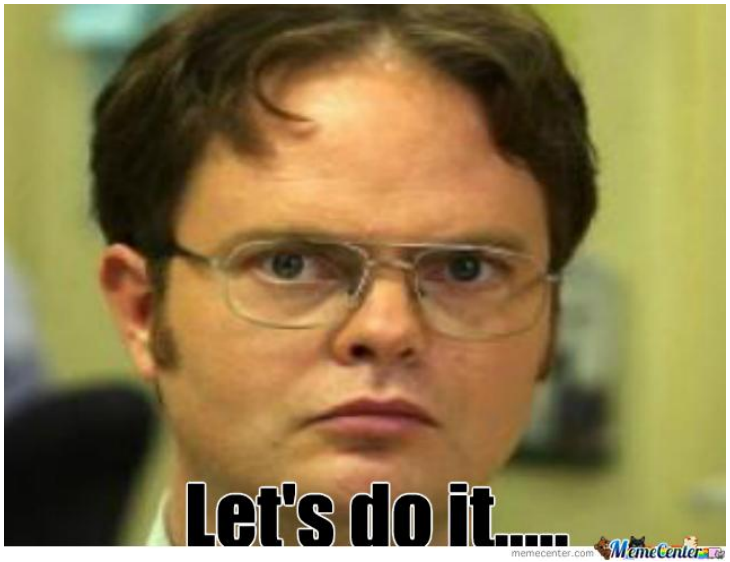
1. Handling RESTful HTTP Endpoints
2. Exception Handling
3. Validations
4. Documentation
5. Logging

All of this saves a developer's time and increase efficiency.

Spring Boot Initializer

Bootstrap your application at https://start.spring.io/

Deploy a H2O model into a RESTful microservice in 10 minutes



Live Demo

https://github.com/serafimpinto/spring-boot-h2o/

Data Science & Machine Learning

Productionizing H2O

geral@performetric.net