

## 在Hadoop 上运行Tensorflow

### 官网帮助

<https://github.com/tensorflow/examples/blob/master/community/en/docs/deploy/hadoop.md>

```
export JAVA_HOME=/usr/java/jdk1.8.0_192
export HADOOP_HOME=/opt/cloudera/parcels/CDH/lib/hadoop
source $HADOOP_HOME/libexec/hadoop-config.sh
export
LD_LIBRARY_PATH=${LD_LIBRARY_PATH}:${JAVA_HOME}/jre/lib/amd64/server:/opt/cloudera/parcels/CDH/lib/hadoop/lib/native:/opt/
export CLASSPATH=${HADOOP_HOME}/bin/hadoop classpath --glob:${CLASSPATH}
```

```
/etc/alternatives/spark-submit \
--master yarn \
--deploy-mode client \
--queue xy_yarn_pool.development \
--num-executors 2 \
--executor-memory 2G \
--conf spark.executorEnv.LD_LIBRARY_PATH=${LD_LIBRARY_PATH} \
--conf spark.executorEnv.CLASSPATH=${CLASSPATH} \
/tmp/keras_inference_.py \
--cluster_size 3 \
--images_labels /tmp/TensorFlowOnSpark/examples/mnist/tfr/test \
--export hdfs://bigdata-1.baofoo.cn:8020/user/xy_hpt/best_model.pb \
--output / \
--test_file /user/xy_hpt/test.txt
```

### #cluster模式

```
/etc/alternatives/spark-submit \
--master yarn \
--deploy-mode cluster \
--queue xy_yarn_pool.development \
--num-executors 2 \
--executor-memory 2G \
--conf spark.executorEnv.LD_LIBRARY_PATH=${LD_LIBRARY_PATH} \
--conf spark.executorEnv.CLASSPATH=${CLASSPATH} \
hdfs:///tmp/keras_inference_.py \
--cluster_size 3 \
--images_labels /tmp/TensorFlowOnSpark/examples/mnist/tfr/test \
--export hdfs://bigdata-1.baofoo.cn:8020/user/xy_hpt/best_model.pb \
--output / \
--test_file /user/xy_hpt/test.txt
```

### #-----线上环境

```
export JAVA_HOME=/usr/java/jdk1.8.0_141-cloudera
export HADOOP_HOME=/opt/cloudera/parcels/CDH/lib/hadoop
source $HADOOP_HOME/libexec/hadoop-config.sh
export
LD_LIBRARY_PATH=${LD_LIBRARY_PATH}:${JAVA_HOME}/jre/lib/amd64/server:/opt/cloudera/parcels/CDH/lib/hadoop/lib/native:/opt/
export CLASSPATH=${HADOOP_HOME}/bin/hadoop classpath --glob:${CLASSPATH}
```

### --client模式

```
spark-submit \
--master yarn \
--deploy-mode client \
--queue xy_yarn_pool.development \
--conf spark.ui.port=5051 \
```

```
--num-executors 2 \  
--executor-memory 2G \  
--conf spark.executorEnv.LD_LIBRARY_PATH=${LD_LIBRARY_PATH} \  
--conf spark.executorEnv.CLASSPATH=${CLASSPATH} \  
/home/xy_yq/keras_inference_.py \  
--cluster_size 3 \  
--export hdfs://ns1/user/xy_yq/best_model.pb \  
--test_file hdfs:///user/xy_yq/test.txt \  
--word_index hdfs://ns1/user/xy_yq/nlp/models/word_index1.json
```

--cluster模式

```
spark-submit \  
--master yarn \  
--deploy-mode cluster \  
--queue xy_yarn_pool.development \  
--conf spark.ui.port=5051 \  
--num-executors 2 \  
--executor-memory 2G \  
--conf spark.executorEnv.LD_LIBRARY_PATH=${LD_LIBRARY_PATH} \  
--conf spark.executorEnv.CLASSPATH=${CLASSPATH} \  
hdfs://ns1/user/xy_yq/keras_inference_.py \  
--cluster_size 3 \  
--export hdfs://ns1/user/xy_yq/best_model.pb \  
--test_file hdfs:///user/xy_yq/test.txt \  
--word_index hdfs://ns1/user/xy_yq/nlp/models/word_index1.json
```