scala\_library(

platform = "java8",

tags = ["bazel-compatible"],

dependencies = [

":tweet\_similarity\_test\_datarecords\_120min-java",

":tweet\_similarity\_test\_datarecords\_30min-java",

":tweet\_similarity\_train\_datarecords\_120min-java",

":tweet\_similarity\_train\_datarecords\_30min-java",

":tweet\_similarity\_unhydrated\_pairs\_120min-scala",

":tweet\_similarity\_unhydrated\_pairs\_30min-scala",

"3rdparty/jvm/com/twitter/storehaus:algebra",

"3rdparty/jvm/com/twitter/storehaus:core",

"dataproducts/insights/common/common",

"snowflake:id",

"src/java/com/twitter/ml/api/constant",

"src/scala/com/twitter/ads/dataservice\_account/snapshot/jobs:db\_snapshots\_promoted\_tweets-scala",

"src/scala/com/twitter/ml/api:api-base",

"src/scala/com/twitter/ml/featurestore/catalog/features/recommendations:aggregate",

"src/scala/com/twitter/ml/featurestore/lib/embedding",

"src/scala/com/twitter/scalding\_internal/dalv2",

"src/scala/com/twitter/scalding\_internal/dalv2/dataset",

"src/scala/com/twitter/scalding\_internal/job",

"src/scala/com/twitter/scalding\_internal/job/analytics\_batch",

"src/scala/com/twitter/scalding\_internal/source",

"src/scala/com/twitter/scalding\_internal/source/lzo\_scrooge",

"src/scala/com/twitter/simclusters\_v2/common",

"src/scala/com/twitter/simclusters\_v2/scalding",

"src/scala/com/twitter/simclusters\_v2/tweet\_similarity",

"src/scala/com/twitter/wtf/scalding/jobs/client\_event\_processing:user\_interaction-scala",

"twadoop\_config/configuration/log\_categories/group/timeline:timeline\_service\_favorites-scala",

"tweetsource/common:unhydrated\_flat-scala",

],

)

hadoop\_binary(

name = "training\_data\_collection-adhoc",

main = "com.twitter.simclusters\_v2.scalding.tweet\_similarity.TrainingDataCollectionAdhocApp",

platform = "java8",

runtime\_platform = "java8",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

"bazel-only",

],

dependencies = [

":tweet\_similarity",

],

)

hadoop\_binary(

name = "training\_data\_collection\_30min",

main = "com.twitter.simclusters\_v2.scalding.tweet\_similarity.TrainingDataCollection30MinScheduledApp",

platform = "java8",

runtime\_platform = "java8",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

"bazel-only",

],

dependencies = [

":tweet\_similarity",

],

)

hadoop\_binary(

name = "training\_data\_collection\_120min",

main = "com.twitter.simclusters\_v2.scalding.tweet\_similarity.TrainingDataCollection120MinScheduledApp",

platform = "java8",

runtime\_platform = "java8",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

"bazel-only",

],

dependencies = [

":tweet\_similarity",

],

)

hadoop\_binary(

name = "unhydrated\_pair\_collection-adhoc",

main = "com.twitter.simclusters\_v2.scalding.tweet\_similarity.UnhydratedPairsCollectionAdhocApp",

platform = "java8",

runtime\_platform = "java8",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

"bazel-only",

],

dependencies = [

":tweet\_similarity",

],

)

hadoop\_binary(

name = "unhydrated\_pair\_collection\_30min",

main = "com.twitter.simclusters\_v2.scalding.tweet\_similarity.UnhydratedPairsCollection30MinScheduledApp",

platform = "java8",

runtime\_platform = "java8",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

"bazel-only",

],

dependencies = [

":tweet\_similarity",

],

)

hadoop\_binary(

name = "unhydrated\_pair\_collection\_120min",

main = "com.twitter.simclusters\_v2.scalding.tweet\_similarity.UnhydratedPairsCollection120MinScheduledApp",

platform = "java8",

runtime\_platform = "java8",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

"bazel-only",

],

dependencies = [

":tweet\_similarity",

],

)

hadoop\_binary(

name = "model\_eval-adhoc",

main = "com.twitter.simclusters\_v2.scalding.tweet\_similarity.ModelEvalAdhocApp",

platform = "java8",

runtime\_platform = "java8",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

"bazel-only",

],

dependencies = [

":tweet\_similarity",

],

)

hadoop\_binary(

name = "dataset\_topk\_analysis-adhoc",

main = "com.twitter.simclusters\_v2.scalding.tweet\_similarity.DatasetTopKAnalysisAdhocApp",

platform = "java8",

runtime\_platform = "java8",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

"bazel-only",

],

dependencies = [

":tweet\_similarity",

],

)

hadoop\_binary(

name = "dataset\_topk\_analysis\_dump-adhoc",

main = "com.twitter.simclusters\_v2.scalding.tweet\_similarity.DatasetTopKAnalysisDumpApp",

platform = "java8",

runtime\_platform = "java8",

tags = [

"bazel-compatible",

"bazel-compatible:migrated",

"bazel-only",

],

dependencies = [

":tweet\_similarity",

],

)

create\_datarecord\_datasets(

base\_name = "tweet\_similarity\_train\_datarecords\_30min",

platform = "java8",

role = "cassowary",

segment\_type = "partitioned",

tags = ["bazel-compatible"],

)

create\_datarecord\_datasets(

base\_name = "tweet\_similarity\_test\_datarecords\_30min",

platform = "java8",

role = "cassowary",

segment\_type = "partitioned",

tags = ["bazel-compatible"],

)

create\_datarecord\_datasets(

base\_name = "tweet\_similarity\_train\_datarecords\_120min",

platform = "java8",

role = "cassowary",

segment\_type = "partitioned",

tags = ["bazel-compatible"],

)

create\_datarecord\_datasets(

base\_name = "tweet\_similarity\_test\_datarecords\_120min",

platform = "java8",

role = "cassowary",

segment\_type = "partitioned",

tags = ["bazel-compatible"],

)

create\_datasets(

base\_name = "tweet\_similarity\_unhydrated\_pairs\_30min",

description = "30min coocurrence training pairs before feature hydration",

java\_schema = "com.twitter.simclusters\_v2.thriftjava.LabelledTweetPairs",

platform = "java8",

role = "cassowary",

scala\_schema = "com.twitter.simclusters\_v2.thriftscala.LabelledTweetPairs",

segment\_type = "partitioned",

tags = ["bazel-compatible"],

java\_dependencies = [

"src/thrift/com/twitter/simclusters\_v2:simclusters\_v2-thrift-java",

],

scala\_dependencies = [

"src/thrift/com/twitter/simclusters\_v2:simclusters\_v2-thrift-scala",

],

)

create\_datasets(

base\_name = "tweet\_similarity\_unhydrated\_pairs\_120min",

description = "120min coocurrence training pairs before feature hydration",

java\_schema = "com.twitter.simclusters\_v2.thriftjava.LabelledTweetPairs",

platform = "java8",

role = "cassowary",

scala\_schema = "com.twitter.simclusters\_v2.thriftscala.LabelledTweetPairs",

segment\_type = "partitioned",

tags = ["bazel-compatible"],

java\_dependencies = [

"src/thrift/com/twitter/simclusters\_v2:simclusters\_v2-thrift-java",

],

scala\_dependencies = [

"src/thrift/com/twitter/simclusters\_v2:simclusters\_v2-thrift-scala",

],

)