package com.twitter.ann.hnsw

import com.twitter.ann.common.RuntimeParams

import com.twitter.ann.common.thriftscala.HnswIndexMetadata

import com.twitter.ann.common.thriftscala.HnswRuntimeParam

import com.twitter.ann.common.thriftscala.{RuntimeParams => ServiceRuntimeParams}

import com.twitter.bijection.Injection

import com.twitter.mediaservices.commons.codec.ThriftByteBufferCodec

import com.twitter.search.common.file.AbstractFile

import scala.util.Failure

import scala.util.Success

import scala.util.Try

object HnswCommon {

private[hnsw] lazy val MetadataCodec = new ThriftByteBufferCodec(HnswIndexMetadata)

private[hnsw] val MetaDataFileName = "hnsw\_index\_metadata"

private[hnsw] val EmbeddingMappingFileName = "hnsw\_embedding\_mapping"

private[hnsw] val InternalIndexDir = "hnsw\_internal\_index"

private[hnsw] val HnswInternalMetadataFileName = "hnsw\_internal\_metadata"

private[hnsw] val HnswInternalGraphFileName = "hnsw\_internal\_graph"

val RuntimeParamsInjection: Injection[HnswParams, ServiceRuntimeParams] =

new Injection[HnswParams, ServiceRuntimeParams] {

override def apply(scalaParams: HnswParams): ServiceRuntimeParams = {

ServiceRuntimeParams.HnswParam(

HnswRuntimeParam(

scalaParams.ef

)

)

}

override def invert(thriftParams: ServiceRuntimeParams): Try[HnswParams] =

thriftParams match {

case ServiceRuntimeParams.HnswParam(hnswParam) =>

Success(

HnswParams(hnswParam.ef)

)

case p => Failure(new IllegalArgumentException(s"Expected HnswRuntimeParam got $p"))

}

}

def isValidHnswIndex(path: AbstractFile): Boolean = {

path.isDirectory &&

path.hasSuccessFile &&

path.getChild(MetaDataFileName).exists() &&

path.getChild(EmbeddingMappingFileName).exists() &&

path.getChild(InternalIndexDir).exists() &&

path.getChild(InternalIndexDir).getChild(HnswInternalMetadataFileName).exists() &&

path.getChild(InternalIndexDir).getChild(HnswInternalGraphFileName).exists()

}

}

/\*\*

\* Hnsw runtime params

\* @param ef: The size of the dynamic list for the nearest neighbors (used during the search).

\* Higher ef leads to more accurate but slower search.

\* ef cannot be set lower than the number of queried nearest neighbors k.

\* The value ef of can be anything between k and the size of the dataset.

\*/

case class HnswParams(ef: Int) extends RuntimeParams {

override def toString: String = s"HnswParams(ef = $ef)"

}