package com.twitter.simclusters\_v2.summingbird.common

import com.twitter.simclusters\_v2.thriftscala.EmbeddingType

import com.twitter.util.Duration

import com.twitter.conversions.DurationOps.\_

import com.twitter.simclusters\_v2.thriftscala.ModelVersion

case class ModelVersionProfile(

modelVersion: ModelVersion,

usingLogFavScore: Boolean,

// redundant in the current models because the above parameter does the same currently.

coreEmbeddingType: EmbeddingType,

favScoreThresholdForUserInterest: Double,

// these values are shared between all profiles so lets set up defaults

halfLife: Duration = 8.hours,

scoreThresholdForEntityTopKClustersCache: Double = 0.2,

scoreThresholdForTweetTopKClustersCache: Double = 0.02,

scoreThresholdForClusterTopKTweetsCache: Double = 0.001,

scoreThresholdForClusterTopKEntitiesCache: Double = 0.001)

object ModelVersionProfiles {

final val ModelVersion20M145KUpdated = ModelVersionProfile(

ModelVersion.Model20m145kUpdated,

usingLogFavScore = true,

coreEmbeddingType = EmbeddingType.LogFavBasedTweet,

favScoreThresholdForUserInterest = 1.0

)

final val ModelVersion20M145K2020 = ModelVersionProfile(

ModelVersion.Model20m145k2020,

usingLogFavScore = true,

coreEmbeddingType = EmbeddingType.LogFavBasedTweet,

favScoreThresholdForUserInterest = 0.3

)

final val ModelVersionProfiles: Map[ModelVersion, ModelVersionProfile] = Map(

ModelVersion.Model20m145kUpdated -> ModelVersion20M145KUpdated,

ModelVersion.Model20m145k2020 -> ModelVersion20M145K2020

)

}