package com.twitter.search.ingester.pipeline.twitter;

import org.apache.commons.pipeline.StageException;

import org.apache.commons.pipeline.validation.ConsumedTypes;

import org.apache.commons.pipeline.validation.ProducesConsumed;

import com.twitter.search.common.relevance.classifiers.TweetQualityFeatureExtractor;

import com.twitter.search.ingester.model.IngesterTwitterMessage;

@ConsumedTypes(IngesterTwitterMessage.class)

@ProducesConsumed

public class ComputeTweetSignatureStage extends TwitterBaseStage

<IngesterTwitterMessage, IngesterTwitterMessage> {

private final TweetQualityFeatureExtractor tweetSignatureExtractor =

new TweetQualityFeatureExtractor();

@Override

public void innerProcess(Object obj) throws StageException {

if (!(obj instanceof IngesterTwitterMessage)) {

throw new StageException(this, "Object is not a TwitterMessage instance: " + obj);

}

IngesterTwitterMessage message = IngesterTwitterMessage.class.cast(obj);

extract(message);

emitAndCount(message);

}

private void extract(IngesterTwitterMessage message) {

tweetSignatureExtractor.extractTweetTextFeatures(message);

}

@Override

protected IngesterTwitterMessage innerRunStageV2(IngesterTwitterMessage message) {

extract(message);

return message;

}

}