Data Processing with U-SQL – Cognitive Imaging

Updated on: 1/31/2017

# Introduction

## Ensure the Advanced Analytics is Installed

In the portal, dgo to your ADLA account.

Click on Sample Scripts.

You may see a notification that you need to install or update the analytics extensions. Click it.

## Copy the Imaging Data

The MegaFace dataset is stored in a shared ADLS account as a single large CSV file.

Run the Script below. It will create a database Called Images that has a table called MegaFace that contains the data,

CREATE DATABASE Images;

DROP TABLE IF EXISTS Images.dbo.MegaFace;

CREATE TABLE Images.dbo.MegaFace (

FileName string, ImgData byte[],

INDEX idx

CLUSTERED (FileName ASC)

PARTITIONED BY HASH (FileName)

);

@rs =

EXTRACT

FileName string,

ImgData byte[]

FROM "adl://adltrainingsampledata.azuredatalakestore.net/MegaFace/MegaFace.tsv"

USING Extractors.Tsv();

INSERT INTO Images.dbo.MegaFace

SELECT \*

FROM @rs;

The 1 AU the script above should finish in about 30 minutes.

The MegaFace table has two columns

* FileName – the name of the image file
* ImgData – the binary representation of the image

REFERENCE ASSEMBLY ImageCommon;

REFERENCE ASSEMBLY FaceSdk;

REFERENCE ASSEMBLY ImageEmotion;

REFERENCE ASSEMBLY ImageTagging;

REFERENCE ASSEMBLY ImageOcr;

@imgs =

SELECT

FileName,

ImgData

FROM Images.dbo.MegaFace;

@objects =

PROCESS @imgs

PRODUCE FileName,

NumObjects int,

Tags string

READONLY FileName

USING new Cognition.Vision.ImageTagger();

OUTPUT @objects

TO "/objects.csv"

USING Outputters.Csv();

# The Challenge

This is a MAJOR CHALLENGE and requires using U-SQL and the SDks.

Summary: using the Megaface dataset, locate all the images that contain a “necktie”. And then place JPB thumbnail versions (no bigger than 300x300 pixels) of those images in the account adltrainingsampledata/users/<yourusername>/neckties

Hints:

* Use U-SQL to perform the image tagging
* Use the ADL SDK to extract the images and place them in the correct destination folder
* Validate that the images are correct by downloading a few and viewing on your local machine
* For thumbnail creation see: <https://msdn.microsoft.com/en-us/library/system.drawing.image.getthumbnailimage%28v=vs.110%29.aspx>