COMP-SCI 5542 (SP17) - Big Data Analytics and Applications

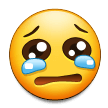
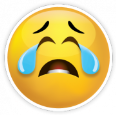
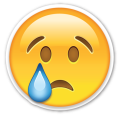
**Tutorial 5 Assignment (Due 02/22/17 by 11:59 PM)**

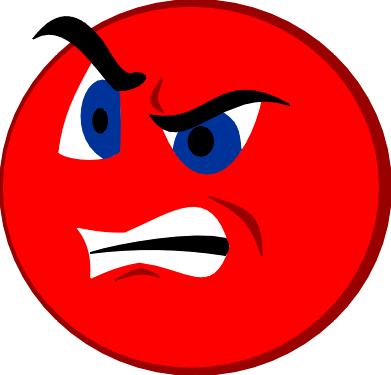
***Dayu Wang* (45)**

1. **Spark Programming - Image Classification**
2. **Description of Dataset**

The dataset consists of **emoji** icons. The data is separated into 3 classes, the *smiling* icons, *angry* icons and *crying* icons. I used the method mentioned in tutorial class and **Random Forest** algorithm. Figure 1 lists some images in my dataset.







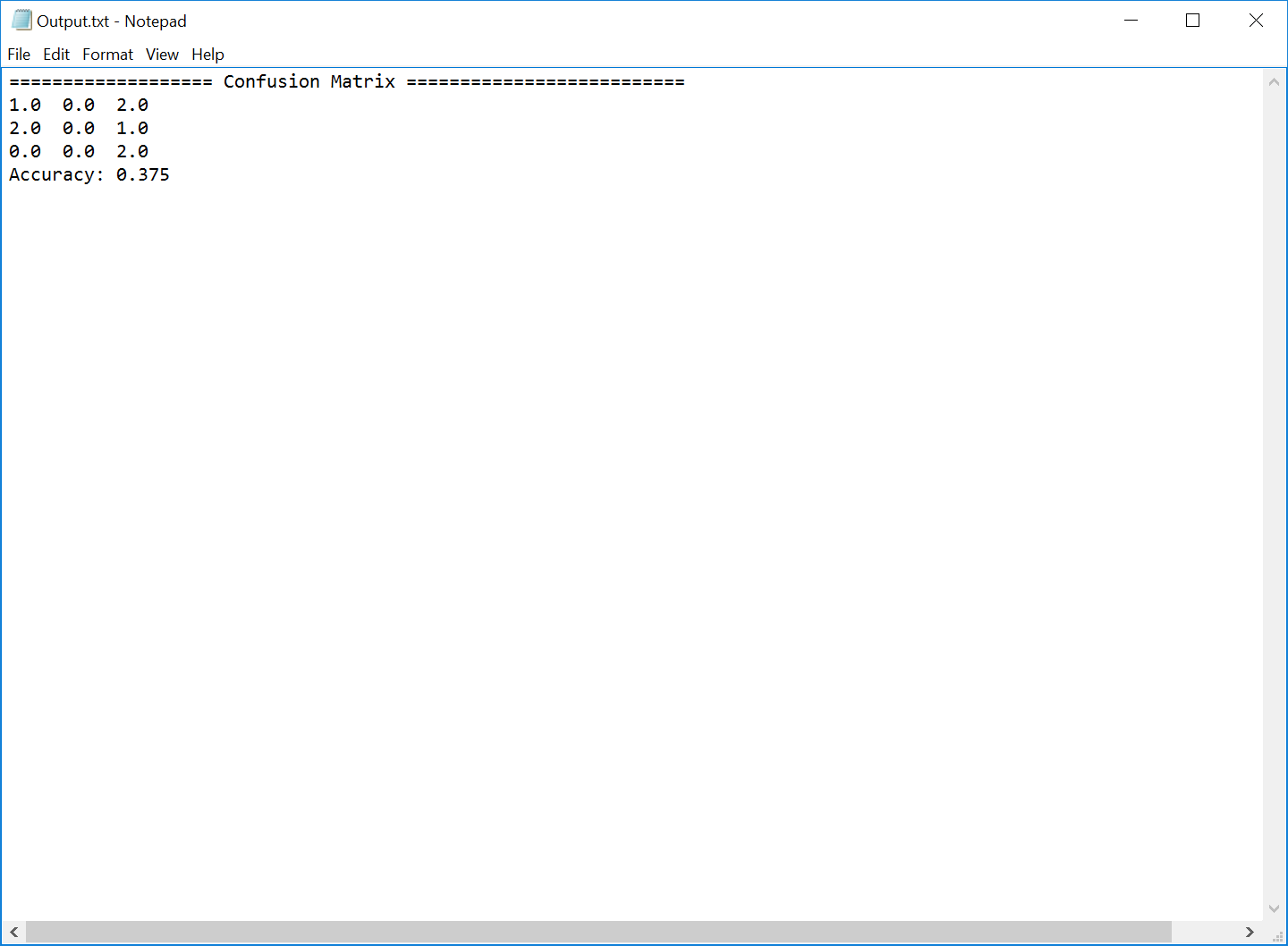
**Figure 1**. Examples of my dataset of **emoji** of smiling, crying, and angry.

There are **60** training images (20 for each class) and **9** test images (3 for each class).

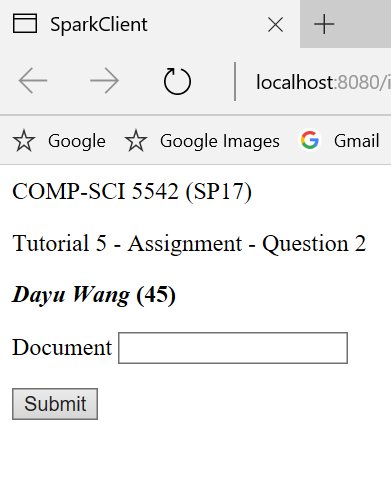
1. **Purpose behind Image Classification**

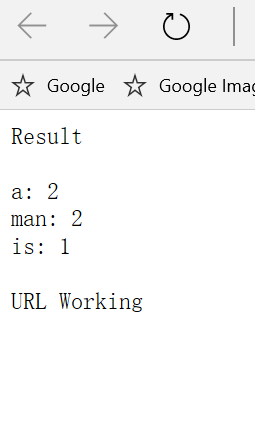
Image classification applied on **emoji** impressions can significantly help researches on cross-platform emoji interpretation area. It opens a novel way to recognize and translate new impressions, with the rapid development of new platform/devices nowadays.

1. **Accuracy and Confusion Matrix**



1. **Spart Client - Web System - Word Count**





1. **Google Conversion**

