Bid Data Mining Homework 2

Team Member:

SID: 109598033 SID: 109598001

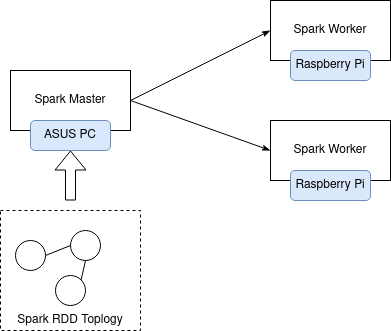
**Spark Platform:**

The platform consists of:  
1. Raspberry Pi 4 Model B x2

* OS: Linux Ubuntu 20.04 Server
* CPU architecture: aarch64
* RAM: 8GB
* CPU: Broadcom BCM2711, Quad core Cortex-A72 (ARM v8) 64-bit SoC @ 1.5GHz
* Number of CPU: 4C (CPU) 1T (Thread Per CPU)

2. Asus-vivobook notebook

* OS: Linux Ubuntu 20.04 LTS
* CPU architecture: x86\_64
* RAM: 8GB
* CPU: Intel(R) Core(TM) i3-8130U CPU @ 2.20GHz
* Number of CPU: 4C (CPU) 2T (Thread Per CPU)

The simple architecture of spark cluster:

**Task arrangement for Team:   
SID: 109598033  
1. Q1  
2. Q4  
SID: 109598001  
1. Q2  
2. Q3**

**The description of tree directory:**

Explain where the files put into and how it does.

* **README file explains in detail for what the steps of spark implementation in this homework.**
* **‘outputs’ directory shows the result of the homework.**
* **park project is ‘hw2.ipynb’ which codes by Python.**

**The generated output:**

**Q1**

一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述

一張含有 桌 的圖片

自動產生的描述

q1\_news\_title\_total

q1\_news\_title\_total\_date

q1\_news\_title\_total\_cate

一張含有 桌 的圖片

自動產生的描述**Q2**

q1\_news\_headline\_total\_cate

q1\_news\_headline\_total\_date

q1\_news\_headline\_total

一張含有 桌 的圖片

自動產生的描述

一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述

q2\_popular\_hour

q2\_popular\_day

**Q3**

一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述

q3\_Title\_sentiment\_score

q3\_Headline\_sentiment\_score

**Q4**

一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述

q4\_Title\_microsoft\_co-occurrence\_matrix

q4\_Title\_economy\_co-occurrence\_matrix

一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述

q4\_Title\_obama\_co-occurrence\_matrix

q4\_Title\_palestine\_co-occurrence\_matrix

一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述

q4\_Headline\_microsoft\_co-occurrence\_matrix

q4\_Headline\_economy\_co-occurrence\_matrix

一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述

q4\_Headline\_palestine\_co-occurrence\_matrix

q4\_Headline\_obama\_co-occurrence\_matrix