**Estimating the Effects of Major PC Components on the Execution Time of an Image Recognition Task**

submitted to:

**Mr. Leonid Shpaner**

University of San Diego

February 5, 2024

submitted by:

**Arup Chakraborty**

**Issa Ennab**

**Geoffrey Fadera**

Final Project Group 1 Technical Report

AAI-500 Spring 2024

1. **Introduction**
2. **Data Cleaning/Preparation**
3. **Exploratory Data Analysis**
4. **Model Selection**

In deriving the model, we have one major assumption – i.e. the execution time of an image processing task is not affected by time. We assume that any PC will always take the same amount of time to execute an image recognition task no matter how long that PC has been running. This is quite contrary to the general perception that PCs slow down over time as explored in { <https://www.usenix.org/legacy/event/sysml07/tech/full_papers/basu/basu.pdf>}

Working on this assumption, we then removed the timestamp from the original data, and assume that each data point is independent of each other.

After removing the time stamp, the execution time density plots for MacBookPro1 and VM show bell-shaped distributions as shown in Fig 4.1 below. This suggests that time had no effect on the execution time.

A graph of a computer

Description automatically generated

Fig 4.1 Density Plots of Execution Time Values Assuming No Time-Relation for VM and MacBookPro1 PCs

However, the density plots for MacBookPro2 and RaspBerry Pi show irregular shapes (Fig 4.2). They appear to be a combination of multiple normal distributions.

A graph of a number of data

Description automatically generated with medium confidence

Fig 4.2 Density Plots of Execution Time Values Assuming No Time-Relation for MacBookPro 2 and Raspberry Pi PCs

With N=1000, the distribution plots for each PC should have taken a normal shape by the Central Limit Theorem (citation). Therefore, there must be other factors that’s causing the execution time to appear to be coming from different population distributions. One factor could be time. A time series analysis could expose any trend or seasonality on the data. Another explanation could be of data collection error - i.e maybe other tasks were executed on these machines resulting in an increase in observed execution time – which could also be exposed by time series analysis.

We attempted to perform time-series analysis on our data, and we present this in Section/Appendix {}…

For the scope of this technical paper, however, we decided to proceed with our assumption and ignore the effect of time. Using the general linear model, we compute the effects of specific PC components to its performance in executing an image recognition task.

where the exploratory variables are:

: CPU Clock Speed

: CPU Core count

: RAM size

1. **Model Analysis**
2. **Conclusions and Recommendations**