**Mata, Algreg M.**

**Toring, Dioven Ian**

**Dataset**

A java Machine learning classifier was trained and tested with the dataset of 5,534 product ratings from Kaggle. The dataset was subdivided from training and testing that has 80-20 proportion respectively. Where 4,427 of dataset was used in training and 1,107 was used in the testing. The ratings were preprocessed and grouped into as “low” or “high” rating.

**RESULTS**

|  |  |
| --- | --- |
| **Total Datasets** | 11,068 |
| **Total Training Dataset** | 8,854 |
| **Total Testing Dataset** | 2,214 |

**Confusion Matrix**

|  |  |  |
| --- | --- | --- |
| **n = 2,214** | **High** | **Low** |
| **High** | 544 | 563 |
| **Low** | 342 | 765 |

**Rates**

|  |  |
| --- | --- |
| **Accuracy** | 59.123% |