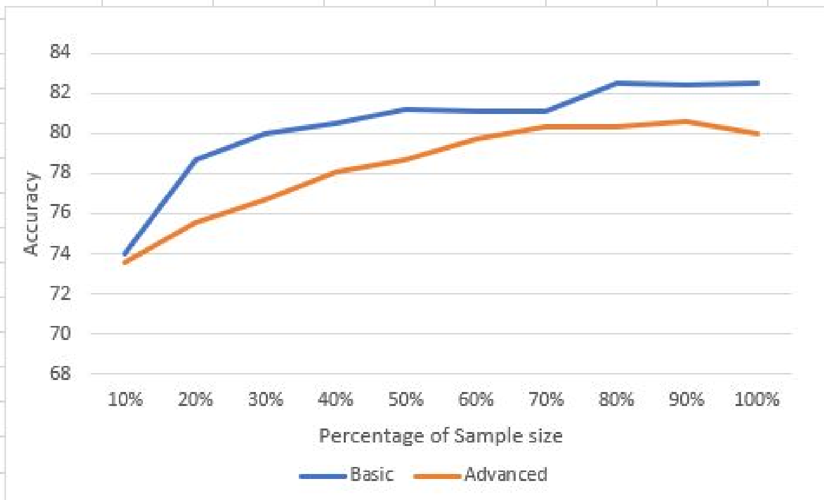
**A.2.**

Given the large number of features drawbacks that could occur from using the standard multiplicative representation are that because there are so many probabilities which would have to be multiplied, it would result in underflow.

**A.5.**

In general, the trend says the more samples you have to scan and draw predictions from, the better your results will be.

| k = 0.00001 | | |
|---|---|---|
| | Basic | Advanced |
| 10% | 74 | 73.6 |
| 20% | 78.7 | 75.6 |
| 30% | 80 | 76.7 |
| 40% | 80.5 | 78.1 |
| 50% | 81.2 | 78.7 |
| 60% | 81.1 | 79.7 |
| 70% | 81.1 | 80.3 |
| 80% | 82.5 | 80.3 |
| 90% | 82.4 | 80.6 |
| 100% | 82.5 | 80 |
| | | |
| AVERAGE | 80.4 | 78.36 |




**A.6.**

The k value we chose was 0.00001 for the laplace smoothing algorithm applied on the data set. Our classifier works by determining the sum of the pixels from the image scanned on top of the log maps of each integer from 0 to 9. The higher the sum, the better the chance of it being the correct number. Result wise, it could be better, if we can extract features in a more refined manner or have more test samples.

| k values | basic at 100% | adv at 100% |
|---|---|---|
| 0.0000001 | 82.4 | 80.1 |
| 0.000001 | 82.4 | 80.1 |
| **0.00001** | **82.5** | **80** |
| 0.0001 | 82.5 | 80 |
| 0.001 | 82.1 | 80.1 |
| 0.01 | 81.8 | 80.1 |
| 0.1 | 81.7 | 79.6 |
| 1 | 81.6 | 79 |
| 2 | 81.6 | 78.8 |
| 5 | 80.9 | 77.9 |
| 10 | 80.1 | 77.2 |
| 20 | 78.9 | 75.6 |
| 40 | 77.1 | 71.9 |
| 100 | 71.3 | 60.9 |
| 500 | 49.9 | 32.1 |

**B.1**
Features V1 = This was an attempt to remove the noise that may exist and cause problems in scanning by having a lower count in the middle of the image. This works by checking if all surrounding values are opposite of current position, and if true, make it match. When active by itself, it results in an 80.6% accuracy.

Features V2 = instead of taking in both 1s and 2s for features, only accept 1s. This was because edges are more prone to error. When active by itself, it results in an 80.7% accuracy

Feature V3 = average of V1 and V2 so pixels that are not meant to exist will have a lower frequency rate in the final testing.

**B.2**
In the effort put in to bring out features better, we could not figure out how we lost accuracy. As shown in the charts above however, both extraction methods yield almost the same results with the Advanced method being less superior.