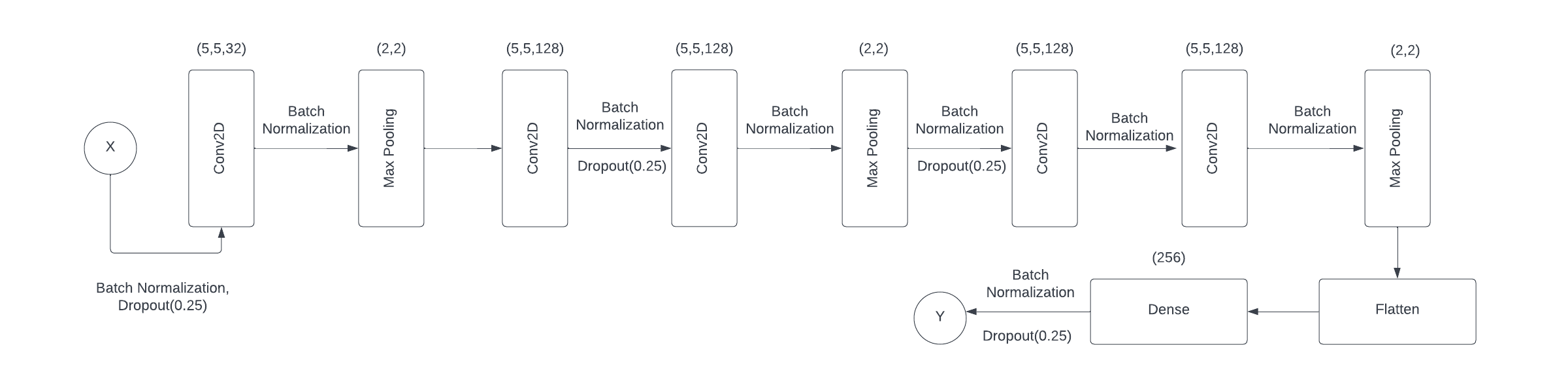
| Model # | Hidden Layers | Optimizer | Batch\_size | Dropout | Train Accuracy | Test Accuracy |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 6  (32, 64, 64, 128, 128, 256) | rmsprop | 128 | 0.25 | 90.20 | 77.63 |
| 2 | 6 (changed filter size)  (32, 64, 64, 128, 128, 256) | rmsprop | 128 | 0.25 | 96.32 | 81.32 |
| 3 | 6  (32, 128, 128, 128, 128, 256) | adam | 128 | 0.25 | 98.45 | 81.52 |
| 4 | 6  (32, 32, 32, 32, 32, 32) | ada | 128 | 0.25 | 51.46 | 50.32 |
| 5 | 8  (256, 256, 512, 512, 128, 128, 64) | Adagrad | 128 | 0.25 | 62.74 | 60.27 |
| 6 | 8  (32, 64, 64, 64, 64, 64, 64, 64) | adagrad | 256 | 0.25 | 40.29 | 30.75 |
| 7 | 10  (32, 64, 64, 64, 64, 64, 64, 128, 128, 256) | rmsprop | 128 | 0.25 | 92.21 | 82.80 |
| 8 | 10  (32, 64, 64, 64, 64, 64, 64, 128, 128, 256) | rmsprop | with data augmentation | 0.25 | 39.85 | 38.95 |
| 9 | 6  (32, 128, 128, 128, 128, 128) | rmsprop | 128 | 0.25 | 97.61 | 81.43 |

## Top 3 Architectures

### Model 3



* In this architecture there is a combination of ‘elu’, ‘relu’ and ‘softmax’ activation functions
* The optimizer used for this model is adam

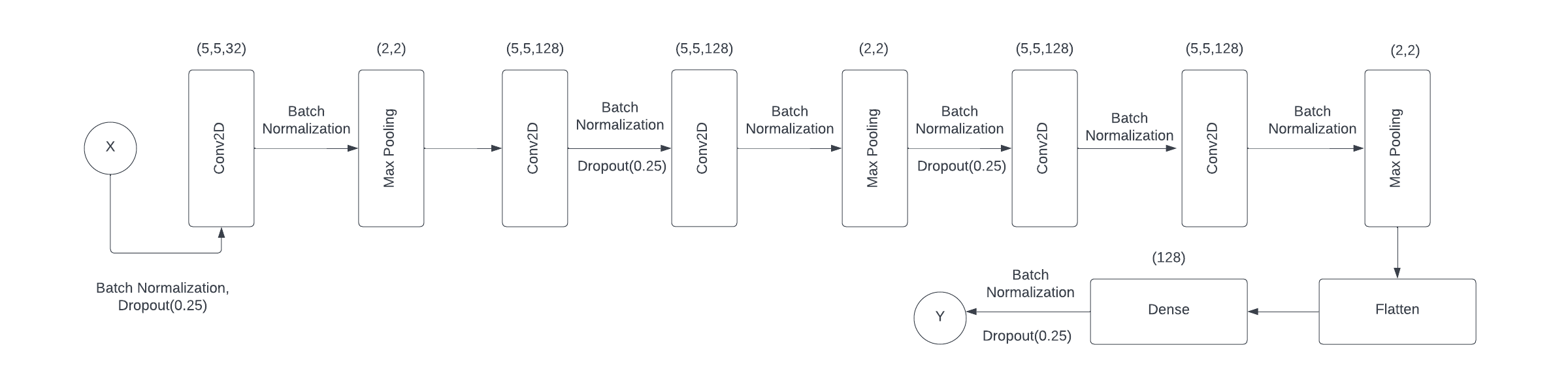
### Model 7

### 

* In this architecture there is a combination of ‘elu’, ‘relu’ and ‘softmax’ activation functions.
* The optimizer used for this model is rmsprop.

### 

### Model 9



* In this architecture there is a combination of ‘elu’, ‘relu’ and sigmoid’ activation functions
* The optimizer used for this model is rmsprop