**Model Evaluation**

**Introduction**

This model is a solution for mushroom classification to classify whether a mushroom is edible or poisonous. The metrics measured in the model is the mean validation accuracy and the mean validation loss. The accuracy is of pinnacle importance to the model as it allows the overall outcome to result in a better decision. The accuracy provides confidence in the consumption of mushrooms as it mitigates the majority of a risk occurring from the potential consumption.

**Data Augmentation**

Data augmentation is used to increase the size of the train set by generating augmented versions of the existing mushroom photos, with the implementation of data augmentation increased the accuracy of the model from 77.20% to 77.84%. Our findings are supported by a paper conducted by Shijie et al. Their study concluded that data augmentation showed an increase in mean validation accuracy in Convolutional Neural Networks due to the increase of train set.

**Activation Functions**

Activation functions determine an output value of the node based on the set of inputs from the previous layer (Apicella, A. et al., 2021). The popular activation functions which are used in deep learning models that are non-linear are sigmoid, ReLU, tanh and SoftMax (Apicella, A. et al., 2021).

A study of the model of the impact on combination activation functions had on the models mean validation accuracy and mean validation loss. The combination of activation functions was measured at a 0.2 dropout rate using the Adam optimizer.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Layer One | Layer Two | Layer Three | Layer Four | Mean Validation Accuracy (%) | Mean Validation Loss (%) | Mean Training Accuracy (%) | Mean Training Loss (%) |
| ReLU | ReLU | ReLU | ReLU | 72.20 | 50.90 | 73.11 | 52.73 |
| tanh | tanh | tanh | ReLU | 78.14 | 49.34 | 72.38 | 55.00 |
| SoftMax | SoftMax | SoftMax | ReLU | 78.39 | 55.38 | 73.30 | 61.94 |
| sigmoid | sigmoid | sigmoid | ReLU | 78.39 | 55.22 | 70.88 | 88.04 |
| ReLU | ReLU | ReLU | tanh | 77.12 | 48.18 | 72.71 | 59.90 |
| tanh | tanh | tanh | tanh | 78.39 | 66.08 | 71.97 | 93.58 |
| SoftMax | SoftMax | SoftMax | tanh | 78.39 | 55.97 | 70.68 | 64.72 |
| sigmoid | sigmoid | sigmoid | tanh | 78.39 | 82.71 | 72.13 | 119.18 |
| ReLU | ReLU | ReLU | SoftMax | 78.39 | 54.86 | 72.26 | 59.80 |
| tanh | tanh | tanh | SoftMax | 78.39 | 55.45 | 72.22 | 59.76 |
| SoftMax | SoftMax | SoftMax | SoftMax | 78.39 | 54.95 | 72.11 | 59.87 |
| sigmoid | sigmoid | sigmoid | SoftMax | 78.39 | 54.93 | 72.26 | 59.84 |
| ReLU | ReLU | ReLU | Sigmoid | 78.31 | 47.72 | 72.97 | 52.77 |
| tanh | tanh | tanh | Sigmoid | 78.56 | 56.37 | 72.18 | 73.19 |
| SoftMax | SoftMax | SoftMax | Sigmoid | 78.39 | 55.00 | 71.47 | 61.18 |
| sigmoid | sigmoid | sigmoid | sigmoid | 78.39 | 57.55 | 72.13 | 73.00 |

The results displayed a range of combinations that yielded the maximum accuracy, however, there was a deciding combination that produced the maximum validation accuracy and the minimum loss. The highlighted row of ReLU for the first three layers and SoftMax for the fourth was chosen in the model as it yielded the highest mean validation accuracy at 78.39% with the lowest mean validation loss at 54.86%.

**Dropout Rate**

Dropout is an algorithm for training neural networks which rely on stochastically “dropping out” neurons during training to avoid the co-adaptation of feature detectors (Baldi, P. et al., 2013). The results below are based on the usage of the ReLU activation functions for the first three layers and the SoftMax activation function for the fourth at a 0.2 dropout rate. A study of the impact of dropout value on the models mean validation accuracy and mean validation loss was conducted; running the model variation which yielded the maximum mean validation accuracy and minimum mean validation loss from the result of the previous study with a range of dropout rates.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dropout value | Mean Validation Accuracy (%) | Mean Validation Loss (%) | Mean Training Accuracy (%) | Mean Training Loss (%) |
| N/A | 78.39 | 54.85 | 72.13 | 59.79 |
| 0.1 | 78.39 | 55.15 | 72.13 | 60.00 |
| 0.2 | 78.39 | 54.84 | 72.05 | 59.82 |
| 0.3 | 78.39 | 54.96 | 72.26 | 59.88 |
| 0.4 | 78.39 | 54.98 | 72.26 | 59.88 |
| 0.5 | 78.39 | 55.00 | 72.17 | 59.90 |
| 0.6 | 78.39 | 54.87 | 72.23 | 59.84 |
| 0.7 | 78.39 | 54.87 | 72.07 | 59.87 |
| 0.8 | 78.39 | 54.96 | 72.07 | 59.88 |
| 0.9 | 78.39 | 54.85 | 72.14 | 59.87 |

The results from the study concluded that the dropout rate which yielded the maximum mean validation accuracy with minimum mean validation loss was 0.2. The results displayed that the change of dropout rate had no impact on the mean validation accuracy, although the 0.2 dropout rate provided for the minimum mean validation loss at 54.84%.

**Optimizers**

Optimizers are algorithms or methods used to change the attributes of the neural network such as weights and learning rate to reduce the losses. Optimizers are used to solve optimization problems by minimizing the loss (Sun, S. et al., 2019).

A further study on the model was conducted based on the impact of optimizers on the mean validation accuracy and mean validation loss. Using the results from the previous studies of the model; the model used is the same as the previous study. The study measured a range of common optimizers used in deep learning models.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Optimizer | Mean Validation Accuracy (%) | Mean Validation Loss (%) | Mean Training Accuracy (%) | Mean Training Loss (%) |
| Adam | 78.39 | 54.84 | 72.05 | 59.82 |
| SGD | 78.39 | 58.02 | 72.26 | 61.82 |
| Adamax | 78.39 | 51.46 | 72.26 | 55.76 |
| Adagrad | 78.39 | 50.71 | 72.26 | 53.25 |

The results displayed a similar trend from the previous study: the optimizer had no impact on the mean validation accuracy, however, the Adagrad optimizer was able to achieve a minimum mean validation loss of 50.71%.

**Conclusion**

This model was able to achieve 78.39% mean validation accuracy with 50.71% validation loss and was able to provide suitable confidence to classify the safety of consumption of a mushroom. The project's outcome was successful in providing a classification of a novel solution. The iterative stage improved the model through the introduction of dropout rates and data augmentation improving the model from 72.20% to 78.39%. Model optimization through changes of the activation functions, dropout rates and optimizers increased the model’s accuracy to 78.39% with the activation functions using ReLU for layers 1-3 and SoftMax for the fourth layer at a 0.2 dropout rate with the Adagrad optimizer.

Future iterations of the implemented model could explore dropout rates in the inner layers. Also, a study into the effect of the number of nodes in the hidden layer impacted the model’s accuracy and loss of the classification of mushroom images to determine the safety of consumption. Zou, W. et al., 2009 concluded that the increase of hidden node increased the validation accuracy of the model for image-based classification. Additionally, future iterations could compare the impact of input image size on the accuracy. A study conducted by Huang, J. et al., concluded that an increase in input image increases accuracy but with the tradeoff of increased time to classify an image.

**References**

Shijie, J., Ping, W., Peiyi, J. and Siping, H., 2017, October. Research on data augmentation for image classification based on convolution neural networks. In *2017 Chinese automation congress (CAC)* (pp. 4165-4170). IEEE.

Apicella, A., Donnarumma, F., Isgrò, F. and Prevete, R., 2021. A survey on modern trainable activation functions. *Neural Networks*, 138, pp.14-32.

Baldi, P. and Sadowski, P.J., 2013. Understanding dropout. *Advances in neural information processing systems*, *26*, pp.2814-2822.

Sun, S., Cao, Z., Zhu, H. and Zhao, J., 2019. A survey of optimization methods from a machine learning perspective. *IEEE transactions on cybernetics*, *50*(8), pp.3668-3681.

Zou, W., Li, Y. and Tang, A., 2009. Effects of the number of hidden nodes used in a structured-based neural network on the reliability of image classification. *Neural Computing and Applications*, *18*(3), pp.249-260.

Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S. and Murphy, K., 2017. Speed/accuracy trade-offs for modern convolutional object detectors. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7310-7311).