Dataset Overview:

|  |  |  |
| --- | --- | --- |
| Mean Popularity | Median Popularity | 75th Percentile Popularity |
| 63.28 | 65 | 77 |

1. Logistic Regression Model with Original Data:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-Score | Accuracy |
| Class 0 | 0.59 | 0.44 | 0.51 | 57% |
| Class 1 | 0.57 | 0.70 | 0.63 |

Logistic Regression Model with Resampled Data:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-Score | Accuracy |
| Class 0 | 0.59 | 0.52 | 0.55 | 59% |
| Class 1 | 0.58 | 0.65 | 0.61 |

Comparing the two models, it appears that the logistic regression model with resampled data performed slightly better than the model with the original data. It achieved a slightly higher accuracy, precision, recall, and F1-score for both classes. Resampling the data helped to balance the class distribution and improve the model's performance in identifying both classes more accurately.

2. Neural Network Model:

cut-off = 65,

layer1 = 9: activation function = relu

layer2 = 18: activation function = relu

|  |  |
| --- | --- |
| Loss | Accuracy |
| 64% | 60% |

A loss value of 65 indicates that the model can be further optimized. The accuracy percent shows that 60% of the model's predicted values align with the true values in the original dataset. A picture containing screenshot, line, plot, text

Description automatically generated

Add hidden layer.

cut-off = 65,

layer1 = 9: activation function = relu

layer2 = 18: activation function = relu

layer3 = 27: activation function = relu

|  |  |
| --- | --- |
| Loss | Accuracy |
| 65% | 61% |

Change activation function.

cut-off = 65,

layer1 = 9: activation function = relu

layer2 = 18: activation function = tanh

layer3 = 27: activation function = tanh

|  |  |
| --- | --- |
| Loss | Accuracy |
| 64% | 62% |

3. Random Forest Model:

|  |
| --- |
| Accuracy |
| 55% |

Conclusion: - The logistic regression model with resampled data showed improved performance compared to the model with original data, with slightly higher precision, recall, and F1-score for both classes. - The neural network model showed varying performance based on changes made, indicating the importance of model architecture and activation functions. - The random forest model achieved an accuracy of 0.5565, indicating moderate performance. - Overall, further model optimization and exploration of different algorithms may be necessary to improve the predictive performance of the models on the given dataset.

**Explain optimization Neural Network Model:**

In the optimization process of the Neural Network Model, several changes were made to improve its performance. Here's an explanation of the optimization steps:

1. Initial Model:
   * Loss: 0.6481
   * Accuracy: 0.6045
2. Change: Add hidden layer
   * An additional hidden layer was added to the neural network architecture.
   * Loss: Loss: 0.6507
   * Accuracy: 0.6153

Explanation: By adding a hidden layer, the model's complexity increased, allowing for the extraction of more intricate patterns in the data. This change resulted in a slightly higher loss and accuracy. It suggests that the added complexity may not have significantly improved the model's performance.

1. Change: Change activation
   * The activation function of the neural network layers was modified.
   * Loss: 0.6425
   * Accuracy: 0.6221

Explanation: Activation functions introduce non-linearity to the neural network, enabling it to learn complex relationships in the data. By changing the activation function, the model might have become more effective in capturing non-linear patterns in the data. This change led to a lower loss and an increase in accuracy compared to the previous step, indicating an improvement in the model's performance.

Overall, the optimization process involved experimenting with different architectural changes and activation functions. The addition of a hidden layer did not yield a significant improvement but changing the activation function resulted in a better-performing model with a lower loss and higher accuracy.

Top of Form

Bottom of Form