**Titanic Survival Prediction Model Evaluation Report**

**Introduction**

The objective of this project was to implement supervised machine learning models to predict the survival of Titanic passengers. The models utilized for comparison were:

* **Logistic Regression**
* **Random Forest**
* **Neural Network** (using TensorFlow)

The dataset was split into training and testing sets using an **80-20** split. The models were evaluated using key performance metrics such as **Accuracy, Precision, Recall, F1-Score, and ROC-AUC**.

**Evaluation Results**

Below is a summary of the model performances:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Accuracy** | **Precision** | **Recall** | **F1-Score** | **ROC-AUC** | **Confusion Matrix** |
| Logistic Regression | 0.810056 | 0.785714 | 0.743243 | 0.763889 | 0.800193 | [[90, 15], [19, 55]] |
| Random Forest | 0.815642 | 0.797101 | 0.743243 | 0.769231 | 0.804955 | [[91, 14], [19, 55]] |
| Neural Network | 0.832402 | 0.854839 | 0.716216 | 0.779412 | 0.815251 | [[96, 9], [21, 53]] |

**Observations and Insights**

1. **Random Forest and Neural Network performed better overall than Logistic Regression**: Both models achieved **higher accuracy and F1-score**, indicating that non-linear models capture the patterns in the Titanic dataset better than Logistic Regression.
2. **Neural Network showed the best performance**: With an **accuracy of 83.24% and an F1-score of 0.779**, the Neural Network outperformed the other two models. However, its **recall score (0.716)** was slightly lower compared to Logistic Regression and Random Forest.
3. **Random Forest was a strong competitor**: Random Forest performed **consistently well across all metrics** and had a slightly better **recall (0.7432) compared to the Neural Network (0.7162)**. This suggests that Random Forest is more stable in predicting survivors correctly.
4. **Impact of Hyperparameter Tuning**:
   * During experimentation, **changes in learning rate and the number of epochs** in the Neural Network **significantly impacted performance**.
   * **Increasing the epochs improved model generalization up to a point** but after a certain threshold, it led to **fluctuations in results**.
   * **Adjusting the learning rate** had a **major influence on accuracy and stability**, highlighting the importance of hyperparameter tuning for Neural Networks.
5. **Fluctuation in Neural Network Performance**:
   * Even when **hyperparameters were kept constant**, there were **slight fluctuations in accuracy and F1-score across multiple runs**.
   * This is likely due to the **limited amount of data available** in the Titanic dataset. Neural Networks tend to require **large datasets to fully leverage their learning potential**, and the Titanic dataset might not have been sufficient to fully utilize its strengths.

**Conclusion**

Based on the evaluation results:

* **Random Forest and Neural Network** outperformed Logistic Regression.
* **Neural Network performed the best** but was **sensitive to hyperparameter changes and dataset size**.
* **Random Forest showed consistent results** and is a **reliable choice with fewer tuning requirements**.
* **For small datasets**, traditional models like **Random Forest may provide more stability compared to Neural Networks**.

Thus, while **Neural Networks have the potential for superior performance**, **Random Forest remains a robust choice for structured datasets like Titanic survival prediction due to its stability and interpretability.**