Model Report

Metrics Summary

The following table summarizes the performance of the three models using both custom and pre-trained Word2Vec embeddings.

Model	Embedding Type	Test Accuracy	F1-Score (Macro Avg)
FFNN	Custom Word2Vec	89.92%	0.90
RNN	Custom Word2Vec	89.92%	0.90
LSTM	Custom Word2Vec	89.92%	0.90
FFNN	Pre-trained Word2Vec	89.14%	0.89
RNN	Pre-trained Word2Vec	89.92%	0.90
LSTM	Pre-trained Word2Vec	89.92%	0.90

Export to Sheets

Note: The F1-score is a balanced metric that considers both precision and recall, making it a reliable measure for multi-class classification. The macro average is used to ensure all classes are weighted equally, regardless of their size.

In-depth Analysis (FFNN with Custom Embeddings)

- **Precision**: The model demonstrates high precision across most categories, indicating that when it predicts a specific category, it is often correct.
- **Recall**: Similarly, the recall is high, showing that the model is effective at identifying most of the complaints for each category.
- **F1-Score**: The F1-scores are consistently high (around 0.90), suggesting a strong balance between precision and recall.
- **Confusion Matrix**: The confusion matrix reveals a few minor misclassifications but no systematic confusion between any two categories. This indicates the models are robust and the categories are well-defined.