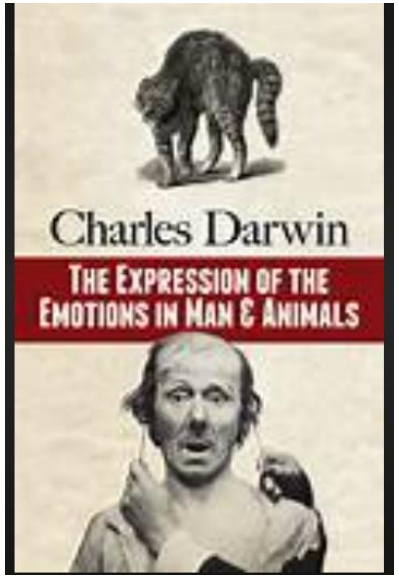
PERFORMANCE ANALYSIS AND HYPERPARAMETER OPTIMIZATION OF STACKED BI-LSTM FOR EMOTION RECOGNITION FROM FACIAL EXPRESSION

By
Epsita Bose
MASTER OF SCIENCE IN DATA SCIENCE
STUDENT ID: 1060770

Introduction

Technology merging with the human brain



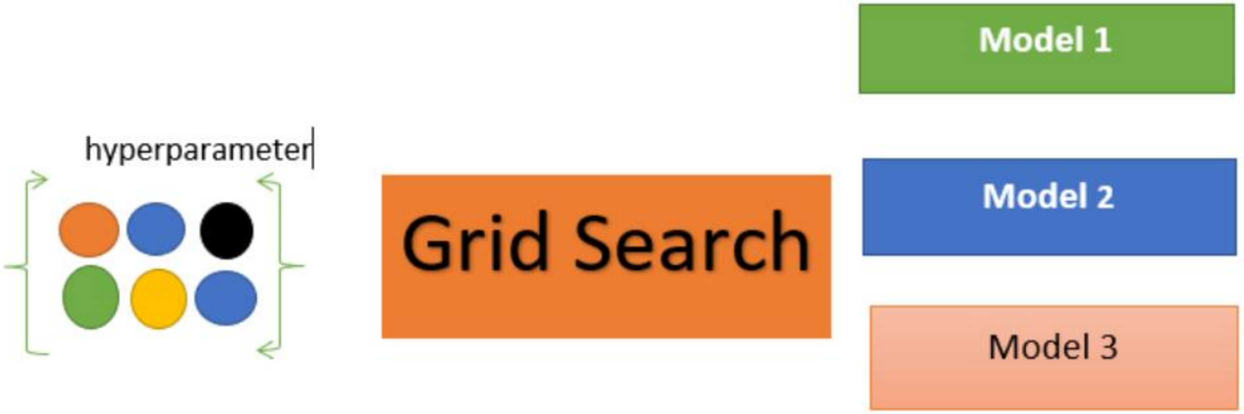
Darwin's contribution to understanding human communication through emotions

Importance of artificial intelligence and emotion detection

Introduction of this study

Analysis the emotion by use of Stacked Bi LSTM JONATHAN OHEIX Dataset

Hyperparameter Choosing



AIM

The aim of this proposed study is to increase the prediction accuracy in detecting emotion from facial expressions by using stacked bidirectional LSTM. This will be achieved by selecting proper hyperparameters with the help of GridSearchCV and analyzing the model's performance on a large dataset in terms of prediction time

OBJECTIVE

- Design and develop a stacked bi-LSTM model for emotion detection from facial expressions.
- Train the model on a large dataset of facial expressions covering a range of emotions.
- Optimize the model architecture and hyperparameters using GridSearchCV for improved accuracy and efficiency.
- Evaluate the model's performance and compare it to existing emotion recognition models.
- Analyze the model's behavior and identify important features for accurate emotion recognition.
- Apply the model to real-world scenarios, such as monitoring emotions in social media or video chats.

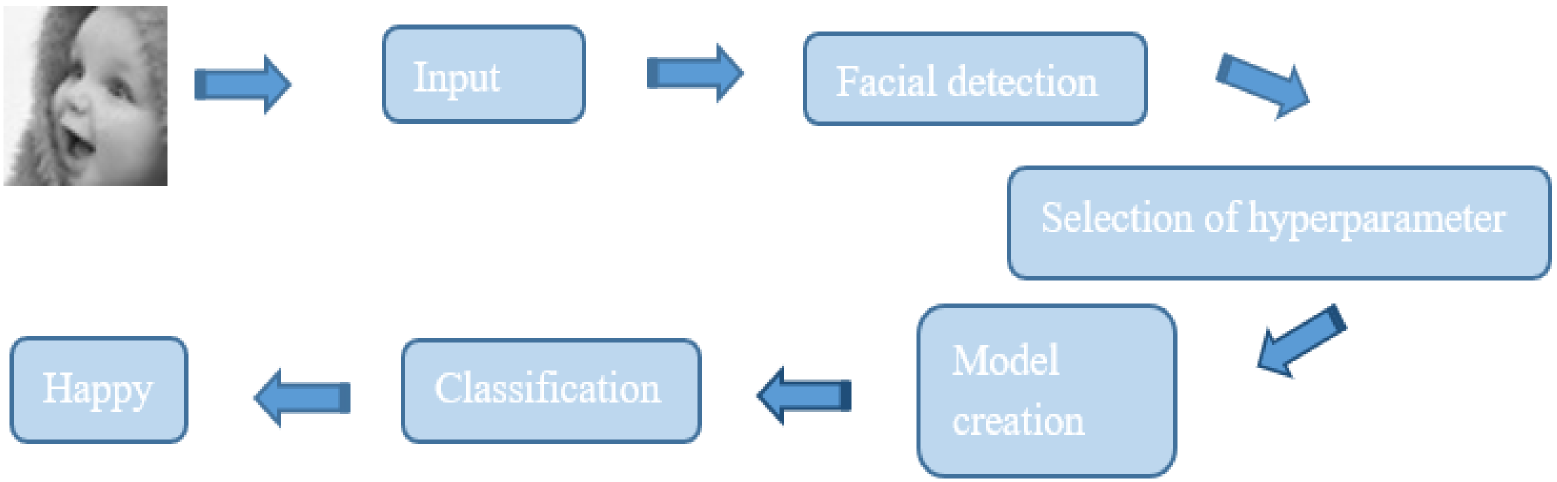
PROBLEM STATEMENT

- In the study, the performance of the stacked bidirectional LSTM algorithm is evaluated on a larger dataset collected from Kaggle, consisting of 35.9k files. The aim is to determine how well the algorithm performs in real-time scenarios when dealing with larger datasets.
- Additionally, the study demonstrates the use of GridSearchCV to select the best hyperparameters for the model. Hyperparameters such as row_hidden, col_hidden, batch_size, and epochs have a significant impact on the model's performance. By systematically searching the hyperparameter space, GridSearchCV helps identify the optimal combination of hyperparameters that maximize the accuracy of the emotion detection model. This optimization is particularly important for real-world applications like healthcare and human-computer interaction.

Literature Review

Healthcare	Pseudo-tumor detection (Dandıl & Karaca, 2021), Medical Image Classification (Pattanaik et al., 2022)
Automobile and mechanical industries	Signal labeling and precise location in a variable parameter milling process (Qiu et al., 2023), Short-term traffic flow prediction based on whale optimization algorithm((Xu et al., 2022)
Price detecting	stock market price prediction (Uddin et al., 2022), Spot price detecting (Chittora & Gupta, 2020)
different type of prediction in daily human life	Melody classification ((Y. Li & Lin, 2020), Detection of human activity from smart phone data ((Ullah et al., 2019)
Emotion detection by facial expression	CNN (Jaiswal et al., 2020, M. Kumar & Srivastava, 2021, Babajee et al., 2020), RNN (K.S. & David, 2020, Said & Barr, 2021), CNN-LSTM (Rajan et al., 2020, Yu et al., 2018, Hung & Tien, 2021, Abdullah et al., 2020)

Methodology

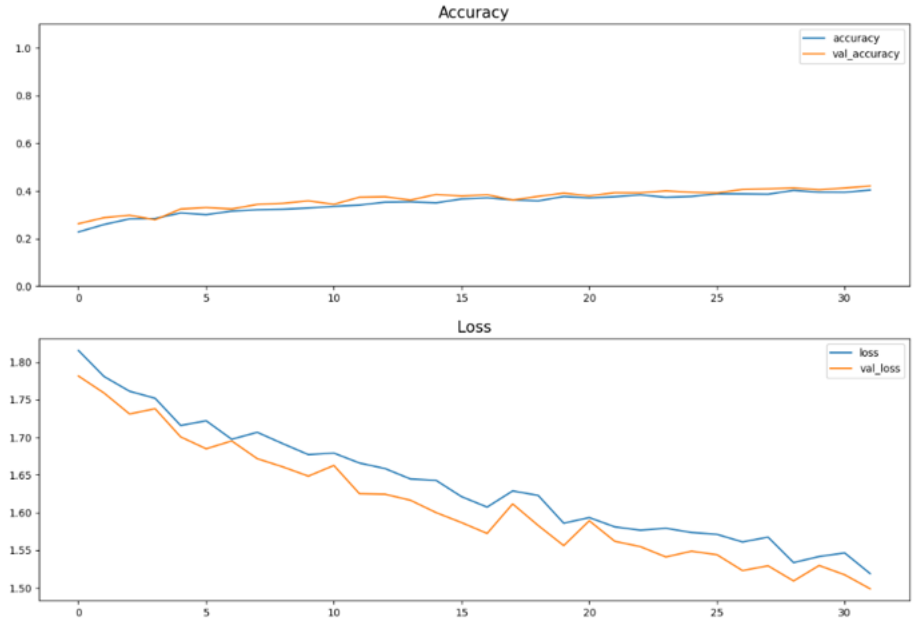


Result and Discussion

```
0.258214 (0.005861) with: {'col_hidden': 16, 'row_hidden': 32}
0.248569 (0.002285) with: {'col_hidden': 16, 'row_hidden': 64}
0.248569 (0.002285) with: {'col_hidden': 16, 'row_hidden': 128}
0.263003 (0.019706) with: {'col_hidden': 32, 'row_hidden': 32}
0.259845 (0.011933) with: {'col_hidden': 32, 'row_hidden': 64}
0.248569 (0.002285) with: {'col_hidden': 32, 'row_hidden': 128}
0.264217 (0.000644) with: {'col_hidden': 64, 'row_hidden': 32}
0.255508 (0.015667) with: {'col_hidden': 64, 'row_hidden': 64}
0.248569 (0.002285) with: {'col_hidden': 64, 'row_hidden': 128}
```

Result of GridSearchCV

Result and Discussion

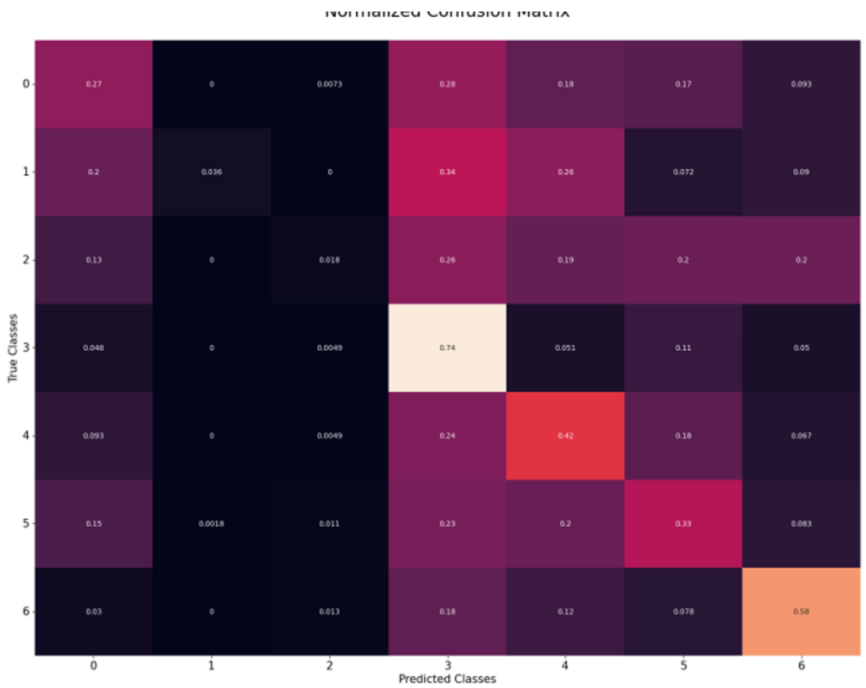


	precision	recall	f1-score	support
0	0.32	0.27	0.29	960
1	0.67	0.04	0.07	111
2	0.29	0.02	0.03	1018
3	0.52	0.74	0.61	1825
4	0.39	0.42	0.40	1216
5	0.31	0.33	0.32	1139
6	0.45	0.58	0.50	797
accuracy			0.42	7066
macro avg	0.42	0.34	0.32	7066
weighted avg	0.40	0.42	0.38	7066

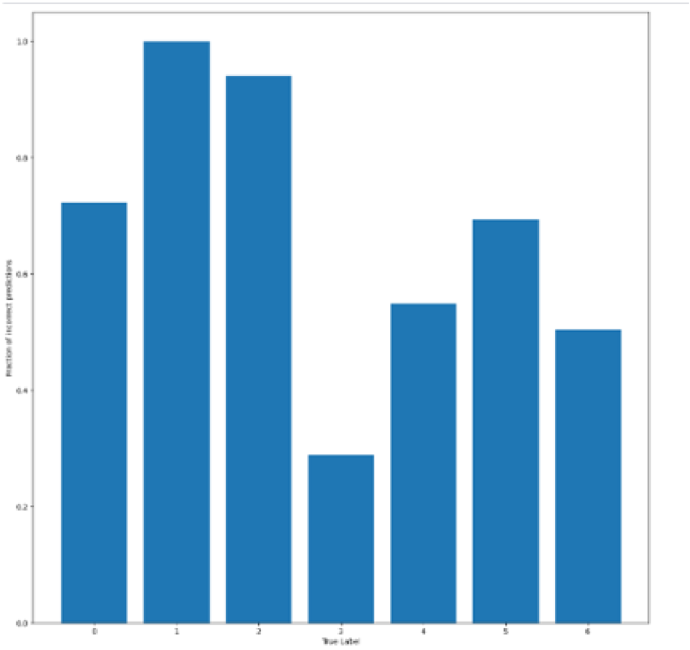
Result of Test

Result and Discussion

Confusion Matrix



Fraction of incorrect prediction



Result of Test

Discussion

- The Jonathan Oheix dataset on Kaggle contains facial expression images with labels for seven emotions. Previous studies achieved good results using this dataset, with CNN showing 53% accuracy (Sidhu et al., 2022), and a CNN-RNN model achieving 95+% accuracy (Singh et al., 2021).
- The model was trained for 32 epochs using the Stochastic Gradient Descent (SGD) optimizer and monitored through training and validation loss and accuracy. Overfitting was observed in some epochs.
- The model's performance on the test set achieved an accuracy of 42.13%. The classification report provided precision, recall, f1-score, and support for each emotion class.

Conclusion

- This study aimed to improve the model's performance by selecting optimal hyperparameters. The initial model achieved 34% accuracy, but after a grid search, the best hyperparameters were determined to be col_hidden = 64 and row_hidden = 32, resulting in a score of 0.264217.
- The model was trained for 32 epochs using the SGD optimizer with a learning rate of 0.001. Although there were some epochs with no improvement or decreased validation accuracy, overall, the model showed progress over time.
- Visualizations were used to track the model's training history, displaying accuracy and loss metrics. The model achieved an accuracy of 42.13% on the test set, with a classification report providing detailed metrics for each class.
- While the model's accuracy on the test set was not high, this study highlighted the importance of pre-processing, EDA, and data augmentation to improve model performance. Further research is needed to enhance the model's classification capabilities and achieve higher accuracy.

Future Work

- To enhance the model, future research can refine the architecture by adding layers or adjusting the neuron count, and explore more complex models like CNNs for better feature extraction.
- Additional hyperparameters, such as learning rate and regularization strength, can be explored to further optimize the model's performance.
- Considering alternative datasets that offer greater diversity in facial expressions can help improve the model's ability to generalize to different expressions.
- Transfer learning, utilizing pre-trained models and fine-tuning them, can be employed to improve performance and reduce training time.
- Incorporating techniques like ensemble learning or active learning can contribute to increased accuracy and efficiency of the model.


Thank You