Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Guadalajara

School of Engineering and Sciences

Comparative Analysis of Rules-based Processing, Neural Networks, and Deep Learning Approaches for Emotion Detection in Textual Content.

A thesis presented by

José Martín Véliz Zambrano

Submitted to the
School of Engineering and Sciences
in partial fulfillment of the requirements for the degree of

Master of Science

in

Mechatronics Engineering

Monterrey, Nuevo León, November, 2023

Comparative Analysis of Rules-based Processing, Neural Networks, and Deep Learning Approaches for Emotion Detection in Textual Content.

by

José Martín Véliz Zambrano

Abstract

The ability to determine feelings from textual data is increasingly critical in fields from customer service to mental health assessment. This research aims to examine the efficacy of numerous computational procedures, specifically Rules based processing, Neural networks, and Deep learning, within the area of emotion detection in text. Therefore, 3 Python-based models were developed, each corresponding to one among the earlier methodologies, and applied to a labelled dataset for prediction accuracy assessment. The rules based processing method utilized a lexicon of words linked to emotions, while the neural network and deep learning models, including more advanced architectures, were trained on the same data. The evaluation involved massive hyperparameter tuning with ten iterations to examine the greatest performance conditions for each iteration. The benchmarks displayed that rules-based processing achieved an average accuracy of 54.02%. At the same time, neural networks and deep learning methods produced greater promising results, with peak accuracies of 85.5% and 89.75%. The study concludes that deep learning outperforms its predecessors, suggesting that advancements in machine learning architectures offer considerable benefits for complex emotional understanding in text analysis. These outcomes show the progress in affective computing and suggest adopting state-of-the-art, adaptable models in practical applications to decorate empathetic interactions among people and machines.

List of Figures

List of Tables

Contents

Chapter 1

Motivation

Emotion detection in text is a multidimensional task on the intersection of natural language processing (NLP) and artificial intelligence. As interactions between humans and machines become more and more usual, accurately interpreting emotions from text is vital. Whether it is to understand purchaser feedback, monitoring mental health through social media posts, or even improving the empathy of chatbots. However, present techniques and procedures with the intent to achieve this goal usually vary significantly in their ability to decode the complexities of emotional expressions in written language.

Current research provides foundational information, showcasing many techniques from rules based processing to neural networks and even deep learning. However, a greater gap in data is still recurrent. Though transparent and easily implemented, rules based processing remains rigid and often misses the complexity of human emotions. Traditional machine learning methods including Neural Networks introduce flexibility and efficiency. However, they struggle with the contextual dependencies of textual content. The latest creation of deep learning has opened doors to improved performance peaks.

The importance of improving emotion detection is manifold. The theoretical implications consist of contributing to the expertise of affective computing processing. Many advancements in this subject promise to improve the quality of service in customer service arrangements, deliver richer information for marketplace evaluation, and improve social media engagement by curating sensitive content that resonates with customers current emotional states.

Therefore, the objectives of this research are clear. To evaluate the effectiveness of current methodologies, compare their overall performance, and sugest procedures for future innovation in emotion detection. This study aims to bridge the gaps between human emotions and affective computing, thus expanding the frontiers of emotion detection. However, addressing this goal is more than academically or technologically motivated; As it is vital for building a future in which machines understand us better, make our society feel heard and understood, and improve services that we use in our everyday lives.

Chapter 2

Literature Review

The growing field of affective computing and natural language processing (NLP) has provoked an increase in research into classifying and detecting human emotions from textual data. This literature review aims to consider current research and how the computational models, from rules based processing to neural networks and deep learning methods, have evolved. The cited research establishes the foundation for the prevailing objective of applying the mentioned computational models to conduct emotion detection in textual datasets.

2.1 Affective Computing Foundations

Firstly, The foundational basis for Affective Computing was established through Rosalind W. Picard, "Affective Computing: From Laughter to IEEE"[6] as it describes the beginning of the development of affective computing. Picard's narrative gives vital insight into the historic scepticism found in this area and its subsequent rise to an interdisciplinary technology with great application ability. The statements made in the research paper pave the way for introducing the effect that emotional records may have on technological development.

Complementing Picard's foundational study on Affective Computing, Tao et al. "Affective Computing: A Review "[7] introduces a comprehensive exam of vital technological areas within affective computing, including emotional speech processing, facial expression evaluation, and multimodal systems. This evaluation introduces the combination of complex situations while combining unique human expressions. This assessment lays the basis for many emerging methodologies in emotion detection tasks.

2.2 Artificial Neural Networks

Secondly, A.D. Dongare et al. provide a detailed overview of the principles governing Artificial Neural Networks in "Introduction to Artificial Neural Networks."[2] This paper mentions many ANN architectures and learning rules, highlighting the importance of pattern recognition abilities that are suitable for emotion detection in texts. Their discussion on training high-quality and application makes it easier to understand how neural networks can be adapted for the task of emotion detection.

In addition to this, in ”Neural Networks and Their Applications”[1] Coban mentions and introduces recent advancements and real-life applications, translating the theoretical complexities into practical implementations. The paper’s focus on distributed systems and advanced computing architectures presents considerable insights into scaling emotion detection models, a consideration that becomes important when processing large text datasets.

2.3 Natural Language Processing

Thirdly, Considering the complexities of NLP, Nadkarni et al. provide a foundational understanding of NLP system design, machine learning applications, and future directions in ”Natural language processing: an introduction.”[5] The authors’ discussion of machine learning techniques like Support Vector Machines and Hidden Markov Models contributes to a greater suggestion of algorithms that could be applied to emotion detection tasks.

Moreover, Elizabeth D. Liddy’s ”Natural Language Processing” [4] offers a summary of NLP’s history, methodologies, and applications. Liddy’s explanation of different approaches, symbolic, statistical, connectionist, and their hybrids, is a vital framework considering the many computational strategies this project employs in implementing the three emotion detection approaches, which are rules based processing, neural networks as mentioned above and deep learning.

Lastly, Lazrig et al. ”Using Machine Learning Sentiment Analysis to Evaluate Learning Impact”[3] provides a closer look at sentiment analysis in educational research. Their analysis displays the strengths and weaknesses of different machine learning algorithms, such as naive Bayes and word-sentiment associations, in handling sentiments and underscores the importance of domain-specific training for algorithm performance.

2.4 Conclusion of the Literature Review

In conclusion, landing the theory to the current study, this literature review is meant to display the evolution and interdisciplinary nature of affective computing, researching and considering many computational methods from basic rules based processing to complex and elaborate machine learning and deep learning approaches. As this research continues, scientific rigor, interdisciplinary approaches, scalability considerations, and computational model complexities discussed in this section will be considered. The gap this study is meant to fill will be covered by directly comparing the capabilities and performances of the three distinct computational approaches in a coherent framework, thus contributing to a complex understanding of the strengths and limitations in the growing field of emotion detection within natural language processing.

Chapter 3

Experimentation

3.1 Methodology

For the methodology section of this investigation, 3 Python programs employing different methodologies for emotion detection were executed alongside a labelled text dataset: The approach involving a rules based processing method was implemented using a lexicon containing words associated with specific emotions. Another implemented method relied on neural networks, a subset of machine learning algorithms modelled after the brain's neural structure. The third program implemented deep learning, an advanced form of neural network that can work with many abstraction layers; moreover, it can learn directly from raw data and recognize complex patterns.

3.2 Experimental Setup

The experimental setup for each of the methods mentioned above implicated separating the dataset into separate subsets for training, testing, and validation. Training sets permitted the models to learn and fit the data, test sets evaluated the performance of the models, and validation sets provided a final assessment of accuracy. The software required included Python with necessary libraries such as PyTorch for neural networks, pandas for data manipulation, sklearn for machine learning tasks and Matplotlib for data visualization.

3.3 Procedure

Each of the python programs used for the experimentation had many hyperparameters that could be adjusted depending on the required results and accuracy. Hyperparameters are important as they directly control the behaviours of the training algorithm and significantly impact the performance and accuracy of the model. For the procedure, each of the programs were executed ten times, each time altering the hyperparameters, including but not limited to learning rate, batch size, layers, embedding dimensions, and hidden dimensions. These parameters affected model complexity and learning process, thus affecting each model's accuracy rate.

Increment Step	Max Word Count
5	50
10	100
15	150
20	200
25	250
30	300
35	350
40	400
45	450
50	500

Table 3.1: Rules Based Processing Hyperparameters

Maximum Sequence Length	Epochs	Embedding Dimensions
25	50	100
30	60	150
35	70	100
40	80	250
45	90	200
50	100	350
55	110	200
60	120	350
65	130	200
70	140	350

Table 3.2: Neural Networks Hyperparameters

3.4 Data Collection and Analysis

Test accuracy and validation accuracy metrics were collected after each run on every program. For analysis purposes, these accuracy values were recorded and graphed to make it easier to understand the changes in accuracy due to hyperparameter variations. Tables 3.1, 3.2 and 3.3 display how altering hyperparameters impacts each algorithm's effectiveness in accurately classifying emotions.

Batch Size	Embedding Dimensions	Hidden Dimensions	Epochs
32	100	128	5
32	100	128	10
32	100	128	15
32	300	256	5
32	300	256	10
32	300	256	15
32	450	256	5
32	450	256	10
32	450	256	15
64	550	128	5

Table 3.3: Deep Learning Hyperparameters

3.4.1 Rules Based Processing Hyperparameters

3.4.2 Neural Networks Hyperparameters

3.4.3 Deep Learning

3.5 Flow Diagrams

Flow diagrams were implemented using the logic from the 3 python programs to represent each program's workflow and the learning process of the algorithms. The diagrams required a high level understanding of the algorithmic procedure, including how data was inputted and processed, and how it conducted emotional predictions. The diagrams illustrate a high level understanding of each methodology. However, they are not intended to be a detailed representation but to show each approach's primary components and data flow.

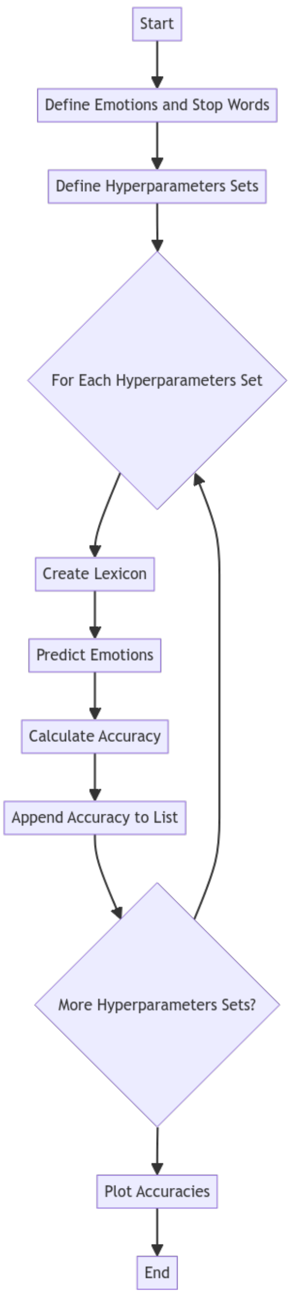

Figure 3.1: Rules Based Flow Diagram

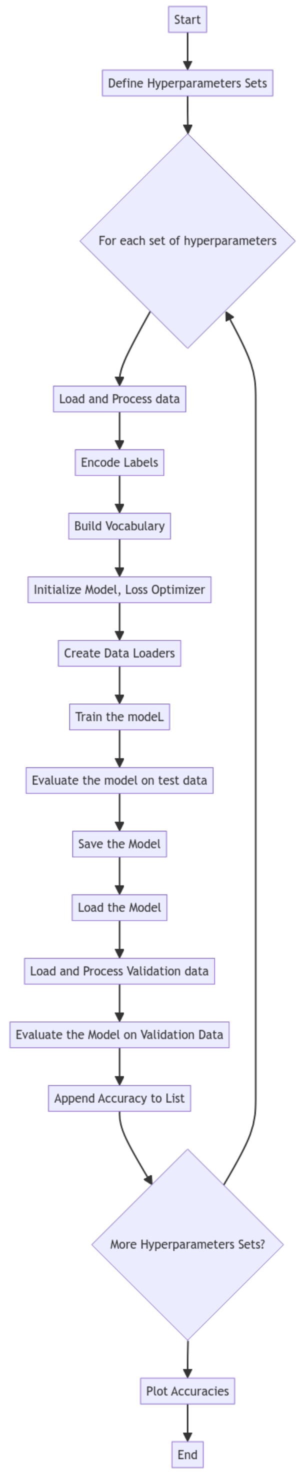

Figure 3.2: Neural Networks Flow Diagram

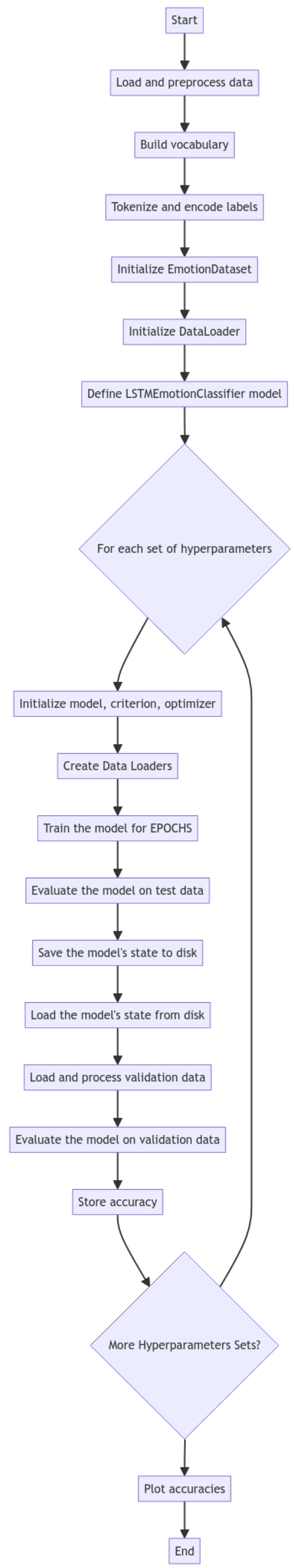

Figure 3.3: Deep Learning Flow Diagram

Chapter 4

Results

4.1 Rules-Based Processing

For this method, the provided dataset was analyzed using a predefined set of lexical rules using the rules-based processing method. After ten iterations, the results, each with a distinct mix of hyperparameters, produced an average accuracy of 54.02%. The peak accuracy was of 60.20%, obtained under a configuration of lexical rules and hyperparameters, indicating the peak performance of the rules-based model under the most favourable conditions.

A detailed performance breakdown, as shown in Table 4.1, reveals the impact of incremental steps in adjusting hyperparameters. Starting from an increment step of 5 words with a subsequent maximum word count of 50, the model showed an initial prediction accuracy of 34.85%. As the increment steps and the corresponding word count increased, the model improved performance, peaking at an increment step of 30 words (maximum count of 300 words), resulting in an accuracy of 60.20%. After that, a decline was noted as the complexity of the rules increased, showing decreased accuracy with maximum word counts reaching 500.

Increment Step	Max Word Count	Prediction Accuracy
5	50	34.85%
10	100	44.25%
15	150	52.20%
20	200	56.35%
25	250	58.80%
30	300	60.20%
35	350	60.20%
40	400	58.80%
45	450	57.60%
50	500	56.95%

Table 4.1: Rules Based Processing Hyperparameters And Results

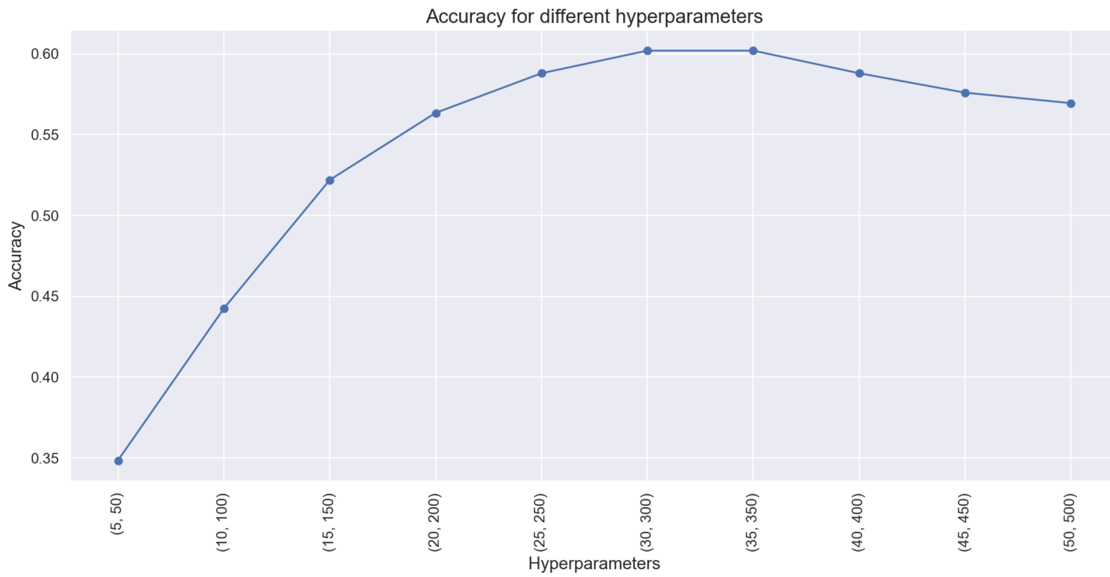

Figure 4.1: Rules Based Processing Prediction Accuracy Variations

4.2 Neural Networks

With the neural networks method execution, the results were more promising, with an average accuracy of 84.875%. A peak accuracy of 85.5% was obtained through careful tuning of hyperparameters, showing the potential of neural networks based emotion detection.

Variations in the model performance as influenced by hyperparameters such as max sequence length, number of epochs, and embedding dimensions are shown in Table 4.2. A max sequence length of 30, with 60 epochs and 150 embedding dimensions, produced the greatest prediction effectiveness, showing a test accuracy of 82.85% and a validation accuracy of 85.5%.

Maximum Sequence Length	Epochs	Embedding Dimensions	Test Accuracy	Validation Accuracy
25	50	100	80.95%	84.15%
30	60	150	82.85%	85.5%
35	70	100	82.8%	85.25%
40	80	250	83.45%	84.95%
45	90	200	82.95%	85.35%
50	100	350	83.15%	84.35%
55	110	200	83.55%	85.35%
60	120	350	83.1%	84.95%
65	130	200	82.9%	84.7%
70	140	350	82.2%	83.9%

Table 4.2: Neural Networks Hyperparameters And Results

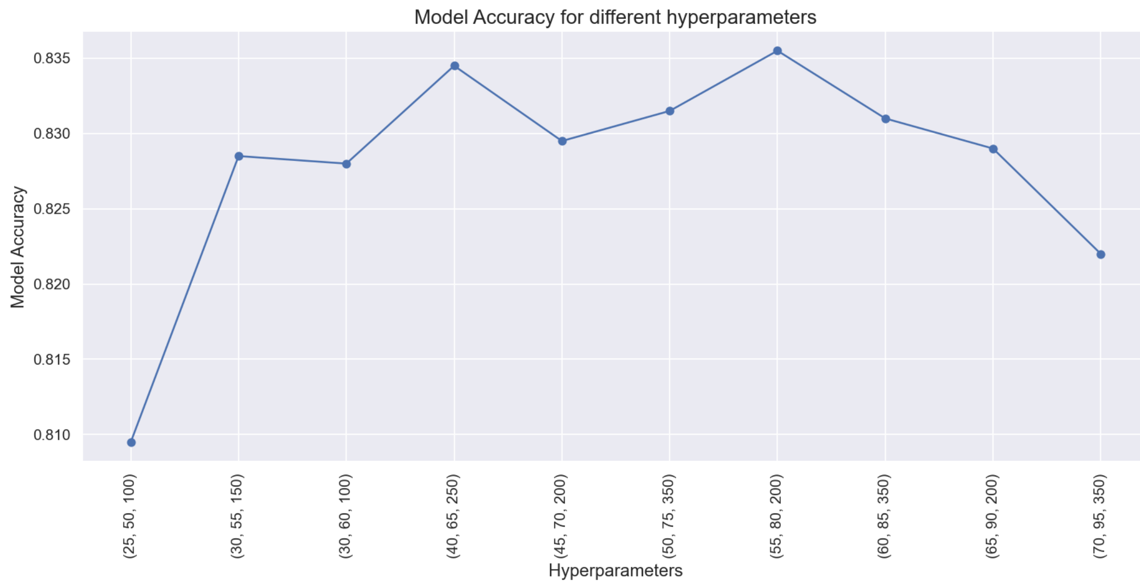

Figure 4.2: Neural Networks Test Accuracy Variations

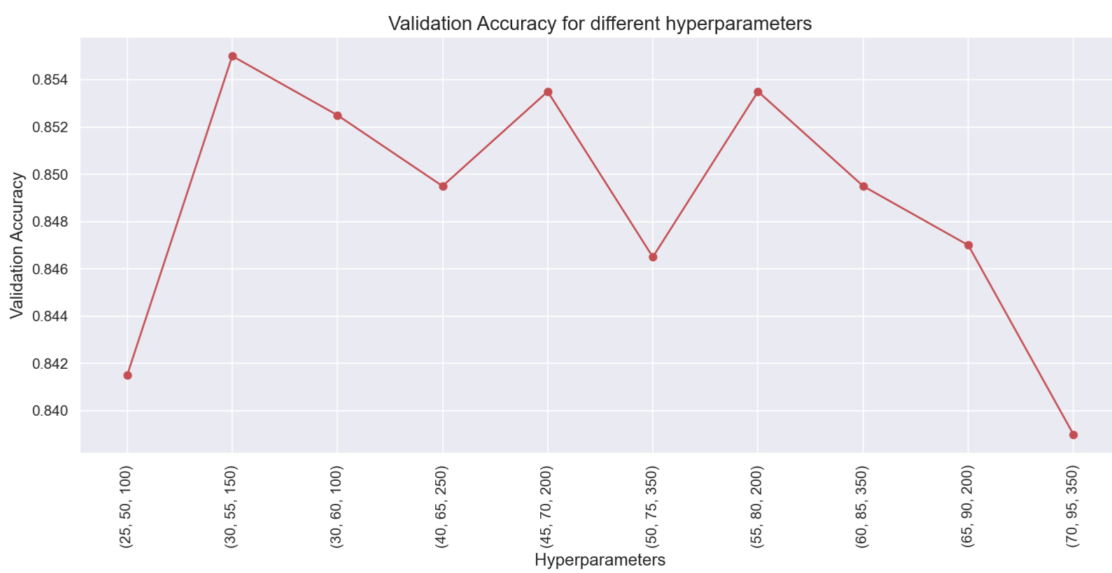

Figure 4.3: Neural Networks Validation Accuracy Variations

Batch Size	Embedding Dimensions	Hidden Dimensions	Epochs	Test Accuracy	Validation Accu
32	100	128	5	33.5%	33.6%
32	100	128	10	62.2%	56.75%
32	100	128	15	86.85%	85.5%
32	300	256	5	34.95%	34.45%
32	300	256	10	88.2%	89.35%
32	300	256	15	88.65%	89.75%
32	450	256	5	34.75%	34.75%
32	450	256	10	81.35%	81.9%
32	450	256	15	89.05%	89.45%
64	550	128	5	32.3%	29.9%

Table 4.3: Deep Learning Hyperparameters And Results

4.3 Deep Learning

The use of the deep learning method produced the highest peak of model performance. Deep learning, through extensive configurations of hyperparameters, displayed an average accuracy of 62.54% with a peak at 89.75%, showing the improved capability of this advanced approach.

As shown in Table 4.3, a presentation of results shows the progress of test and validation accuracies relative to the changes in batch size, embedding dimensions, hidden dimensions, and number of epochs. A batch size of 32, embedding dimensions of 300, hidden dimensions of 256, and 15 epochs produced a test and validation accuracy of 88.65% and 89.75%, respectively, denoting the model's peak of accuracy, as obtained from the deep learning experiments.

4.4 Comparative Analysis

This comparative analysis is meant to contrast the three computational techniques employed for emotion detection in textual data: rules based processing, neural networks, and deep learning methodologies.

Figures 4.1, 4.3 and 4.5 encapsulate the validation accuracy metrics of each approach presented in a line graph, offering a visual representation of the data collected. This visual comparison has been made to analyze and reflect on the precision of each method.

In evaluating these line graphs, several observations emerge:

4.4.1 Rules Based Processing Performance

This method kept a recurring average performance, reaching the 54% threshold. The line graph shows the accuracy levels across different hyperparameter conditions, with peaks corresponding to the efficiency of the logic and lexical rules applied.

4.4.2 Neural Networks Efficacy

The transition from rules-based to neural networks shows a peak performance improvement, with neural networks's peak accuracy reaching 85.5%. Therefore, neural networks suggest a

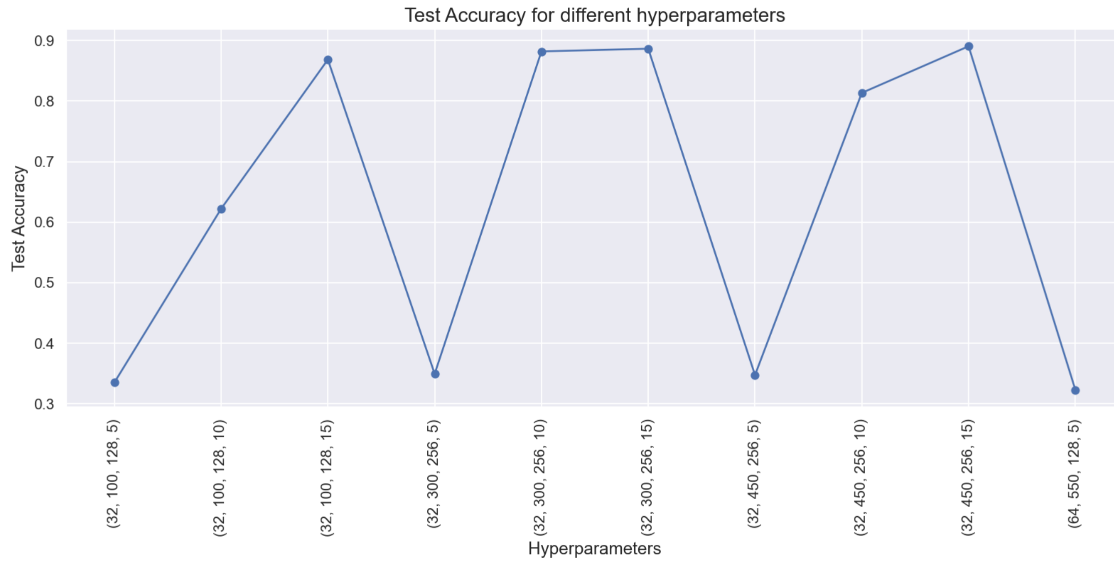

Figure 4.4: Deep Learning Test Accuracy Variations

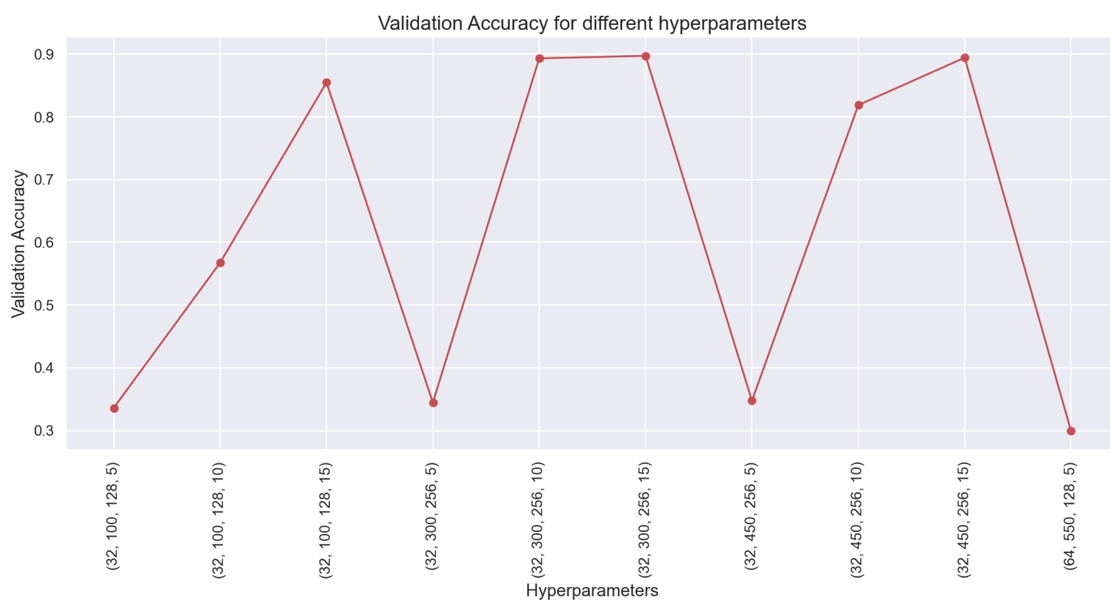

Figure 4.5: Deep Learning Validation Accuracy Variations

significant improvement, reflecting the potency of artificial neural computation models.

4.4.3 Deep Learning Superiority

The deep learning technique outperforms its predecessors with an even more pronounced peak accuracy result of 89.75%. Moreover, the highest accuracies for each method: rule-based (60.20%), neural networks (85.5%), and deep learning (89.75%) define the extent to which algorithm complexity is proportional to the capacity for complex language and emotion comprehension. This comparative analysis allows one to infer that while rules based processing methods are foundational, they are outdated and not complex enough, and methods such as neural networks and deep learning are producing greater benchmark results.

In addition to this, the comparison considers the scale of adaptability of each method under varying conditions, providing a comprehensive overview of each algorithm's complexity. How each model responds to changes in their respective hyperparameters, such as lexical rule refinement or neural architecture adjustments, is critically informative for future applications and improvements in emotion detection.

Lastly, this analysis is meant to clarify the trade-offs between each method, setting insights when choosing an approach based on required accuracy and application-specific restrictions as the deep learning method tends to be more computer heavy. These line graphs eventually serve as a practical basis to compare each method's prediction accuracy and its convenience for practical utilization in related tasks.

Chapter 5

Discussion

5.1 Interpretation of Results

This comparative analysis is meant to provide insight into the effectiveness of rules-based processing, neural networks, and deep learning methods to approach emotion detection. The obtained results indicate that the rule-based approach develops a straightforward application. However, it needs more dynamism and sophistication to be able to process the complexities found in textual data, explaining why it remains with an average accuracy of 54.02%. Finally, the deep learning model, which outputted a peak at 89.75%, shows the advanced capacity of multi-layered architectures to understand and predict better the complexities of human emotion.

5.2 Context with Literature

The obtained results resonate with the literature review that indicates the superior performance of deep learning models in handling complex patterns within textual data. The prediction accuracy improved as we progressed from rules-based processing to neural networks and finally towards deep learning, which aligns with the trajectory mentioned in affective computing and natural language processing advancements. Therefore, our results confirm the growth of methodologies in current literature and expand the understanding of applying various learning models to emotion detection.

5.3 Discussion of Limitations

However, the mentioned results and methods come with their limitations. A significant restriction lies in the specific context of the dataset used for the training, which might not accurately represent the diversity of linguistic expression across different domains, languages or even slang. Moreover, hyperparameter optimization was utilized for the training of these models, it is a process that can become computationally expensive and time-consuming, therefore restricting the extent of finding ideal configurations for larger datasets or real-time applications.

5.4 Hyperparameters

Hyperparameters are vital in defining the learning structure and dynamics within machine learning models. They are responsible for the prediction accuracy. However, infinitely increasing these parameters is not convenient. The law of diminishing returns applies as benefits gained from hyperparameters overoptimization decrease with each successive increment after a specific point. The cause of this is called overfitting, which means that models become so complex that they start memorizing instead of learning from the data, reducing their prediction accuracy.

Chapter 6

Conclusions

This research provides an examination of three computational methodologies to detect human emotions: rules-based processing, neural networks, and deep learning. Rules-based processing offers simplicity and ease of application and training. However, it fails to handle the complex nature of linguistic data, as shown in its peak accuracy of only 60.20%. The neural networks approach surpassed this performance with a greater peak accuracy of 85.5%, showing its superior capacity to analyze emotions. However, the results benefited the superiority of the deep learning model, achieving a peak accuracy of 89.75%. These outcomes displayed the notable advantage of deep learning in processing complex layers of human emotion.

During this research, various emotion detection techniques were compared. By doing this, the practical implications of choosing one method over another for specific application requirements were revealed. For instance, deep learning has been mentioned to be the method that works better to detect emotions with the given data, as it excels in the validation accuracy metrics. Therefore, it is vital to highlight the importance of continuing the research of advanced computational methods.

However, the benchmark results are based on the data that was provided to the models, and it is important to mention that several limitations were found. Firstly, the model training depends on the scale of the dataset and the variation in the hyperparameters. Moreover, optimizations and limitations by computational resources and time have to be considered, too. Secondly, when interpreting the results, these factors are also being considered, and as the field of emotion detection encompasses different contexts due to the complexity of language, it might require implementing different algorithms in diverse scenarios.

This research is open to intents of future research, more specifically in terms of applying these algorithms to broader and different datasets, therefore experimenting with different contexts. Further exploration into different machine learning models might prove beneficial in the quest to understand the capabilities of emotion detection methods.

To sum up, this comparative study aims to reinforce the use of advanced machine learning techniques in the domain of affective computing and emotion detection. Through continued research and technological advancement, the capability to comprehend and understand human emotions is guaranteed to evolve, paving the way for innovations that can transform our digital interactions.

Bibliography

[1] COBAN, E. B. *Neural Networks and Their Applications*. 2016.

[2] DONGARE, A. D., K.-R. R. . K. A. D. *Introduction to Artificial Neural Network*. 2012.

[3] LAZRIG, I., . H.-S. L. Using machine learning sentiment analysis to evaluate learning impact.

[4] LIDDY, E. . natural language processing. in encyclopedia of library and information science, 2001.

[5] NADKARNI, P. M., O.-M. L. . C. W. W. Natural language processing: an introduction. *Journal of the American Medical Informatics Association* (2011), 544—-551.

[6] PICARD, R. W. *Affective computing: From laughter to IEEE*. 2010.

[7] TAO, J., . T.-T. Affective computing: A review.