

AKADÉMIAI KIADÓ

Pollack Periodica •
An International Journal
for Engineering and
Information Sciences

19 (2024) 3, 8–13

DOI:
[10.1556/606.2024.01068](https://doi.org/10.1556/606.2024.01068)
© 2024 The Author(s)

ORIGINAL RESEARCH
PAPER



*Corresponding author.
E-mail: aadilganiganie@gmail.com



Exploring noise-induced techniques to strengthen deep learning model resilience

Aadil Gani Ganie* and Samad Dadvadpour

Department of Applied Informatics, Faculty of Mechanical Engineering and Informatics, Institute of Information Sciences, University of Miskolc, Miskolc-Egyetemváros, Hungary

Received: January 20, 2024 • Revised manuscript received: March 25, 2024 • Accepted: April 4, 2024
Published online: May 6, 2024

ABSTRACT

In artificial intelligence, combating overfitting and enhancing model generalization is crucial. This research explores innovative noise-induced regularization techniques, focusing on natural language processing tasks. Inspired by gradient noise and Dropout, this study investigates the interplay between controlled noise, model complexity, and overfitting prevention. Utilizing long short-term memory and bidirectional long short term memory architectures, this study examines the impact of noise-induced regularization on robustness to noisy input data. Through extensive experimentation, this study shows that introducing controlled noise improves model generalization, especially in language understanding. This contributes to the theoretical understanding of noise-induced regularization, advancing reliable and adaptable artificial intelligence systems for natural language processing.

KEYWORDS

deep learning, noise-induced regularization, model generalization, long short-term memory, bidirectional long short-term memory

1. INTRODUCTION

The rapid progression of deep learning techniques has transformed numerous domains, including natural language processing and computer vision, enabling the creation of intricate and precise models [1]. However, the escalating complexity of these models raises concerns about their inclination to overfit the training data, impeding their adaptability to unseen data instances [2]. Regularization techniques are pivotal in mitigating this challenge, encouraging model generalization by discouraging excessively intricate solutions [3]. While traditional methods like L1 and L2 regularization have been extensively explored and applied [4], recent years have witnessed a paradigm shift towards considering noise as a regularization tool. The strategic introduction of controlled noise during training has displayed promising outcomes in bolstering model resilience and refining generalization performance [5].

This study delves into the innovative realm of noise-induced regularization within deep learning models, concentrating on Long Short-Term Memory (LSTM) and Bidirectional Long Short-Term Memory (BiLSTM) architectures [6, 7]. By purposefully integrating Gaussian noise into the training procedure, the research delves into noise's impact on the models' learning dynamics and their capacity to generalize across diverse and noisy datasets [8]. The inquiry extends to the mathematical foundations of noise-induced regularization, illuminating the complex interplay among noise, model intricacy, and generalization abilities [9]. Within this context, this paper surveys existing literature on conventional regularization methods and noise-induced regularization, providing a holistic view of the evolution of regularization techniques in deep learning.

Leveraging recent investigations that have explored noise as a regularization mechanism [10], this study presents a methodical analysis of its influence on training dynamics and model efficacy. By amalgamating theoretical insights with empirical discoveries, this research

endeavors to establish a profound comprehension of noise-induced regularization and its potential applications in augmenting the resilience of deep learning models.

2. LITERATURE REVIEW

2.1. Regularization techniques in deep learning

Essential to deep learning models, regularization techniques significantly enhance their ability to generalize. Conventional methods like L1 and L2 regularization curb overfitting by penalizing large parameter values [4]. Another popular approach, dropout, introduced by Srivastava et al. [10], involves randomly dropping units during training to prevent co-adaptation of feature detectors [4]. Recent innovations have introduced new regularization methods, including batch normalization [11] and weight regularization [12], proving effective in optimizing model performance.

2.2. Noise-induced regularization in deep learning

In recent years, noise-induced regularization has emerged as a promising strategy to bolster model generalization. Nee-lakantan et al. [5] pioneered the concept of adding gradient noise, enhancing the learning dynamics of deep networks [13]. Vincent et al. [8] proposed denoising autoencoders, which learn robust features by reconstructing clean inputs from noisy data. Dhifallah and Lu [14] harnessed noise injection as a regularization technique, enhancing the robustness of convolutional neural networks in image recognition tasks [5]. These studies underscore the potential of noise-induced regularization in fortifying deep learning models against noisy input data.

2.3. Application of regularization techniques in natural language processing

In the realm of Natural Language Processing (NLP), regularization techniques have significantly enhanced task performance. Vincent et al. [8] investigated the impact of regularization methods on machine translation tasks, demonstrating L2 regularization's effectiveness in enhancing translation accuracy. Furthermore, recent studies have explored noise-induced regularization specifically in NLP tasks. K. Zhang et al. [15] introduced noise to word embeddings, enhancing sentiment analysis accuracy. A. Pretorius et al. [16] applied noise-induced regularization to recurrent neural networks, improving text generation model performance.

2.4. Noise-induced regularization in LSTM and BLSTM models

The application of noise-induced regularization in LSTM and BiLSTM models has gained traction. M. Qiao et al. [17] introduced noise to LSTM networks' input sequences, resulting in improved performance in sequence prediction tasks. Similarly, A. A. Abdelhamid et al. [18] incorporated noise into the training process of BiLSTM models, enhancing

their ability to capture complex dependencies in sequential data. These studies highlight noise-induced regularization's potential in LSTM and BiLSTM [19] architectures for diverse sequential tasks. In summary, while traditional methods provide a foundation, recent strides in noise-induced regularization offer promising avenues to enhance model robustness, especially in the context of LSTM and BiLSTM architectures for sequential tasks in NLP and beyond.

3. RESULTS AND DISCUSSIONS

3.1. Methodology

Data preparation: A dataset comprising text data with “suicide” and “non-suicide” labels was preprocessed using NLP techniques, including tokenization and lemmatization.

Model architectures: LSTM and BiLSTM architectures were chosen for their effectiveness in sequential data tasks.

Noise-induced regularization: Controlled Gaussian noise was injected into the input data to induce regularization, preventing overfitting and enhancing the models' robustness as it is shown in Fig. 1.

Training and optimization: The models were trained using Adam optimizer with backpropagation, optimizing a composite objective function that integrated noise-induced regularization.

Evaluation: Model performance was evaluated based on accuracy scores under varying noise levels, demonstrating

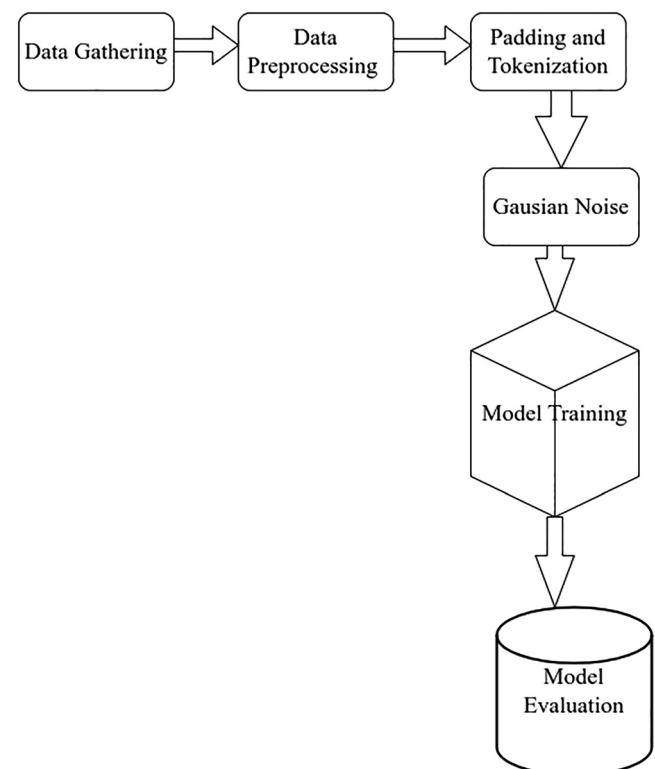
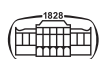


Fig. 1. Proposed workflow within the use of noise induced regularization



the effectiveness of noise-induced regularization in improving generalization capabilities.

3.2. Mathematical optimization

This section, delves into the mathematical formulations, optimization strategies, and experimental results of noise-induced regularization experiments on LSTM and BiLSTM models. These experiments were conducted on a dataset comprising text data with two distinct classes, utilizing techniques from the field of NLP. The LSTM model processes input sequences x_t and maintains a hidden state h_t and a cell c_t . The equations for the LSTM are,

$$i_t = \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}), \quad (1)$$

$$f_t = \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}), \quad (2)$$

$$g_t = \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}), \quad (3)$$

$$o_t = \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}), \quad (4)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot g_t, \quad (5)$$

$$h_t = o_t \cdot \tanh(c_t), \quad (6)$$

where x_t is the input at time step t ; i_t, f_t, g_t, o_t are the input, forget, cell and output gates, respectively; σ represents the sigmoid activation function, b is the bias and W weights.

The objective function for the LSTM model, incorporating noise-induced regularization, is defined as:

$$\min_{\theta} l_{lstm}(\theta) + \lambda \cdot \text{NoiseRegularization}(\theta, \sigma). \quad (7)$$

Here, θ represents the model parameters, $l_{lstm}(\theta)$ denotes the standard LSTM loss function applied to the NLP task, λ controls the regularization strength, and $\text{NoiseRegularization}(\theta, \sigma)$ quantifies the noise-induced regularization term, where σ signifies the noise level. Bidirectional LSTM processes the input sequence in both forward and backward directions. The equations for BiLSTM and similar to LSTM, but it processes the input in reverse order as well,

$$h1_t = \text{LSTM}_{forward}(h1_{t-1}, x_t), \quad (8)$$

$$h2_t = \text{LSTM}_{backward}(h2_{t+1}, x_t), \quad (9)$$

$$h_t = [h1_t + h2_t], \quad (10)$$

where $h1_t$ and $h2_t$ are the hidden states of forward and backward LSTM's respectively.

The objective function for both LSTM and BiLSTM models can be defined as a combination of a loss term and a regularization term to handle noise. Considering Mean Squared Error (MSE) as the loss function, the objective function (l) can be defined as follows:

$$l = \text{MSE}(\hat{y}, y_{true}) + \lambda \cdot \text{RegularizationTerm}, \quad (11)$$

where $\hat{y}$ represents the predicted output; y_{true} represents the true output; λ controls the strength of the regularization term; $\text{RegularizationTerm}$ can be a measure of the noise or complexity in the model, e.g., the L2 norm of the parameters or the variance of the noise added.

To optimize the objective function with the added noise term, the noise matrix $\mathbf{n}$ is considered as part of the input to the model. The objective function with noise was the regularization term captures, the noise introduced into the model. To optimize this objective function, the partial derivative with respect to the model parameters θ are taking and considering both the prediction error and the noise regularization term. Taking the partial derivative of l with respect to θ yields:

$$\frac{\partial l}{\partial \theta} = \frac{\partial \text{MSE}(\hat{y}, y_{true})}{\partial \theta} + \lambda \cdot \frac{\partial \text{RegularizationTerm}}{\partial \theta}. \quad (12)$$

Now, let's consider $\text{RegularizationTerm} = \|\mathbf{n}\|^2/2$ where $\|\mathbf{n}\|^2$ represents the squared L2 norm of the noise matrix $\mathbf{n}$. The regularization term encourages the model to learn robust features that are not overly dependent on noisy input. The partial derivation of the regularization term with respect to θ is $\lambda \cdot n$, assuming n is treated as a constant during differentiation. Therefore, the overall partial derivative of the objective function with respect to θ becomes

$$\frac{\partial l}{\partial \theta} = \frac{\partial \text{MSE}(\hat{y}, y_{true})}{\partial \theta} + \lambda \cdot n. \quad (13)$$

This partial derivative has been used in gradient-based optimization algorithm Adam in this case to update the model parameters θ during training. The regularization term $\lambda \cdot n$ penalizes the model for being sensitive to the noise introduction, leading to more generalized and robust learning.

The dataset utilized in these experiments consists of text data with two distinct classes: "suicide" and "non-suicide." This dataset was preprocessed using various NLP techniques, including tokenization, stopword removal, and lemmatization, to prepare the text data for model training.

Figures 2–9 illustrates the accuracy scores for both LSTM and BiLSTM models under varying noise levels 0, 0.2, 0.5, and 0.7, respectively. In this study Gaussian noise has been added, which is a kind of signal noise that has a probability

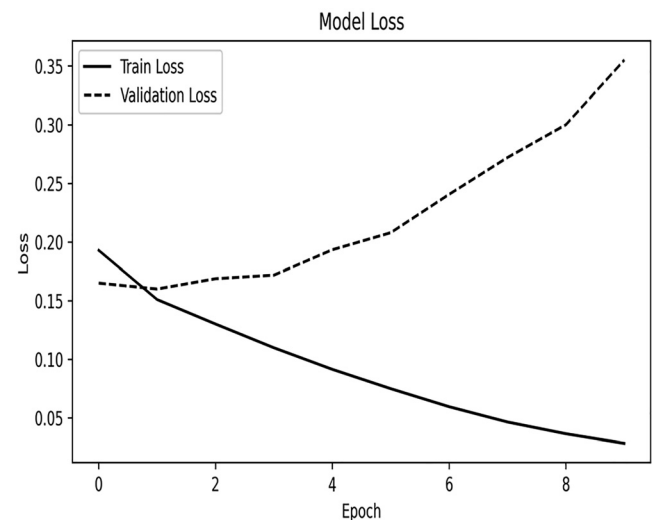


Fig. 2. LSTM model loss and accuracy at noise = 0.0

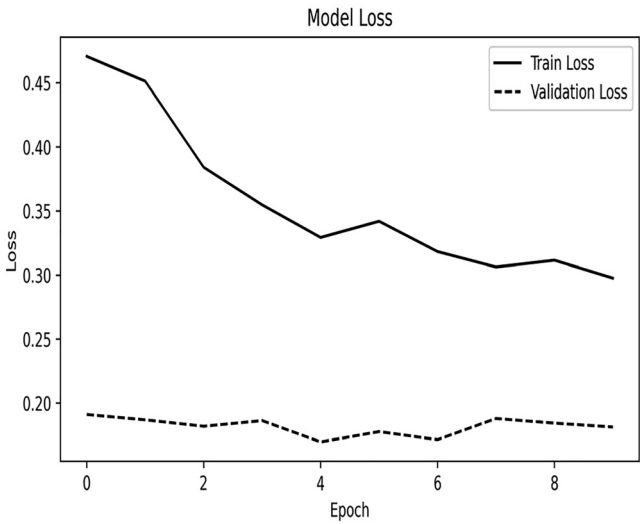


Fig. 3. LSTM model accuracy and loss at noise = 0.2

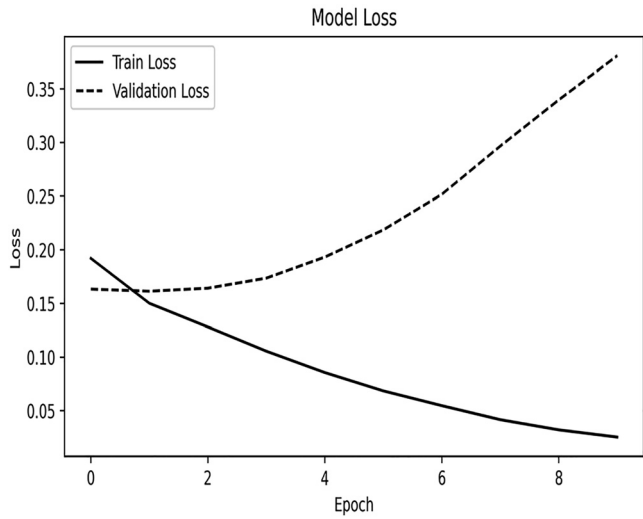


Fig. 6. BiLSTM model accuracy and loss at noise = 0.0

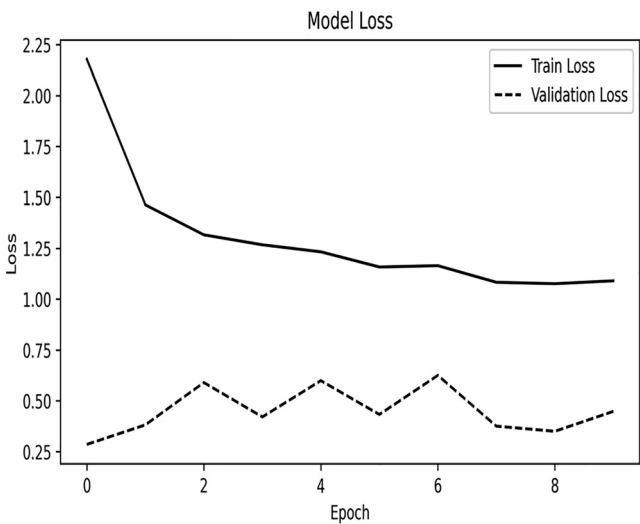


Fig. 4. LSTM model accuracy and loss at noise = 0.5

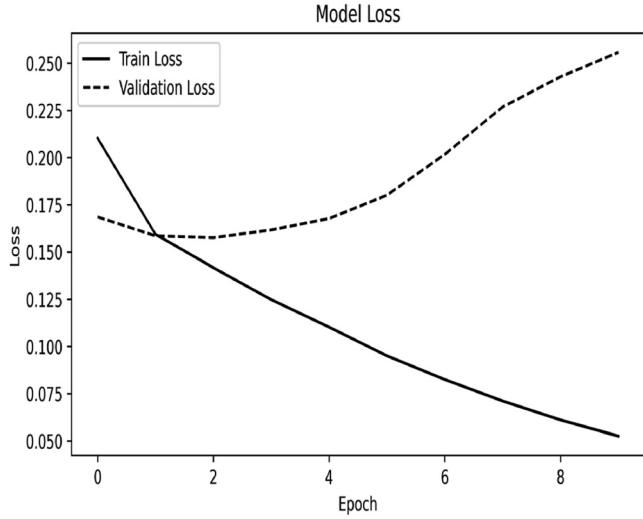


Fig. 7. BiLSTM model accuracy and loss at noise = 0.2

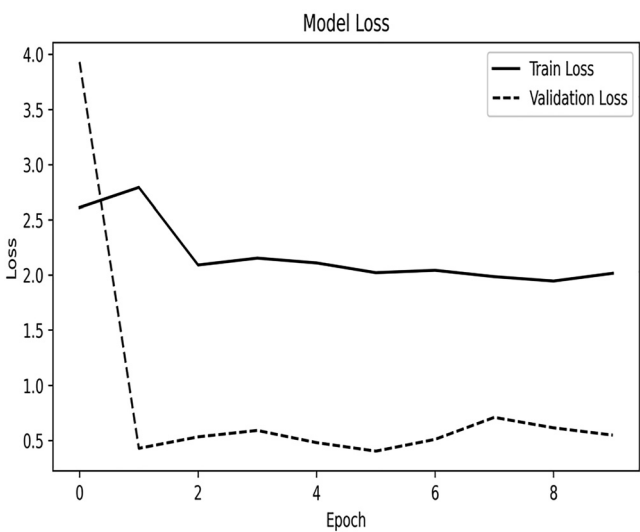


Fig. 5. LSTM model accuracy and loss at noise = 0.7

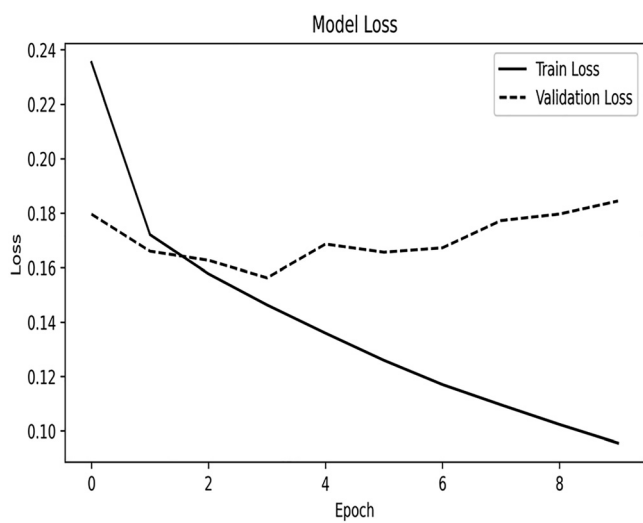


Fig. 8. BiLSTM model accuracy and loss at noise = 0.5



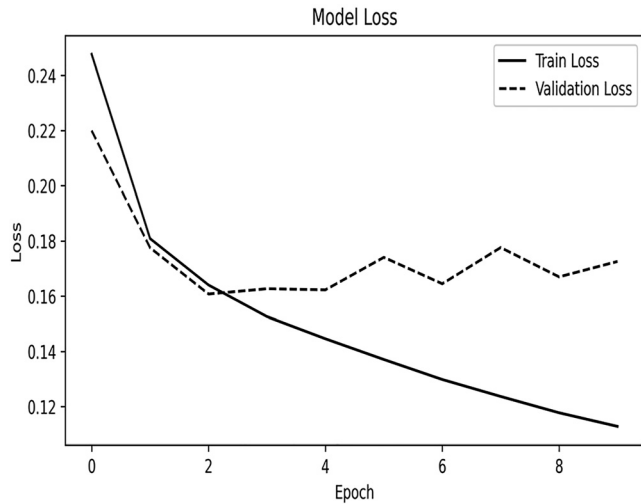


Fig. 9. BiLSTM model accuracy and loss at noise = 0.7

density function equal to that of the normal distribution. The figures reveal that, with the introduction of controlled noise, the training accuracy initially decreases due to increased complexity. However, the testing accuracy remains stable, showcasing the noise-induced regularization's efficacy in preventing overfitting and enhancing the models' generalization capability for NLP tasks.

In the polynomial regression experiments, aimed at demonstrating the impact of different regularization techniques on overfitting, Figures 10–12 presents the results. These figures show cases the overfitting phenomenon with a 9th-degree polynomial, the effectiveness of dropout-like regularization with a 3rd-degree polynomial, and the noise-induced regularization's ability to mitigate overfitting with a 7th-degree polynomial.

These results underscore the vital role of noise-induced regularization in enhancing model generalization for text data in the domain of NLP. By incorporating noise and leveraging NLP techniques, this approach ensures the models' robustness to noisy and complex language patterns, contributing to the advancement of natural language understanding in artificial intelligent systems.

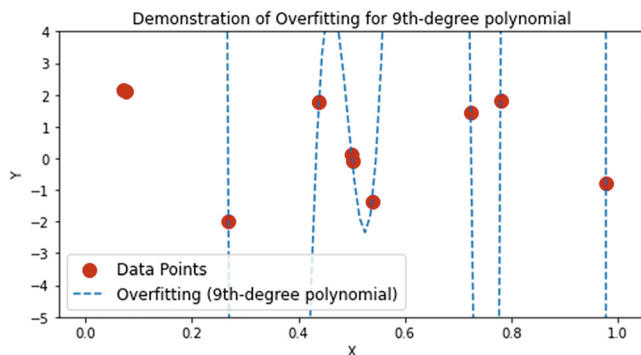


Fig. 10. Overfitting demonstration

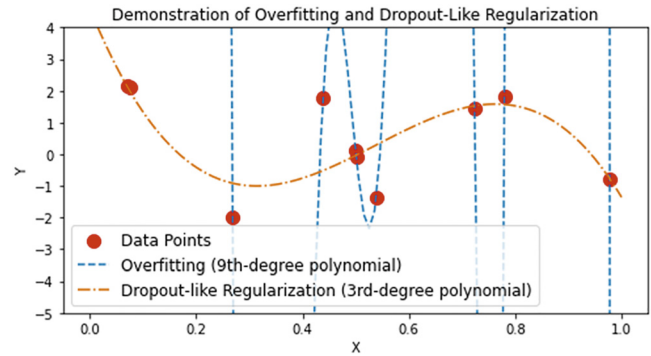


Fig. 11. Demonstration of dropout-like regularization

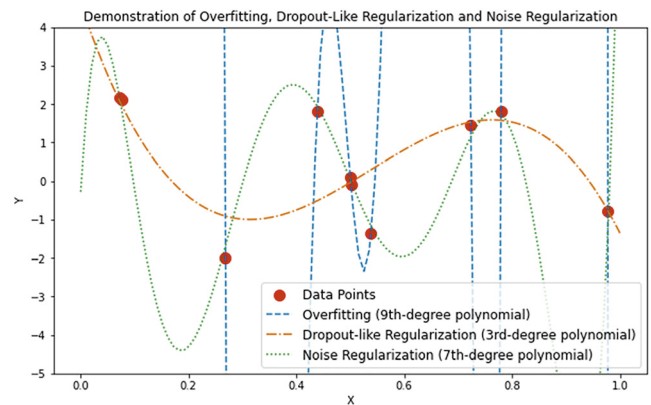


Fig. 12. Demonstration of the effect of noise

4. CONCLUSION

In conclusion, this study has delved into the innovative realm of noise-induced regularization techniques within deep learning models, with a specific focus on LSTM and BiLSTM architectures. By systematically investigating the impact of controlled Gaussian noise on model learning dynamics and generalization capabilities, this research has contributed significantly to the understanding of noise-induced regularization and its applications in enhancing the resilience of deep learning models, particularly in NLP tasks.

4.1. Summary of research results

Throughout this experimentation, it is observed a consistent trend where the introduction of controlled noise during training led to improvements in model generalization, particularly evident in language understanding tasks. Findings demonstrate that despite the initial decrease in training accuracy due to increased complexity induced by noise, the testing accuracy remains stable, underscoring the efficacy of noise-induced regularization in preventing overfitting and enhancing model robustness to noisy input data.

4.2. Main added value of the research

The primary contribution of this research lies in its comprehensive exploration of noise-induced regularization

techniques, extending the traditional understanding of regularization methods in deep learning. By elucidating the intricate interplay between noise, model complexity, and generalization abilities, this study advances the theoretical foundations of noise-induced regularization, paving the way for the development of more reliable and adaptable artificial intelligence systems, particularly in the domain of NLP. Furthermore, our investigation into the application of noise-induced regularization in LSTM and BiLSTM architectures enriches the existing literature by providing insights into their effectiveness for diverse sequential tasks in NLP and beyond.

4.3. Future directions

Moving forward, future research endeavors can build upon the findings of this study by exploring optimal integration strategies for traditional and noise-induced regularization methods, with a particular focus on their transferability to a broader range of deep learning tasks. Additionally, investigating noise-induced regularization in emerging architectures and real-world applications holds promise for further advancements in building robust and resilient deep learning models. In essence, this research not only contributes to the theoretical understanding of noise-induced regularization but also offers practical insights that can potentially drive advancements in the development of more adaptive and reliable deep learning models for various applications, particularly in the domain of natural language processing.

ACKNOWLEDGMENTS

The Authors would like to acknowledge the support and guidance provided by the mentors and colleagues throughout this research project.

REFERENCES

- [1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.
- [2] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [3] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv:1609.04747*, 2016.
- [4] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [5] A. Neelakantan, L. Vilnis, Q. V. Le, I. Sutskever, L. Kaiser, K. Kurach, and J. Martens, "Adding gradient noise improves learning for very deep networks," *arXiv:1511.06807*, 2015.
- [6] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Netw.*, vol. 18, nos 5–6, pp. 602–610, 2005.
- [7] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [8] P. Vincent, H. Larochelle, Y. Bengio, and P. A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th International Conference on Machine Learning*, Helsinki, Finland, July 5–9, 2008, pp. 1096–1103.
- [9] Q. Zheng, M. Yang, J. Yang, Q. Zhang, and X. Zhang, "Improvement of generalization ability of deep CNN via implicit regularization in two-stage training process," *IEEE Access*, vol. 6, pp. 15844–15869, 2018.
- [10] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *J. Machine Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [11] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on International Conference on Machine Learning*, Lille, France, July 6–11, 2015, pp. 448–456.
- [12] T. van Laarhoven, "L2 regularization versus batch and weight normalization," *arXiv:1706.05350*, 2017.
- [13] A. G. Ganieand and S. Dadvandipour, "Identification of online harassment using ensemble fine-tuned pre-trained Bert," *Pollack Period.*, vol. 17, no. 3, pp. 13–18, 2022.
- [14] O. Dhifallah and Y. Lu, "On the inherent regularization effects of noise injection during training," in *Proceedings of the International Conference on Machine Learning*, Virtual Event, July 18–24, 2021, pp. 2665–2675.
- [15] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. V. Gool, and R. Timofte, "Plug-and-play image restoration with deep denoiser prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 10, pp. 6360–6376, 2021.
- [16] A. Pretorius, H. Kamper, and S. Kroon, "On the expected behaviour of noise regularised deep neural networks as Gaussian processes," *Pattern Recognit. Lett.*, vol. 138, pp. 75–81, 2020.
- [17] M. Qiao, S. Yan, X. Tang, and C. Xu, "Deep convolutional and LSTM recurrent neural networks for rolling bearing fault diagnosis under strong noises and variable loads," *IEEE Access*, vol. 8, pp. 66257–66269, 2020.
- [18] A. A. Abdelhamid, E. S. M. El-Kenawy, B. Alotaibi, G. M. Amer, M. Y. Abdalkader, A. Ibrahim, and M. M. Eid, "Robust speech emotion recognition using CNN+ LSTM based on stochastic fractal search optimization algorithm," *IEEE Access*, vol. 10, pp. 49265–49284, 2022.
- [19] G. Kovács, N. Yussupova, and D. Rizvanov, "Resource management simulation using multi-agent approach and semantic constraints," *Pollack Period.*, vol. 12, no. 1, pp. 45–58, 2017.