

## **Abstract**

Hyperparameters are the most essential part of a deep learning model. They have a big impact for the performance of the model. Recent works show that if the hyperparameters of a Long Short Term Memory (LSTM) are carefully adjusted, its performance achieves the same performance as the more complex LSTM model. Hence, it opens opportunities for Swarm Intelligence (SI) algorithms, such as Grey Wolf Optimizer (GWO), that have promising performance in optimization problems to improve the LSTM performance by optimizing the best combination of its hyperparameters. In this paper, the GWO is exploited to optimize the LSTM hyperparameters for a language modeling task. Evaluation for the Penn Tree Bank dataset shows that GWO is capable of giving an optimum hyperparameters of the LSTM.

Keywords: *Grey Wolf Optimizer, Language Modeling, Long Short-Term Memory*