

Deep Learning Project 2

Dutt Salveen Singh, Patryk Prusak

November 18, 2024

Warsaw University of Technology

Table of contents

1. Introduction
2. Experiments
3. Dataset without background noise
4. Dataset with split background noise
5. Simpler LSTM
6. Final accuracies

Introduction

Problem

test and compare different network architectures (preferably Transformers) with different hyperparameters on a Speech Commands Dataset[1].

Hyperparameters

For the hyper-parameters, we chose to investigate batch size, learning rate, type of optimizer, number of epochs, and train-test splits

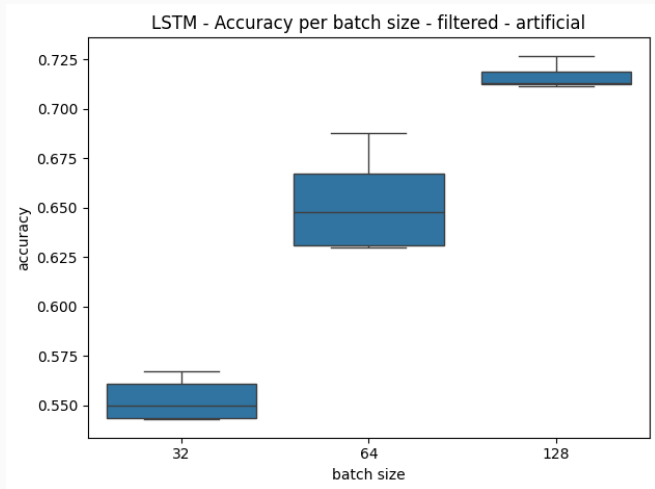
Experiments

Procedure

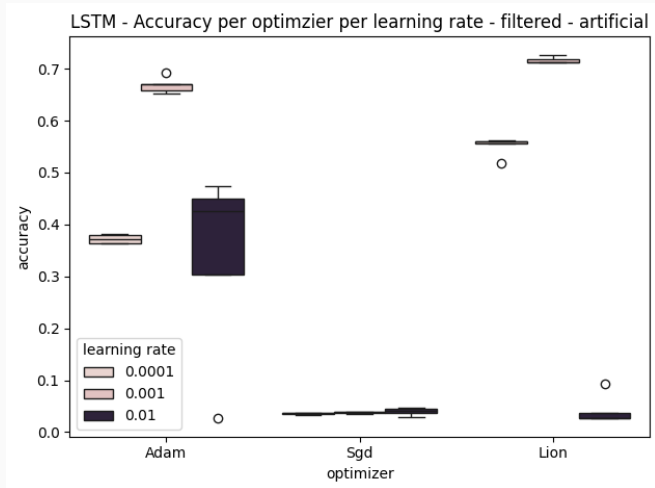
1. Choose a starting parameter X (such as batch size) and answer the following question: given a set of basic (commonly seen in literature) parameter values (such as a hyperparameter) excluding parameter X , what value for X results in best accuracy?
2. Choose the best-found value for parameter X and replace the parameter's basic initial value, then choose a different parameter X and start from 1.

1. Experiments performed on artificially removed background noise class
2. Experiments performed on a normalized dataset with split background noise

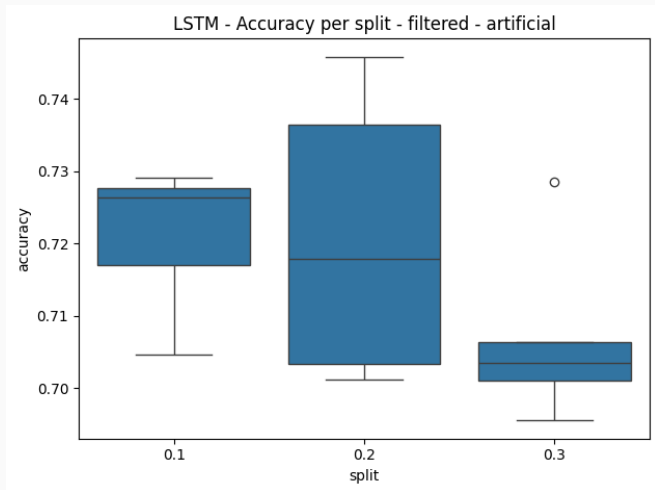
Dataset without background
noise



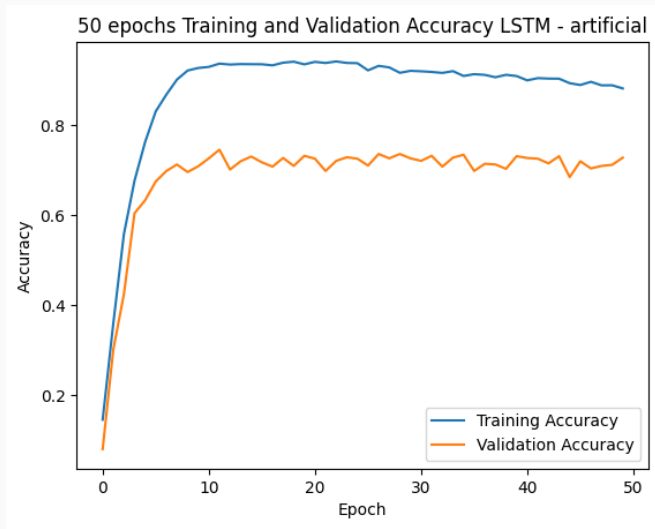
LSTM



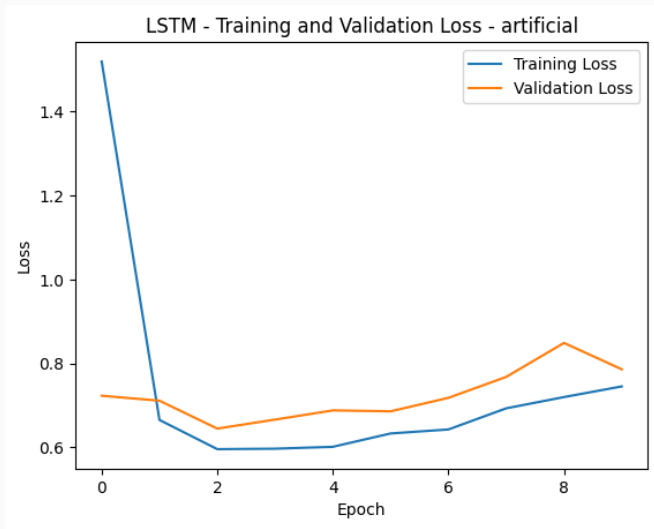
LSTM



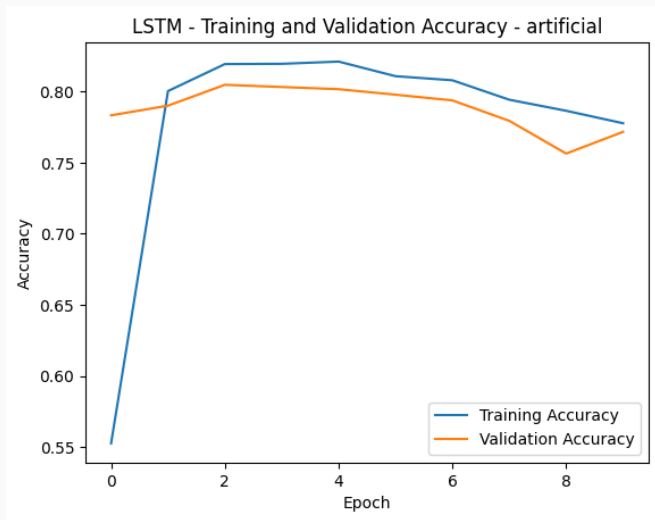
LSTM



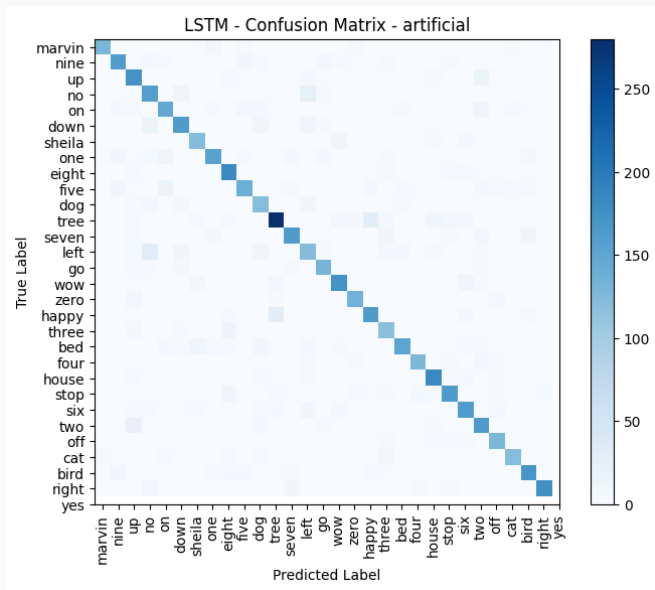
LSTM (Lion)



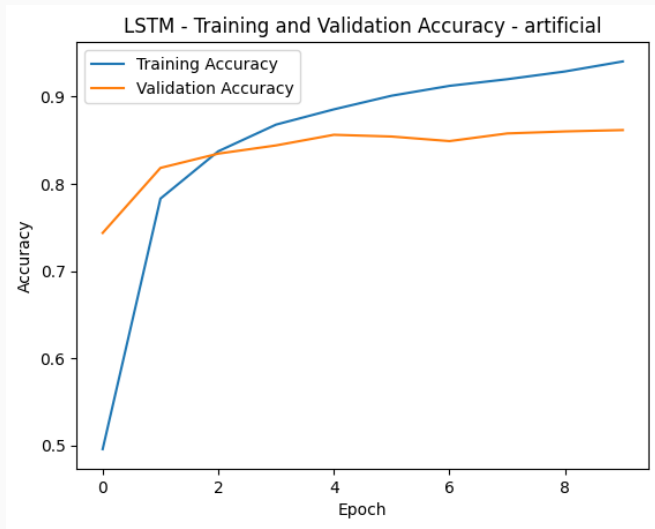
LSTM (Lion)



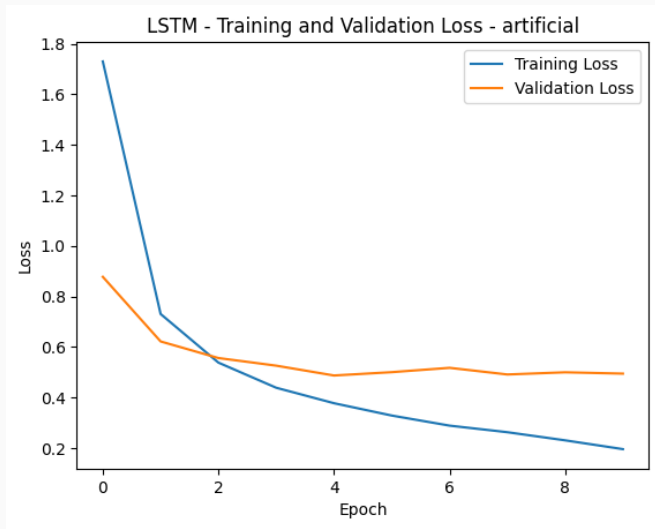
LSTM (Lion)



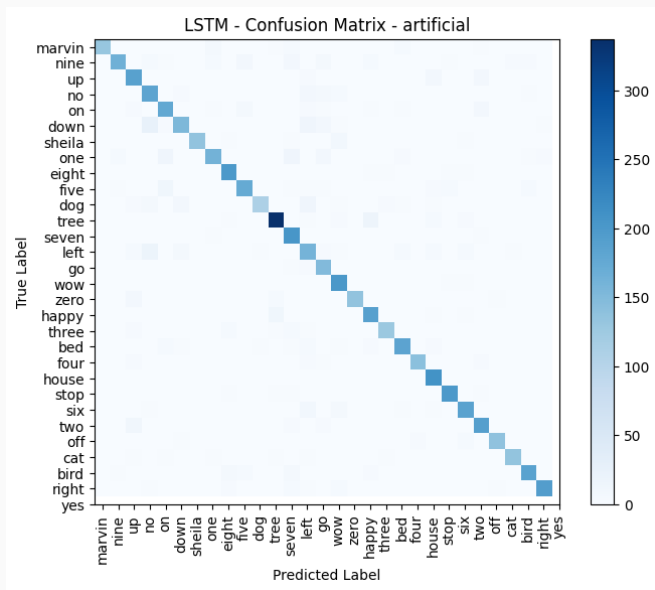
LSTM (Adam)



LSTM (Adam)

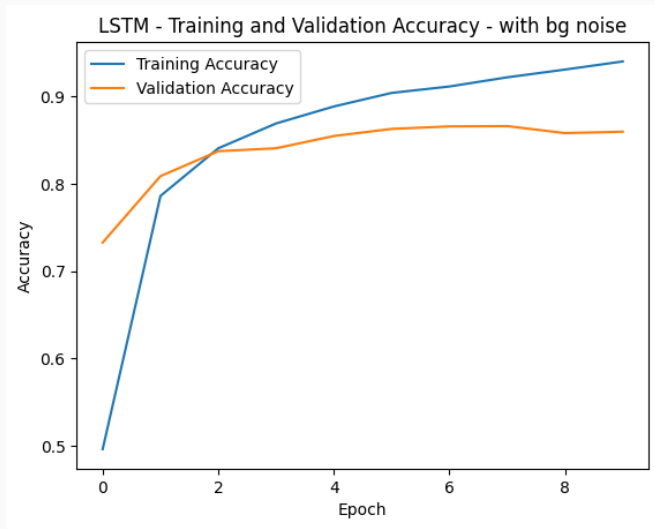


LSTM (Adam)

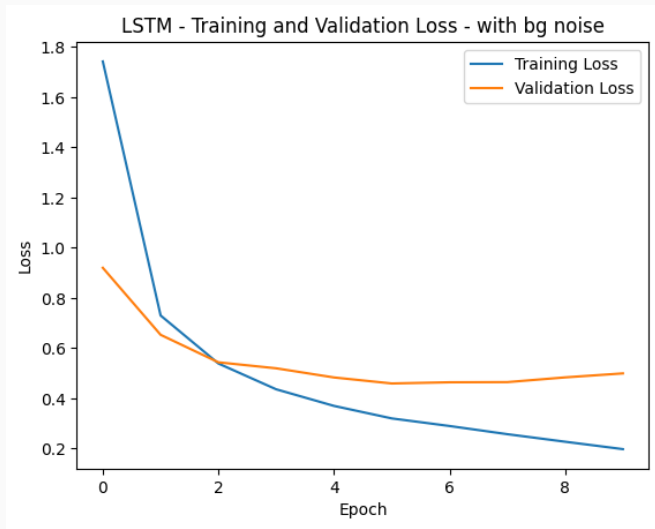


Dataset with split background noise

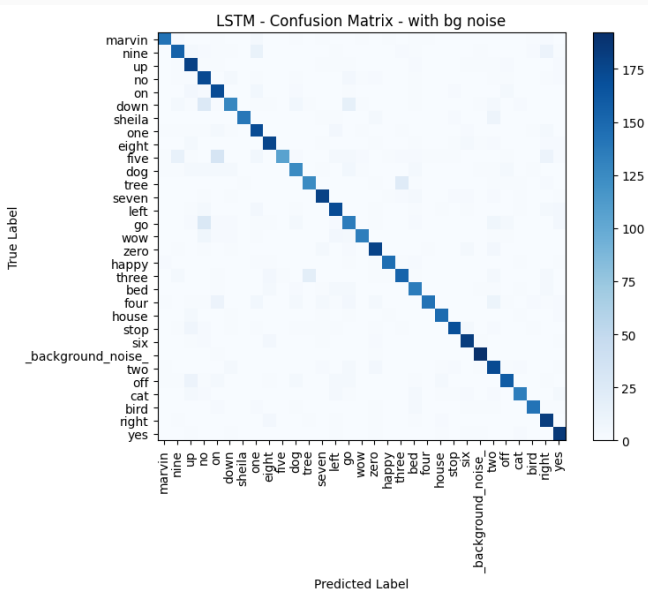
LSTM (Adam)



LSTM (Adam)

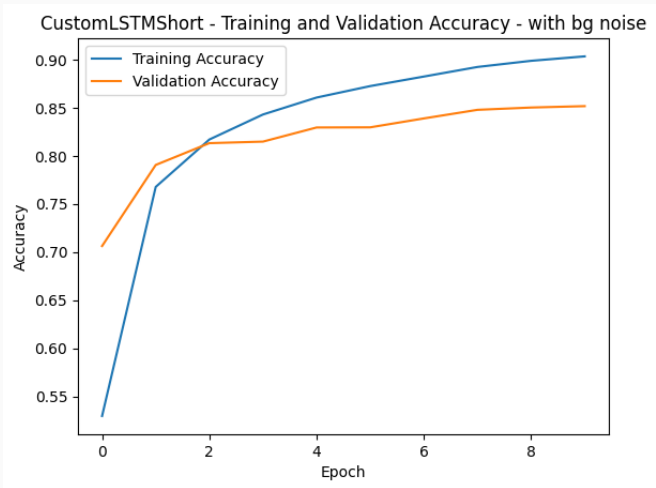


LSTM (Adam)

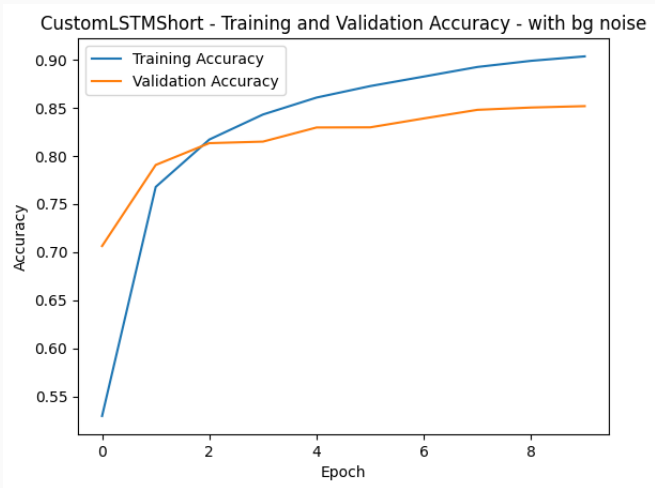


Simpler LSTM

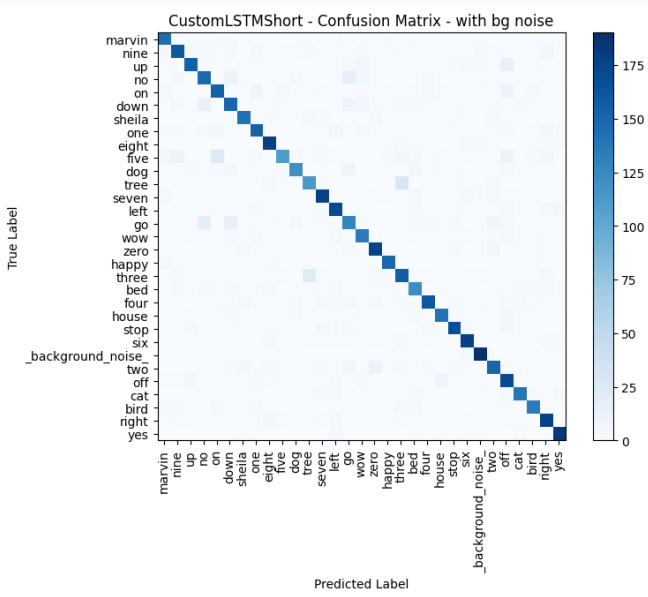
Simpler LSTM



Simpler LSTM



Simpler LSTM



Final accuracies

Final accuracies

	dataset without "_background_noise_" class	dataset with split "_background_noise_" class
Whisper	86.5%	84.2%
LSTM	86.6%	85.6%
LSTM simple	85.4%	83.8%

Thank You

References



Speech commands dataset.