# Provided MLP and LSTM architectures

## Description of MLP

* Input layer: 64 (8\*8 – board size)
* Two hidden layers: 128 and 128 Linear layers
* Output layer: 64 (8\*8 – board size)
* Optimizer: Adam
* Learning rate: 0.001
* Dropout: 0.1
* Epoch: 200 (Early stopping – 20)
* Batch size: 1000
* Len Samples: 1 (one to one)
* Number of parameters:
* The best score on DEV:

## Description of LSTM

* Input layer: 64 (8\*8 – board size)
* One hidden layer: 128 Linear layer
* Output layer: 64 (8\*8 – board size)
* Optimizer: Adam
* Learning rate: 0.005
* Dropout: 0.1
* Epoch: 200 (Early stopping – 20)
* Batch size: 1000
* Len Samples: 5 (sequence to one)
* Number of parameters:
* The best score on DEV:

# MLP optimizing

# LSTM optimizing

# Testing different optimizers

# Optimizing learning rate

# The impact of different epochs and batch size

# Learning curves

# Result analyses

# Evaluation metric

# CNN

# New data

# Conclusion