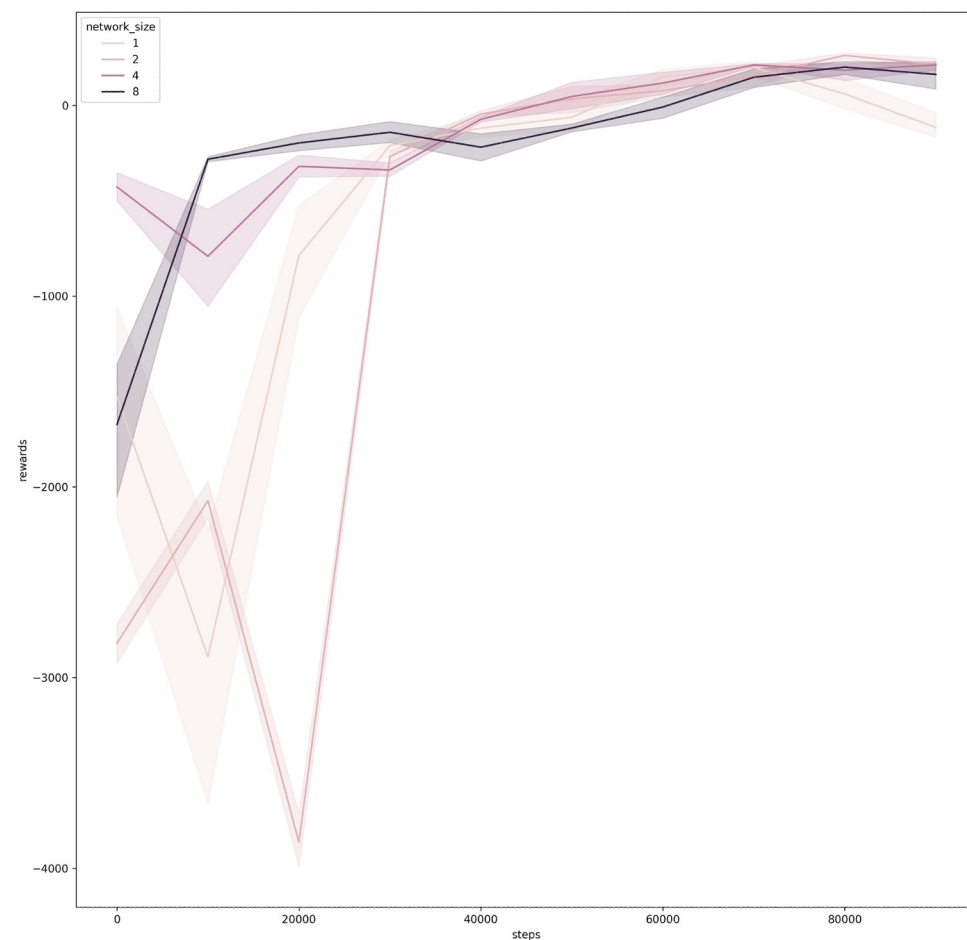


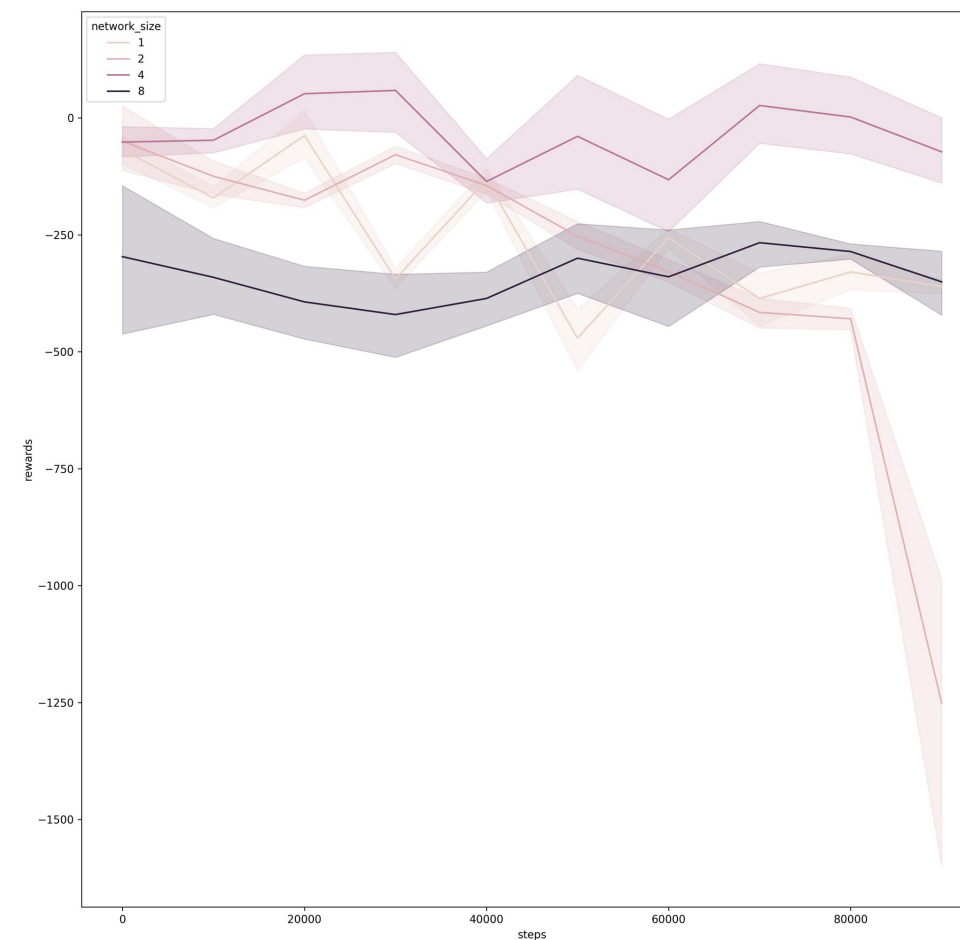
Investigating Scaling Laws in RL

- Environment: LunarLander-v2
- Algorithm: PPO with default hyperparameters and 100,000 training steps
- Same architecture for policy and value network
- Tested different network depths with 1, 2, 4 and 8 building blocks (+ 1)
- Building block:
 - Linear layer with (64, 64)
 - Batch normalization
 - ReLU activation (tanh in final layer)

Investigating Scaling Laws in RL - Resu



Without batch normalization



With batch normalization