1. Default_LSTM: The Model detail

(model = {"use_lstm": True,

"max_seq_len": 32,

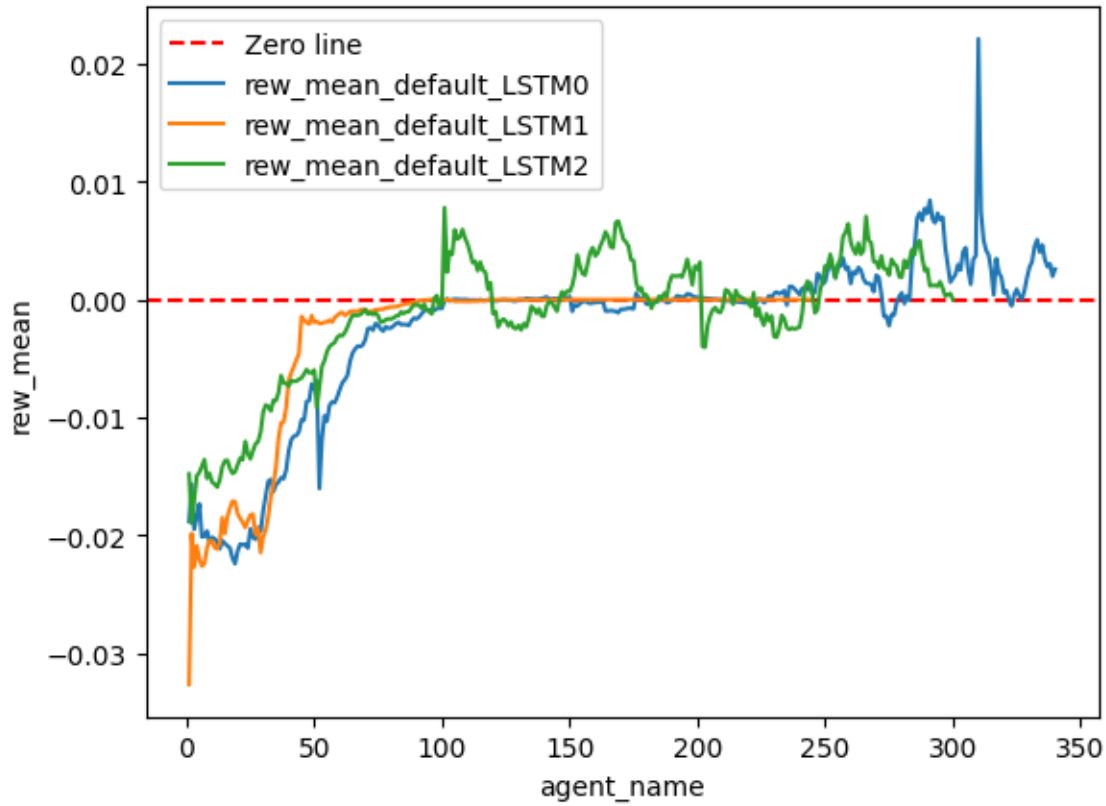
"lstm_cell_size": 256,

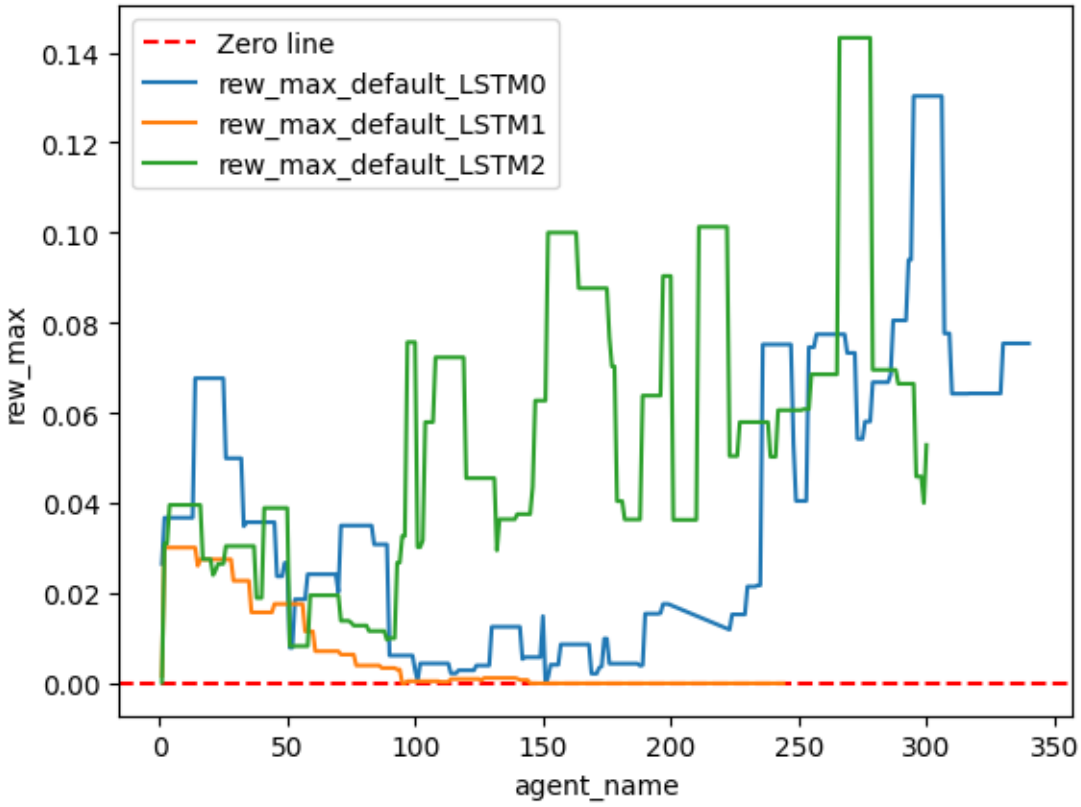
"fcnet_hiddens": [256, 256],

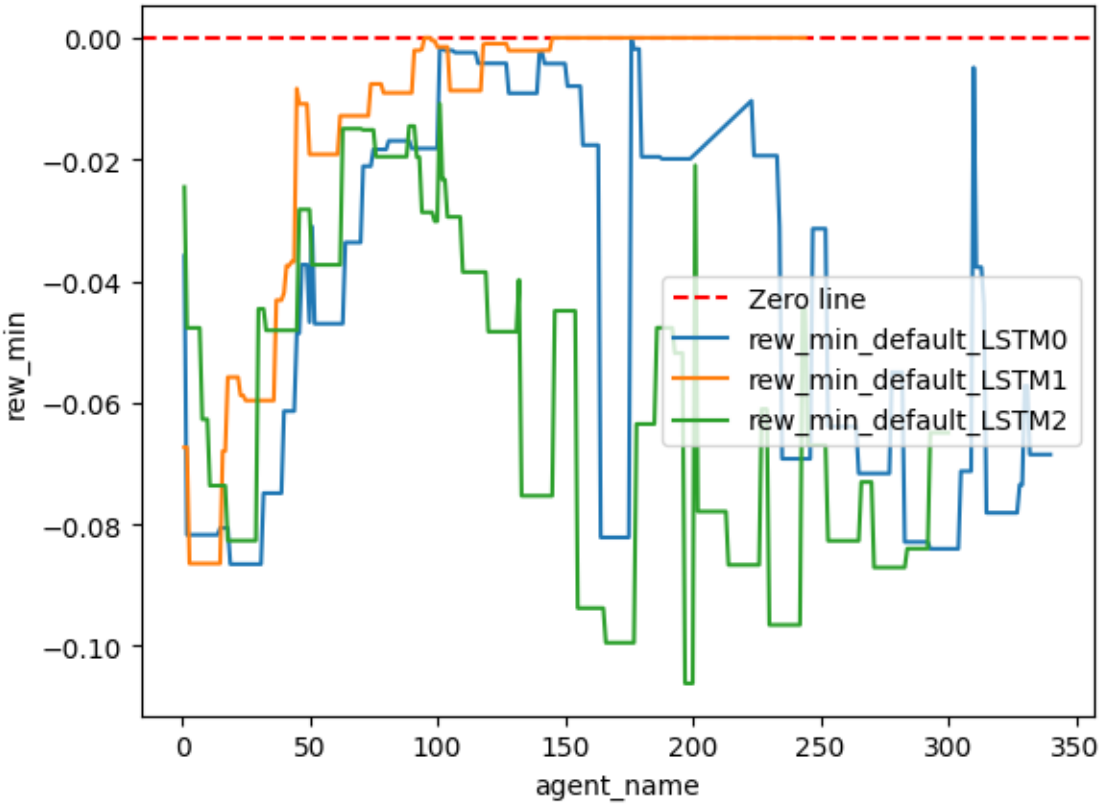
"fcnet_activation": "relu"

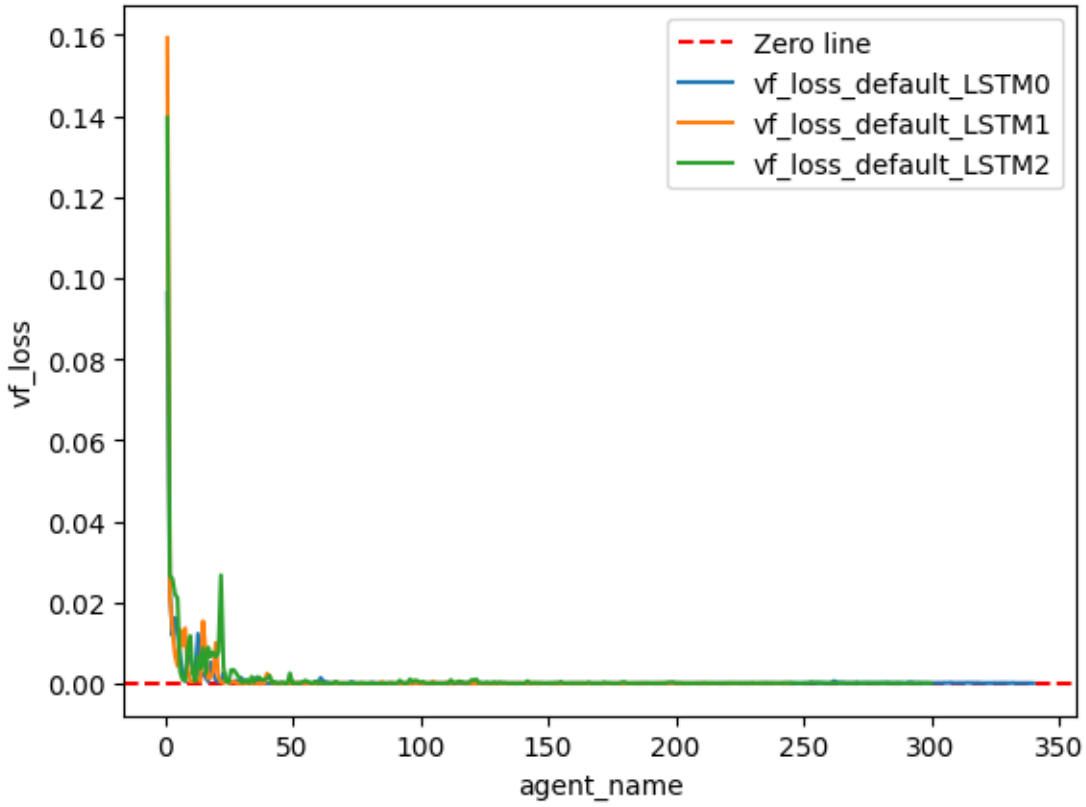
})

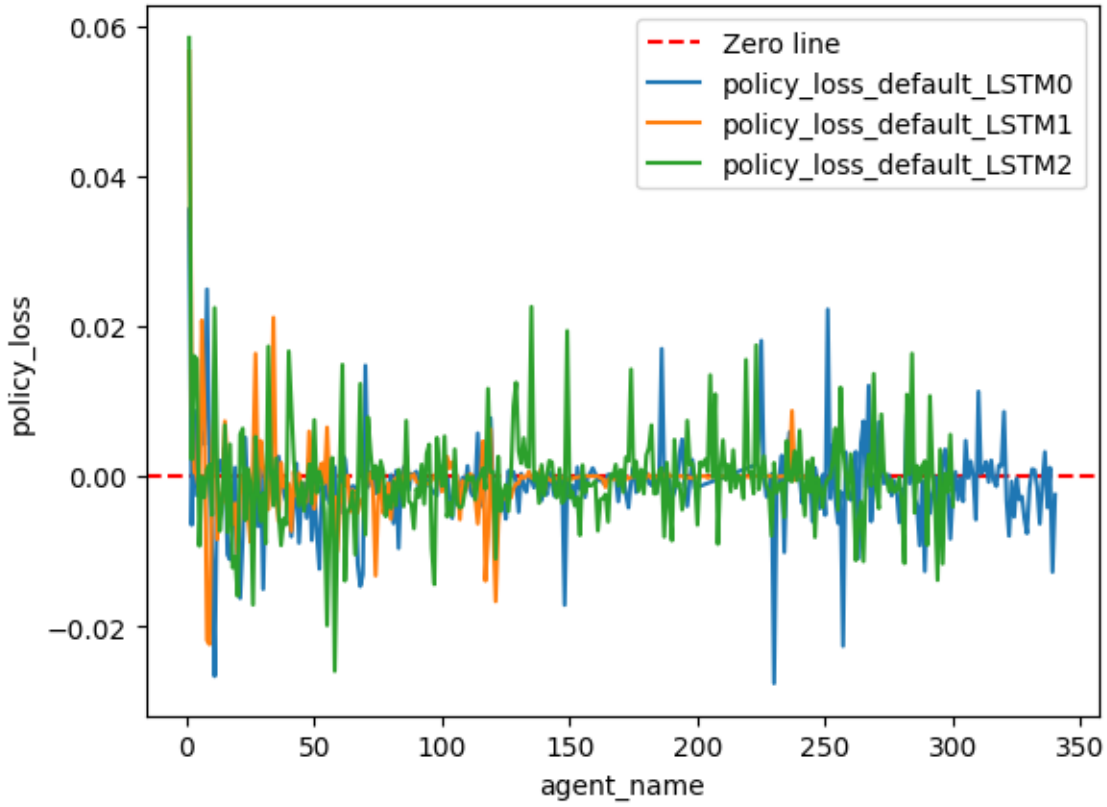
Results:




Zero line
rew_max_default_LSTM0
rew_max_default_LSTM1
rew_max_default_LSTM2
rew_max
agent_name
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
50
100
150
200
250
300
350


Zero line
rew_min_default_LSTM0
rew_min_default_LSTM1
rew_min_default_LSTM2
rew_min
agent_name
0.00
−0.02
−0.04
−0.06
−0.08
−0.10
0
50
100
150
200
250
300
350


Zero line
vf_loss_default_LSTM0
vf_loss_default_LSTM1
vf_loss_default_LSTM2
vf_loss
agent_name
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
50
100
150
200
250
300
350


Zero line
policy_loss_default_LSTM0
policy_loss_default_LSTM1
policy_loss_default_LSTM2
policy_loss
agent_name
0.06
0.04
0.02
0.00
−0.02
0
50
100
150
200
250
300
350

2. Default_LSTM_v2: A small modification is done over Default_LSTM as follows

(model = {"use_lstm": True,

"lstm_use_prev_action": True,

"lstm_use_prev_reward": True,

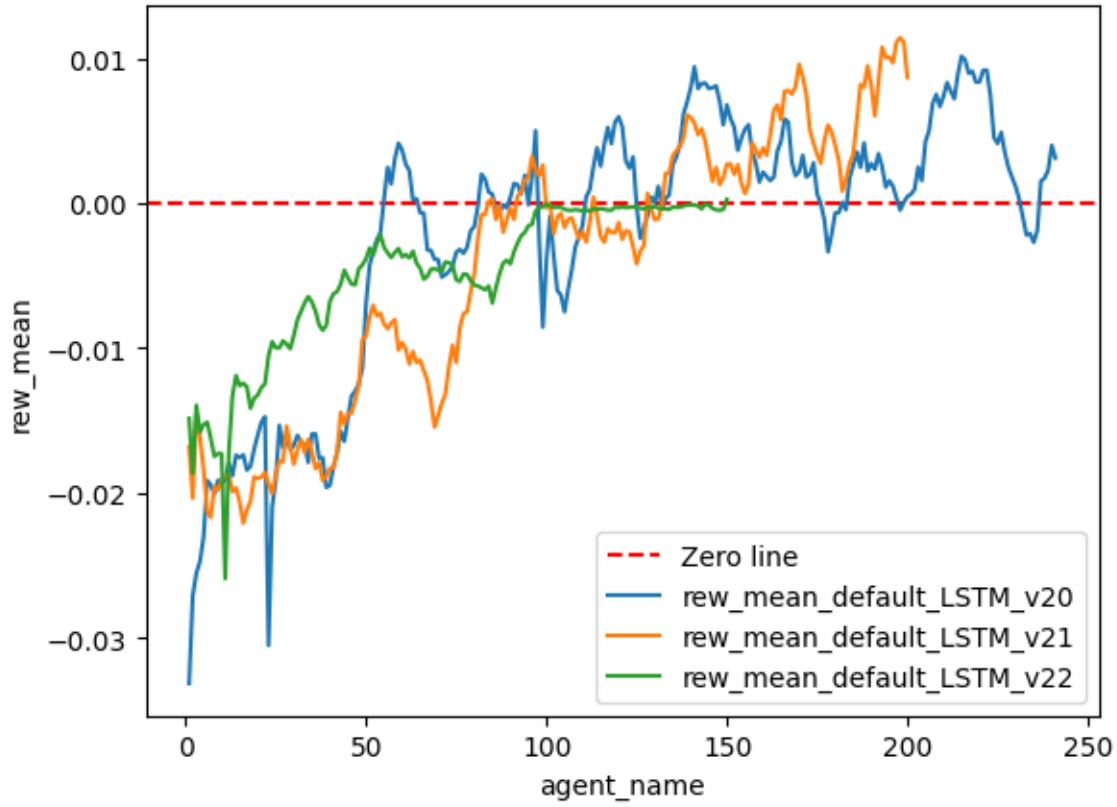
"max_seq_len": 32,

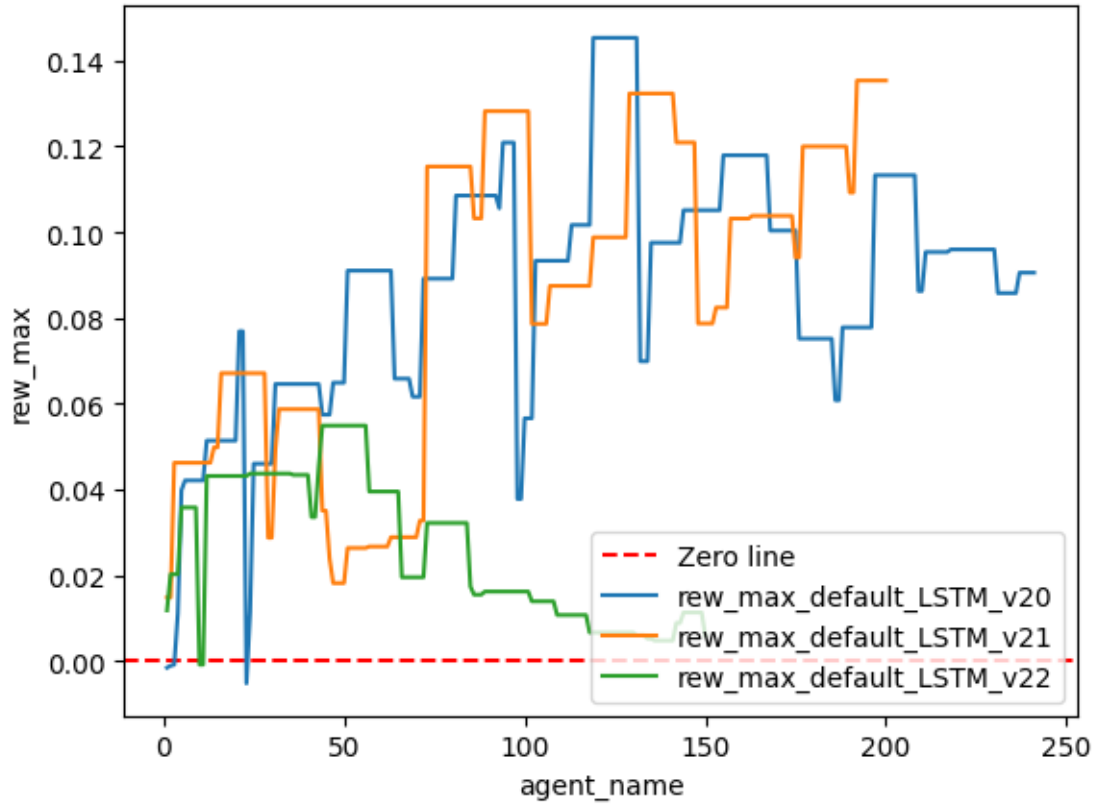
"lstm_cell_size": 256,

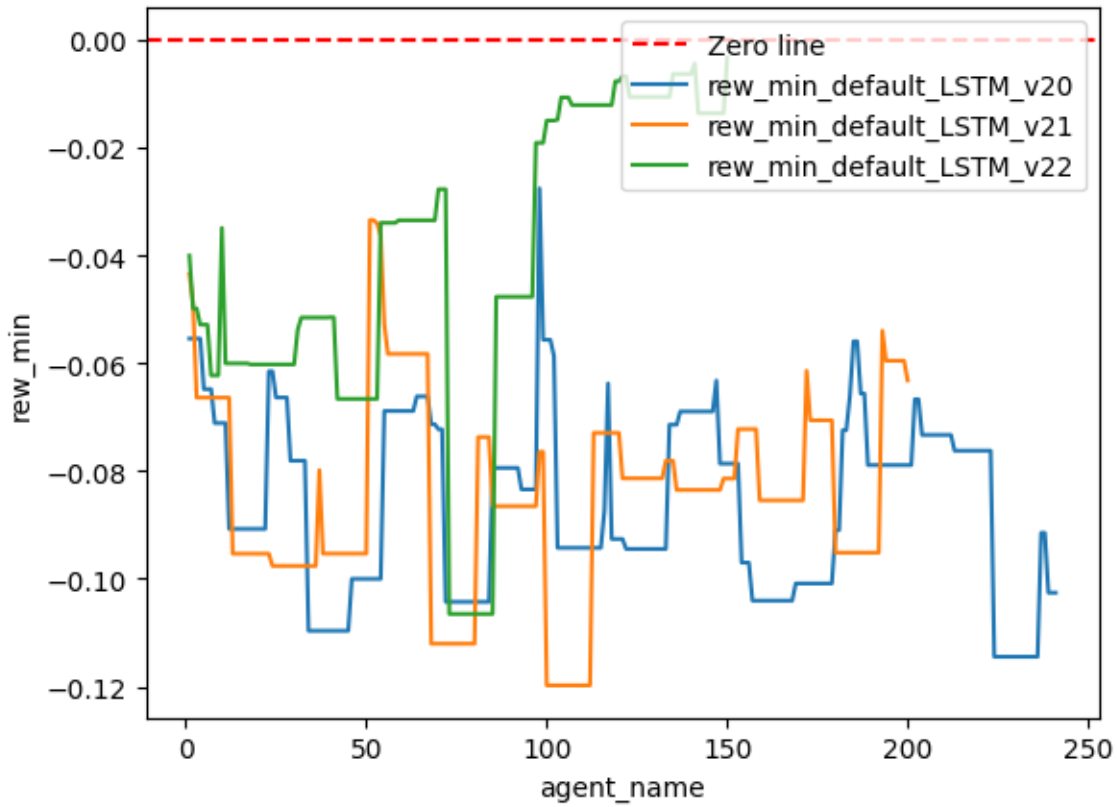
"fcnet_hiddens": [256, 256],

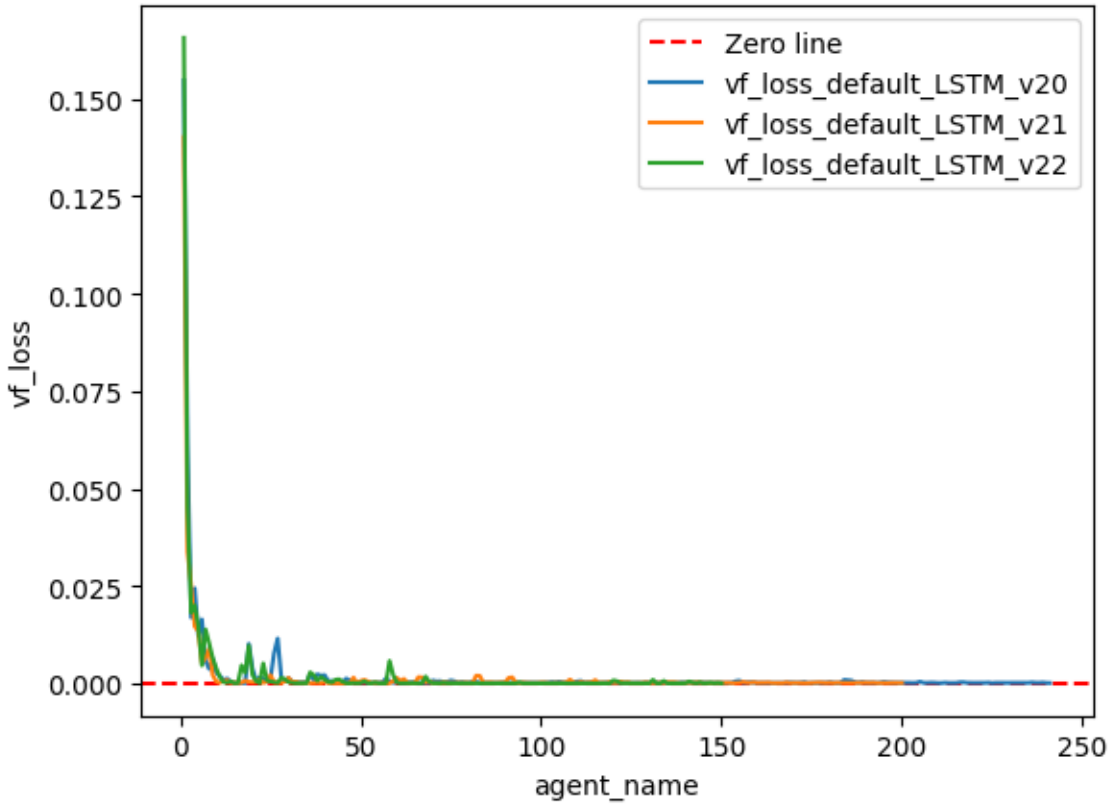
"fcnet_activation": "relu"

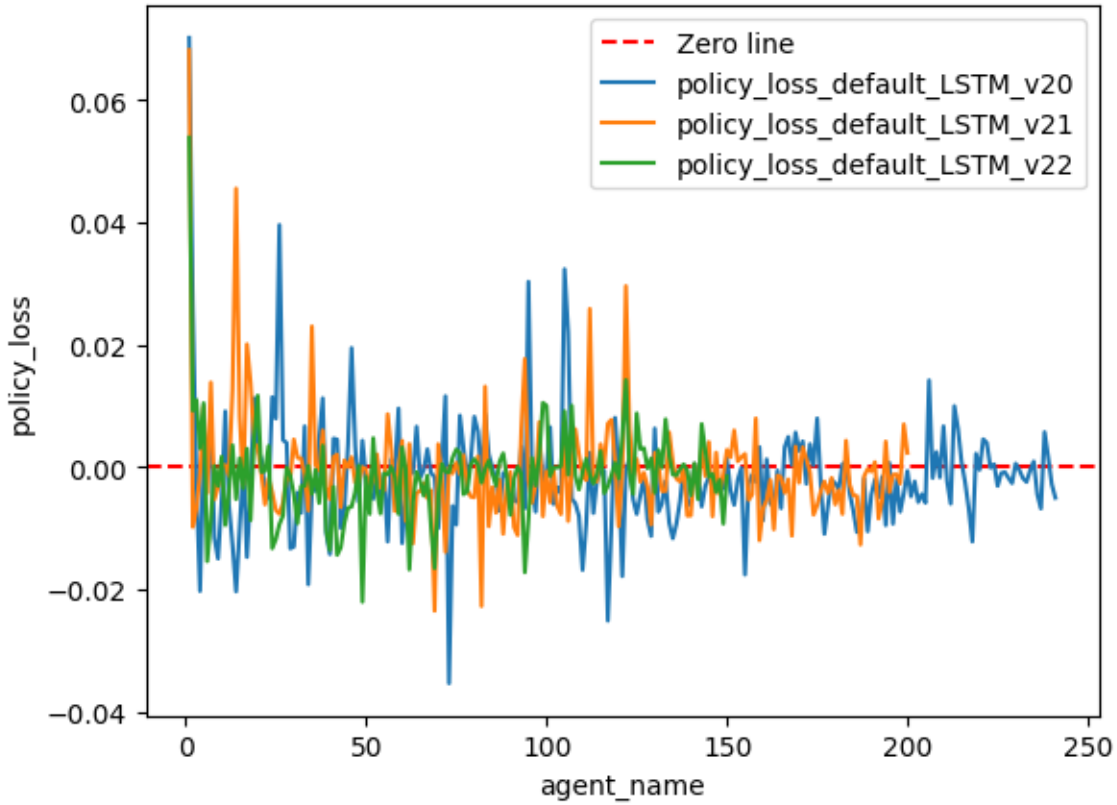
})




0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
rew_max
0
50
100
150
200
250
agent_name
Zero line
rew_max_default_LSTM_v20
rew_max_default_LSTM_v21
rew_max_default_LSTM_v22


0.00
−0.02
−0.04
−0.06
−0.08
−0.10
−0.12
rew_min
0
50
100
150
200
250
agent_name
Zero line
rew_min_default_LSTM_v20
rew_min_default_LSTM_v21
rew_min_default_LSTM_v22


Zero line
vf_loss_default_LSTM_v20
vf_loss_default_LSTM_v21
vf_loss_default_LSTM_v22
vf_loss
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0
50
100
150
200
250
agent_name


Zero line
policy_loss_default_LSTM_v20
policy_loss_default_LSTM_v21
policy_loss_default_LSTM_v22
policy_loss
0.06
0.04
0.02
0.00
−0.02
−0.04
0
50
100
150
200
250
agent_name

3. Default_LSTM_v3: A small modification is done over Default_LSTM as follows

(model = {"use_lstm": True,

"lstm_use_prev_action": True,

"lstm_use_prev_reward": True,

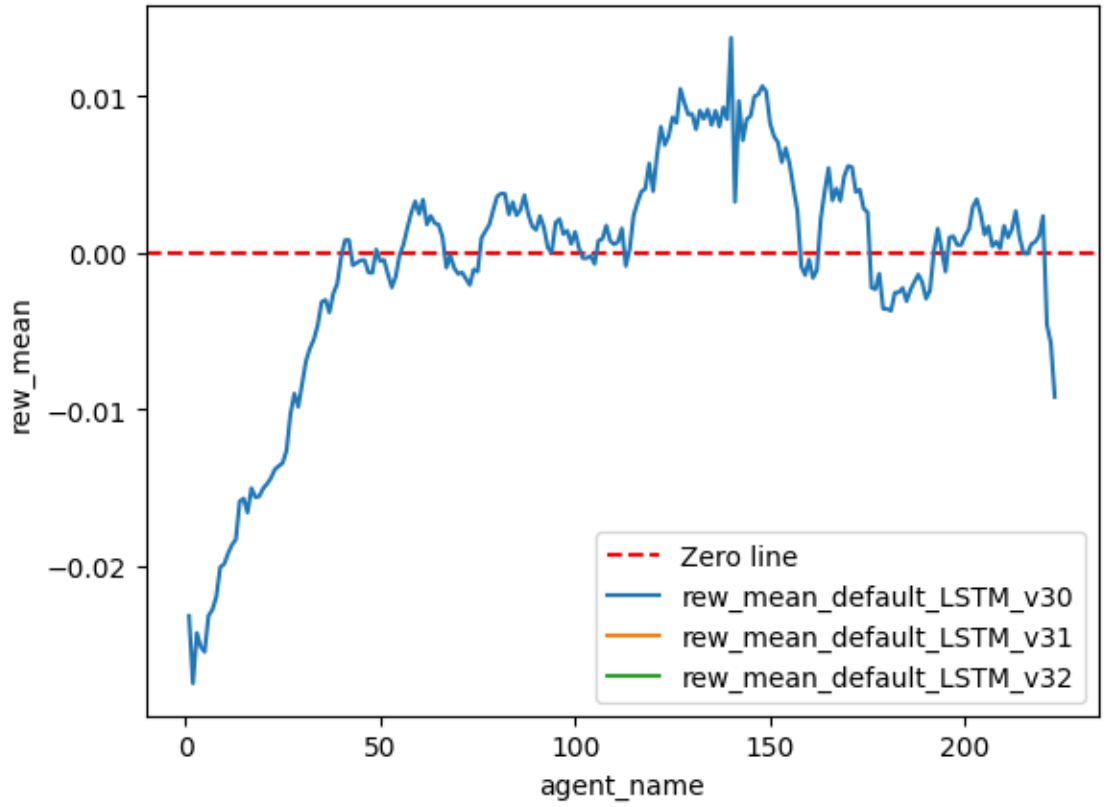
"max_seq_len": 128,

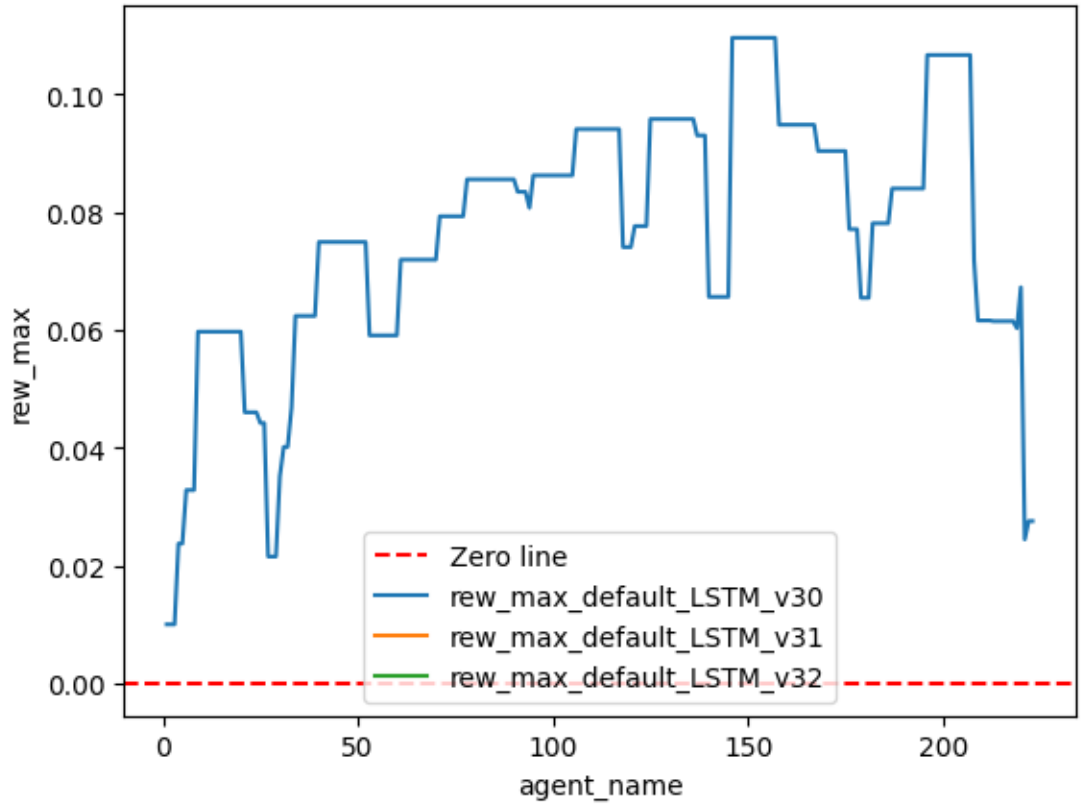
"lstm_cell_size": 256,

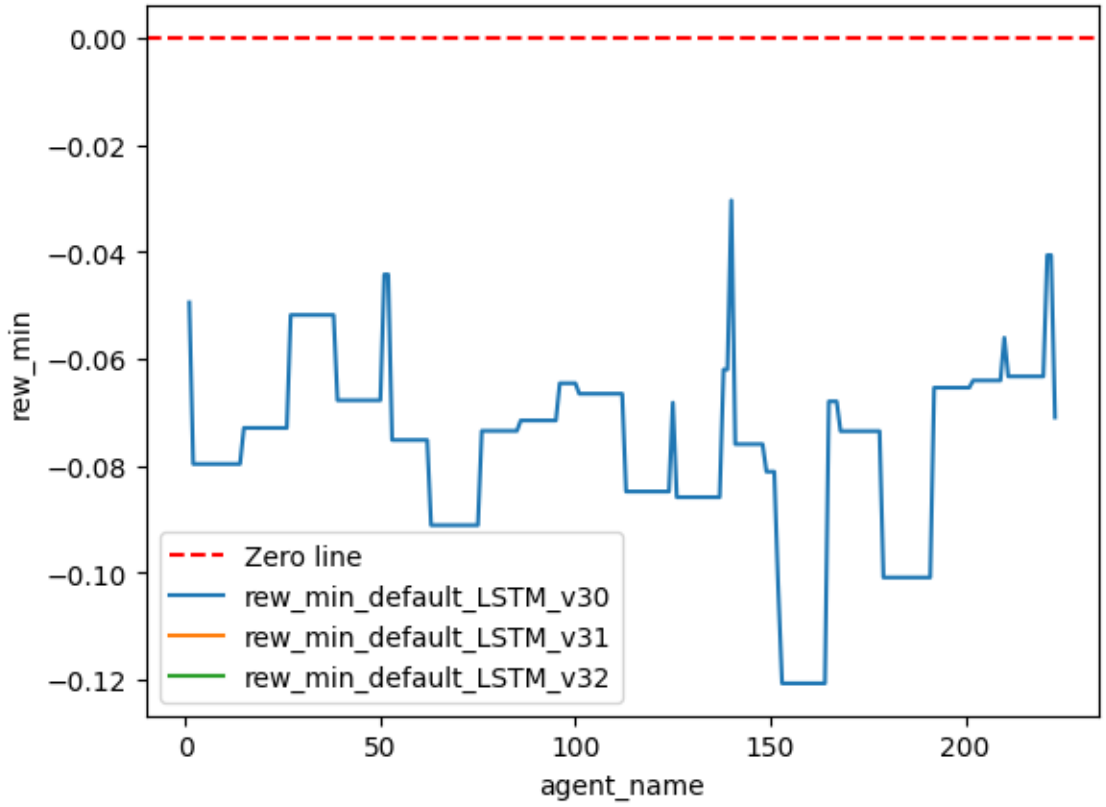
"fcnet_hiddens": [256, 256],

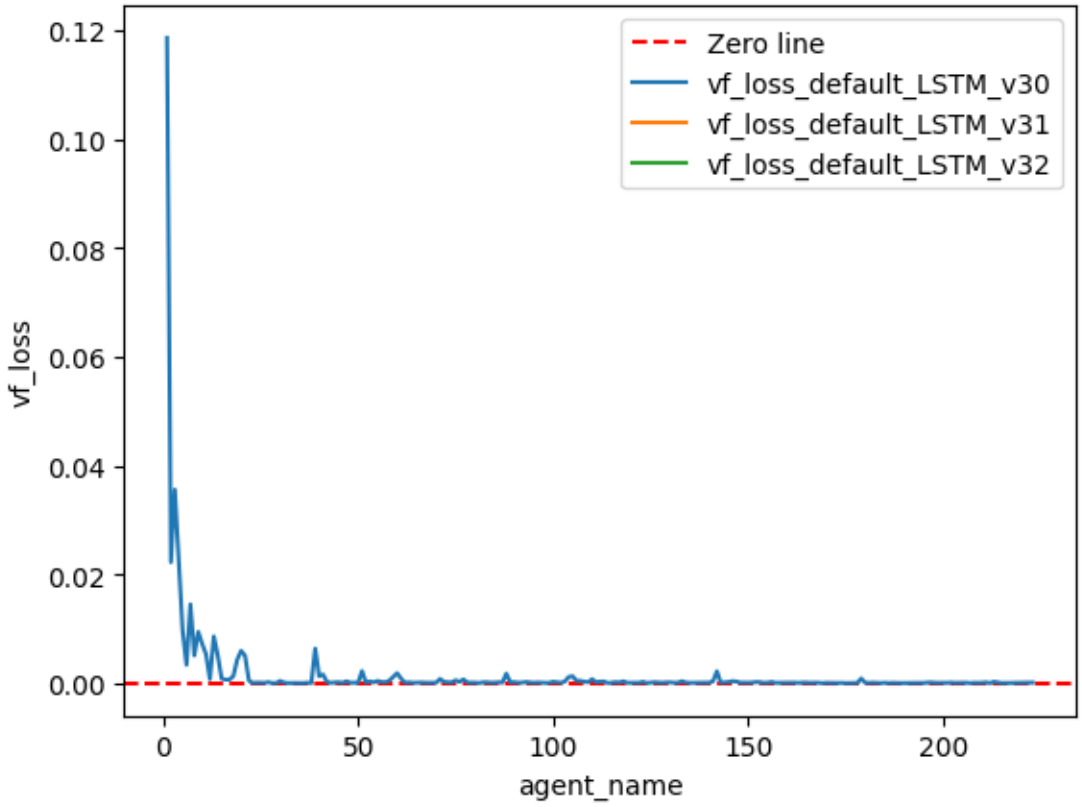
"fcnet_activation": "relu"

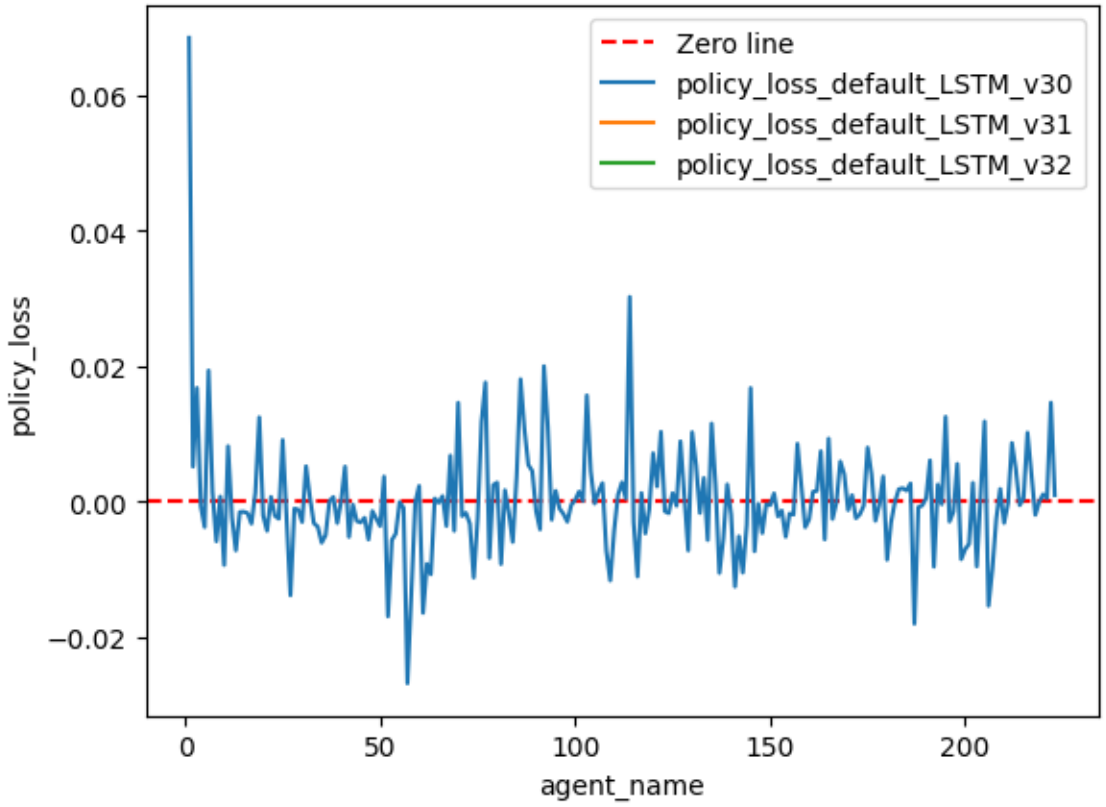
})




0.10
0.08
0.06
0.04
0.02
0.00
rew_max
0
50
100
150
200
agent_name
Zero line
rew_max_default_LSTM_v30
rew_max_default_LSTM_v31
rew_max_default_LSTM_v32


0.00
−0.02
−0.04
−0.06
−0.08
−0.10
−0.12
rew_min
0
50
100
150
200
agent_name
Zero line
rew_min_default_LSTM_v30
rew_min_default_LSTM_v31
rew_min_default_LSTM_v32


Zero line
vf_loss_default_LSTM_v30
vf_loss_default_LSTM_v31
vf_loss_default_LSTM_v32
vf_loss
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
50
100
150
200
agent_name


Zero line
policy_loss_default_LSTM_v30
policy_loss_default_LSTM_v31
policy_loss_default_LSTM_v32
policy_loss
0.06
0.04
0.02
0.00
−0.02
0
50
100
150
200
agent_name

4. Default_LSTM_v4:

(model = "use_lstm": True,

"lstm_use_prev_action": True,

"lstm_use_prev_reward": True,

"max_seq_len": 128,

"lstm_cell_size": 512,

"fcnet_hiddens": [],

"fcnet_activation": "relu"

})

The Agent is not getting trained at all. The training is getting hanged every time while other agents are getting trained fined.