Train/Val losses: 100ms hidden: 20, batch: 2048, dropout: 0.1, layers: 1