

Training & Validation Loss - {'n\_heads': 8, 'num\_layers': 3, 'ff\_hidden\_dim': 1024, 'embed\_dim': 5

