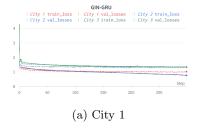
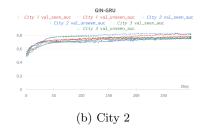
## Appendix A Convergence plots





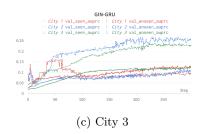
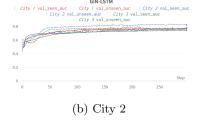


Fig. 1: GIN-GRU





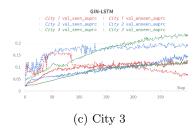
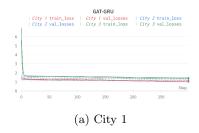


Fig. 2: GIN-LSTM





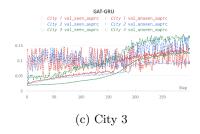
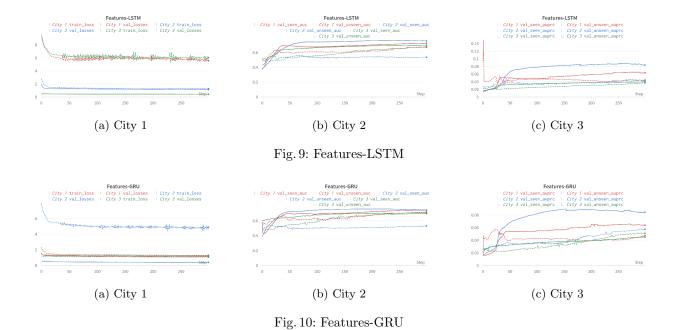


Fig. 3: GAT-GRU



Fig. 8: Features+PR-LSTM  $\,$ 



Appendix B Best model configurations

Table 1: Model configurations

Model	#layers GNN	l #emb.	GNN #hid.dim.	GNN SMOTE	${\bf rate\ \# layers}$	RNN #hid.dim.	RNN Loss
GIN-LSTM	2	512	512	0	1	256	Focal
GIN-GRU	2	512	512	0	1	512	Focal
GAT-LSTM	2	256	256	0	1	512	Focal
GAT-GRU	2	256	256	0	2	256	Focal
Features+PR-LSTM	[ /	/	/	0	1	512	BCE
Features+PR-GRU	/	/	/	0.5	1	512	Focal
Features-LSTM	/	/	/	0.5	1	256	Focal
Features-GRU	/	/	/	0	1	512	BCE
Static GAT	4	256	512	0.5	/	/	Focal
Static GIN	2	512	512	0.5	/	/	BCE