max recall accuracy, weight constraint35
pointer_network vs transformer

Legend:
- transformer_same, best val-max: 0.97
- transformer_lr1-e4_decay05, best val-max: 0.9
- transformer_lr1-e6, best val-max: 0.98
- transformer_lr1-e8, best val-max: 0.94
- pointer_network, best val-max: 0.99