Results on QM9

Adarsh Jamadandi

QM9 dataset and current SOTA

```
Dataset: QM9(130831):
=====================
Number of graphs: 130831
Number of features: 11
Number of classes: 19

Data(x=[15, 11], edge_index=[2, 28], edge_attr=[28, 4], y=[1, 19], pos=[15, 3], idx=[1], name='gdb_134', z=[15])
========================================================
Number of nodes: 15
Number of edges: 28
Average node degree: 1.87
Has isolated nodes: False
Has self-loops: False
Is undirected: True
mean 2.672952651977539
std 1.5034793615341187
```

Table 1: Comparison of MAEs of targets on QM9 for different models.

Target	SchNet	PhysNet	MEGNet-f	Cormorant	MGCN	DimeNet	MXMNet BS=32 d_g=5Å	MXMNet BS=128 d_g=5Å	MXMNet BS=128 d_g=10Å
μ (D)	**0.021**	0.0529	0.040	0.038	0.056	0.0286	0.0396	0.0382	0.0255
$\alpha(a_0^3)$	0.124	0.0615	0.083	0.085	**0.030**	0.0469	0.0447	0.0482	0.0465
ϵ_{HOMO} (meV)	47	32.9	38	34	42.1	27.8	24.7	23.0	**22.8**
ϵ_{LUMO} (meV)	39	24.7	31	38	57.4	19.7	19.7	19.5	**18.9**
$\Delta\epsilon$ (meV)	74	42.5	61	61	64.2	34.8	32.6	31.2	**30.6**
$\langle R^2 \rangle (a_0^2)$	0.158	0.765	0.265	0.961	0.11	0.331	0.512	0.506	**0.088**
ZPVE (meV)	1.616	1.39	1.40	2.027	**1.12**	1.29	1.15	1.16	1.19
U_0 (meV)	12	8.15	9	22	12.9	8.02	**5.90**	6.10	6.59
U (meV)	12	8.34	10	21	14.4	7.89	**5.94**	6.09	6.64
H (meV)	12	8.42	10	21	16.2	8.11	**6.09**	6.21	6.67
G (meV)	13	9.40	10	20	14.6	8.98	**7.17**	7.30	7.81
$c_v(\frac{\text{cal}}{\text{molK}})$	0.034	0.0280	0.030	0.026	0.038	0.0249	**0.0224**	0.0228	0.0233
std. MAE (%)	1.78	1.37	1.57	1.61	1.89	1.05	1.06	1.02	**0.93**

Vanilla GNN + QM9

- Regression task - Target - μ - Dipole moment.

- LR - 0.0001 and decay to 0.00001

- Epochs = 300

- Metric - MAE

- Model -

```
GCN(
    (conv1): GCNConv(11, 64)
    (conv2): GCNConv(64, 64)
    (conv3): GCNConv(64, 64)
    (lin): Linear(in_features=64, out_features=19, bias=True)
)
```

Results - Vanilla GNN + QM9

```
Epoch: 294, LR: 0.000010, MSELoss: 80.3882695, Test MAE: 12.0164788
number of distinguishable graphs:  394267  of  395789
Epoch: 295, LR: 0.000010, MSELoss: 80.3782898, Test MAE: 11.6712106
number of distinguishable graphs:  394139  of  395789
Epoch: 296, LR: 0.000010, MSELoss: 80.3760804, Test MAE: 11.8250868
number of distinguishable graphs:  394174  of  395789
Epoch: 297, LR: 0.000010, MSELoss: 80.3846005, Test MAE: 11.9154141
number of distinguishable graphs:  394177  of  395789
Epoch: 298, LR: 0.000010, MSELoss: 80.3873209, Test MAE: 11.8555786
number of distinguishable graphs:  394139  of  395789
Epoch: 299, LR: 0.000010, MSELoss: 80.3757762, Test MAE: 11.9482946
number of distinguishable graphs:  394090  of  395789
Epoch: 300, LR: 0.000010, MSELoss: 80.3805613, Test MAE: 11.8287799
```

1699 Non distinguishable graphs

Principal Neighbourhood Aggregation + QM9 Spectra

- Regression task - Target - spectrum

- Number of graphs - 131460

- LR - 0.001 and decay to 0.00001

- Epochs = 300

- Metric - MAE

- Model -

```
Net(
  (convs): ModuleList(
    (0): PNAConv(1, 1, towers=1, edge_dim=None)
    (1): PNAConv(1, 1, towers=1, edge_dim=None)
    (2): PNAConv(1, 1, towers=1, edge_dim=None)
    (3): PNAConv(1, 1, towers=1, edge_dim=None)
    (4): PNAConv(1, 1, towers=1, edge_dim=None)
  )
  (batch_norms): ModuleList(
    (0): BatchNorm(1)
    (1): BatchNorm(1)
    (2): BatchNorm(1)
    (3): BatchNorm(1)
    (4): BatchNorm(1)
  )
  (mlp): Sequential(
    (0): Linear(in_features=1, out_features=64, bias=True)
    (1): ReLU()
    (2): Linear(in_features=64, out_features=150, bias=True)
  )
)
```

Vanilla GNN+ QM9 Spectra

- Regression task - Target - spectrum

- LR - 0.001 and decay to 0.00001

- Number of graphs - 131460

- Epochs = 300

- Metric - MAE

- Model -

```
GCN(
  (conv1): GCNConv(1, 64)
  (conv2): GCNConv(64, 64)
  (conv3): GCNConv(64, 64)
  (conv4): GCNConv(64, 64)
  (conv5): GCNConv(64, 64)
  (mlp_1): Sequential(
    (0): Linear(in_features=64, out_features=64, bias=True)
    (1): ReLU()
    (2): Linear(in_features=64, out_features=128, bias=True)
    (3): ReLU()
    (4): Linear(in_features=128, out_features=150, bias=True)
  )
)
```
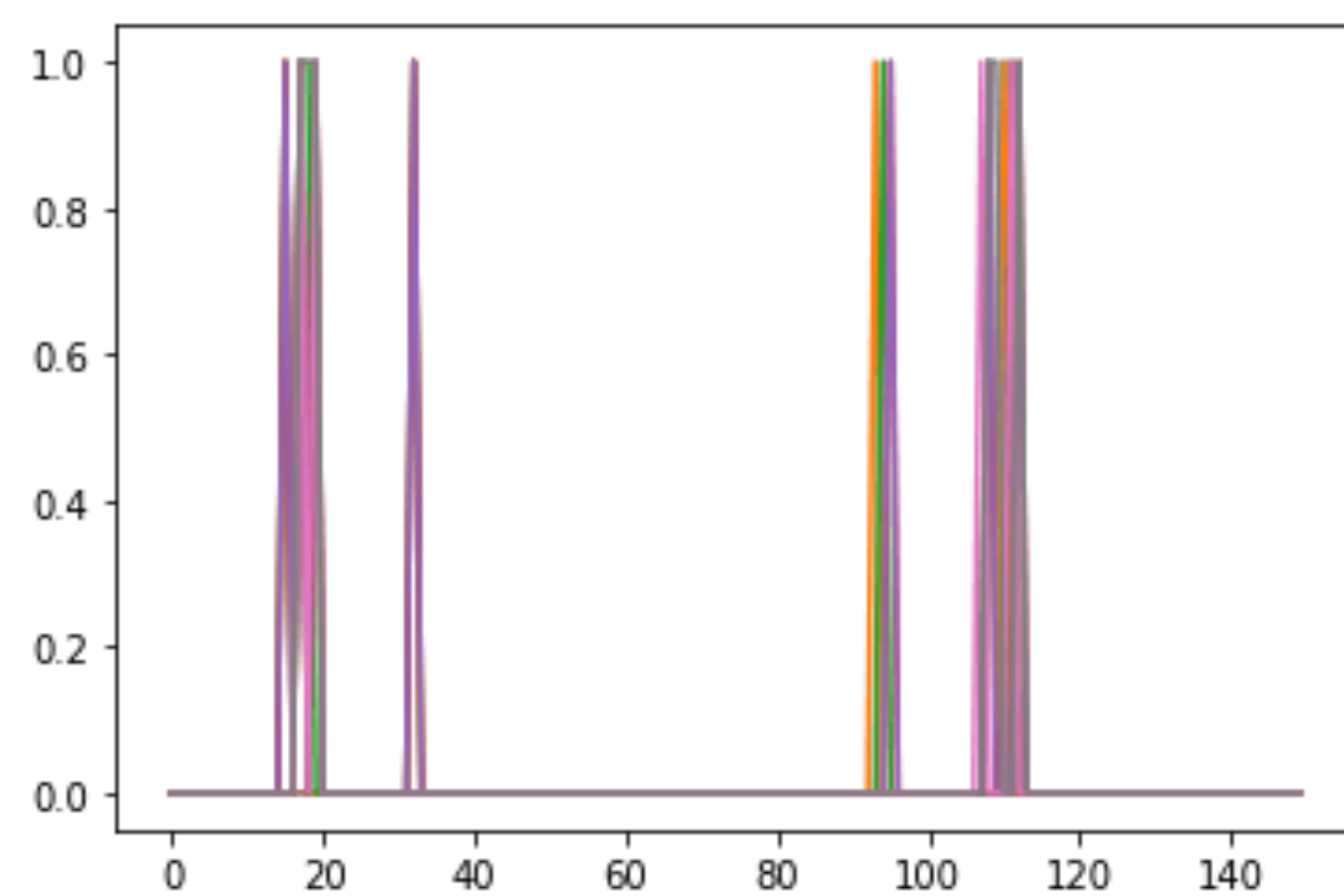
Results - QM9 spectra

Method	MSE Loss	Test MAE
Vanilla GNN	1.9323	1.9353
PNA	0.0345	1.1209

Groundtruth

PNA-QM9 Spectra Prediction

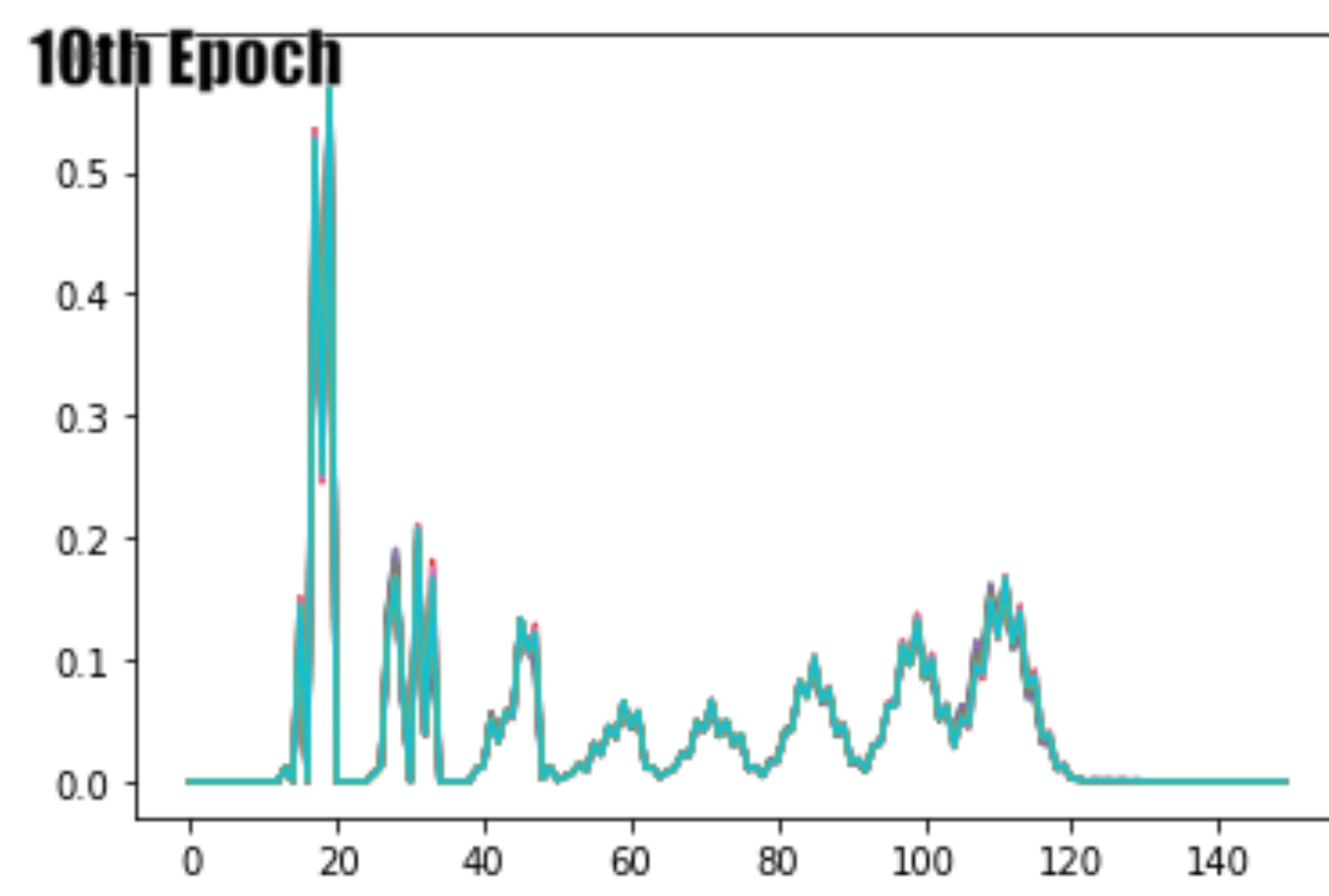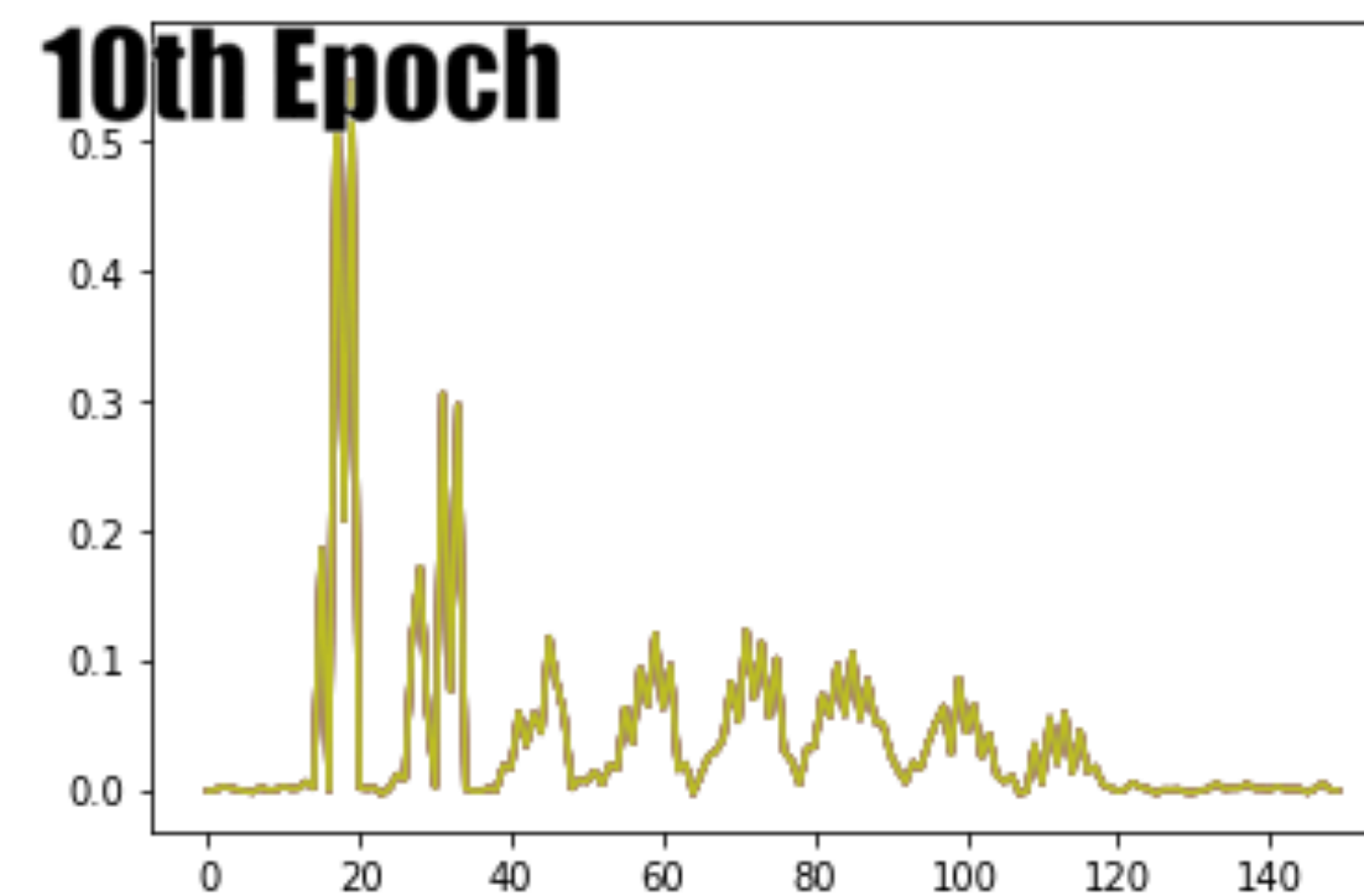Vanilla-GNN QM9 Spectra Prediction
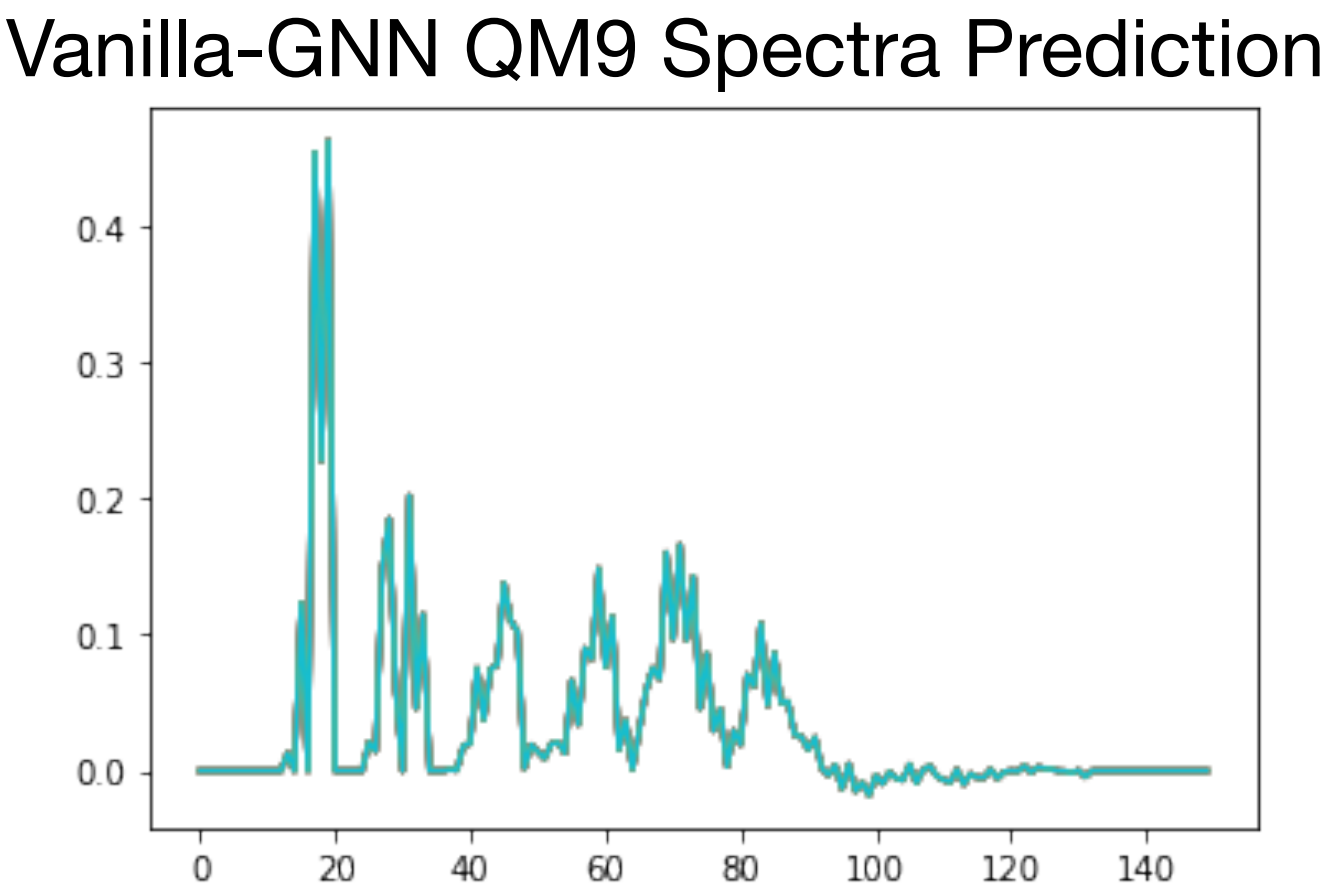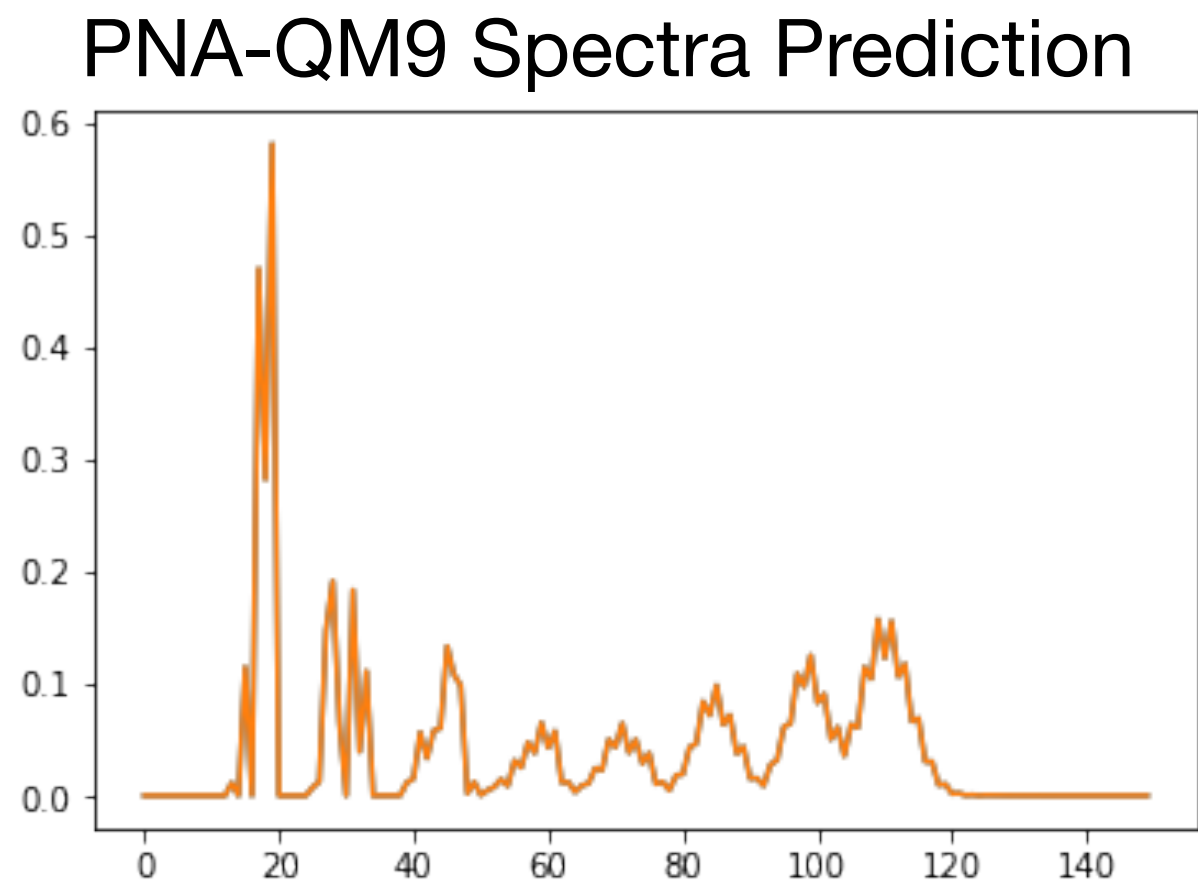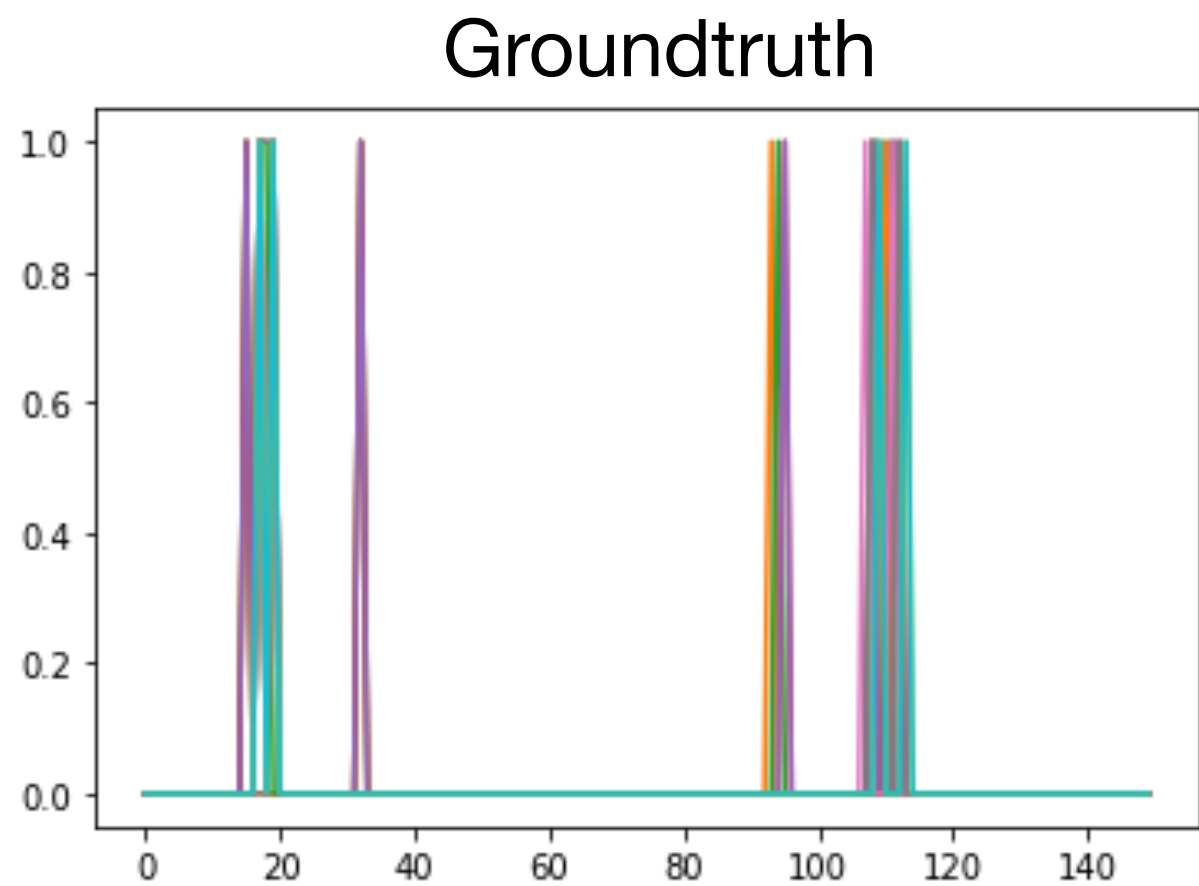
More results

Ground truth

PNA-QM9 Spectra Prediction

Vanilla-GNN QM9 Spectra Prediction

Results on Mol-HIV - Predicting the number of peaks

Adarsh Jamadandi

Method	MSE Loss	Train MAE	Test MAE	Validation MAE
PNA	0.5093	0.5455	1.5671	NA
PNA + Edgefeatures	2.1267	0.8208	1.7472	NA
ChemProp	3.7535	0.4324	2.002	NA
InfoGraph - Supervised Loss	2.045	NA	2.4834	2.2809
InfoGraph* - UnsupervisedLoss	1.920	NA	2.966	2.646

Results on Mol-HIV - Predicting spectra

PNA + Cosine similarity Loss + Not cropping the Y-vector
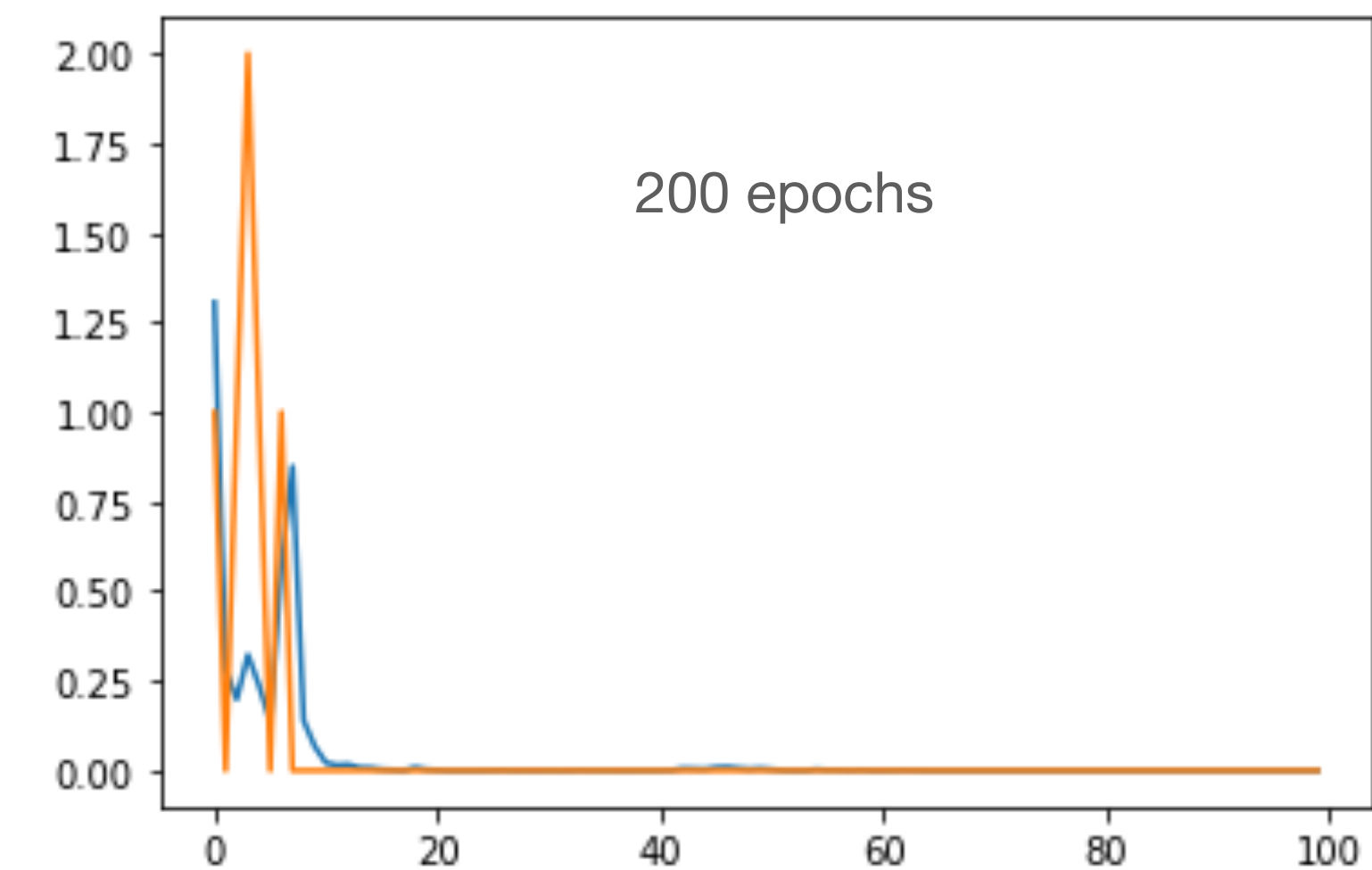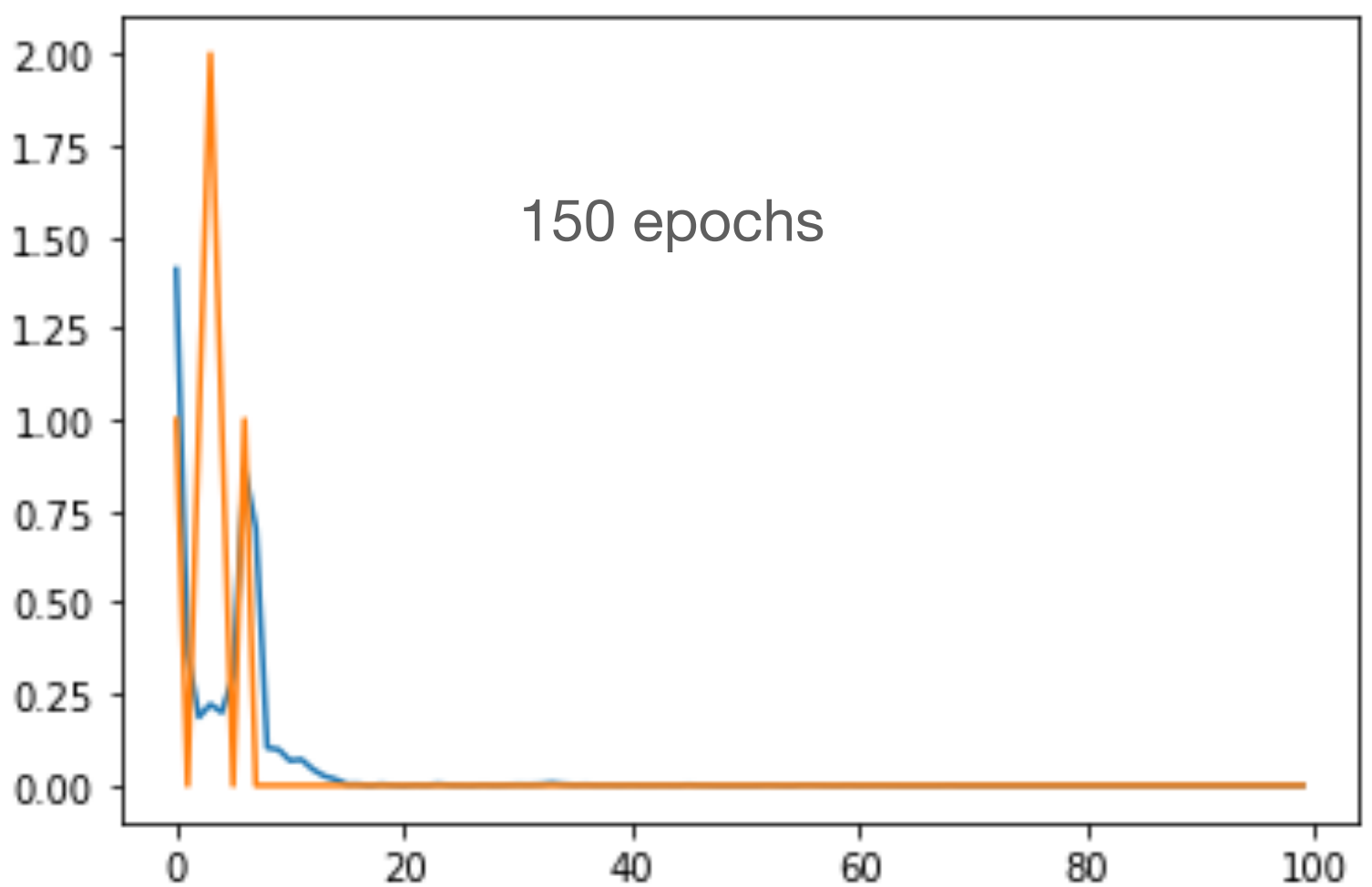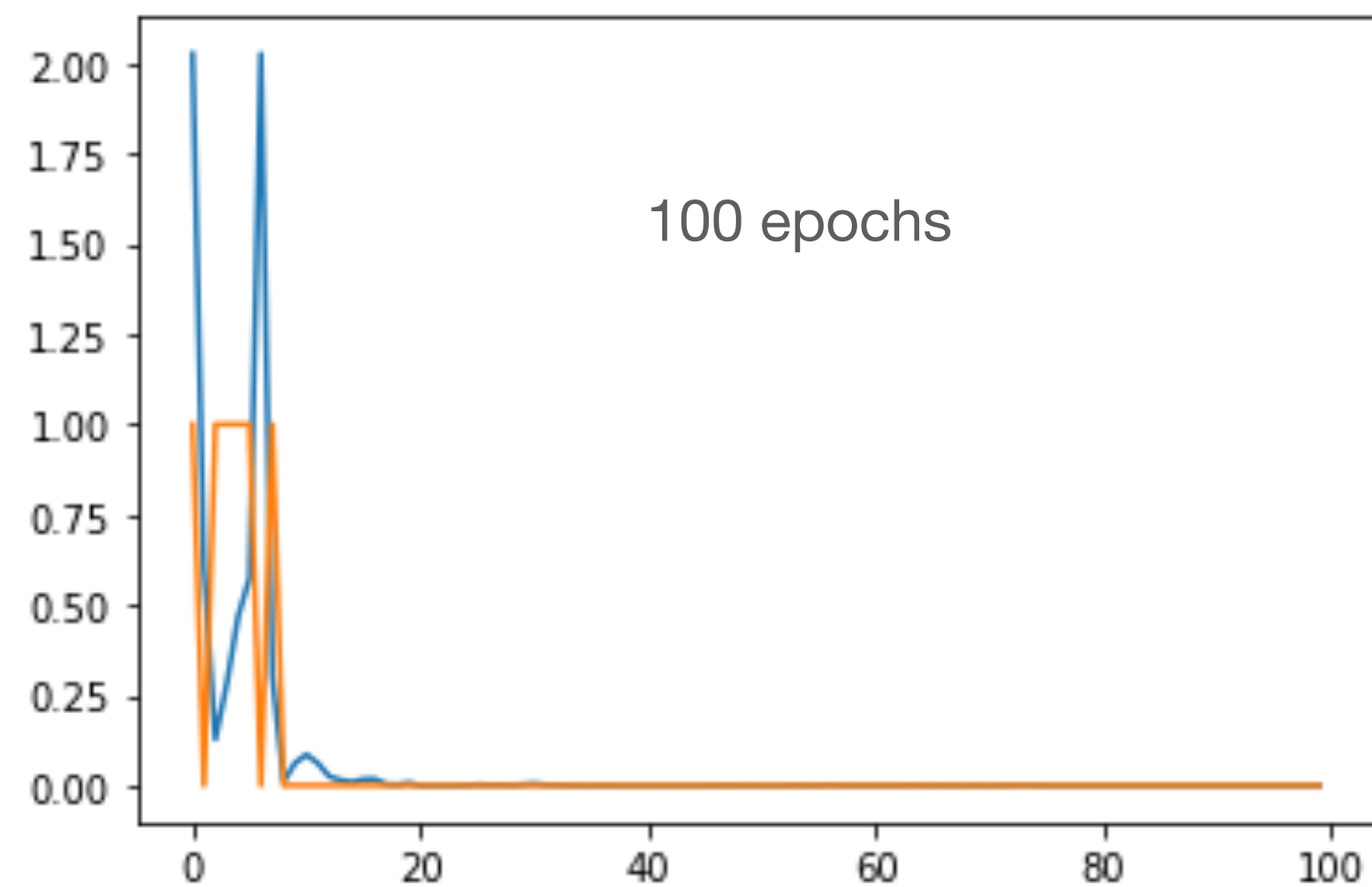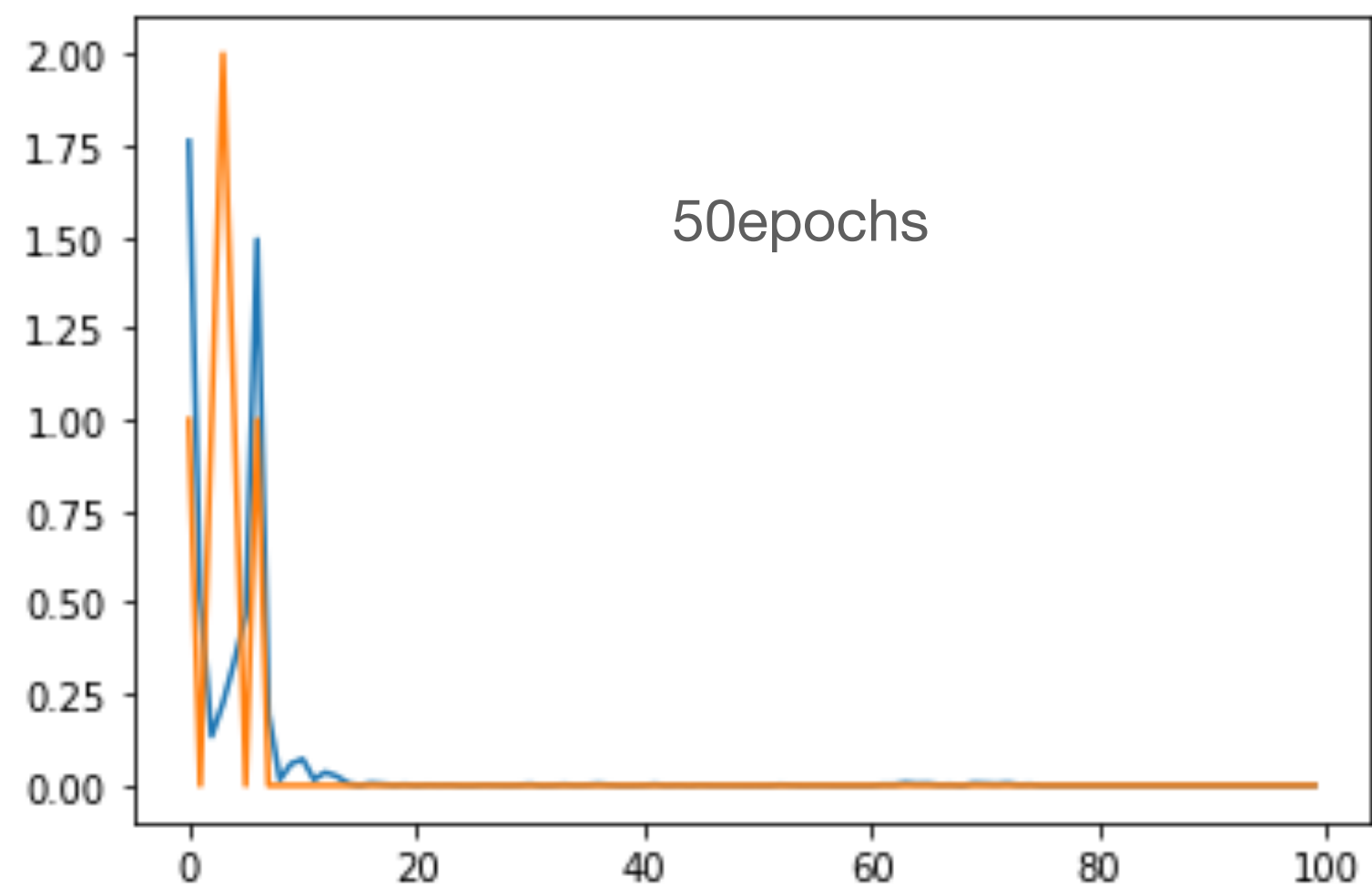
Predicted
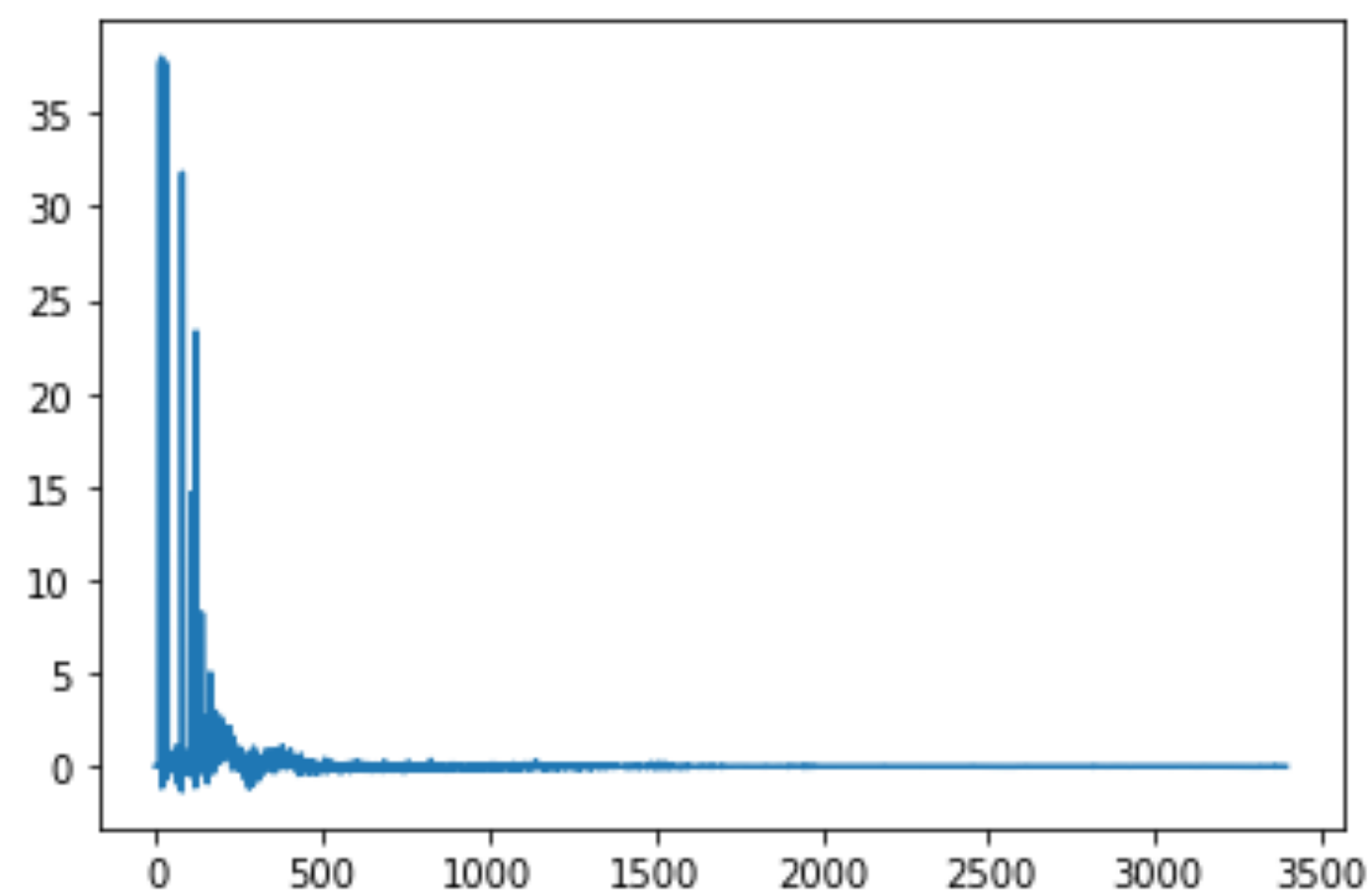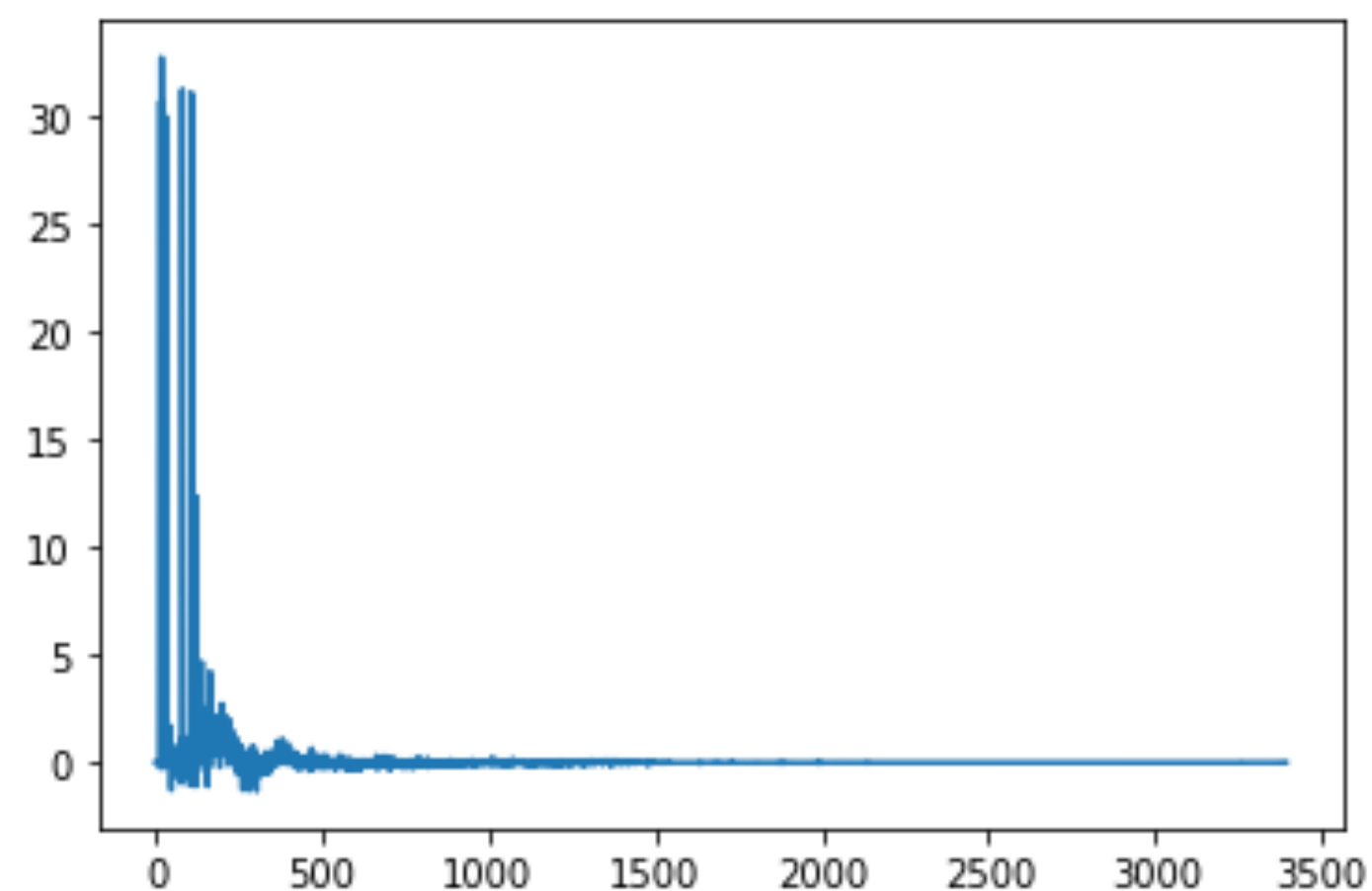
GT
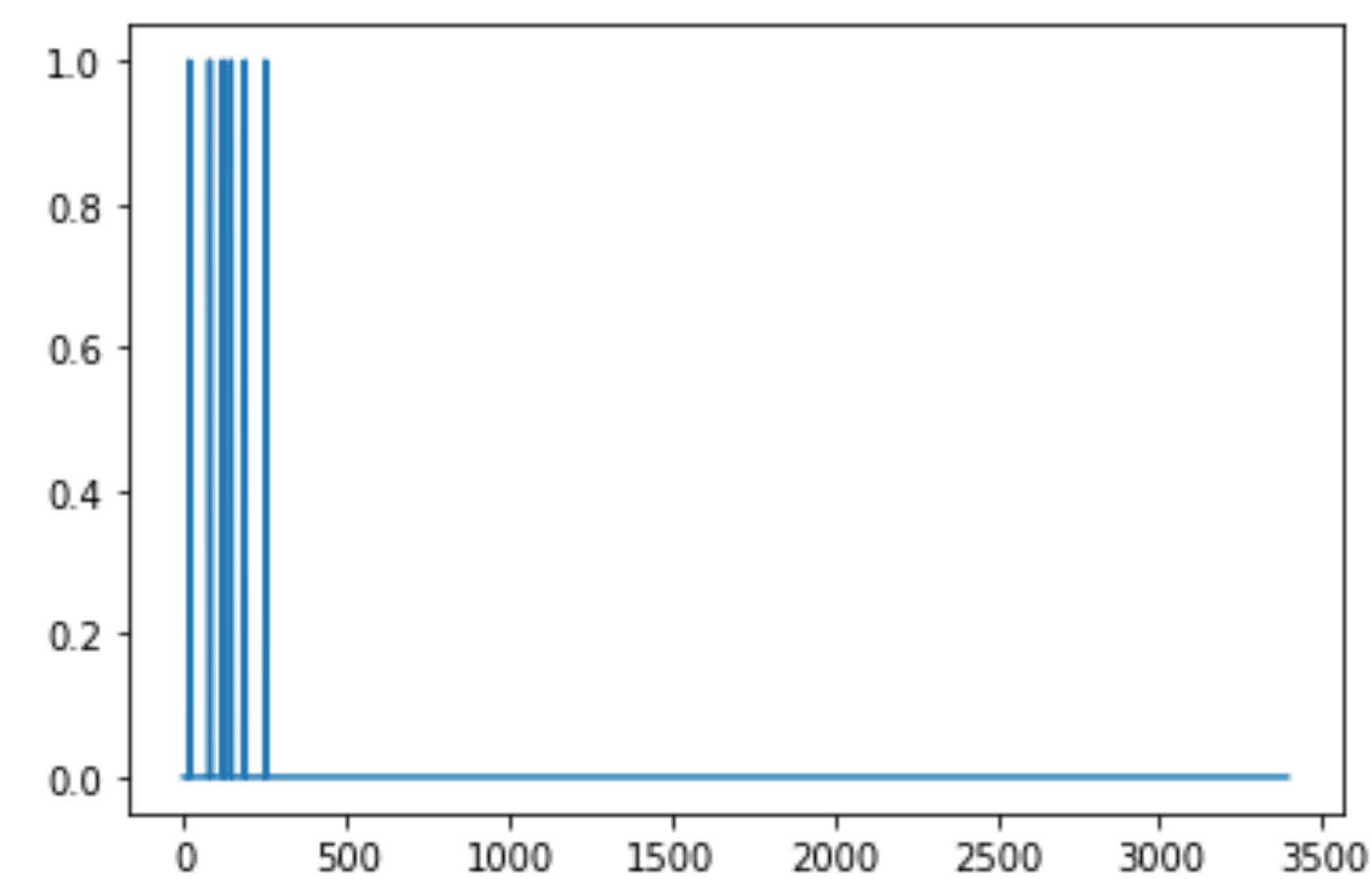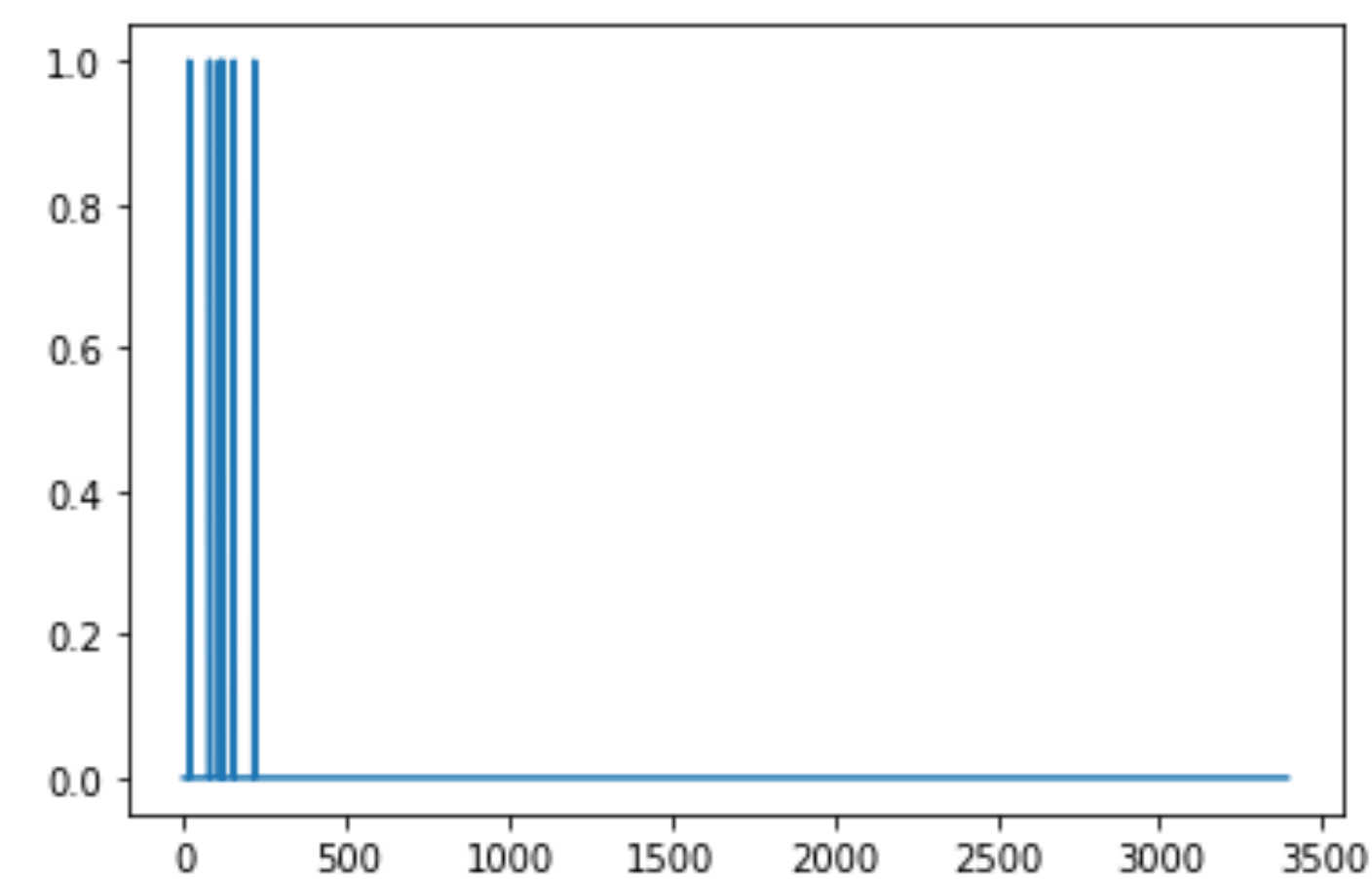
PNA + Cosine similarity Loss + Cropped
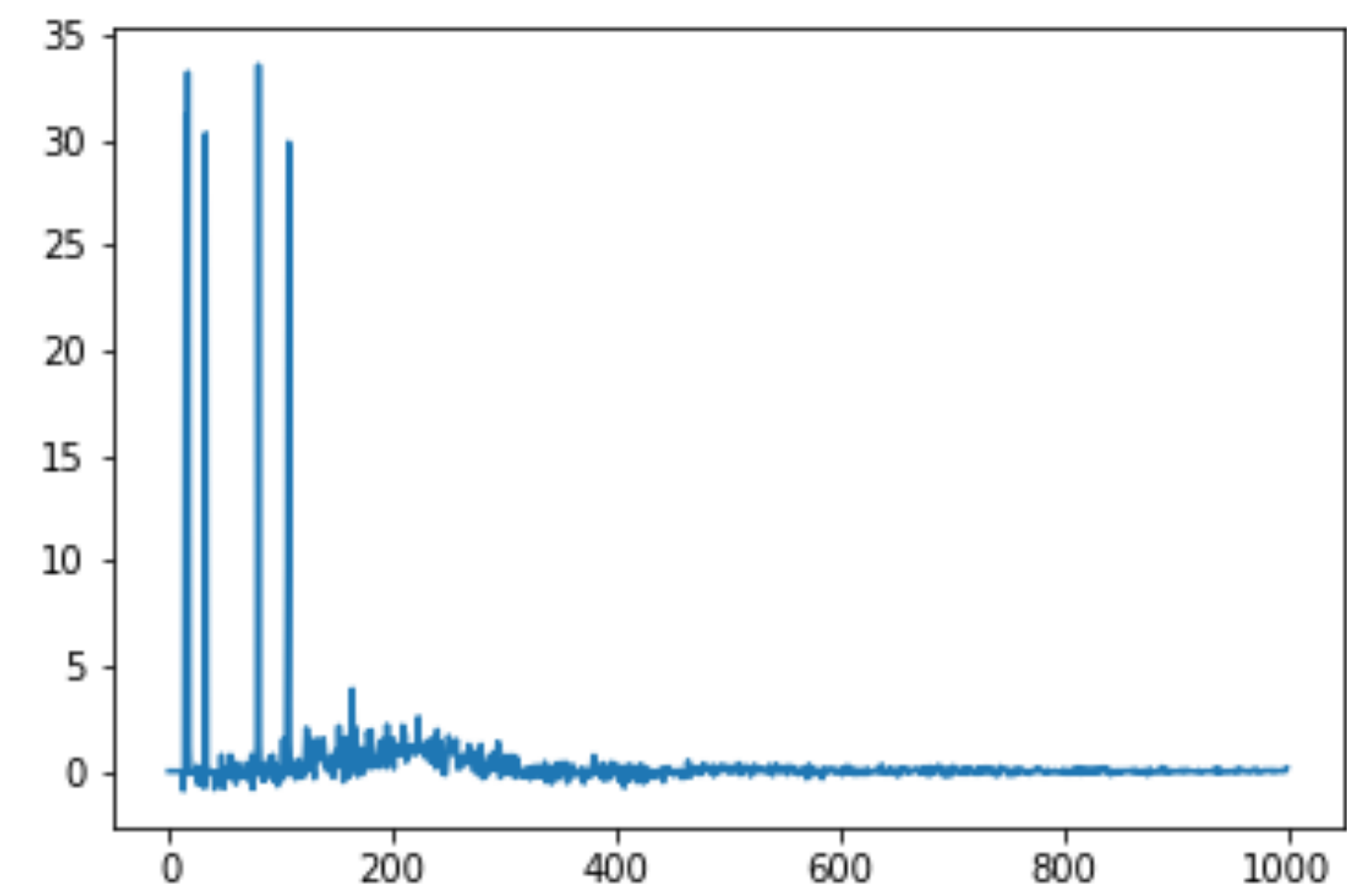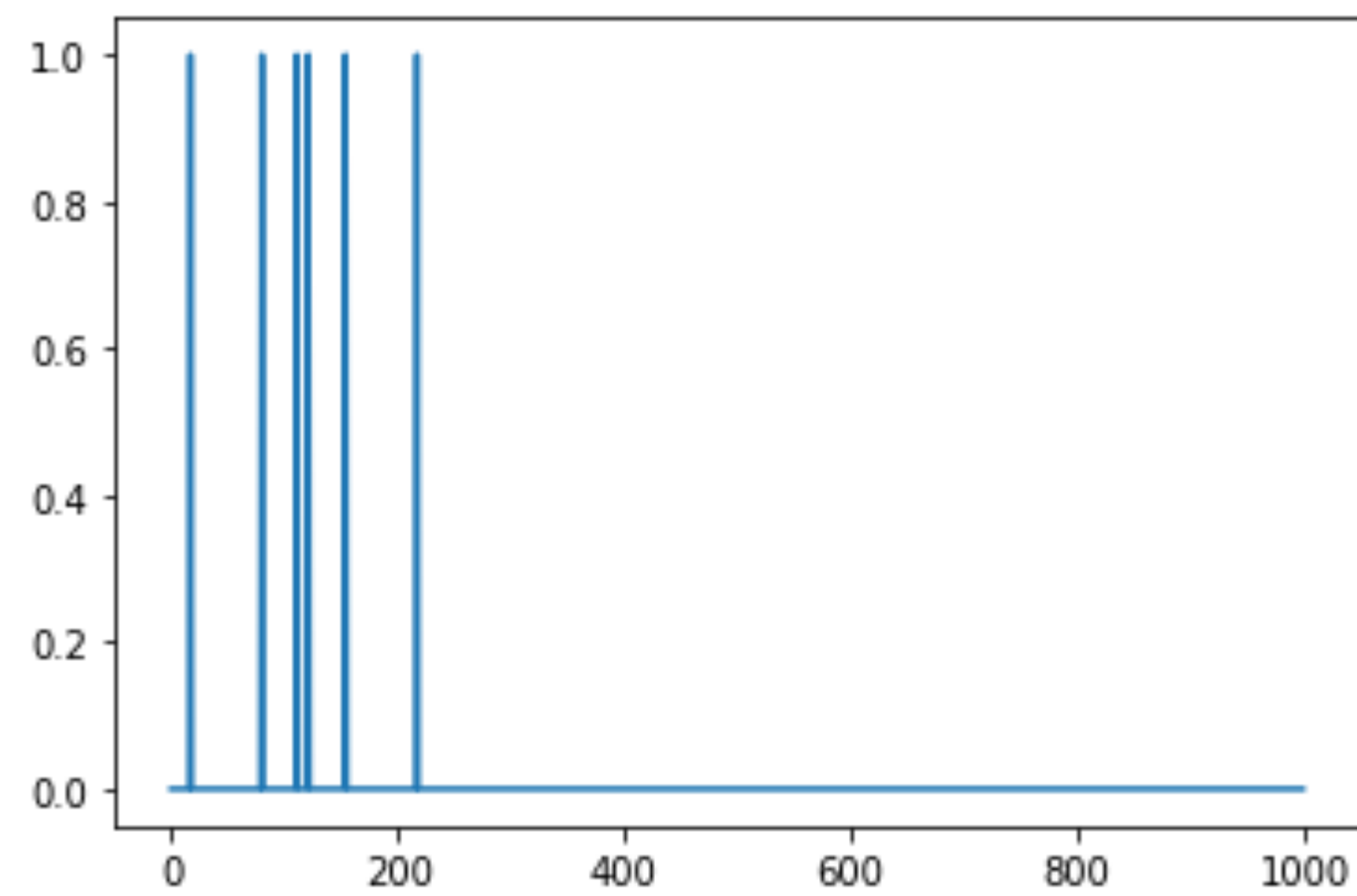
PDNConv + SampleLoss + Cropped
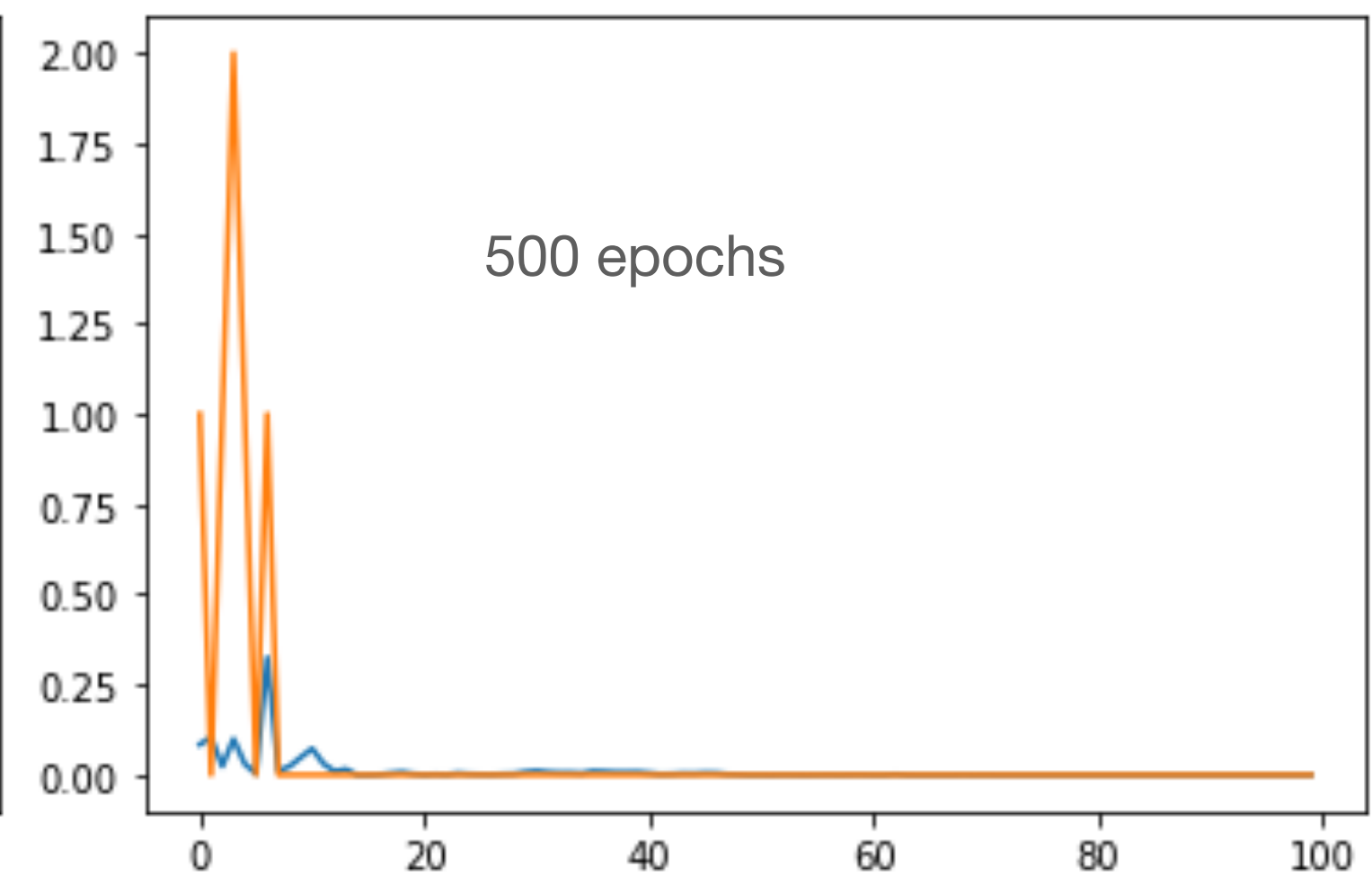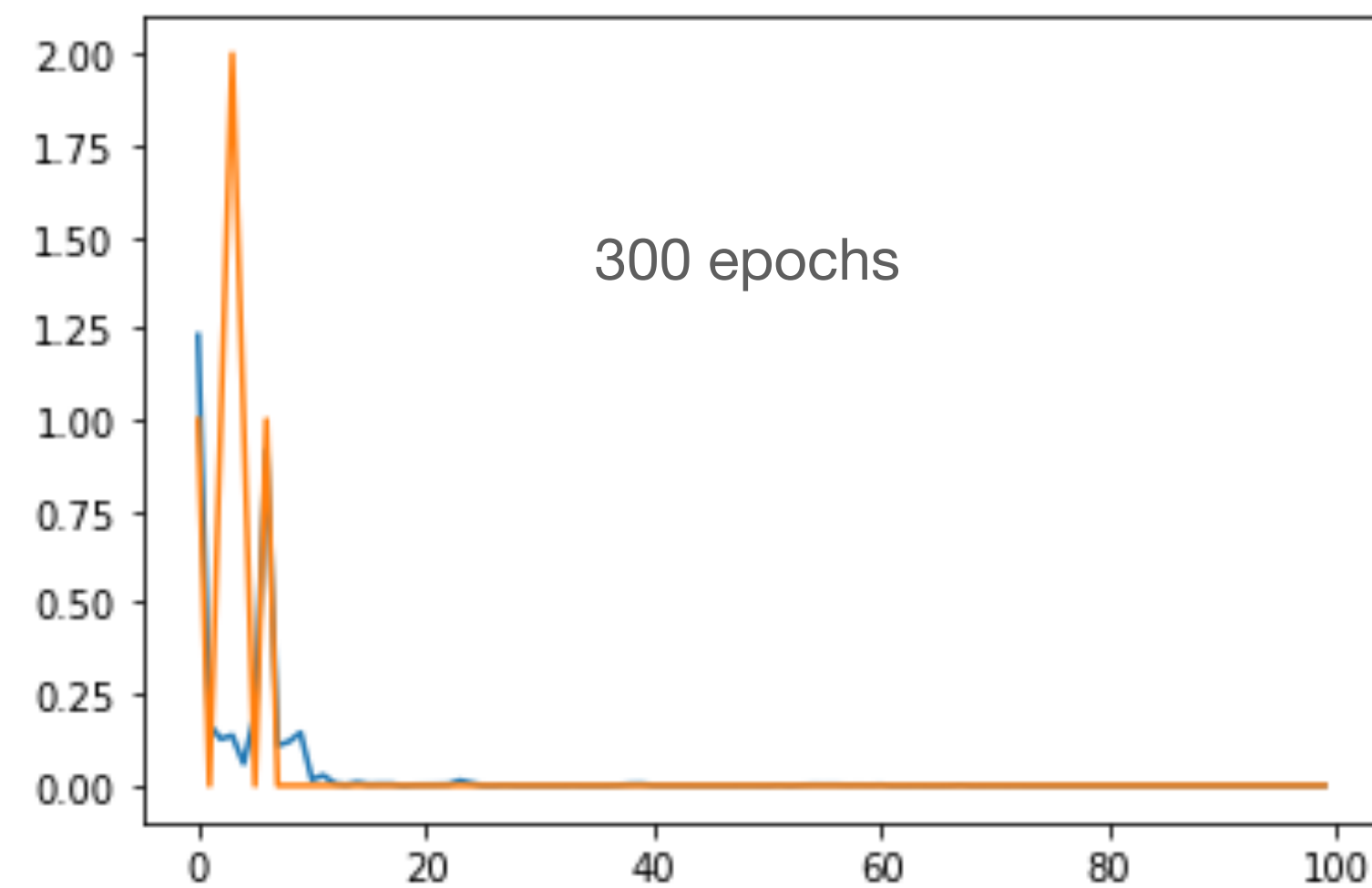
PDNConv + Energy minimisation + Cropped