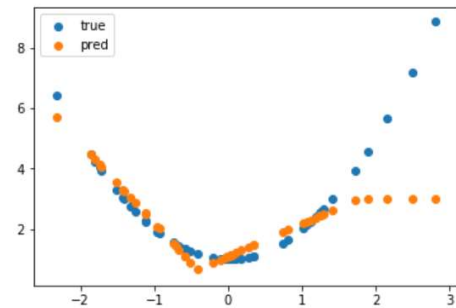


Problem: Finding a best fit line for  $f(x) = x^2 + 1$

Base line:

```
model = nn.Sequential(  
    nn.Linear(in_dim, 2),  
    nn.ReLU(),  
    nn.Linear(2, out_dim)  
)
```

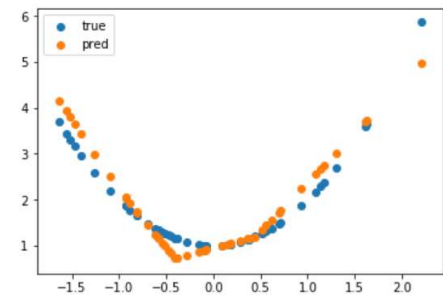
```
Epoch [500/10000], Loss: 1.6070  
Epoch [1000/10000], Loss: 1.3746  
Epoch [1500/10000], Loss: 1.2662  
Epoch [2000/10000], Loss: 1.1870  
Epoch [2500/10000], Loss: 1.1191  
Epoch [3000/10000], Loss: 1.0523  
Epoch [3500/10000], Loss: 0.9818  
Epoch [4000/10000], Loss: 0.9089  
Epoch [4500/10000], Loss: 0.8388  
Epoch [5000/10000], Loss: 0.7750  
Epoch [5500/10000], Loss: 0.7176  
Epoch [6000/10000], Loss: 0.6668  
Epoch [6500/10000], Loss: 0.6224  
Epoch [7000/10000], Loss: 0.5838  
Epoch [7500/10000], Loss: 0.5504  
Epoch [8000/10000], Loss: 0.5215  
Epoch [8500/10000], Loss: 0.4962  
Epoch [9000/10000], Loss: 0.4741  
Epoch [9500/10000], Loss: 0.4545  
Epoch [10000/10000], Loss: 0.4371
```



Attempt # 1: 3 neurons

```
model = nn.Sequential(  
    nn.Linear(in_dim, 3),  
    nn.ReLU(),  
    nn.Linear(3, out_dim)  
)
```

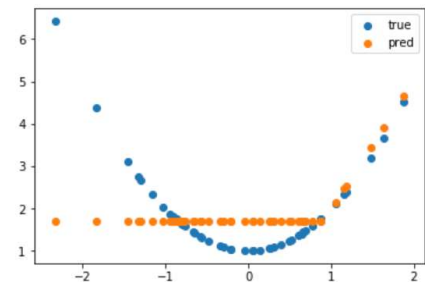
```
Epoch [500/10000], Loss: 2.0115  
Epoch [1000/10000], Loss: 1.3341  
Epoch [1500/10000], Loss: 0.8181  
Epoch [2000/10000], Loss: 0.5314  
Epoch [2500/10000], Loss: 0.3878  
Epoch [3000/10000], Loss: 0.3177  
Epoch [3500/10000], Loss: 0.2832  
Epoch [4000/10000], Loss: 0.2609  
Epoch [4500/10000], Loss: 0.2434  
Epoch [5000/10000], Loss: 0.2293  
Epoch [5500/10000], Loss: 0.2168  
Epoch [6000/10000], Loss: 0.2059  
Epoch [6500/10000], Loss: 0.1943  
Epoch [7000/10000], Loss: 0.1830  
Epoch [7500/10000], Loss: 0.1704  
Epoch [8000/10000], Loss: 0.1593  
Epoch [8500/10000], Loss: 0.1495  
Epoch [9000/10000], Loss: 0.1410  
Epoch [9500/10000], Loss: 0.1339  
Epoch [10000/10000], Loss: 0.1276
```



Attempt # 2: more layer but still 2 neurons. Not too good

```
model = nn.Sequential(  
    nn.Linear(in_dim, 2),  
    nn.ReLU(),  
    nn.Linear(2, 2),  
    nn.Linear(2, out_dim)  
)
```

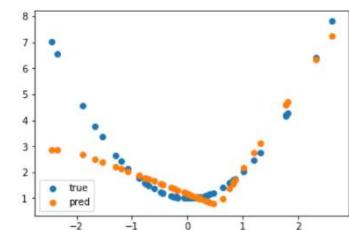
```
Epoch [500/10000], Loss: 1.8526  
Epoch [1000/10000], Loss: 1.7389  
Epoch [1500/10000], Loss: 1.6512  
Epoch [2000/10000], Loss: 1.5604  
Epoch [2500/10000], Loss: 1.4779  
Epoch [3000/10000], Loss: 1.4080  
Epoch [3500/10000], Loss: 1.3504  
Epoch [4000/10000], Loss: 1.3039  
Epoch [4500/10000], Loss: 1.2689  
Epoch [5000/10000], Loss: 1.2405  
Epoch [5500/10000], Loss: 1.2189  
Epoch [6000/10000], Loss: 1.2019  
Epoch [6500/10000], Loss: 1.1892  
Epoch [7000/10000], Loss: 1.1792  
Epoch [7500/10000], Loss: 1.1716  
Epoch [8000/10000], Loss: 1.1660  
Epoch [8500/10000], Loss: 1.1617  
Epoch [9000/10000], Loss: 1.1585  
Epoch [9500/10000], Loss: 1.1559  
Epoch [10000/10000], Loss: 1.1539
```



Attempt # 3: A little better with 3

```
model = nn.Sequential(  
    nn.Linear(in_dim, 3),  
    nn.ReLU(),  
    nn.Linear(3, 3),  
    nn.Linear(3, out_dim)  
)
```

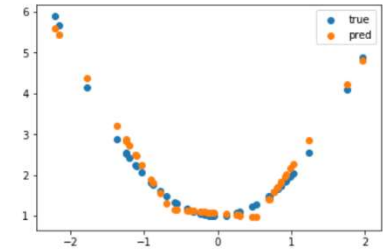
```
Epoch [500/10000], Loss: 1.3753  
Epoch [1000/10000], Loss: 0.8219  
Epoch [1500/10000], Loss: 0.6392  
Epoch [2000/10000], Loss: 0.5378  
Epoch [2500/10000], Loss: 0.4699  
Epoch [3000/10000], Loss: 0.4217  
Epoch [3500/10000], Loss: 0.3839  
Epoch [4000/10000], Loss: 0.3488  
Epoch [4500/10000], Loss: 0.3151  
Epoch [5000/10000], Loss: 0.2826  
Epoch [5500/10000], Loss: 0.2576  
Epoch [6000/10000], Loss: 0.2414  
Epoch [6500/10000], Loss: 0.2296  
Epoch [7000/10000], Loss: 0.2216  
Epoch [7500/10000], Loss: 0.2158  
Epoch [8000/10000], Loss: 0.2108  
Epoch [8500/10000], Loss: 0.2062  
Epoch [9000/10000], Loss: 0.2017  
Epoch [9500/10000], Loss: 0.1974  
Epoch [10000/10000], Loss: 0.1948
```



#### Attempt # 4: More layers help

```
model = nn.Sequential(
    nn.Linear(in_dim, 2),
    nn.ReLU(),
    nn.Linear(2, 2),
    nn.Linear(2, 2),
    nn.Linear(2, out_dim)
)
```

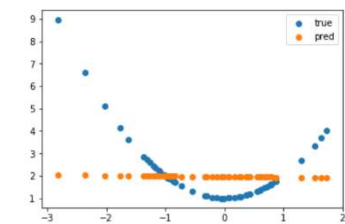
```
Epoch [500/10000], Loss: 1.2079
Epoch [1000/10000], Loss: 0.5616
Epoch [1500/10000], Loss: 0.2736
Epoch [2000/10000], Loss: 0.2348
Epoch [2500/10000], Loss: 0.2256
Epoch [3000/10000], Loss: 0.2165
Epoch [3500/10000], Loss: 0.2061
Epoch [4000/10000], Loss: 0.1947
Epoch [4500/10000], Loss: 0.1827
Epoch [5000/10000], Loss: 0.1709
Epoch [5500/10000], Loss: 0.1592
Epoch [6000/10000], Loss: 0.1481
Epoch [6500/10000], Loss: 0.1377
Epoch [7000/10000], Loss: 0.1285
Epoch [7500/10000], Loss: 0.1204
Epoch [8000/10000], Loss: 0.1131
Epoch [8500/10000], Loss: 0.1067
Epoch [9000/10000], Loss: 0.1015
Epoch [9500/10000], Loss: 0.0971
Epoch [10000/10000], Loss: 0.0932
```



#### Attempt # 5: Don't know what the heck is going on with this one

```
model = nn.Sequential(
    nn.Linear(in_dim, 3),
    nn.ReLU(),
    nn.Linear(3, 3),
    nn.Linear(3, 3),
    nn.Linear(3, out_dim)
)
```

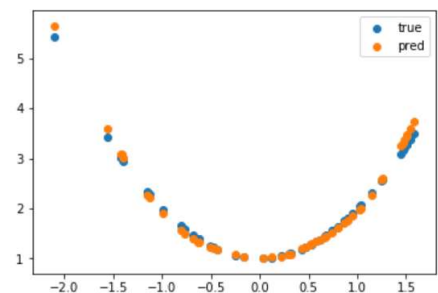
```
Epoch [500/10000], Loss: 1.9161
Epoch [1000/10000], Loss: 1.8940
Epoch [1500/10000], Loss: 1.8874
Epoch [2000/10000], Loss: 1.8850
Epoch [2500/10000], Loss: 1.8839
Epoch [3000/10000], Loss: 1.8832
Epoch [3500/10000], Loss: 1.8829
Epoch [4000/10000], Loss: 1.8827
Epoch [4500/10000], Loss: 1.8826
Epoch [5000/10000], Loss: 1.8826
Epoch [5500/10000], Loss: 1.8826
Epoch [6000/10000], Loss: 1.8826
Epoch [6500/10000], Loss: 1.8826
Epoch [7000/10000], Loss: 1.8826
Epoch [7500/10000], Loss: 1.8826
Epoch [8000/10000], Loss: 1.8826
Epoch [8500/10000], Loss: 1.8826
Epoch [9000/10000], Loss: 1.8826
Epoch [9500/10000], Loss: 1.8826
Epoch [10000/10000], Loss: 1.8826
```



#### Attempt # 6:

```
model = nn.Sequential(
    nn.Linear(in_dim, 100),
    nn.ReLU(),
    nn.Linear(100, 100),
    nn.Linear(100, out_dim)
)
```

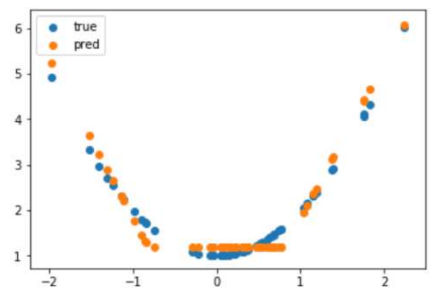
```
Epoch [500/10000], Loss: 0.1840
Epoch [1000/10000], Loss: 0.0710
Epoch [1500/10000], Loss: 0.0606
Epoch [2000/10000], Loss: 0.0534
Epoch [2500/10000], Loss: 0.0475
Epoch [3000/10000], Loss: 0.0426
Epoch [3500/10000], Loss: 0.0385
Epoch [4000/10000], Loss: 0.0350
Epoch [4500/10000], Loss: 0.0320
Epoch [5000/10000], Loss: 0.0295
Epoch [5500/10000], Loss: 0.0274
Epoch [6000/10000], Loss: 0.0255
Epoch [6500/10000], Loss: 0.0238
Epoch [7000/10000], Loss: 0.0224
Epoch [7500/10000], Loss: 0.0211
Epoch [8000/10000], Loss: 0.0199
Epoch [8500/10000], Loss: 0.0189
Epoch [9000/10000], Loss: 0.0179
Epoch [9500/10000], Loss: 0.0171
Epoch [10000/10000], Loss: 0.0163
```



#### Attempt # 7: more layers doesn't necessarily mean it's better if there isn't sufficient neurons

```
model = nn.Sequential(
    nn.Linear(in_dim, 2),
    nn.ReLU(),
    nn.Linear(2, 2),
    nn.Linear(2, 2),
    nn.Linear(2, 2),
    nn.Linear(2, out_dim)
)
```

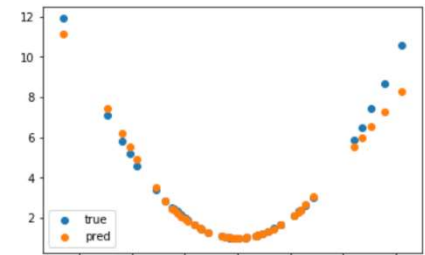
```
Epoch [500/10000], Loss: 1.7683
Epoch [1000/10000], Loss: 1.2087
Epoch [1500/10000], Loss: 0.1498
Epoch [2000/10000], Loss: 0.1184
Epoch [2500/10000], Loss: 0.1015
Epoch [3000/10000], Loss: 0.0904
Epoch [3500/10000], Loss: 0.0848
Epoch [4000/10000], Loss: 0.0818
Epoch [4500/10000], Loss: 0.0806
Epoch [5000/10000], Loss: 0.0799
Epoch [5500/10000], Loss: 0.0796
Epoch [6000/10000], Loss: 0.0796
Epoch [6500/10000], Loss: 0.0796
Epoch [7000/10000], Loss: 0.0795
Epoch [7500/10000], Loss: 0.0795
Epoch [8000/10000], Loss: 0.0795
Epoch [8500/10000], Loss: 0.0795
Epoch [9000/10000], Loss: 0.0795
Epoch [9500/10000], Loss: 0.0795
Epoch [10000/10000], Loss: 0.0795
```



## Attempt # 8: start testing with 1000+ neurons

```
model = nn.Sequential(  
    nn.Linear(in_dim, 1000),  
    nn.ReLU(),  
    nn.Linear(1000, out_dim)  
)
```

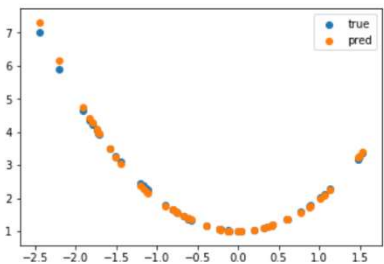
```
Epoch [500/10000], Loss: 0.0337  
Epoch [1000/10000], Loss: 0.0217  
Epoch [1500/10000], Loss: 0.0166  
Epoch [2000/10000], Loss: 0.0139  
Epoch [2500/10000], Loss: 0.0123  
Epoch [3000/10000], Loss: 0.0112  
Epoch [3500/10000], Loss: 0.0102  
Epoch [4000/10000], Loss: 0.0095  
Epoch [4500/10000], Loss: 0.0089  
Epoch [5000/10000], Loss: 0.0083  
Epoch [5500/10000], Loss: 0.0078  
Epoch [6000/10000], Loss: 0.0074  
Epoch [6500/10000], Loss: 0.0070  
Epoch [7000/10000], Loss: 0.0066  
Epoch [7500/10000], Loss: 0.0063  
Epoch [8000/10000], Loss: 0.0060  
Epoch [8500/10000], Loss: 0.0058  
Epoch [9000/10000], Loss: 0.0056  
Epoch [9500/10000], Loss: 0.0053  
Epoch [10000/10000], Loss: 0.0052
```



## Attempt # 9:

```
model = nn.Sequential(  
    nn.Linear(in_dim, 1000),  
    nn.ReLU(),  
    nn.Linear(1000, 1000),  
    nn.Linear(1000, out_dim)  
)
```

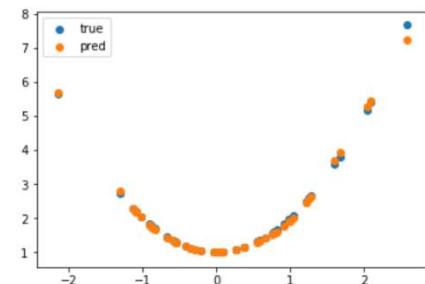
```
Epoch [500/10000], Loss: 0.0671  
Epoch [1000/10000], Loss: 0.0443  
Epoch [1500/10000], Loss: 0.0325  
Epoch [2000/10000], Loss: 0.0258  
Epoch [2500/10000], Loss: 0.0216  
Epoch [3000/10000], Loss: 0.0187  
Epoch [3500/10000], Loss: 0.0164  
Epoch [4000/10000], Loss: 0.0145  
Epoch [4500/10000], Loss: 0.0129  
Epoch [5000/10000], Loss: 0.0116  
Epoch [5500/10000], Loss: 0.0105  
Epoch [6000/10000], Loss: 0.0095  
Epoch [6500/10000], Loss: 0.0086  
Epoch [7000/10000], Loss: 0.0079  
Epoch [7500/10000], Loss: 0.0072  
Epoch [8000/10000], Loss: 0.0067  
Epoch [8500/10000], Loss: 0.0062  
Epoch [9000/10000], Loss: 0.0058  
Epoch [9500/10000], Loss: 0.0054  
Epoch [10000/10000], Loss: 0.0050
```



## Attempt # 10:

```
model = nn.Sequential(  
    nn.Linear(in_dim, 1000),  
    nn.ReLU(),  
    nn.Linear(1000, 1000),  
    nn.Linear(1000, 1000),  
    nn.Linear(1000, out_dim)  
)
```

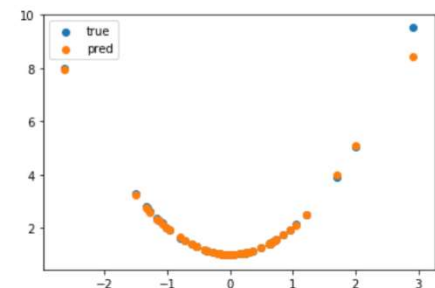
```
Epoch [500/10000], Loss: 0.0352  
Epoch [1000/10000], Loss: 0.0230  
Epoch [1500/10000], Loss: 0.0158  
Epoch [2000/10000], Loss: 0.0117  
Epoch [2500/10000], Loss: 0.0093  
Epoch [3000/10000], Loss: 0.0080  
Epoch [3500/10000], Loss: 0.0071  
Epoch [4000/10000], Loss: 0.0065  
Epoch [4500/10000], Loss: 0.0060  
Epoch [5000/10000], Loss: 0.0056  
Epoch [5500/10000], Loss: 0.0053  
Epoch [6000/10000], Loss: 0.0050  
Epoch [6500/10000], Loss: 0.0047  
Epoch [7000/10000], Loss: 0.0044  
Epoch [7500/10000], Loss: 0.0042  
Epoch [8000/10000], Loss: 0.0040  
Epoch [8500/10000], Loss: 0.0038  
Epoch [9000/10000], Loss: 0.0036  
Epoch [9500/10000], Loss: 0.0035  
Epoch [10000/10000], Loss: 0.0033
```



## Attempt # 11: Took couple hrs to run

```
model = nn.Sequential(  
    nn.Linear(in_dim, 10000),  
    nn.ReLU(),  
    nn.Linear(10000, 10000),  
    nn.Linear(10000, out_dim)  
)
```

```
Epoch [500/10000], Loss: 0.0116  
Epoch [1000/10000], Loss: 0.0073  
Epoch [1500/10000], Loss: 0.0053  
Epoch [2000/10000], Loss: 0.0040  
Epoch [2500/10000], Loss: 0.0032  
Epoch [3000/10000], Loss: 0.0026  
Epoch [3500/10000], Loss: 0.0022  
Epoch [4000/10000], Loss: 0.0019  
Epoch [4500/10000], Loss: 0.0016  
Epoch [5000/10000], Loss: 0.0014  
Epoch [5500/10000], Loss: 0.0013  
Epoch [6000/10000], Loss: 0.0011  
Epoch [6500/10000], Loss: 0.0010  
Epoch [7000/10000], Loss: 0.0009  
Epoch [7500/10000], Loss: 0.0008  
Epoch [8000/10000], Loss: 0.0008  
Epoch [8500/10000], Loss: 0.0007  
Epoch [9000/10000], Loss: 0.0007  
Epoch [9500/10000], Loss: 0.0006  
Epoch [10000/10000], Loss: 0.0006
```



## Attempt #12:

```
a = 20

model = nn.Sequential(
    nn.Linear(in_dim, a),
    nn.ReLU(),
    nn.Linear(a, a),
    nn.Linear(a, a),
    nn.Linear(a, a),
    nn.Linear(a, a),
    nn.Linear(a, out_dim)
)
```

```
Epoch [500/10000], Loss: 1.0819
Epoch [1000/10000], Loss: 0.0608
Epoch [1500/10000], Loss: 0.0685
Epoch [2000/10000], Loss: 0.0590
Epoch [2500/10000], Loss: 0.0509
Epoch [3000/10000], Loss: 0.0440
Epoch [3500/10000], Loss: 0.0382
Epoch [4000/10000], Loss: 0.0332
Epoch [4500/10000], Loss: 0.0290
Epoch [5000/10000], Loss: 0.0253
Epoch [5500/10000], Loss: 0.0221
Epoch [6000/10000], Loss: 0.0193
Epoch [6500/10000], Loss: 0.0170
Epoch [7000/10000], Loss: 0.0151
Epoch [7500/10000], Loss: 0.0136
Epoch [8000/10000], Loss: 0.0122
Epoch [8500/10000], Loss: 0.0110
Epoch [9000/10000], Loss: 0.0100
Epoch [9500/10000], Loss: 0.0090
Epoch [10000/10000], Loss: 0.0083
```

