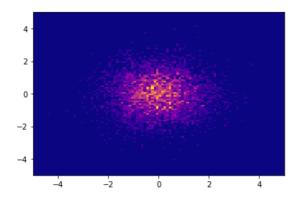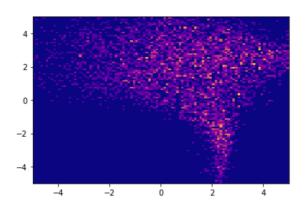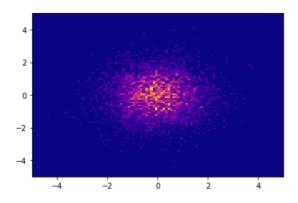# Task 1

For the random samples, given the distribution:

Generated the following distribution on Forward flow:
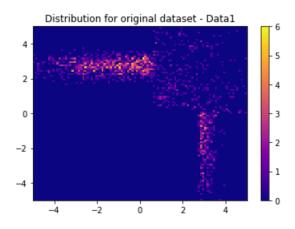

Inverse Flow gave back the original distribution:
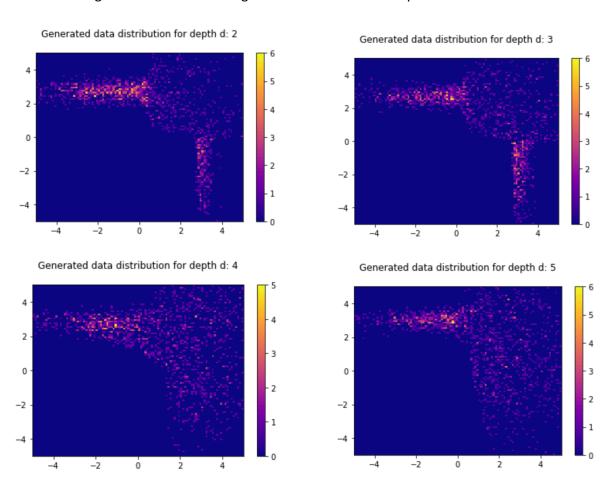
Task 2

Results on Data1 dataset:

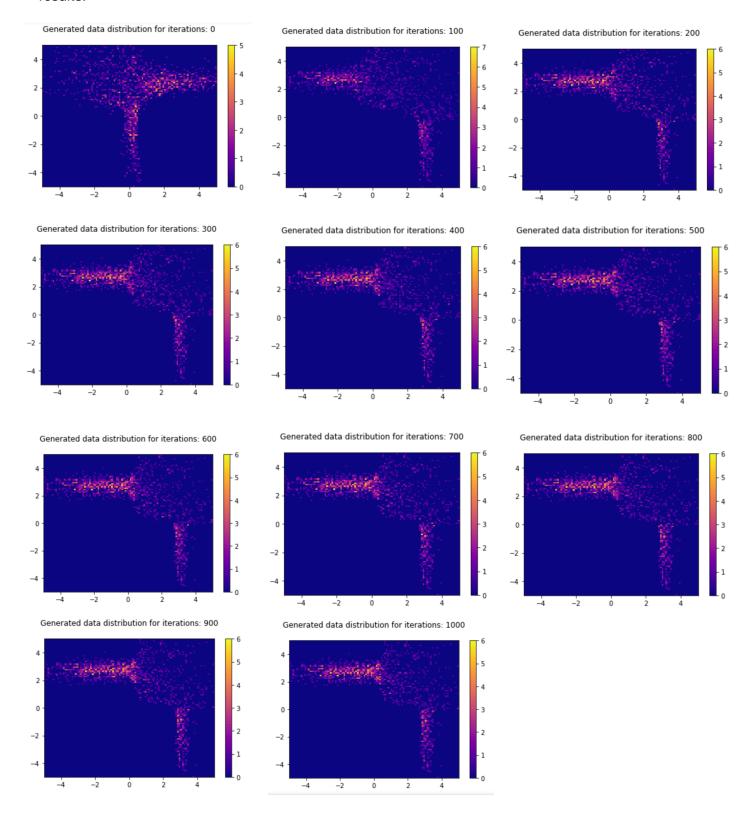Distribution for original dataset:

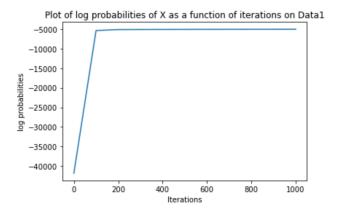Following are the distributions generated for different depth d:

From the above results we can see that, for depth d = 2, 3, the distributions are more similar to the original distribution. And further increase in depth, leads to overfitting the dataset. And also from the experiments with dataset data2 shown in the report, we can see that similarity of distribution obtained depends on random seed used for initialization.

By picking depth d = 2, we perform the training on different iterations and obtain the following results:

Generated data distribution for iterations: 0

Generated data distribution for iterations: 100

Generated data distribution for iterations: 200

Generated data distribution for iterations: 300

Generated data distribution for iterations: 400

Generated data distribution for iterations: 500

Generated data distribution for iterations: 600

Generated data distribution for iterations: 700

Generated data distribution for iterations: 800

Generated data distribution for iterations: 900

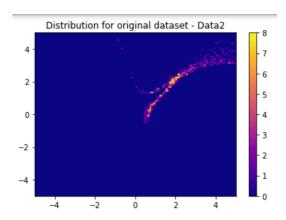Generated data distribution for iterations: 1000

From the above figures we can see that the approximate distribution improves with increase in number of iterations. Below is the plot of log probability of data X as a function of iterations. We can see that the log probability of data increases with the training.

```
log probabilities on Data1:
[tensor(-41779.7227), tensor(-5407.3677), tensor(-5153.8540), tensor(-5136.9263), tensor(-5123.6685), tensor(-5113.
0664), tensor(-5104.3857), tensor(-5097.1304), tensor(-5090.9668), tensor(-5085.6641), tensor(-5081.0547)]
```
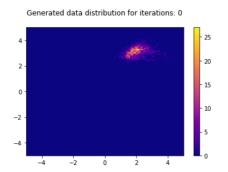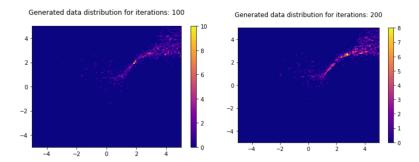

Plot of log probabilities of X as a function of iterations on Data1

Results on Data2 dataset:

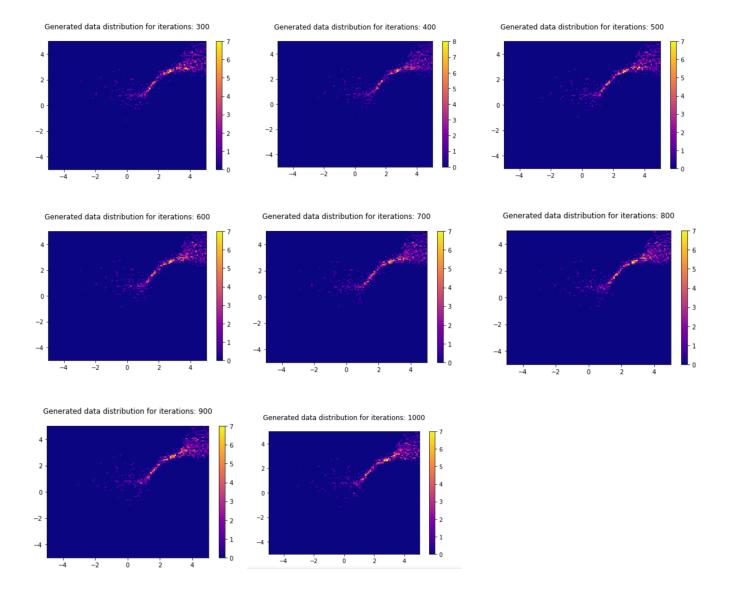Distribution for original dataset:

Distribution for original dataset - Data2

By picking depth d = 5 and seed = 0 for initialization, we perform the training on different iterations and obtain the following results:

Generated data distribution for iterations: 0

Generated data distribution for iterations: 100

Generated data distribution for iterations: 200

Generated data distribution for iterations: 300

Generated data distribution for iterations: 400

Generated data distribution for iterations: 500

Generated data distribution for iterations: 600

Generated data distribution for iterations: 700

Generated data distribution for iterations: 800

Generated data distribution for iterations: 900

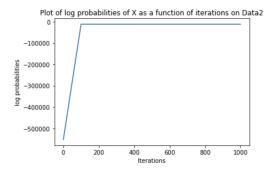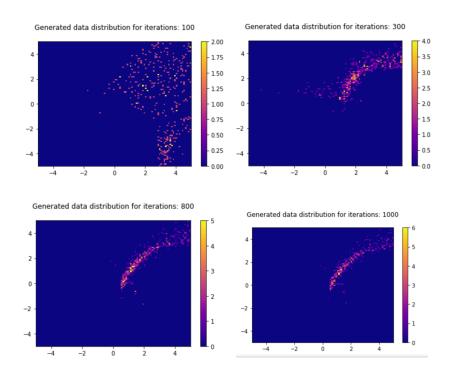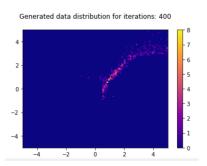**Generated data distribution for iterations: 1000**

From the above figures we can see that the approximate distribution improves with increase in number of iterations. Below is the plot of log probability of data X as a function of iterations. We can see that the log probability of data increases with the training.

```
log probabilities on Data2:
[tensor(-552505.4375), tensor(-11608.2617), tensor(-11590.4629), tensor(-11584.8848), tensor(-11580.2969), tensor(-
11575.6934), tensor(-11570.8301), tensor(-11569.2666), tensor(-11558.5947), tensor(-11553.1562), tensor(-11543.3604
)]
```
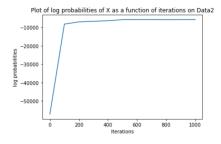

Plot of log probabilities of X as a function of iterations on Data2

For seed = 325, following results were obtained:

Generated data distribution for iterations: 100

Generated data distribution for iterations: 300

Generated data distribution for iterations: 400

Generated data distribution for iterations: 800

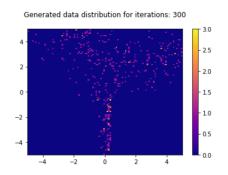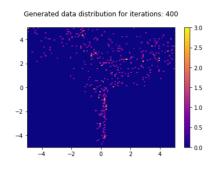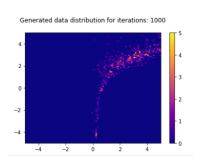Generated data distribution for iterations: 1000

log probabilities on Data2:
[tensor(-57248.1094), tensor(-8198.1709), tensor(-6991.9932), tensor(-6677.4756), tensor(-6360.1562), tensor(-5782.6221), tensor(-5770.4980), tensor(-5773.7212), tensor(-5794.9971), tensor(-5779.9375), tensor(-5776.4067)]

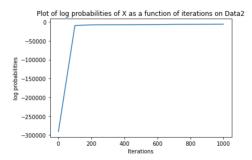Plot of log probabilities of X as a function of iterations on Data2

For seed = 4567, following results were obtained:

Generated data distribution for iterations: 300

Generated data distribution for iterations: 400

Generated data distribution for iterations: 1000

```
log probabilities on Data2:
[tensor(-290547.5938), tensor(-9666.3174), tensor(-7845.8848), tensor(-7697.9399), tensor(-7616.9526), tensor(-7471
.1357), tensor(-6921.5986), tensor(-6498.4639), tensor(-6381.4121), tensor(-6259.9199), tensor(-6032.8013)]
```

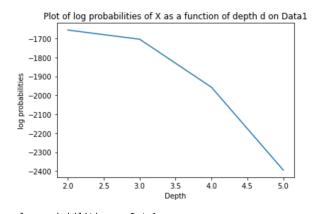Plot of log probabilities of X as a function of iterations on Data2

From the above experimentation, we can see that the similarities to distribution to the original distribution depends on seeds during initialization of parameters. Produces different similarities for different seeds.
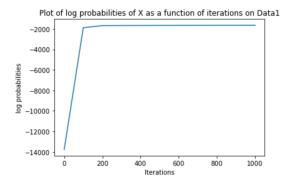
Extra Fun and Credits:

Dataset was split into train and test data, 2/3 of the dataset is used as training data and 1/3rd as test data. Log Probabilities were calculated on test data for various depth and iterations.

Results on Data1 Dataset:

```
log probabilities on Data1:
[tensor(-1655.6016), tensor(-1704.0582), tensor(-1957.4718), tensor(-2394.4875)]
```
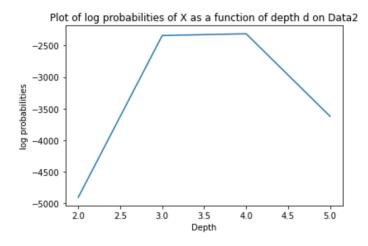
Plot of log probabilities of X as a function of depth d on Data1

```
log probabilities on Data1:
[tensor(-13768.1455), tensor(-1885.5620), tensor(-1683.3137), tensor(-1675.3910), tensor(-1670.0334), tensor(-1666.
1263), tensor(-1663.1188), tensor(-1660.7136), tensor(-1658.7301), tensor(-1657.0519), tensor(-1655.6016)]
```

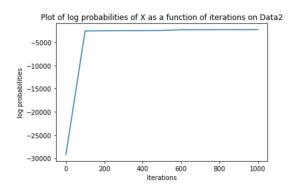Plot of log probabilities of X as a function of iterations on Data1

Experiment on dataset data1 shows that log probability is high at d=2 and decreases further due to overfitting of data and also log probability increases with increase in number of iterations. Therefore for we can see that quantitative results agree with qualitative(Visual) results.

Results on Dataset Data2:

```
log probabilities on Data2:
[tensor(-4904.8975), tensor(-2342.2649), tensor(-2314.8252), tensor(-3622.0884)]
```


Plot of log probabilities of X as a function of depth d on Data2

```
log probabilities on Data2:
[tensor(-29236.0605), tensor(-2604.7759), tensor(-2546.4441), tensor(-2514.4326), tensor(-2511.1411), tensor(-2486.
1794), tensor(-2330.3091), tensor(-2327.0747), tensor(-2321.5166), tensor(-2317.4138), tensor(-2314.8252)]
```


Plot of log probabilities of X as a function of iterations on Data2

Experiment on dataset data2 shows that log probability increases with depth and is highest at d=4 and decreases further due to overfitting of data and also log probability increases with increase in number of iterations. Therefore for we can see that quantitative results agree with qualitative(Visual) results.
(Didn't get time to explore second part of the Extra credit part, will definitely experiment on it)