```
(venv) christos_kyriakou@CK:~/Erg2$ python nlsh_build.py \
  -d /home/christos_kyriakou/project2/data/SIFT/sift_learn.fvecs \
  -i sift_index \
  -type sift \
  --knn 10 \
  -m 100 \
  --epochs 5 \
  --batch_size 256
[BUILD] Loading dataset...
[INFO] Loading SIFT (.fvecs)
[BUILD] Loaded dataset: 100000 vectors, dim=128
[BUILD] Type: sift
[BUILD] Building KNN graph using C IVFFlat (k=10)...
[IVFFLAT] Running C IVFFlat binary...
Loading SIFT: 100000 vectors (dimension=128)
  Loaded 100000 / 100000
SIFT dataset loaded successfully

Loaded 100000 vectors of dimension 128
Building IVFFlat index with 256 clusters...
[KMEANS] Iteration 0/100 (0.0%)
[KMEANS] Iteration 5/100 (5.0%)
[KMEANS] Iteration 10/100 (10.0%)
[KMEANS] Iteration 15/100 (15.0%)
[KMEANS] Iteration 20/100 (20.0%)
[KMEANS] Iteration 25/100 (25.0%)
[KMEANS] Iteration 30/100 (30.0%)
[KMEANS] Iteration 35/100 (35.0%)
[KMEANS] Iteration 40/100 (40.0%)
[KMEANS] Iteration 45/100 (45.0%)
[KMEANS] Iteration 50/100 (50.0%)
```

[KMEANS] Iteration 55/100 (55.0%)

[KMEANS] Iteration 60/100 (60.0%)

[KMEANS] Iteration 65/100 (65.0%)

[KMEANS] Iteration 70/100 (70.0%)

[KMEANS] Iteration 75/100 (75.0%)

[KMEANS] Iteration 80/100 (80.0%)

[KMEANS] Iteration 85/100 (85.0%)

[KMEANS] Iteration 90/100 (90.0%)

[KMEANS] Iteration 95/100 (95.0%)

[KMEANS] Done.

IVFFlat build completed.

Processed 0 / 100000

Processed 1000 / 100000

Processed 2000 / 100000

Processed 3000 / 100000

Processed 4000 / 100000

Processed 5000 / 100000

Processed 6000 / 100000

Processed 7000 / 100000

Processed 8000 / 100000

Processed 9000 / 100000

Processed 10000 / 100000

Processed 11000 / 100000

Processed 12000 / 100000

Processed 13000 / 100000

Processed 14000 / 100000

Processed 15000 / 100000

Processed 16000 / 100000

Processed 17000 / 100000

Processed 18000 / 100000

Processed 19000 / 100000

Processed 20000 / 100000

Processed 21000 / 100000

Processed 22000 / 100000

Processed 23000 / 100000

Processed 24000 / 100000

Processed 25000 / 100000

Processed 26000 / 100000

Processed 27000 / 100000

Processed 28000 / 100000

Processed 29000 / 100000

Processed 30000 / 100000

Processed 31000 / 100000

Processed 32000 / 100000

Processed 33000 / 100000

Processed 34000 / 100000

Processed 35000 / 100000

Processed 36000 / 100000

Processed 37000 / 100000

Processed 38000 / 100000

Processed 39000 / 100000

Processed 40000 / 100000

Processed 41000 / 100000

Processed 42000 / 100000

Processed 43000 / 100000

Processed 44000 / 100000

Processed 45000 / 100000

Processed 46000 / 100000

Processed 47000 / 100000

Processed 48000 / 100000

Processed 49000 / 100000

Processed 50000 / 100000

Processed 51000 / 100000

Processed 52000 / 100000

Processed 53000 / 100000

Processed 54000 / 100000

Processed 55000 / 100000

Processed 56000 / 100000

Processed 57000 / 100000

Processed 58000 / 100000

Processed 59000 / 100000

Processed 60000 / 100000

Processed 61000 / 100000

Processed 62000 / 100000

Processed 63000 / 100000

Processed 64000 / 100000

Processed 65000 / 100000

Processed 66000 / 100000

Processed 67000 / 100000

Processed 68000 / 100000

Processed 69000 / 100000

Processed 70000 / 100000

Processed 71000 / 100000

Processed 72000 / 100000

Processed 73000 / 100000

Processed 74000 / 100000

Processed 75000 / 100000

Processed 76000 / 100000

Processed 77000 / 100000

Processed 78000 / 100000

Processed 79000 / 100000

Processed 80000 / 100000

Processed 81000 / 100000

Processed 82000 / 100000

Processed 83000 / 100000

Processed 84000 / 100000

Processed 85000 / 100000

Processed 86000 / 100000

Processed 87000 / 100000

Processed 88000 / 100000

Processed 89000 / 100000

Processed 90000 / 100000

Processed 91000 / 100000

Processed 92000 / 100000

Processed 93000 / 100000

Processed 94000 / 100000

Processed 95000 / 100000

Processed 96000 / 100000

Processed 97000 / 100000

Processed 98000 / 100000

Processed 99000 / 100000

Done. Wrote k-NN graph to knn_out.bin

[IVFFLAT] Reading knn_out.bin (100000 x 10) ...

[IVFFLAT] KNN graph build complete.

[BUILD] Building undirected weighted graph...

[BUILD] Running KaHIP (m=100, imbalance=0.03)...

[KaHIP] Graph: 100000 vertices, 1692124 edges

[KaHIP] Partitioning into 100 parts (imbalance=0.03)...

[KaHIP] Edgecut: 359555

[BUILD] KaHIP finished.

[BUILD] Sample partitions: [33 8 1 52 15 66 81 29 18 34 29 80 4 50 23 46 47 9 16 27]

[BUILD] Saved partitions → sift_index/partitions.npy

[BUILD] Starting MLP training...

[MLP] Using device: cpu

[MLP] Training...

[MLP] Epoch 1/5 — loss=4046176.9238

[MLP] Epoch 2/5 — loss=698.6420

[MLP] Epoch 3/5 — loss=222.1851

[MLP] Epoch 4/5 — loss=91.4091

[MLP] Epoch 5/5 — loss=44.7784

[MLP] Training finished.

[BUILD] Building inverted index...

[BUILD] Saved model → sift_index/model.pth

[BUILD] Saved inverted index → sift_index/inverted_index.npy

[BUILD] All tasks completed successfully.

(venv) christos_kyriakou@CK:~/Erg2$

MNIST

```
(venv) christos_kyriakou@CK:~/project2$ python nlsh_build.py \
  -d /home/christos_kyriakou/project2/data/SIFT/sift_groundtruth.ivecs \
  -i sift_test_index \
  -type sift \
  --knn 3 \
  -m 20 \
  --epochs 3 \
  --batch_size 128
```

[BUILD] Loading dataset...

[INFO] Loading GT (.ivecs)

[BUILD] Loaded dataset: 10000 vectors, dim=100

[BUILD] Type: gt

[BUILD] Building KNN graph using C IVFFlat (k=3)...

[IVFFLAT] Running C IVFFlat binary...

Loading SIFT: 10000 vectors (dimension=100)

  Loaded 10000 / 10000

SIFT dataset loaded successfully

Loaded 10000 vectors of dimension 100

Building IVFFlat index with 256 clusters...

[KMEANS] Iteration 0/100 (0.0%)

[KMEANS] Iteration 5/100 (5.0%)

[KMEANS] Iteration 10/100 (10.0%)

[KMEANS] Iteration 15/100 (15.0%)

[KMEANS] Iteration 20/100 (20.0%)

[KMEANS] Iteration 25/100 (25.0%)

[KMEANS] Iteration 30/100 (30.0%)

[KMEANS] Iteration 35/100 (35.0%)

[KMEANS] Iteration 40/100 (40.0%)

[KMEANS] Iteration 45/100 (45.0%)

[KMEANS] Iteration 50/100 (50.0%)

[KMEANS] Iteration 55/100 (55.0%)

[KMEANS] Iteration 60/100 (60.0%)

[KMEANS] Iteration 65/100 (65.0%)

[KMEANS] Iteration 70/100 (70.0%)

[KMEANS] Iteration 75/100 (75.0%)

[KMEANS] Iteration 80/100 (80.0%)

[KMEANS] Iteration 85/100 (85.0%)

[KMEANS] Iteration 90/100 (90.0%)

[KMEANS] Iteration 95/100 (95.0%)

[KMEANS] Done.

IVFFlat build completed.

Processed 0 / 10000

Processed 1000 / 10000

Processed 2000 / 10000

Processed 3000 / 10000

Processed 4000 / 10000

Processed 5000 / 10000

Processed 6000 / 10000

Processed 7000 / 10000

Processed 8000 / 10000

Processed 9000 / 10000

Done. Wrote k-NN graph to knn_out.bin

[IVFFLAT] Reading knn_out.bin (10000 x 3) ...

[IVFFLAT] KNN graph build complete.

[BUILD] Building undirected weighted graph...

[BUILD] Running KaHIP (m=20, imbalance=0.03)...

[KaHIP] Graph: 10000 vertices, 59988 edges

[KaHIP] Partitioning into 20 parts (imbalance=0.03)...

[KaHIP] Edgecut: 28418

[BUILD] KaHIP finished.

[BUILD] Sample partitions: [0 1 5 5 0 2 1 0 4 16 6 0 0 0 15 0 6 0 14 0]

[BUILD] Saved partitions → sift_test_index/partitions.npy

[BUILD] Starting MLP training...

[MLP] Using device: cpu

[MLP] Training...

[MLP] Epoch 1/3 — loss=8316.0854

[MLP] Epoch 2/3 — loss=32.8010

[MLP] Epoch 3/3 — loss=16.4522

[MLP] Training finished.

[BUILD] Building inverted index...

[BUILD] Saved model → sift_test_index/model.pth

[BUILD] Saved inverted index → sift_test_index/inverted_index.npy

[BUILD] All tasks completed successfully.