

Comparison of Intrinsic Dimension Estimators: Choosing the Best Intrinsic Dimension Estimator

Aim: Compare the performance of different ID estimators. We will evaluate these estimators based on their mean and standard deviation.

Experiments Setup:

Model: DistilBert

Dataset: Multiclass Sentiment Analysis Dataset

Number of samples: 1000

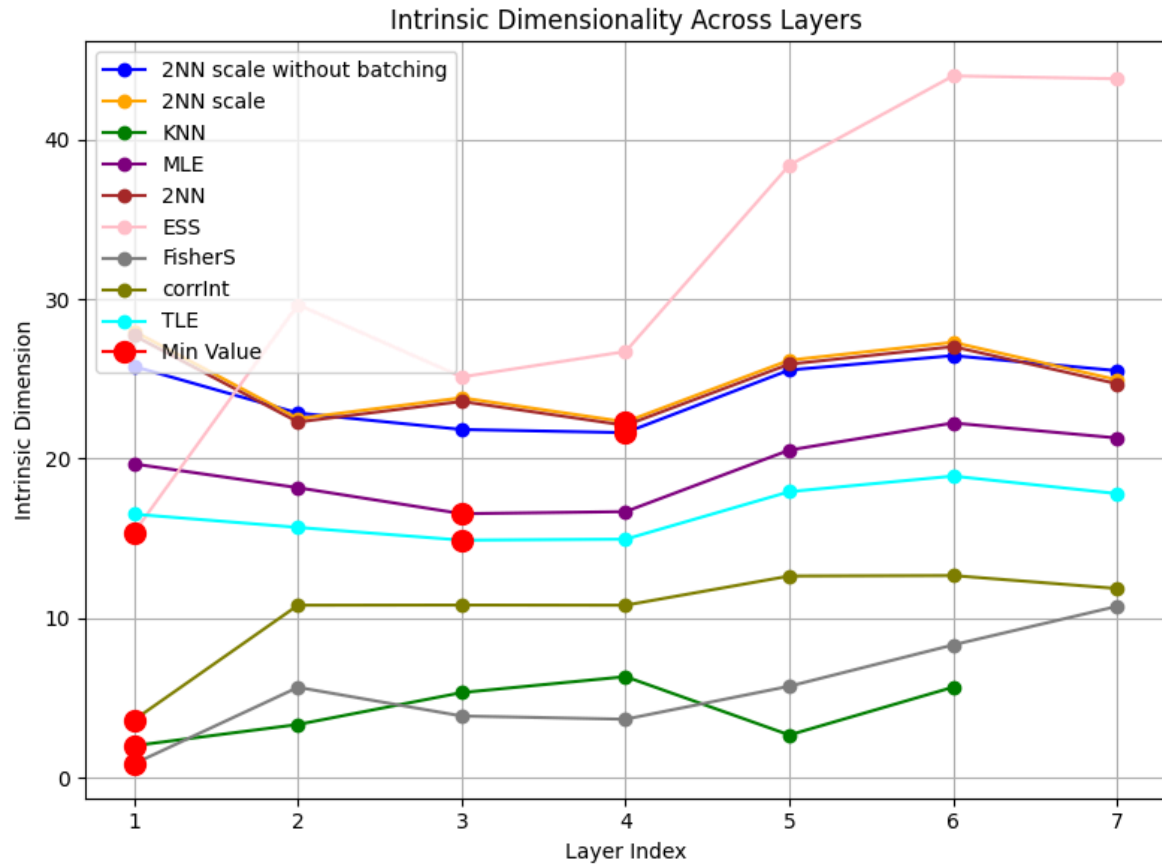
Batches: 3

Evaluation Method: Mean Intrinsic Dimension and Standard Deviation (Calculated from the intrinsic dimensions obtained in each batch).

Results:

Intrinsic Dimension Means across 3 batches

Layer ID	1	2	3	4	5	6	7
2NN no batch	25.78	22.86	21.83	21.63	25.56	26.47	25.52
2NN scale	27.99	22.52	23.82	22.31	26.18	27.3	24.95
KNN	2.	3.33	5.33	6.33	2.67	5.67	258.67
MLE	19.67	18.18	16.55	16.68	20.54	22.24	21.3
2NN	27.72	22.3	23.59	22.09	25.93	27.03	24.7
ESS	15.38	29.66	25.13	26.72	38.41	44.01	43.83
FisherS	0.87	5.66	3.86	3.66	5.73	8.32	10.74
CorrInt	3.62	10.81	10.82	10.81	12.64	12.67	11.86
TLE	16.52	15.69	14.89	14.95	17.93	18.91	17.81



Approaches:

1. Intrinsic Dimension estimation using 2NN Algorithm with Scaling on 1000 samples without batching

- Method: 'return_id_scaling_2NN' [[source](#)]

Layer ID	Intrinsic Dimension
1	25.783684711530256
2	22.862326971615232
3	21.83180717509225
4	21.63306150660067
5	25.56115256440055
6	26.465031360283753
7	25.521369064933594

2. Intrinsic Dimension estimation using 2NN Algorithm with Scaling

- Method: 'return_id_scaling_2NN' [\[source\]](#)

Approach:

Intrinsic dimensionality is estimated using the 2NN algorithm across various scales, from the entire dataset down to smaller subsets. For each layer in the transformer, we select the ID associated with the scale with the minimum error, ensuring the most accurate representation of the data's dimensionality.

In this experiment, the intrinsic dimension at the first scale, which is the full subset, is considered. At this scale, the error is always zero.

Layer ID	Mean Intrinsic Dimension	Standard Deviation
1	27.99	0.83
2	22.52	0.66
3	23.82	2.13
4	22.31	0.16
5	26.18	2.33
6	27.3	1.71
7	24.95	2.71

3. Intrinsic Dimension estimation using kNN algorithm

- Method: 'skdim.id.KNN' [\[source\]](#)

Layer ID	Mean Intrinsic Dimension	Standard Deviation
1	2.	0.82
2	3.33	1.89
3	5.33	4.71
4	6.33	4.78
5	2.67	0.94
6	5.67	2.87
7	258.67	360.15

4. Intrinsic dimension estimation using the Maximum Likelihood algorithm

- Method: 'skdim.id.MLE' [[source](#)]

Layer ID	Mean Intrinsic Dimension	Standard Deviation
1	19.67	0.47
2	18.18	0.32
3	16.55	0.39
4	16.68	0.09
5	20.54	0.48
6	22.24	0.72
7	21.3	0.7

5. Intrinsic dimension estimation using the TwoNN algorithm.

- Method: 'skdim.id.TwoNN' [[source](#)]

Layer ID	Mean Intrinsic Dimension	Standard Deviation
1	27.72	0.83
2	22.3	0.66
3	23.59	2.11
4	22.09	0.16
5	25.93	2.31
6	27.03	1.7
7	24.7	2.69

6. Intrinsic dimension estimation using the Expected Simplex Skewness algorithm.

- Method: 'skdim.id.ESS' [[source](#)]

Layer ID	Mean Intrinsic Dimension	Standard Deviation
1	15.38	0.95
2	29.66	0.65
3	25.13	0.56
4	26.72	0.5
5	38.41	1.02
6	44.01	1.44
7	43.83	0.46

7. Intrinsic dimension estimation using the Fisher Separability algorithm.

- Method: 'skdim.id.FisherS' [[source](#)]

Layer ID	Mean Intrinsic Dimension	Standard Deviation
1	0.87	0.01
2	5.66	0.71
3	3.86	0.36
4	3.66	0.03
5	5.73	0.32
6	8.32	0.83
7	10.74	0.57

8. Intrinsic dimension estimation using the Correlation Dimension.

- Method: 'skdim.id.CorrInt' [[source](#)]

Layer ID	Mean Intrinsic Dimension	Standard Deviation
1	3.62	0.79
2	10.81	0.34
3	10.82	0.2
4	10.81	0.27
5	12.64	0.62
6	12.67	0.83
7	11.86	0.16

9. Intrinsic dimension estimation using the Tight Local intrinsic dimensionality Estimator algorithm.

- Method: 'skdim.id.TLE' [[source](#)]

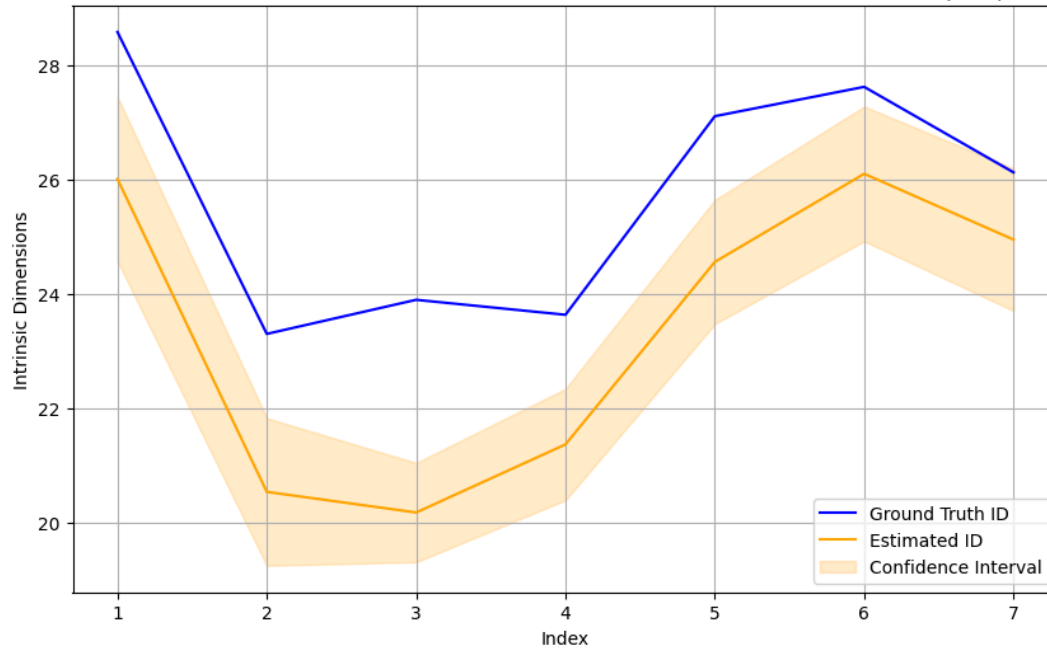
Layer ID	Mean Intrinsic Dimension	Standard Deviation
1	16.52	0.32
2	15.69	0.26
3	14.89	0.32
4	14.95	0.11
5	17.93	0.42
6	18.91	0.63
7	17.81	0.48

Results: Intrinsic Dimensions with Confidence Intervals

Ground Truth ID - Intrinsic Dimension estimation using 2NN Algorithm with scaling

Estimated ID - Intrinsic Dimension estimation using 2NN Algorithm with scaling averaged (weighted) over number of samples

Intrinsic Dimensions with Confidence Intervals -- DistilBert Model -- SP Dataset (4500 samples | 10 batches)



Intrinsic Dimensions with Confidence Intervals -- DistilBert Model -- SP Dataset (4500 samples | 4 batches)

