Appendix – Dataset-Adaptive Dimensionality Reduction

A THEOREMS AND PROOFS

A.1 Theoretical Verifications for the Complexity Metrics

Theorem 1. *Let $X, Y \subset \mathbb{R}^d$ ($d \in \mathbb{N}$) be random variables such that $\mathbf{x} \in X$ and $\mathbf{y} \in Y$ follow i.i.d. (independent and identically distributed) density functions. $\sigma(X)$ and $\mathsf{E}(X)$ refer to the standard deviation and expectation value of X. We also define $\|X\| = \{\|\mathbf{x}\| \mid \mathbf{x} \in X\}$, where $\|\mathbf{x}\|$ is the Euclidean norm of $\mathbf{x}$. Suppose there exists $Z \subset \mathbb{R}^d$, such that $Z = \{\mathbf{x} - \mathbf{y} \mid \mathbf{x} \in X, \mathbf{y} \in Y\}$. Then,*

$$\frac{\sigma(\|Z\|)}{\mathsf{E}(\|Z\|)} = \Theta(1/\sqrt{d}).$$

proof According to Demartines [3],

$$\mathsf{E}(\|Z\|) = \sqrt{ad - b} + \Theta(1/d) \text{ and } \sigma(\|Z\|) = b + \Theta(1/\sqrt{d}),$$

where $a, b \in \mathbb{R}$ and $a > 0$.
This implies that $\mathsf{E}(\|Z\|) = \Theta(\sqrt{d})$ and $\sigma(\|Z\|) = \Theta(1)$ since $1/d$ is negligible compared to $\sqrt{d}$. Hence,

$$\frac{\sigma(\|Z\|)}{\mathsf{E}(\|Z\|)} = \Theta(1/\sqrt{d}).$$

□

Theorem 1-1. *Given a random variable $X \subset \mathbb{R}^d$ such that $\mathbf{x} \in X$ follow i.i.d. (independent and identically distributed) density function. Then,*

$$\frac{\sigma(\|X\|)}{\mathsf{E}(\|X\|)} \to 0 \text{ as } d \to \infty.$$

.

proof For any $\varepsilon > 0$, we should show that $\exists \delta$ such that $\forall d > \delta$,

$$\frac{\sigma(\|X\|)}{\mathsf{E}(\|X\|)} < \varepsilon.$$

By Theorem 1, there exists $C > 0$ and $d_0 > 0$ satisfying

$$\frac{\sigma(\|X\|)}{\mathsf{E}(\|X\|)} \leq \frac{C}{\sqrt{d}}$$

$\forall d > d_0$.
Let

$$\delta = \max\left(\left(\frac{C}{\varepsilon}\right)^2, d_0\right).$$

Then, $\forall d > \delta$,

$$\frac{\sigma(\|X\|)}{\mathsf{E}(\|X\|)} < \frac{C}{\sqrt{\delta}} = C \cdot \frac{\varepsilon}{C} = \varepsilon.$$

□

Theorem 2. *Let $\overline{X}_N$ be a set of N data samples from a multidimensional dataset $X \subset \mathbb{R}^d$. Let $i, j, k, \in N$. Two points $p_i, p_j \in \overline{X}_N$ are said to be in a k-nearest neighbor (k-NN) relationship if one is among the k nearest neighbors of the other, and are in an SNN relationship if the SNN similarity between p_i and p_j is above 0. Suppose there exists a constant integer $k > 0$ where $k \ll N$. Then, $\forall p_i, p_j \in \overline{X}_N$, the probability in which p_i and p_j are both in kNN and SNN relationship approximates to k^3/N^2 as $d \to \infty$.*

proof According to Theorem 1-1,

$$\frac{\sigma(\|X\|)}{\mathsf{E}(\|X\|)} \to 0 \text{ as } d \to \infty.$$

Under such equidistance, the probability of p_j to become kNN of p_i approximates to:

$$\mathbb{P}\Big[p_j \in k\text{NN}(p_i)\Big] \approx \frac{k}{N-1} \sim \frac{k}{N}. \tag{1}$$

and vice versa.
Let

$$A_i = \{a_i | a_i \in k\text{NN of } p_i\}, \quad A_j = \{a_j | a_j \in k\text{NN of } p_j\}.$$

An estimate for the probability that A_i and A_j have a nonempty intersection is obtained as follows. As mentioned above, $\forall p \in \overline{X}$ where $p \neq p_i, p \neq p_j$, the probability of being in A_i and A_j is approximately k/N. Hence, the probability of p_j and p_j to be in SNN relationship approaches:

$$\mathbb{P}\Big[A_i \cap A_j \neq \varnothing\Big] \approx (N-2)\left(\frac{k}{N}\right)^2$$

Using Taylor expansion:

$$\left(1 - \frac{k}{N}\right)^k = 1 - \frac{k^2}{N} + O\left(\frac{k^4}{N^2}\right),$$

we get

$$\mathbb{P}\Big[A_i \cap A_j \neq \varnothing\Big] \approx 1 - \left(1 - \frac{k}{N}\right)^k \sim \frac{k^2}{N}.$$

Thus, the probability in which p_i and p_j is both kNN and SNN approximates to $k/N \cdot k^2/N = k^3/N^2$. □

Theorem 4. Given $\overline{X}_N$ and $0 < \varepsilon < 1$, suppose there exists a linear mapping $f : \mathbb{R}^d \to \mathbb{R}^2$ such that for all distinct $x, y \in \overline{X}_N$,

$$(1-\varepsilon)\|x-y\|^2 \leq \|f(x) - f(y)\|^2 \leq (1+\varepsilon)\|x-y\|^2.$$

Then there exists a constant $c > 0, c \in \mathbb{R}$ such that

$$\varepsilon \geq c\sqrt{\ln d}.$$

proof By the Johnson–Lindenstrauss lemma [7], for any $0 < \varepsilon < 1$ and $\overline{X}_N$, there exists a linear mapping

$$f : \mathbb{R}^d \to \mathbb{R}^m$$

satisfying

$$(1-\varepsilon)\|x-y\|^2 \leq \|f(x) - f(y)\|^2 \leq (1+\varepsilon)\|x-y\|^2$$

for all $x, y \in \overline{X}_N$, provided that the target dimension m satisfies

$$m \geq C\frac{\ln d}{\varepsilon^2},$$

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

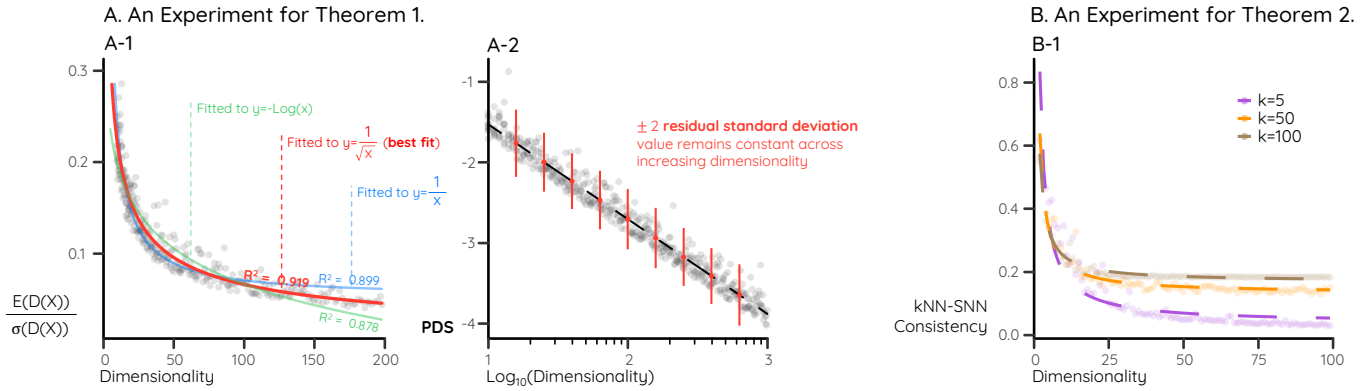

Fig. 1: Empirical verifications for grounding our metric design. We generate HD random distributions with diverse dimensionality following i.i.d. density functions randomly determined by the mixture of Gaussian distributions, denoted as X, and investigate how (A) $\frac{\sigma(D(X))}{\mathbf{E}(D(X))}$ and (B) average cosine similarity between the rows of kNN and SNN (i.e., kNN-SNN consistency) changes over dimensionality.

where $C > 0$ is a constant.
Our theorem's goal is to prove the inequality when $m = 2$. Defining $c = \sqrt{C/2}$, we have

$$2 \geq C\frac{\ln d}{\varepsilon^2} \Rightarrow \varepsilon \geq \sqrt{\frac{C\ln d}{2}} \Rightarrow \varepsilon \geq \sqrt{\frac{C}{2}}\sqrt{\ln d} = c\sqrt{\ln d}.$$

□

A.2 Empirical Verifications for Theorem 1 and 3

We empirically verify Theorem 1 and 3, further justifying the validity of our metric design.

Empirical verification for Theorem 1. We generate HD random Gaussian distribution with diverse dimensionality following i.i.d. density function, denoted as X. We then calculate all pairwise Euclidean distances between distinct data points in X, represented as $D(X) = \{\delta(p_i, p_j) \mid p_i, p_j \in X, p_i \neq p_j\}$, where $\delta(p_i, p_j)$ is the Euclidean distance between points p_i and p_j. We investigate how the ratio $\sigma(D(X))/E(D(X))$—degree of distance shift—changes with dimensionality. Our experimental results (Fig. 1A) show that this ratio decreases proportionally to $1/\sqrt{d}$, which aligns with our theoretical prediction. Furthermore, we observe a linear decrease when we plot the logarithm of this ratio against the logarithm of the dimensionality. This linear relationship reaffirms our theoretical findings as $\log(1/\sqrt{d}) = -0.5\log d$.

Empirical validation for Theorem 2. We prepare HD random Gaussian distribution X and examine how the average cosine similarity between the rows of kNN and SNN matrices (i.e., kNN-SNN consistency) changes over dimensionality, where the (i, j)-th cell of kNN and SNN matrices is 1 if the i-th and j-th points are in kNN or SNN relationship, respectively, and otherwise 0. The results (Fig. 1B) show that consistency converges to a certain value as d increases, supporting the theorem. Additionally, the fact that the converged value decreases as k gets small supports our theoretical finding that the consistency converges to k^3/N^2.

A.3 Theorems on the Violation of R3

Theorem 5. *Let $X \subset \mathbb{R}^d$ and arbitrary small number ε, the correlation integral of X is defined as:*

$$C_X(\varepsilon) = \frac{2 \cdot g}{|X|(|X|-1)},$$

where g is the number of pairs in X that has distance smaller than ε. We also define correlation dimension of X ν_X as the value satisfying:

$$C_X(\varepsilon) \sim \varepsilon^{\nu_X}.$$

Then, the correlation dimension is not scale invariant.

proof Let $\alpha > 0$ and define $\alpha X := \{\alpha x : x \in X\}$. For $C_X(\varepsilon)$ as in the statement, note that

$$C_{\alpha X}(\varepsilon) = C_X\left(\tfrac{\varepsilon}{\alpha}\right),$$

since scaling by α shrinks distances by α^{-1} when comparing to the same ε. Here, ν_X can be represented as:

$$\nu_X = \lim_{\varepsilon\to 0}\frac{C_X(\varepsilon)}{\phi(\varepsilon)}$$

for some function ϕ. Then

$$\nu_{\alpha X} = \lim_{\varepsilon\to 0}\frac{C_{\alpha X}(\varepsilon)}{\phi(\varepsilon)} = \lim_{\varepsilon\to 0}\frac{C_X\left(\frac{\varepsilon}{\alpha}\right)}{\phi(\varepsilon)}.$$

By substituting $\delta = \varepsilon/\alpha$, we have $\varepsilon = \alpha\,\delta$ and $\delta \to 0$ as $\varepsilon \to 0$. Hence

$$\nu_{\alpha X} = \lim_{\delta\to 0}\frac{C_X(\delta)}{\phi(\alpha\,\delta)} \neq \lim_{\delta\to 0}\frac{C_X(\delta)}{\phi(\delta)} = \nu_X.$$

□

Theorem 6. *Let $X \subset \mathbb{R}^d$ and define Box-counting dimension*

$$BC_X = \lim_{\varepsilon\to 0}\frac{N_X(\varepsilon)}{\log(1/\varepsilon)},$$

where $N_X(\varepsilon)$ is the number of d-dimensional cubes of side ε needed to cover X. Then, box-counting dimension is not scale invariant.

proof For any $\alpha > 0$,

$$N_{\alpha X}(\varepsilon) = N_X\left(\tfrac{\varepsilon}{\alpha}\right),$$

where $\alpha X := \{\alpha x : x \in X\}$. Hence,

$$BC_{\alpha X} = \lim_{\varepsilon\to 0}\frac{N_{\alpha X}(\varepsilon)}{\log(1/\varepsilon)} = \lim_{\varepsilon\to 0}\frac{N_X\left(\frac{\varepsilon}{\alpha}\right)}{\log(1/\varepsilon)}.$$

Now substitute $\delta = \varepsilon/\alpha$, so $\varepsilon = \alpha\,\delta$ and $\varepsilon \to 0$ implies $\delta \to 0$. Then,

$$BC_{\alpha X} = \lim_{\delta\to 0}\frac{N_X(\delta)}{\log(1/(\alpha\,\delta))} = \lim_{\delta\to 0}\frac{N_X(\delta)}{\log(1/\delta)+\log(1/\alpha)}.$$

Meanwhile,

$$BC_X = \lim_{\delta\to 0}\frac{N_X(\delta)}{\log(1/\delta)}.$$

Thus, $BC_{\alpha X} \neq BC_X$. □

Theorem 7. *Steadiness & Cohesiveness is not scale invariant.*

proof Given HD dataset X and its low-dimensional projection Y, Steadiness & Cohesiveness is computed by finding clusters in X (or Y) using HDBSCAN [2] and checking their dispersion in Y (or X).
Here, HDBSCAN is not scale invariant by design [2]. Thus, Steadiness & Cohesiveness is not also scale invariant. □

Theorem 8. *Let $X \subset \mathbb{R}^d$ and we define KL divergence*

$$KL_{\text{div}}(X,Y) \;=\; \sum_{i=1}^{N} f_\sigma^X(x_i)\,\log\Big(\tfrac{f_\sigma^X(x_i)}{f_\sigma^Y(y_i)}\Big),$$

where

$$f_\sigma^X(x_i) \;=\; \sum_{x_j\in X}\exp\big(-\tfrac{d(x_i,x_j)^2}{\sigma}\big), \quad f_\sigma^Y(y_i) \;=\; \sum_{y_j\in Y}\exp\big(-\tfrac{d(y_i,y_j)^2}{\sigma}\big),$$

and $\sigma > 0$ is a kernel bandwidth parameter. Then, KL divergence is not scale invariant)

proof Consider the rescaled set $\alpha X = \{\alpha x \mid x \in X\}$ for some $\alpha > 0$. Then $d(\alpha x_i, \alpha x_j) = \alpha\, d(x_i,x_j)$. Consequently,

$$f_\sigma^{\alpha X}(\alpha x_i) \;=\; \sum_{x_j\in X}\exp\Big(-\tfrac{\big(\alpha\, d(x_i,x_j)\big)^2}{\sigma}\Big) \;=\; \sum_{x_j\in X}\exp\Big(-\tfrac{\alpha^2 d(x_i,x_j)^2}{\sigma}\Big),$$

thus

$$KL_{\text{div}}(\alpha X, Y) \;\neq\; KL_{\text{div}}(X,Y),$$

unless Y is also rescaled by the same factor α. □

Theorem 9. *Let $X \subset \mathbb{R}^d$ and we define DTM*

$$DTM(X,Y) \;=\; \sum_{i=1}^{N}\big|\, f_\sigma^X(x_i) \;-\; f_\sigma^Y(y_i)\big|,$$

where f_σ^X and f_σ^Y are given as in Theorem 4. Then, DTM is not scale invariant.)

Proof. Under a global scaling $\alpha X = \{\alpha x \mid x \in X\}$, the pairwise distances $d(x_i,x_j)$ become $\alpha\, d(x_i,x_j)$, so

$$f_\sigma^{\alpha X}(\alpha x_i) \;=\; \sum_{x_j\in X}\exp\Big(-\tfrac{\alpha^2 d(x_i,x_j)^2}{\sigma}\Big).$$

Hence

$$\big|\, f_\sigma^{\alpha X}(\alpha x_i) \;-\; f_\sigma^Y(y_i)\big| \;\neq\; \big|\, f_\sigma^X(x_i) \;-\; f_\sigma^Y(y_i)\big|,$$

and thus $DTM(\alpha X,Y) \neq DTM(X,Y)$ unless Y is also rescaled. □

B Technical Settings of the Experiments

The following are the experimental settings that are shared across the experiments.

DR techniques and hyperparameters. We use `scikit-learn` [9] implementation for t-SNE, LLE, Isomap, and PCA, and used authors' implementation for UMAP and UMATO. The hyperparameter range we used to optimize the techniques is as follows:

- UMAP
  - `n_neighbors`: (2, 200)
  - `min_dist`: (0.01, 0.99)
- t-SNE
  - `perplexity`: (2, 500)
- LLE
  - `n_neighbors`: (2, 200)
- Isomap
  - `n_neighbors`: (2, 200)
- UMATO
  - `n_neighbors`: (2, 200)
  - `min_dist`: (0.01, 0.99)
  - `hub_num`: (50, 500)

DR evaluation metrics and hyperparameters. We execute five DR evaluation metrics (T&C, MRREs, L-T&C, S-ρ, and P-r) using ZADU [6] library. We set $k = 25$ for T&C and MRREs as they rely on kNN structure. We use the Distance consistency function [11] to examine cluster structure in L-T&C, as it is the default setting of the ZADU library. S-ρ and P-r have no hyperparameter.

Regression models and hyperparameters. We optimize the hyperparameters of regression models to find their ideal performance in predicting the surrogate ground truth structural complexity. We leverage five-fold cross-validation and Bayesian optimization. The following is the hyperparameter range we use. We omit linear and polynomial regressions as they have no hyperparameter that significantly affects regression results. In the polynomial regression, we set the degree to be two. Note that we use `scikit-learn` implementations for all regression models.

- k-Nearest Neighbors Regression
  - `n_neighbors`: (1, 20),
  - `weights`: `uniform | distance`
- Random Forest Regression
  - `n_estimators`: (5, 200)
  - `max_depth`: (1, 10)
  - `criterion`: `squared_error | friedman_mse | absolute_error | poisson`
- Gradient Boosting Regression
  - `n_estimators`: (5, 200)
  - `max_depth`: (1, 10)
  - `learning_rate`: (0.01, 0.5)

C Intrinsic Dimensionality Metrics

We detail the procedure of computing projection-based and geometric intrinsic dimensionality metrics. For the former, we use the number of principal components required to explain more than 95% of data variance in PCA, as they are validated to correlate with accuracy scores by a recent survey on DR [4]. For the latter, we use the fractal dimension as they are widely used in literature [8, 12]. We use the correlation dimension method [5] to calculate the fractal dimension because it satisfies R3. Formally, given dataset X and an arbitrary small number ε, the correlation dimension is computed by first computing the correlation integral:

$$C(\varepsilon) = \frac{2\cdot g}{|X|(|X|-1)},$$

where g is the number of pairs in X that has distance smaller than ε. Then, we find correlation dimension ν as the value satisfying:

$$C(\varepsilon) \sim \varepsilon^{\nu}.$$

Following Grassberger and Procaccia [5], we compute ν by computing $C(\varepsilon)$ over many ε values and obtaining the slope represented by a log-log plot of $C(\varepsilon)$ versus ε.

D Efficiency of Structural Complexity Metrics in Larger Datasets

We want to check whether our structural complexity metrics work scalably for large datasets.

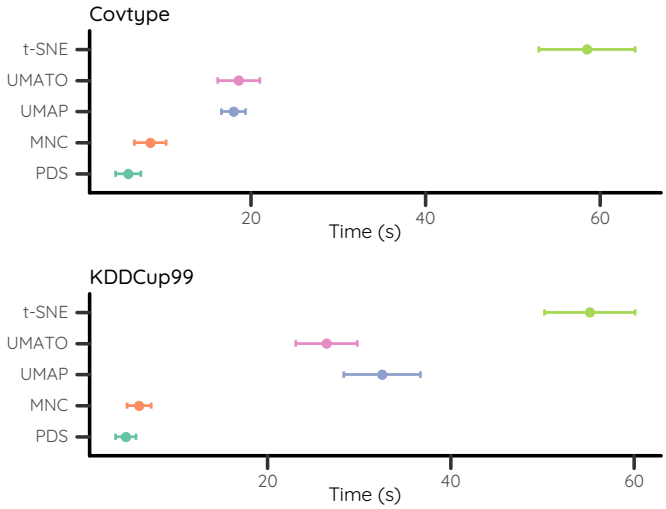

Fig. 2: The runtime of sctructural complexity metrics compared to the widely used DR techniques (Appendix D.1).

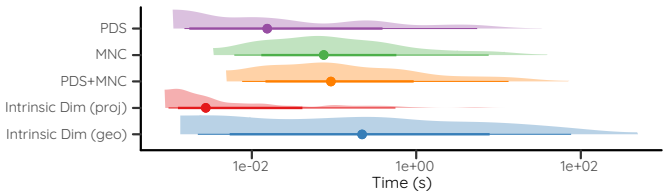

Fig. 3: The runtime of our structural complexity metrics and baselines without sampling data points (Appendix D.2).

D.1 Validating the Computational Benefit of PDS and MNC

We aim to validate that PDS and MNC are computationally beneficial (P3), which means that their runtime is faster than the single execution of DR techniques that are applied for the dataset-adaptive workflow. To do so, we apply our metrics to two large datasets: kddcup99 (494,021 $\times$ 41) [1] and covtype (581,012 $\times$ 54) [1]. To scrutinize the scalability, we randomly sample the datasets from 0.5% to 10% with an interval 0.5% and apply PDS and MNC five times each and record the runtime. To verify the computational benefit of two metrics, we also compute three scalable DR techniques (UMAP, t-SNE, and UMATO) five times and record the runtime. We utilize these three DR techniques as they are top-3 fast techniques that we test in our main experiments.

By examining a average runtime over the sampled datasets (Fig. 2), we find that PDS and MNC are substantially faster than DR techniques we compared for all datasets. The result verifies that these metrics are computationally beneficial.

It is worth noting that we are not able to test the scalability of our structural complexity metrics on full datasets as our current implementation stores all required memory in a single GPU. We find that large datasets thus result in an out-of-memory error (empirically about 70,000–100,000 data points, depending on the dimensionality, on TITAN RTX GPU). Improving our implementation to alleviate this issue, e.g., by using multiple GPUs or utilizing chunk operation, will be a critical future direction to pursue to further enhance the applicability of our metrics and the dataset-adaptive workflow.

D.2 Replicating Experiment 1 without Sampling

We compare the runtime of PDS, MNC, PDS+MNC, and intrinsic dimensionality metrics on 96 datasets without sampling the number of points (original experiment in Section 6). We do not test the DR ensemble as it needs more than weeks of computations, already having a substantially longer runtime compared to structural complexity metrics and intrinsic dimensionality metrics.

The study shows consistent results with our initial analysis in the main manuscript (Fig. 3): PDS and MNC outperform the geometric intrinsic dimensionality metric but are slower than the projection-based one. Still, even with large datasets, the average time to compute PDS and MNC are below one second (1e+00), further verifying their efficiency.

Table 1: Reproduction of experiment 1 using MNC with different k values. The results confirm that the correlation between MNC and ground truth structural complexity stays robust regardless of k value.

		Structural Complexity				
		Local		Cluster	Global	
		T&C	MRREs	L-T&C	S-ρ	P-r
MNC ($k=25$)	LR	.8452	.7047	.3205	.5280	.5105
	PR	.8803	.8035	.2813c	.5171	.5198
	kNN	.8188	.7429	.3452	.5589	.5045
	RF	.8449	.7780	.2489	.5270	.5232
	GB	.8296	.7525	.1986	.5105	.4677
MNC ($k=75$)	LR	.8510	.6975	.4345	.5954	.5872
	PR	.8822	.7452	.4484	.6115	.5513
	kNN	.8386	.6753	.3592	.6171	.6073
	RF	.8508	.6654	.4273	.6182	.5720
	GB	.8240	.6126	.3355	.5886	.6130
MNC ($k=100$)	LR	.8300	.6952	.4730	.6071	.6023
	PR	.8704	.7035	.4441	.6162	.5940
	kNN	.8352	.6861	.4262	.6421	.6311
	RF	.8346	.6867	.4230	.6239	.6135
	GB	.8068	.5759	.4512	.6218	.5991

1. ■ / ■ / ■: very strong ($R^2 \geq 0.8$) / strong ($0.6 \leq R^2 < 0.8$) / moderate ($0.4 \leq R^2 < 0.6$) correlation [10]
2. ***Bold and italic*** refers to the top score of the column

E Reproduction of Experiment 1 with Different k values

In our main manuscript, we test MNC with $k = 50$. In Table 1, we report the reproduction of our correlation analysis with MNC with k values. The table clearly depicts that the performance of MNC stays still or even enhanced with different k values.

References

[1] A. Asuncion, D. Newman, et al. Uci machine learning repository, 2007. 4

[2] R. J. G. B. Campello, D. Moulavi, and J. Sander. Density-based clustering based on hierarchical density estimates. In *Advances in Knowledge Discovery and Data Mining*, pp. 160–172. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. doi: 10.1007/978-3-642-37456-2_14 3

[3] P. Demartines. *Analyse de données par réseaux de neurones auto-organisés*. PhD thesis, Grenoble INPG, 1994. 1

[4] M. Espadoto, R. M. Martins, A. Kerren, N. S. T. Hirata, and A. C. Telea. Toward a quantitative survey of dimension reduction techniques. *IEEE Transactions on Visualization and Computer Graphics*, 27(3):2153–2173, 2021. doi: 10.1109/TVCG.2019.2944182 3

[5] P. Grassberger and I. Procaccia. Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9(1):189–208, 1983. doi: 10.1016/0167-2789(83)90298-1 3

[6] H. Jeon, A. Cho, J. Jang, S. Lee, J. Hyun, H.-K. Ko, J. Jo, and J. Seo. Zadu: A python library for evaluating the reliability of dimensionality reduction embeddings. In *2023 IEEE Visualization and Visual Analytics (VIS)*, pp. 196–200, 2023. doi: 10.1109/VIS54172.2023.00048 3

[7] W. B. Johnson, J. Lindenstrauss, et al. Extensions of lipschitz mappings into a hilbert space. *Contemporary mathematics*, 26(189-206):1, 1984. 1

[8] R. Karbauskaitė and G. Dzemyda. Fractal-based methods as a technique for estimating the intrinsic dimensionality of high-dimensional data: a survey. *Informatica*, 27(2):257–281, 2016. 3

[9] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn:

Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011. 3

[10] K. Sarjana, L. Hayati, and W. Wahidaturrahmi. Mathematical modelling and verbal abilities: How they determine students' ability to solve mathematical word problems? *Beta: Jurnal Tadris Matematika*, 13(2):117–129, Nov. 2020. doi: 10.20414/betajtm.v13i2.390 4

[11] M. Sips, B. Neubert, J. P. Lewis, and P. Hanrahan. Selecting good views of high-dimensional data using class consistency. *Computer Graphics Forum*, 28(3):831–838, 2009. doi: 10.1111/j.1467-8659.2009.01467.x 3

[12] J. Theiler. Estimating fractal dimension. *J. Opt. Soc. Am. A*, 7(6):1055–1073, Jun 1990. doi: 10.1364/JOSAA.7.001055 3