On the impact of the fitness lanscape of clustering problems on the performance of neural-network based optimizers

ABSTRACT

CCS CONCEPTS

•**Computer systems organization** → **Embedded systems;** *Redundancy;* Robotics; •**Networks** → Network reliability;

KEYWORDS

ACM proceedings, LaTeX, text tagging

ACM Reference format:

. 2023. On the impact of the fitness lanscape of clustering problems on the performance of neural-network based optimizers. In *Proceedings of ACM, , July 2023 (GECCO-2023),* 3 pages.

DOI: 10.475/123_4

1 INTRODUCTION

Clustering problems are ubiquitous in real-world applications. Solving this type of problems is a required step for addressing many machine learning tasks. Grouping a set of points or solutions into a number of clusters can be posed as an optimization problem and several heuristics have been defined for solving it. Clustering problems can also be used to evaluate the efficiency of heuristic optimization algorithms [6, 14, 16], and the impact of the parameters that define the clustering in the fitness landscape of the objective function to be solved has been also analyzed [8]. Since clustering problems are both relevant from the point of view of real-world applications and challenging from the point of view of optimization, they are useful benchmark for understanding how the characteristics of problems influence the performance of the optimization algorithms.

Recently, an increasing number of papers report on the use of deep neural networks as a component of optimization methods [2, 17]. While the application of neural networks to combinatorial optimization is not new [11, 19], the emergence of novel neural network paradigms has consolidated this research trend. Among the proposed approaches, some consider the problem of evolving neural network models able to efficiently solve a class of optimization problem [10]. The evolved neural network solvers can be seen not as a solution to a particular instance of the optimization problem, but as a model able to produce solutions for different instances.

In this paper we investigate the behavior of neural networks when solving instances of the clustering problem with well defined, and previously investigated characteristics. The prediction task to be tackled by the neural network is the one traditionally solved by optimization methods, given a set of points that can be grouped into k clusters, predict the centroid for each of the clusters. We assume that k is known and focus on the analysis of the Euclidean sum of squared error (SSE) clustering variant, which can be defined as a continuous optimization problem. In previous work [6, 7], a characterization of the fitness landscape of the SSE clustering problem in terms of key parameters of the problem class has been proposed. We study the performance or neural networks trained with backpropagation, and consider both, neural networks with randomly generated architectures and network architectures evolved to maximize the performance for the SSE clustering problem.

The goals of the research are: i) To investigate the performance of neural networks for the SSE clustering problem and their sensitivity to the parameters that define the problem instances[1]. ii) To evaluate the suitability of neuro-evolutionary techniques to create general solvers of the SSE clustering problem [2]. iii) To assess the ability of the neural networks to capture patterns of the clustering instances and their capacity to generalize solutions between clustering instances with different points distributions [3].

The paper is organized as follows:

2 BACKGROUND

In this section we introduce the the Euclidean sum of squared error (SSE) clustering problem.

2.1 Sum of squares clustering

Given a finite dataset $\mathcal{Y} = \{\mathbf{y}_1, \dots \mathbf{y}_n\}\mathcal{R}^d$ of of n data points located in p-dimensional continuous space together with a second set of k points (aka cluster centers), $C = \{\mathbf{c}_1, \dots \mathbf{c}_n\}\mathcal{R}^d$, the clustering problem is to determine the positions of the cluster centers such that the mean sum of Euclidean distances (L2 Norm) between each data point and its nearest cluster center is minimized:

$$\min_{C} f(C|\mathcal{Y}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k} b_{i,j} \parallel \mathbf{y}_i - \mathbf{c}_j \parallel^2 \tag{1}$$

where

$$b_{i,j} = \begin{cases} 1, & \text{if } \parallel \mathbf{y}_i - \mathbf{c}_j \parallel = min_j \parallel \mathbf{y}_i - \mathbf{c}_j \parallel \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

[1]The idea here would be to try to re-use the findings from your previous paper where you identified combinations of parameters that were "more difficult" for the (k-means) optimizer to solve and determining if the characteristics that make a problem difficult for a "traditional" solver like k-means are equally difficult for a neural network

[2]In the experiments conducted I have seen that evolution can progressively generate better architectures in terms of the accuracy of the predicted centroids. Therefore, we can also present the research from the poing of view of neuro-evolution and I have not found papers that address it from this perspective.

[3]This is related to what we discussed about the capacity of the neural networks to capture patterns of the centroid solutions. For example, for the seven circles within a circle, to capture the feature that centroids should be within a given distance of each other. In the preliminary examples, using only two clusters, the neural network *seems* to be learning some sort of dependency between the location of the centroids for the two clusters

DOI: 10.475/123_4

This is an unconstrained, continuous optimization problem of dimensionality kd. A candidate solution vector, x' can be represented by concatenating the d-dimensional coordinates of the cluster centers, i.e, $x^{'} = (c_1^1, \ldots, c_1^d, c_2^{1}`, \ldots, c_2^d, \ldots, c_k^{1}`, \ldots, c_k^d)$.

2.2 Sum of squares clustering as a prediction task

3 SOLVING THE CLUSTERING PROBLEM WITH NEURAL NETWORKS

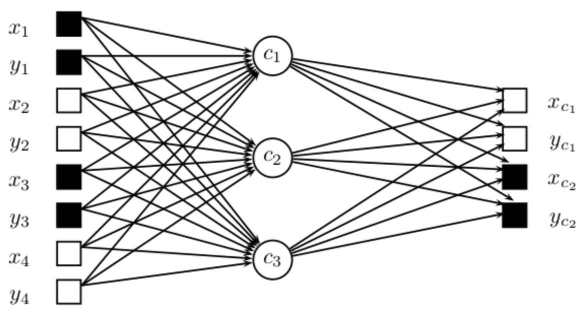

Figure 1: Example of a neural network for a clustering problem in 2-d, four points and $k = 2$.

4 RELATED WORK

4.1 Clustering problem characterization and instance generation

- Fitness landscape analysis in data-driven optimization: An investigation of clustering problems [8]
- Generating Diverse Clustering Datasets with Targeted Characteristics [5]
- Towards understanding clustering problems and algorithms: an instance space analysis [6]
- A new data characterization for selecting clustering algorithms using meta-learning [14]
- Fitness landscape analysis of circles in a square packing problems [13]

4.1.1 Circle packing. The SSE problem is also related to the circle packing problem, another optimization problem for which research on the evolutionary algorithms has been conducted [3, 13].

The circle packing problem consists of arranging N arbitrary sized circles inside a container in such a way that there are no overlapping between them. The problem is relevant for real world applications since a variety of practical problems (e.g., communication networks, dashboard layout problems, cylinder packing, etc.) can be approached as a circle packing problem [4].

4.2 Evolutionary approaches to clustering and packing problems

- Evolutionary Algorithms in Clustering: Challenging Problem Generation and Search Space Adaptation [18]
- Exploratory Analysis of Clustering Problems Using a Comparison of Particle Swarm Optimization and Differential Evolution [16].
- Beware the parameters: estimation of distribution algorithms applied to circles in a square packing [7]
- The importance of implementation details and parameter settings in black-box optimization: a case study on Gaussian estimation-of-distribution algorithms and circles-in-a-square packing problems [3]

An early use of a particular type of neural network for circle packing problem was presented in [15]. In the case of neural-network based evolutionary algorithm, the neural network is applied to learn the distribution of the selected solution, in an approach similar to the one used by estimation of distribution algorithms. This type of EDAs have been extensively investigated [1, 9, 12], however they have not been applied to the circle packing problem in constrast to traditional EDAs that have been evaluated for the circles-in-a-square packing problem [3].

4.3 Neural networks for clustering and optimization

- Combinatorial optimization with physics-inspired graph neural networks [17]
- Neural network methods in combinatorial optimization [11]
- Neural networks for combinatorial optimization: a review of more than a decade of research [19]
- Machine learning for combinatorial optimization: a methodological tour d'horizon [2]

In [10], the NeuroEvolution of Augmenting Topologies (NEAT) algorithm [20] is used to evolve ANNs that solve allocation and sequencing problems in a hybrid flow shop environment.

5 SUM OF SQUARES CLUSTERING AS A PREDICTION PROBLEM

6 INSTANCE DIFFICULTY FOR THE SUM OF SQUARES CLUSTERING

7 EXPERIMENTS

7.1 Experimental design

7.2 Benchmark

- Explanation of the probability distributions benchmark.

8 CONCLUSIONS

8.1 Further work

-

REFERENCES

[1] Shumeet Baluja. 2017. Learning deep models of optimization landscapes. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 1–10.

[2] Yoshua Bengio, Andrea Lodi, and Antoine Prouvost. 2021. Machine learning for combinatorial optimization: a methodological tour d'horizon. *European Journal of Operational Research* 290, 2 (2021), 405–421.

[3] Peter AN Bosman and Marcus Gallagher. 2018. The importance of implementation details and parameter settings in black-box optimization: a case study on

Gaussian estimation-of-distribution algorithms and circles-in-a-square packing problems. *Soft Computing* 22, 4 (2018), 1209–1223.

[4] Ignacio Castillo, Frank J Kampas, and János D Pintér. 2008. Solving circle packing problems by global optimization: numerical results and industrial applications. *European Journal of Operational Research* 191, 3 (2008), 786–802.

[5] Luiz Henrique dos Santos Fernandes, Kate Smith-Miles, and Ana Carolina Lorena. 2022. Generating Diverse Clustering Datasets with Targeted Characteristics. In *Brazilian Conference on Intelligent Systems*. Springer, 398–412.

[6] Luiz Henrique dos Santos Fernandes, Ana Carolina Lorena, and Kate Smith-Miles. 2021. Towards understanding clustering problems and algorithms: an instance space analysis. *Algorithms* 14, 3 (2021), 95.

[7] Marcus Gallagher. 2012. Beware the parameters: estimation of distribution algorithms applied to circles in a square packing. In *International Conference on Parallel Problem Solving from Nature*. Springer, 478–487.

[8] Marcus Gallagher. 2019. Fitness landscape analysis in data-driven optimization: An investigation of clustering problems. In *2019 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2308–2314.

[9] Unai Garciarena, Roberto Santana, and Alexander Mendiburu. 2018. Evolved GANs for generating Pareto set approximations. In *Proceedings of the 2018 on Genetic and Evolutionary Computation Conference*. ACM, 434–441. https://dl.acm.org/doi/abs/10.1145/3205455.3205550

[10] Sebastian Lang, Tobias Reggelin, Johann Schmidt, Marcel Müller, and Abdulrahman Nahhas. 2021. NeuroEvolution of augmenting topologies for solving a two-stage hybrid flow shop scheduling problem: A comparison of different solution strategies. *Expert Systems with Applications* 172 (2021), 114666.

[11] Chee-Kit Looi. 1992. Neural network methods in combinatorial optimization. *Computers & Operations Research* 19, 3-4 (1992), 191–208.

[12] L. Marti, J. García, A. Berlanga, and J. M. Molina. 2008. Introducing MONEDA: scalable multiobjective optimization with a neural estimation of distribution algorithm. In *Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation GECCO-2008*. ACM, New York, NY, USA, 689–696. DOI: http://dx.doi.org/10.1145/1389095.1389228

[13] Rachael Morgan and Marcus Gallagher. 2014. Fitness landscape analysis of circles in a square packing problems. In *Asia-Pacific Conference on Simulated Evolution and Learning*. Springer, 455–466.

[14] Bruno Almeida Pimentel and Andre CPLF De Carvalho. 2019. A new data characterization for selecting clustering algorithms using meta-learning. *Information Sciences* 477 (2019), 203–219.

[15] Jiri Pospichal. 2016. Solving Circle Packing Problem by Neural Gas. In *International Conference on Soft Computing-MENDEL*. Springer, 281–289.

[16] Sobia Saleem and Marcus Gallagher. 2017. Exploratory Analysis of Clustering Problems Using a Comparison of Particle Swarm Optimization and Differential Evolution. In *Australasian Conference on Artificial Life and Computational Intelligence*. Springer, 314–325.

[17] Martin JA Schuetz, J Kyle Brubaker, and Helmut G Katzgraber. 2022. Combinatorial optimization with physics-inspired graph neural networks. *Nature Machine Intelligence* 4, 4 (2022), 367–377.

[18] Cameron William Benjamin Shand. 2020. *Evolutionary Algorithms in Clustering: Challenging Problem Generation and Search Space Adaptation*. Ph.D. Dissertation. The University of Manchester (United Kingdom).

[19] Kate A Smith. 1999. Neural networks for combinatorial optimization: a review of more than a decade of research. *INFORMS Journal on Computing* 11, 1 (1999), 15–34.

[20] Kenneth O Stanley and Risto Miikkulainen. 2002. Evolving neural networks through augmenting topologies. *Evolutionary computation* 10, 2 (2002), 99–127.