Report for Project 2

Negatively Correlated Search

and OLMP Problems'

Hyperparameter Adjustment Algorithm

description (40%)

Main Idea of NCS（10%）	In the later stage, the traditional population search algorithm will shrink the scope of the search, resulting in the values in the range being repeatedly calculated, and the variation cannot be obtained well. The evaluation function used by the negative correlation search algorithm will consider the value and The distance between other values and the degree of goodness of the acquired solution, and sharing information among individuals, encouraging individuals to explore the location outside the existing exploration range of the population, thereby expanding the effective search range of the population and promoting the diversity of the population, and thus Improve search efficiency and obtain more effective search results than traditional algorithms..
Applications of NCS（10%） (i.e. what problems NCS can solve)	Almost all genetic algorithms, simulated annealing algorithms, etc. in a fixed solution space, using random methods to iteratively obtain the optimal value can be optimized using the NCS algorithm. Specifically, for example, to obtain a complex multi-peak (non-convex) continuous Function optimization maximum and minimum, sensor model optimization, mathematical modeling optimization, scheduling problems, fast positioning, path planning, adjusting neural network parameters, combinatorial optimization, automatic control, image processing, artificial life, data mining, game experience optimization, mode Identification, adaptive control, etc.
Main Idea of OLMP （10%）	Deep neural networks are very easy to use, but the models are too complex to be migrated to lower-profile devices. To solve this problem, some methods are needed to cut off the connection between each layer of the neural network, thus reducing the volume of the model. Compared with the runtime consumption, but the method of judging becomes a problem when the connection between each layer is removed, the MP algorithm uses the method of retraining the model after cutting off the smaller amplitude connection, and the LMP algorithm further The MP algorithm is used to simplify the model between layers, but these are too expensive. The OLMP algorithm uses the non-derivative optimization method (specifically, the NCS algorithm) to perform parameters based on the two algorithms. Adjustment, using a random sample set

	to evaluate, so that the parameters are automatically adjusted.
Applications of OLMP（10） (i.e. what problems OLMP can solve)	Automatic adjustment of deep neural network parameters, model optimization, thus facilitating the operation of DNN results on low-profile equipment. Automating complex manual adjustment parameters, ensuring automation of the process, turning the problem of tuning parameters into optimizable problems and The NCS algorithm is applied to this problem. The fields of face detection, image recognition, pattern migration, graphic transformation and so on are helpful for the simplification of the model and the improvement of the model's interpretability.

1. Parameter description(40%)

Parameters	Your final values and results（10per cell）			Summary 1. What the role of the parameter 2. The effect about different values of the parameter on the final performance The best range for the parameter
	F6	F12	OLMP	
lambda	0.9877166748046875	0.9953641891479492	1.1145141919810	Lambda: The meaning of this parameter is to adjust the average of the normal distribution of the value of xi' that can be updated each time the judgment is generated. The best range is about 1±0.1 in F6 and about 1±0.1 in F12,in OLMP has no obvious meaning, down to 0.0000000000000000000000000000000000000000000000000000000000000000000001, up to 9.99999999999999999999999999999999999999999999999999999999999999999999, the same value can be obtained if other parameters are unchanged.
r	0.70368092126464843	0.88295650482177733	1.1919810114514	r: The meaning of this parameter is the adjustment range at each adjustment of the variance of the normal distribution that generates the random disturbance. If the improvement rate of each acquisition is high, the amplitude can be increased (expressed by dividing by rr less than 1, then increasing Large) (The reason is that

				the adjustment rate we want to achieve is a stable value, so expanding the range helps to reduce the adjustment rate, which in turn controls the adjustment rate to around 20%), and the lower improvement rate reduces the magnitude (also hopes that The regulation rate is controlled at around 20%).H.-G. Beyer and H.-P. Schwefel, "Evolution strategies—A comprehensive introduction," Nat. Comput., vol. 1, no. 1, pp. 3–52, 2002.As mentioned in this paper, the best range is probably [0.85, 1); in practice, I've probably realized 0.9±0.1 is better; among the better solutions, the r of F6 is about 0.70±0.02, and the F12 is about 0.90±0.05, the r parameter in OLMP does not show obvious significance, down to 0.00000000000000000000000000000000000000000000000000000000000001, up to 1145141919810(without other parameters change) can get the same result.
epoch	18	26	1	Epoch: The meaning of this parameter is the number of single iteration trainings performed in each period of the iterative process. In the algorithm, the main reason is to increase or decrease the parameters of the search range by judging the improvement rate. After trying, In F6 and F12, this parameter is more reasonable in the range of 0-30. By calculation, the obtained solution has F6 of 18 and F12 of 26 is the most suitable. In OLMP, epoch is preferably kept at a small value, and the range is about Trying between 0-30, 1 can also get the optimal solution.
n	3	3	92	N: The meaning of this parameter is the number of population individuals used in the initial and each iteration of the population. In the two problems of F6 and F12, the optimal range of N is about 3-5, and iteratively tries through genetic algorithm. The result is 3 is the optimal value, although the population is small but the number of iterations is more. In the problem of OLMP, N obtains the optimal value at 92.
Final Result	390.006561998098556	-459.7287763845455 4	0.9889914106747684	

Running Time	100.2510380744934 1 s	47.31123304367065 4 s	62.87s	

2. Tuning procedure(15%)

F6 && F12:

I use three steps to adjust the parameters, Random produce, Genetic algorithm and manual adjustment.

In the First step, I use absolutely random way to get all the parameters, in this step, for every I get 1000 examples.

In the second step, I use Genetic algorithm to produce better parameters. In the begging,

Use the random parameters to get a better basic data. Then run an infinite cycle and use genetic algorithm to produce better solutions.

Finally, because I use several threads to run the codes at the same time, I select the better solutions to be the basic values to repeat the genetic algorithm and learn the better range of parameters to adjust the Hyperparameter in my own's Genetic algorithm.

OLMP:

First of all, the environment is not difficult to build. I used the 10 minutes to fully understand the construction process and successfully set up.

I use three steps to adjust the parameters, Random produce, Genetic algorithm and manual adjustment.

In the First step, I use absolutely random way to get all the parameters, In this step, for every I get 1000 examples (because of I have an Nvidia GTX1080Ti,so it is not so hard).

In the second step, I use Genetic algorithm to produce better parameters. In the begging,

Use the random parameters to get a better basic data. Then run an infinite cycle and use genetic algorithm to produce better solutions.

Then, because I use several threads to run the codes at the same time, I select the better solutions to be the basic values to repeat the genetic algorithm and learn the better range of parameters to adjust the Hyperparameter in my own's Genetic algorithm.

Finally, In the case where the epoch is small and n is 92, no matter what value the lambda and r take, the result does not change at all. In other words, it does not show any meaning at all. The epoch clearly shows that the result is better in a smaller case. Verify that the minimum value is obtained when n=92