

# Investigation on the Performance of Surrogate-assisted Genetic Algorithms According to Changes in Learning Model and Data

## 기계학습 모델과 데이터의 변화에 따른 대리 모델의 성능 관찰

Seo-Ah Yu<sup>1</sup>, Yong-Hyuk Kim<sup>2</sup>

유서아<sup>1</sup>, 김용혁<sup>2</sup>

<sup>1</sup> Master Student, Dept. Comp. Sci, KwangWoon University, Korea, 9380187@naver.com

<sup>2</sup> Professor, Dept. Comp. Sci, KwangWoon University, Korea, yhdgly@kw.ac.kr

Corresponding author: Yong-Hyuk Kim

**Abstract:** Optimization problems are used in various fields, but it is difficult to use optimization techniques such as genetic algorithms when calculating fitness takes a very long time. In this context, many studies have been conducted to reduce the computation time of black box optimization problems by predicting the quality of solutions by creating surrogate models with machine learning algorithms such as gaussian process and support vector regression. In this study, we compared the difference in approximate performance of machine learning according to the type of learning model, the amount of training data, and the sampling method of training data in binary, permutational, real, and periodic-function encodings. Therefore, the simpler the problem is, the better the performance of the gaussian process is, also, the larger the number of training data is, the better the approximation performance is, The results showed that unbiased training through sampling is advantageous for high performance. Based on these results, experiments were performed by replacing the fitness function of the genetic algorithm with a machine-learning model on binary, permutation, and real number encoding. When the machine learning approximation was well done, on average, the performance of genetic algorithms with surrogate models was slightly lower or similar to that of genetic algorithms without surrogate models. In rare cases, it showed better performance, such as finding an optimal solution within about 1.6 times fewer generations, or finding an optimal solution with a higher fit than the genetic algorithm that does not use a surrogate model.

**Keywords:** Optimization, Genetic Algorithm, Sampling Method, Surrogate Model

**요약:** 최적화 문제는 실세계의 다양한 분야에서 사용되지만, 적합도를 계산하는 시간이 매우 오래 걸리는 경우 유전 알고리즘과 같은 최적화 기법을 사용하기 어렵다. 이러한 맥락에서, 가우시안 프로세스, 서포트 벡터 회귀와 같은 기계학습 알고리즘으로 대리 모델(surrogate model)을 만들어 해의 품질을 예측해 블랙박스 최적화 문제의 계산 시간을 줄이기 위한 많은 연구가 수행되어 왔다. 본 연구에서는 이진, 순열, 실수, 주기함수 인코딩 상황에서 학습 모델의 종류, 학습 데이터의 양, 학습 데이터의 샘플링 방법에 따른 기계학습의 근사성능 차이를 비교하였다. 그리하여 단순한 문제일수록 가우시안 프로세스의 성능이 우수하고, 학습

Received: February 05, 2023; 1<sup>st</sup> Review Result: March 21, 2023; Accepted: May 31, 2023

데이터의 수가 많을수록 근사성능이 좋았으며, 샘플링을 통한 편향되지 않은 학습이 높은 성능을 내기에 유리하다는 결과를 보였다. 이 결과를 바탕으로 이진, 순열, 실수 인코딩에서 한가지씩 뽑아 유전 알고리즘의 적합도 함수를 기계학습 모델로 대체하여 실험을 수행하였고, 기계학습 근사가 잘 이루어졌을 때 성능이 평균적으로 대리 모델 유전 알고리즘이 대리 모델을 사용하지 않은 유전 알고리즘보다 약간 낮거나 비슷했다. 드물게는 약 1.6배 적은 세대 안에 최적해를 찾아내거나, 대리 모델이 아닌 유전 알고리즘보다 더 적합도가 높은 최적해를 찾는 등 더 나은 성능을 보이기도 하였다.

**핵심어:** 최적화, 유전 알고리즘, 샘플링 기법, 대리 모델

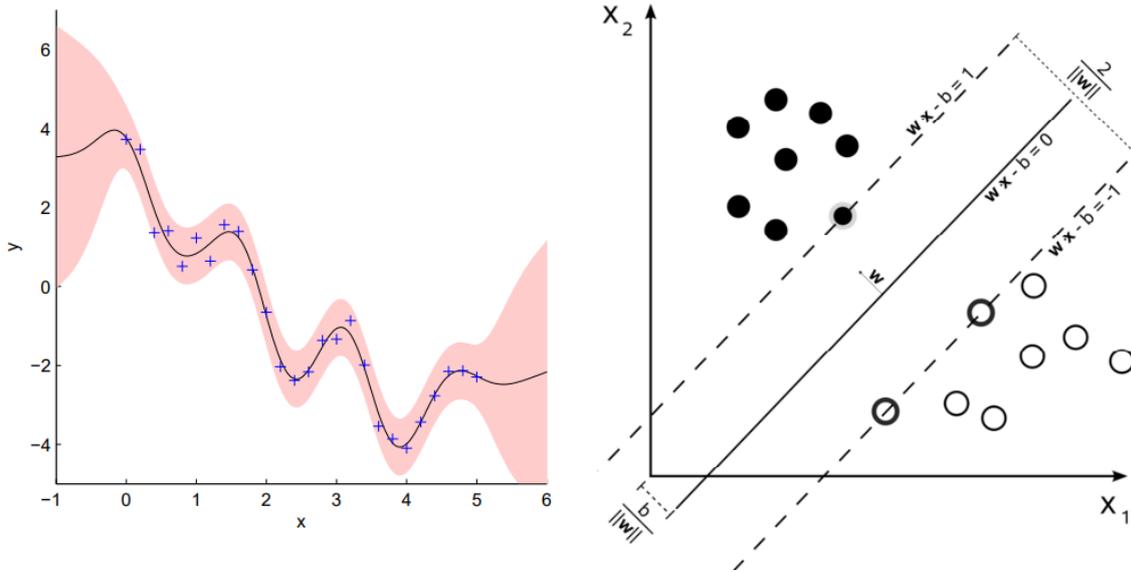
## 1. 서론

최적화 문제는 물류, 공학 설계, 에너지 생성과 같은 많은 응용 분야에서 사용되고 있다. 이러한 문제는 일반적으로 NP-난해(NP-hard)이며 해결하기 어렵다. 유전 알고리즘과 같은 진화 알고리즘은 적자생존으로 자연적 과정을 모방한다. 초기 모집단은 무작위로 생성되며, 교배를 위해 가장 적합한 개체가 선택된다. 후손은 교차 및 돌연변이 연산에 의해 생성된다. 하지만 유전 알고리즘은 최적의 근사치를 찾으려면 많은 수의 기능 평가를 수행해야 한다는 한계가 있다. 따라서 유전 알고리즘은 계산비용이 많이 드는 최적화 문제를 해결하는 데 적합하지 않을 수 있다[1]. 최근 최적화 과정에서 유전 알고리즘의 적합도 함수 부분을 기계학습 모델로 대체한 대리 모델(surrogate model)을 사용하는 방법에 관한 많은 연구가 진행되었다[2-5]. 하지만 대리 모델의 근사성능이 낮으면 유전 알고리즘이 좋은 해를 찾기 어렵다. 따라서 우리는 모델의 종류, 데이터의 양, 샘플링 방법을 변환하며 비교하여 최종적으로 대리 모델의 성능을 높이고자 하였다.

본 논문은 기계학습 모델의 근사성능을 높여 궁극적으로 유전 알고리즘 대리 모델이 어떤 성능을 보이는지 확인하는 것에 관심이 있다. 따라서 각 조건에서 평균 제곱근 오차(root mean square error, RMSE)의 평균이 작은 조건으로 기계학습 모델을 학습시킨다. 우선 적은 수의 학습 데이터를 이용하여 서포트 벡터 회귀와 가우시안 프로세스의 성능을 비교하여 이후 연구에 어떤 모델로 학습할지 결정한다. 결정된 모델로 학습 데이터의 양이 적을 때와 많을 때의 성능 차이를 비교한다. 마지막으로 학습 데이터를 생성할 때 이진 공간의 경우 1의 개수가 고르게 분포하도록 설정하고, 순열 공간에서는 역순의 개수가 고르게 나오도록 학습 데이터를 제작하여 샘플링의 품질을 높이고자 한다. 본 논문의 구성은 다음과 같다. 2절에서는 사용된 근사 모델의 종류를 소개한다. 3절에서는 이용한 각 인코딩의 함수에 대하여 설명한다. 4절에서는 다양한 조건에 따른 근사성능을 보여주고 3개의 공간의 문제를 선별하여 대리 모델 실험을 진행하고, 전통 유전 알고리즘의 성능과 비교한다. 5절에서는 결론을 내린다.

## 2. 근사 모델

본 실험에 사용한 기계학습 모델은 대표적으로 널리 사용되는 서포트 벡터 회귀(support vector regression, SVR)와 결과가 확률 분포를 가져 유연성이 있는 가우시안 프로세스(gaussian process, GP)이다.



(a) Gaussian process example plot[6].

(b) Support vector regression example plot[7].

[그림 1] 가우시안 프로세스와 서포트 벡터 회귀의 그래프 예시

[Fig. 1] Example Plots of Gaussian Process and Support Vector Regression

## 2.1 가우시안 프로세스

GP는 지도학습의 일종으로, 매개변수가 없는 회귀 문제에 대한 베이저안 접근이다. GP는 이론적으로 무한한 수의 매개변수를 활용하고 베이저안 추론을 통해 데이터가 복잡성 수준을 결정하도록 하여 입력과 출력 간의 다양한 관계를 포착할 수 있다. 하지만 입력공간이 커지면 계산하기가 어렵고, 시간이 오래 걸린다는 단점이 있다[8]. 하지만 가우시안 프로세스는 [그림 1](a)와 같이 각 점에서 가우시안 분포를 가지게 되어, 특정 데이터 하나가 아닌 확률 분포를 가지게 된다는 점에서 다른 모델들보다 유연성을 갖는 이점이 있다.

## 2.2 서포트 벡터 회귀

SVR의 손실함수는 회귀계수의 크기를 작게 하여 회귀 식을 평평하게 만들되, 오차가 작아지도록 하는 선을 찾는 것을 목표로 한다[9]. SVR은 [그림 1](b)와 같이 회귀 식이 추정되면 회귀 식에 대칭적으로  $(-\epsilon, \epsilon)$ 만큼 튜브를 생성하고 특정 임계 값 미만의 오차 절대값은 튜브 밖에서 모두 무시된다. 이러한 방식으로 튜브 외부의 값은 벌점을 받지만, 튜브 내부 지점은 벌점을 받지 않는다. SVR의 이점으로는 계산 복잡성이 입력공간의 차원에 의존하지 않는다는 것이다. 또한, 예측 정확도가 높고 일반화 기능이 뛰어나다[10].

## 3. 테스트 문제

One-max는 일련의 비트(0 또는 1)에서 1의 수를 최대화하는 문제이다. 길이  $n$ 일 경우

최적해는  $n$  개의 1이 있는 것이다. 이는 가장 간단한 최적화 문제이다[11]. 유전 알고리즘의 진화 과정에서 스키마가 어떻게 처리되는지 알아보기 위해 왕도(royal-road) 함수를 사용하였고 스키마의 길이는 4로 설정하였다[12]. NK-지형(NK-landscape) 문제는 다양한 차원과 상위(epistasis)로 적합도 함수를 정의하기 위해 구성되었다. 적합도 함수는 두 개의 매개변수  $n$ 과  $k$ 에 의해 조정된다. 여기서  $n$ 은 문제 공간의 차원(dimension,  $d$ )을 정의하고  $k$ 는 염색체를 구성하는 유전자 간의 상위 정도를 결정한다.  $k$ 를 0부터  $n - 1$ 까지 변화시키면서 NK-지형은 부드럽거나 거칠게 변할 수 있다[13]. 실험에서  $k$ 값은 2, 12로 설정하였다. 이후  $n$ 은  $d$ 로 통일된다. 순열 문제로 난수 순열이 오름차순으로 정렬되기 위해 교환이 일어나는 최대 횟수를 반환하는 IversionMax 함수를 고안하여 사용하였다. 실수 공간 문제로 Sphere, Ackley, Schwefel 함수를 사용하였고, 고차원 실수 문제로 Rosen-Brock, Rastrigin 함수를 사용하였다. 이 함수들은 CEC2013에서 제공하는 실수를 매개변수로 가진 최적화 함수들이다[14]. 각각의 함수는 식 (1), (2), (3), (4), (5)와 같다.

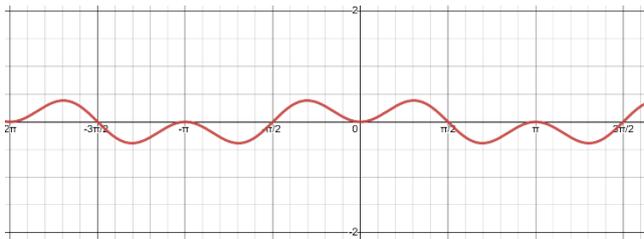
$$\text{Sphere}(x) = \sum_{i=1}^d x_i^2 \quad (1)$$

$$\text{Schwefel}(x) = 418.98 - \sum_{i=1}^d x_i \sin(\sqrt{|x_i|}) \quad (2)$$

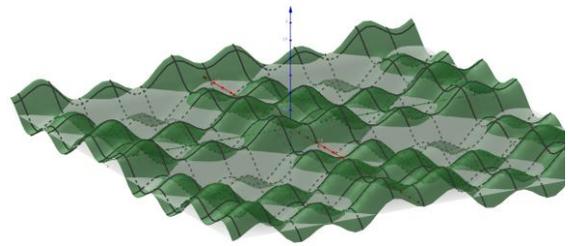
$$\text{Ackley}(x) = -20 \cdot \exp\left(-0.2 \sqrt{\frac{1}{d} \sum_{i=1}^d x_i^2}\right) - \exp\left(\frac{1}{d} \sum_{i=1}^d \cos(2\pi x_i)\right) + 20 + e \quad (3)$$

$$\text{Rastrigin}(x) = 10d + \sum_{i=1}^d [x_i^2 - 10 \cos(2\pi x_i)] \quad (4)$$

$$\text{RosenBrock}(x) = \sum_{i=1}^{d-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2] \quad (5)$$



(a) Periodic function (6) when  $d=2$



(b) Periodic function (6) when  $d=3$

[그림 2] 주기함수 (6)의 그래프

[Fig 2] Plots of Periodic Function (6)

경계가 없는 실수 공간을 가정하기 위하여 주기  $2\pi$ 를 가지는 함수를 고안하였다. 이는 식 (6)과 같고 [그림 2]의 형태를 가지고 있다.

$$f(x) = \sum_{i=1}^d (\sin^2(x_i) \cdot \cos(x_i)) \quad (6)$$

이산 공간 인코딩은 최대화 문제이고, 실수 공간 인코딩은 최소화 문제이다. 모델의 RMSE 또는 이를 정규화 한 RMSE를 차원으로 나눈 것(RMSE/ $d$ )을 사용하며, RMSE의

정의는 식 (7)과 같다.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum(\hat{y}-y)^2}{n}} \quad (\hat{y}=\text{예측 값}, y=\text{실제 값}, \bar{y}=\text{예측 값의 평균}, n=\text{샘플의 개수}) \quad (7)$$

#### 4. 실험과 분석

[표 1] SVR과 GP, 학습 데이터 1,000개와 10,000개의 RMSE 비교

[Table 1] Comparison between SVR and GP and between 1,000 and 10,000 as the number of Train samples

Test problem		$d$	SVR with 1,000 training samples	GP with 1,000 training samples	GP with 10,000 training samples
Binary encoding	One-max	8	5.04e-02	2.36e-11	5.45E-12
		50	1.15e-01	1.42e-06	6.65E-10
		100	3.25e-01	1.03e-05	2.60E-08
	Royal-road, Schema = 4	8	2.50e-01	2.40e-04	2.18e-11
		40	6.61e-01	6.23e-01	3.80e-01
		96	1.10e+00	1.08e+00	8.69e-01
	NK-landscape, $k=2, n=d$	12	4.80e-02	9.57e-06	7.57e-09
		50	3.86e-02	1.94e-02	5.34e-03
		100	2.51e-02	1.83e-02	6.82e-03
	NK-landscape, $k=12, n=d$	12	8.48e-02	8.40e-02	2.23e-01
		50	4.12e-02	4.12e-02	4.12e-02
		100	2.84e-02	2.78e-02	2.87e-02
Permutation encoding	InversionMax	5	8.80e-01	6.87e-01	5.58e-01
		30	6.18e+00	6.10e+00	5.90e+00
		50	1.03e+01	9.84e+00	9.89e+00
Real encoding	Sphere	5	2.50e-01	1.05e-03	2.88e+00
		30	2.49e+01	3.08e-02	3.23e+01
		50	5.16e+01	5.88e+01	5.56e+01
	Schwefel	5	4.42e+02	4.32e+02	4.20e+02
		30	9.46e+02	9.26e+02	1.01e+03
		50	1.24e+03	1.20e+03	1.32e+03
	Ackley	5	1.12e+00	1.12e+00	1.19e+00
		30	2.32e+01	2.33e+01	2.02e+01
		50	9.09e+01	9.10e+01	1.68e+02
	Rastrigin	5	2.21e+01	2.27e+01	1.99e+01
		30	6.19e+01	6.19e+01	5.59e+01
		50	7.01e+01	7.02e+01	7.44e+01
	Rosen-Brock	5	9.05e+02	6.04e+02	3.49e+02
		30	4.24e+03	3.82e+03	2.72e+03
		50	5.09e+03	4.92e+03	4.82e+03
Periodic encoding	Periodic Function	5	4.84e-01	4.25e-01	1.29e-01
		30	1.44e+00	1.44e+00	1.24e+00
		50	1.81e+00	1.81e+00	1.76e+00
<b>Average</b>			<b>4.01E+02</b>	<b>3.71E+02</b>	<b>3.36E+02</b>

### 4.1 학습 모델에 따른 근사성능 비교

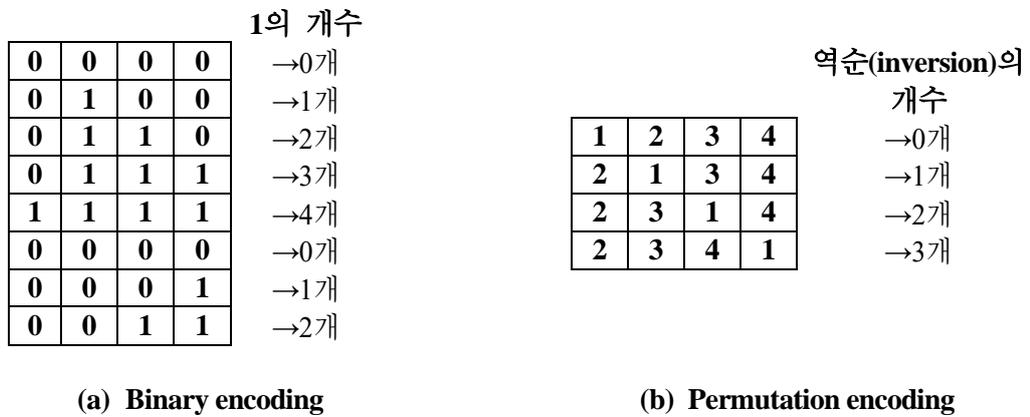
[표 1]의 데이터 1,000개로 학습한 SVR과 GP의 RMSE의 평균으로 GP와 SVR의 성능을 비교하였을 때, GP가 더 작은 오차를 내었다. 두 모델의 오차의 간극은 단순한 문제일수록, 차원이 낮을수록 컸고, 차원이 높고 복잡한 문제에서는 GP의 성능이 급격히 떨어지며 큰 차이를 보이지 못했다. GP는 고차원에 적합하지 않다는 2절에서의 설명을 참고하여 각 문제에 적합한 모델을 사용하는 것이 좋다. 본 논문에서는 이후 실험에 GP를 사용하였다.

### 4.2 학습 데이터양에 따른 근사성능 비교

[표 1]의 데이터 1,000개와 10,000개로 학습한 GP의 RMSE의 평균으로 보았을 때, 학습 데이터의 수가 많을수록 오차가 적었다. 드물게는 NK-지형,  $k=12$ 의 경우와, 실수 공간 문제에서 오히려 10,000개 학습 데이터의 오차가 더 크게 나타나기도 하였다. 하지만 1,000개와 10,000개의 오차 차이가 작아, 학습 데이터의 수는 근사성능 향상에 큰 영향을 미치지 못하는 것으로 보인다. 본 실험에는 RMSE의 평균 결과를 따라 10,000개의 학습 데이터를 사용하였다.

### 4.3 학습 데이터 샘플링 방식에 따른 근사성능 비교

#### 4.3.1 이산 공간



[그림 3] 이산 공간에서의 고른 분포 생성 방법

[Fig. 3] How to Generate Uniform Distribution in Discrete Space

#### 4.3.1.1 이진 인코딩에서의 학습 데이터 생성 방법

이진 인코딩에서는 훈련/테스트 데이터를 무작위로 0, 1로 생성하여 “R”(Random)으로 지칭하였다. 반대로 [그림 3](a)와 같이 훈련/테스트 데이터를 1의 개수 분포가 고르게 나오도록 생성한 것을 “U”(Uniform)으로 명명하였다.

#### 4.3.1.2 순열 인코딩에서의 학습 데이터 생성 방법

순열 인코딩에서는 전체를 무작위 하게 난수를 생성하여 만든 데이터 세트를 “R”로

지칭하고, “R”세트 절반과 [그림 3](b)와 같이 오름차순으로 정렬하기 위해 교환이 이루어지는 횟수(역순(inversion)의 개수)가 고르게 나오도록 인위적으로 만든 세트 절반으로 이루어진 데이터 세트를 “U”로 명명하였다.

#### 4.3.1.3 이산 공간 샘플링 별 기계학습 근사 결과

[표 2] 이산 공간에서 샘플링 별 정규화 된 RMSE/d

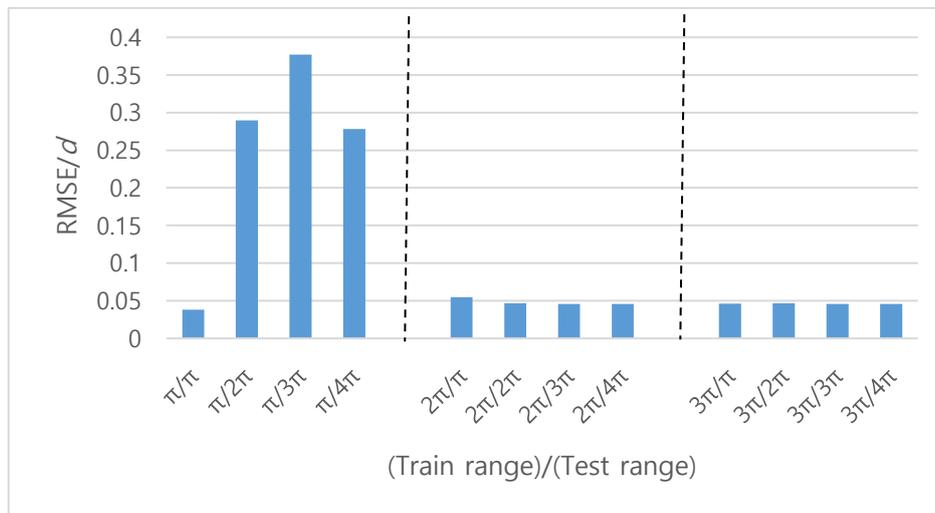
[Table 2] Normalized RMSE/d by Sampling in Discrete Space

	One-max	NK-landscape, $k=2$	NK-landscape, $k=12$	Royal-road Schema=4	InversionMax	Average
Train on “R”	8.91e-02	1.88e+00	1.59e+00	7.73e-01	1.25e+00	<b>1.12e+00</b>
Train on “U”	1.91e+00	1.17e-01	4.10e-01	1.23e+00	7.54e-01	<b>8.84e-01</b>

오차가 큰 함수의 경우, 오차 평균에 큰 영향을 미치고, 오차가 작은 함수의 경우는 오차 평균에 적은 영향을 미치게 될 것이므로, 단순히 RMSE/d 로만 평균을 구하는 것이 아니라, 각 함수별로 전체 샘플링의 평균을 구하고 그 평균을 각 샘플링에다 나누어 주어 정규화를 진행하였다. [표 2]의 평균을 확인하면, 이산 공간에서 인위적으로 만든 데이터 세트로 학습시킨 모델의 근사성능은 무작위 하게 생성한 데이터 세트로 학습한 모델보다 조금 더 좋은 성능을 보였다. 이는 편향되지 않은 고른 분포의 학습 데이터세트가 근사성능을 향상시킬 수 있다는 가능성을 보여준다.

#### 4.3.2 경계가 없는 실수 공간

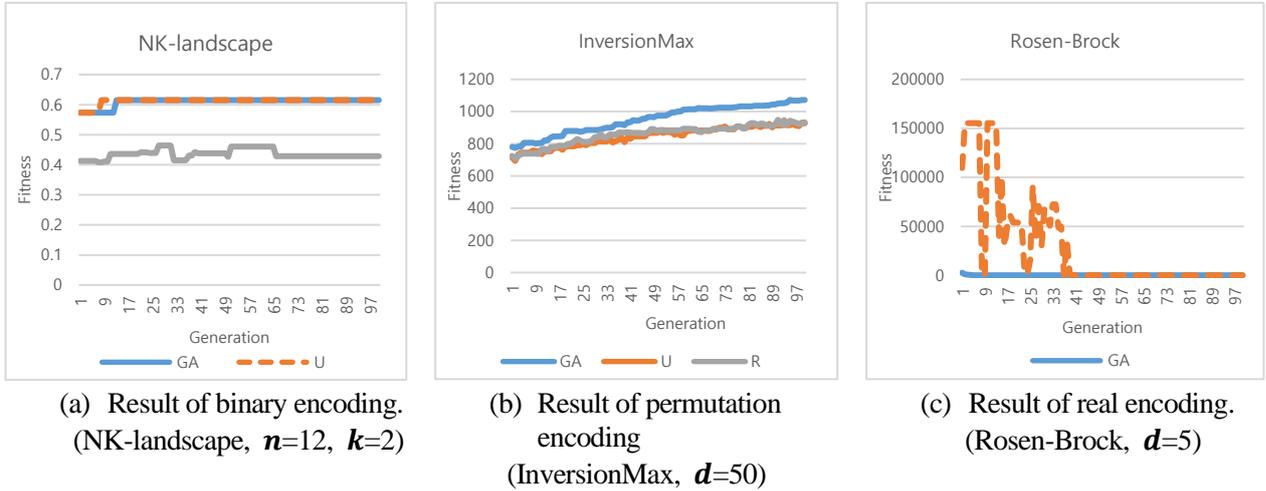
[그림 4]를 보면, 1주기인  $[-\pi, \pi]$ (이하  $\pi$ ) 범위로 학습하고, 더 큰 범위를 테스트하는 경우 높은 오차율을 보였다. 하지만, 2주기인  $[-2\pi, 2\pi]$ (이하  $2\pi$ ) 범위로 학습시킨 모델은 2주기보다 더 큰 범위의 테스트를 진행하여도 오차율의 큰 차이를 보이지 않았으며, 1주기로 학습하였을 때 보다 약 2배의 좋은 근사성능을 내었다. 단, 3주기인  $[-3\pi, 3\pi]$ (이하  $3\pi$ )로 학습한 것이 1주기와 2주기의 성능 차이만큼 월등한 차이를 보이지는 않았다. 이는 1주기에서는 함수의 주기성을 학습하지 못했지만, 2주기 이상에서는 함수의 주기성을 학습하여 더 넓은 범위도 예측할 수 있음을 시사한다.



ex)  $\pi = [-\pi, \pi]$ ,  $2\pi = [-2\pi, 2\pi]$

[그림 4] 훈련/테스트 데이터의 범위 별 주기함수의 근사 결과  
 [Fig. 4] Approximate Results of Periodic Function by Train/Test Range

4.4 일반 유전 알고리즘과 대리 모델을 사용한 유전 알고리즘의 성능 차이



[그림 5] 대리 모델을 사용한 유전 알고리즘과 그렇지 않은 유전 알고리즘의 세대에 따른 적합도 값

[Fig. 5] Fitness Values according to Generations of Genetic Algorithms using Surrogate Models and those without Surrogate Model

[그림 5](a)를 보면 고르게 생성한 데이터 세트에 학습한 모델로 적합도 함수를 대체한 대리 모델 유전 알고리즘의 경우, 무작위 하게 생성한 데이터 세트에 학습한 모델로 적합도 함수를 대체한 대리 모델 유전 알고리즘보다 최적해를 찾는 능력이 높음을 볼 수 있다. 하지만 [그림 5](b)의 InversionMax 함수의 경우 고른 학습 세트와 무작위 학습 세트의 성능이 비슷하고, 일반 유전 알고리즘의 성능이 가장 좋은 것으로 나타나는데, 이는 NK-지형의 경우 고른 학습 세트 근사성능이 무작위 학습 세트보다 월등히 높았고, InversionMax의 경우에는 근사성능이 비슷하였기 때문에, 이 영향이 대리 모델 유전 알고리즘의 성능에도 영향을 미친 것으로 보인다. 실수 공간 문제인 Rosen-Brock 문제 [그림 5](c)의 경우 세대 초반에는 불안정한 모습을 보이지만, 결국 최적해를 찾아냄으로써 기계학습으로 학습한 모델로 유전 알고리즘의 적합도를 대체할 수 있을 것으로 보인다.

5. 결론

본 연구는 유전 알고리즘의 적합도 함수 부분을 기계학습으로 학습한 모델의 대체 가능성을 살펴보았다. 이진 인코딩에서 4가지, 순열 인코딩에서 1가지, 실수 인코딩에서 5가지, 주기함수 1가지를 이용해 기계학습의 근사성능에 미치는 요소들을 확인했다. 그 결과 GP와 SVR을 비교했을 때, 특히 저차원과 쉬운 문제에서 GP가 월등히 좋은 성능을

보였고, 학습 데이터의 수는 큰 영향을 미치지 않았다. 또한, 편향되지 않은 데이터로 학습시키는 것이 기계학습 근사성능 향상에 긍정적인 영향을 주었다. 이진, 순열, 실수 인코딩에서 문제 한 개씩을 선정하여 유전 알고리즘의 최적해 부분을 기계학습 근사 모델로 대체한 대리 모델로 실험을 진행한 결과, 높은 근사성능을 보인 대리 모델일수록 최적해를 찾는 능력도 높았다. 근사성능이 높은 경우 대리 모델을 사용하는 것이 충분히 경쟁력 있는 방법이다. 본 연구에서는 여러 도메인에서 대리 모델의 성능향상에 영향을 미치는 요소에 대하여 확인하였고, 대리 모델을 사용하는 경우도 올바른 최적 해로 이끌 수 있음을 실험적으로 확인했다. 이 외에도 향후 실 세계 문제를 대상으로 실험을 진행해 볼 필요가 있고, 랜덤 포레스트, 신경망 등의 다양한 학습 모델을 사용하여 비교해 볼 수 있다. 또한, 단순히 결과의 분포를 고르게 하는 샘플링 방법 외에, 어떤 함수에도 적용할 수 있는 범용적인 샘플링 방법을 고안한다면 문제마다 그에 적합한 샘플링 방법을 생성해야 하는 번거로움을 줄 일 수 있다. GP의 단점인 오래 걸리는 시간과, 고차원에 취약한 점을 보완하기 위해 차원 축소 및 효율적 인코딩 방법에 대하여 연구할 필요가 있다.

## 6. 감사의 글

이 논문은 2023년도 광운대학교 우수연구자 지원 사업에 의해 연구되었음. 또한, 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2021R1F1A1048466).

## References

- [1] Juliane Muller, Cristien A Shoemaker, Robert Piche, SO-MI: a surrogate model algorithm for computationally expensive nonlinear mixed-integer black-box global optimization problems, *Computers & Operations Research*, (2013), Vol.40, No.5, pp.1383-1400.  
DOI: <http://dx.doi.org/10.1016/j.cor.2012.08.022>
- [2] Alison Cozad, Nikolaos V. Sahinidis, David C. Miller, Learning surrogate models for simulation-based optimization, *AIChE Journal*, (2014), Vol.60, No.6, pp.1956-2372.  
DOI: <https://doi.org/10.1002/aic.14418>
- [3] Wencan Zhang, Zhicheng Liang, Weixiong Wu, Guozhi Ling, Ruixin Ma, Design and optimization of a hybrid battery thermal management system for electric vehicle based on surrogate model, *International Journal of Heat and Mass Transfer*, (2021), Vol.174, 121318.  
DOI: <https://doi.org/10.1016/j.ijheatmasstransfer.2021.121318>
- [4] Saman Razavi, Bryan A. Tolson, Donald H. Burn, Review of surrogate modeling in water resources, *Water resources research*, (2012), Vol.48, No.7.  
DOI: <https://doi.org/10.1029/2011WR011527>
- [5] Bo Liu, Hao Yang, Michael J. Lancaster, Global optimization of microwave filters based on a surrogate model-assisted evolutionary algorithm, *Transactions on Microwave Theory and Techniques*, (2017), Vol.65, No.6, pp.1976-1985.  
DOI: <https://doi.org/10.1109/TMTT.2017.2661739>
- [6] Mark Ebdon, Gaussian processes : a quick introduction, *ArXiv*, (2015)  
DOI: <https://doi.org/10.48550/arXiv.1505.02965>
- [7] Sunflower Soft Company, Tutorial on Support vector machine, Some novel algorithms for global optimization and relevant subjects, (2017), Vol.6, No.4-1, pp.1-15.

DOI: 10.11648/j.acm.s.2017060401.11

- [8] Eric Schulz, Maarten Speekenbrink, Andreas Krause, A tutorial on gaussian process regression: modelling, exploring, and exploiting functions, *Journal of Mathematical Psychology*, (2018), Vol.85, pp.1-16.  
DOI: <https://doi.org/10.1016/j.jmp.2018.03.001>
- [9] <https://leejiyoon52.github.io/Support-Vector-Regression/>, Nov 20 (2018)
- [10] Mariette Awad and Rahul Khanna, *Efficient Learning Machines*, Apress, pp.67-80, (2015)  
Available from: [https://link.springer.com/chapter/10.1007/978-1-4302-5990-9\\_4](https://link.springer.com/chapter/10.1007/978-1-4302-5990-9_4)
- [11] Yong-Hoon Kim, Yong-Hyuk Kim, Finding a better basis on binary representation through DNN-based epistasis estimation, *Proceeding of the Genetic and Evolutionary Computation Conference Companion*, (2020), pp.229-230.  
DOI: <https://doi.org/10.1145/3377929.3390076>
- [12] Dong-Pil Yu, Yong-Hyuk Kim, Is it worth to approximate fitness by machine learning?: investigation on the extensibility according to problem size, *Proceeding of the Genetic and Evolutionary Computation Conference Companion*, (2018), pp.77-78.  
DOI: <https://doi.org/10.1145/3205651.3208773>
- [13] Sebastien Verel, Gabriela Ochoa, Marco Tomassini, Local optima networks of NK landscapes with neutrality, *Transactions on Evolutionary Computation*, (2018), Vol.15, No.6, pp.783-797.  
DOI: <https://doi.org/10.1109/TEVC.2010.2046175>
- [14] Mohammed Qais, Zeyad AbdulWahid, A new method for improving particle swarm optimization algorithm (TriPSO), *Proceeding of the 5th International Conference on Modeling, Simulation and Applied Optimization*, (2013), pp.1-6.  
DOI: <https://doi.org/10.1109/ICMSAO.2013.6552560>