

[ML/DL] 베이지안 최적화(Bayesian Optimization)란? — 나무늘보의 개발 블로그

노트북: 첫 번째 노트북

만든 날짜: 2021-06-09 오후 10:17

URL: <https://continuous-development.tistory.com/244?category=736685>

---

Data scientist/Machine Learning

## [ML/DL] 베이지안 최적화(Bayesian Optimization)란?

2021. 6. 9. 00:08 수정 삭제 공개

# 베이지안 최적화란?

"Bayesian Optimization의 핵심은 사전 정보를 최적 값 탐색에 반영하는 것이다!"

여기서 제일 중요한 것 두 가지가 나온다. 첫 번째는 사전 정보이고 두 번째는 최적 값 탐색이다.

이제 이말을 좀 더 풀어써보겠다.

베이지안 최적화에서는 **Surrogate model**과 **Acquisition function**이 두가지를 통해 결과를 내고 다시 학습하는 것처럼 계속 돌아간다. (내 생각에는 부스팅 학습 같은 방식 같다.)

**Surrogate model**이란 것은 지금까지의 데이터를 통해 내가 예상하는 값의 모델링이 이렇다 라고 만들게 된다.

수학적으로는 목적함수에 대해 확률적으로 추정한 결과라고 하는 것 같다.

이제 이렇게 만들어진 모델을 기반으로 다음 탐색 지점을 결정하게 된다.

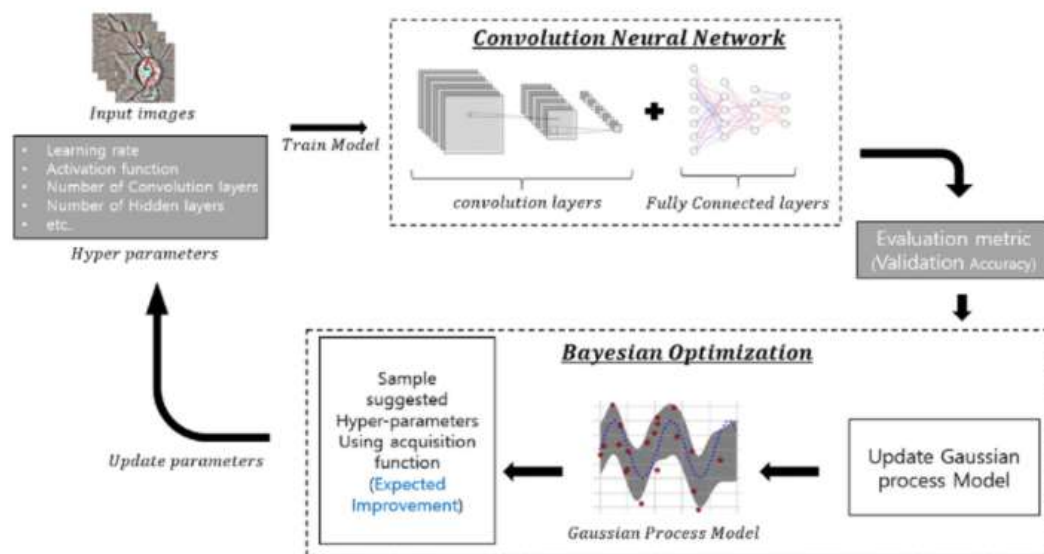
이때 탐색을 하는데 있어서 모델을 검증하는데 최적의 입력 값을 (정보량이 많은 곳이라고 생각하면 될 것 같다.) 추천해주는 함수를 통해 값을 추천하게 된다.

예를들어 설명하자면 수학 문제가 있다고 가정하겠다. 이때 초등학생들에게 고등학교 수학 문제를 출제하게 된다면 이 문제는 학생들의 수준을 평가하는데 변별력을 가지는 좋은 문제라고 할 수 없다. 왜냐하면 모두 배우지 않고 풀 수 없는 문제이기 때문이다.

이와 반대로 고등학생에게 덧셈, 뺄셈 수학을 출제하는 것은 학생들에게 변별력을 평가할 수 있는 문제가 아니다. 이처럼 고등학생에게는 고등학교 때 배우는 난이도의 수학 문제를, 초등학생에게는 초등학교 때 배우는 난이도의 수학 문제를 출제하는 것이 중요하다.

이처럼 이 모델을 검증하는데 있어서 최적의 값을 선택해서 추천함으로써 모델을 검증하고 다음 모델을 만드는 데 있어서 가장 효과적이고 좋은 값을 추천하는 것이 **Acquisition function** 이라고 생각하면 될 것 같다.

즉 **Bayesian Optimization**은 실험 결과를 반영해가면서 효율적으로 **hyperparameter**를 찾을 수 있는 알고리즘이다!



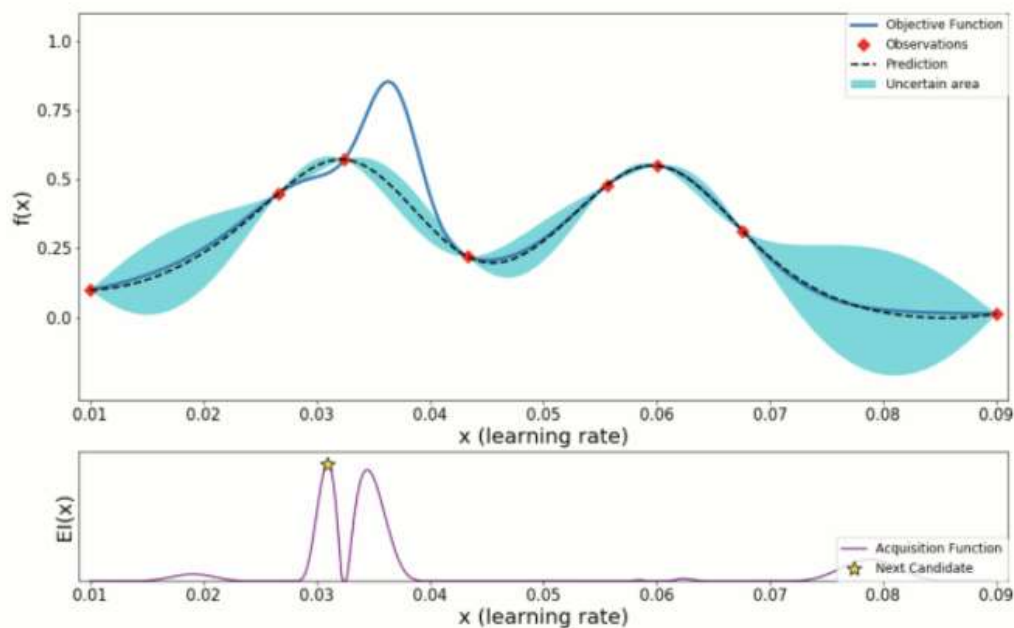
[https://www.researchgate.net/figure/Schematic-view-of-Bayesian-optimization-S-even-hyper-parameters-were-tuned-using-Bayesian\\_fig1\\_335082635](https://www.researchgate.net/figure/Schematic-view-of-Bayesian-optimization-S-even-hyper-parameters-were-tuned-using-Bayesian_fig1_335082635)

그림으로 보면 이런 느낌일 것 같다.

다시 위로 돌아가서 사전 정보를 통해 모델링을 만드는 것이 **Surrogate model** 이고 이 모델링에 최적 값을 탐색하고 추천하는 것이 **Acquisition function** 이라고 생각하면 될 것 같다.

Surrogate Model은 주로 Gaussian Process를 주로 사용하고 Acquisition Function은 주로 Expected Improvement를 사용한다고 한다.  
더 자세한 내용은 밑에 참조를 추천한다.

## Bayesian Optimization 수행 과정



위의 파란색 선은 우리가 찾으려고 하는 목적함수(unknown black box function)를 나타내고,  
검은색 점선은 지금까지 관측한 데이터를 바탕으로 우리가 예측한 estimated function을 의미한다.

검은색 점선 주변에 있는 파란 영역은, 목적함수  $f(x)$ 가 존재할만한 confidence bound(function의 variance)를 의미한다.

밑에 있는  $EI(x)$ 는 위에서 언급한 Acquisition function을 의미하며 다음 입력값 후보를 추천해준다.

Acquisition function 값이 컸던 지점의 function 값을 관측하고 estimation을 update 한다.한다.

계속 update를 진행하면 estimation과 실제 function이 흡사해진다.

관측한 지점 중 best point를  $\operatorname{argmin} f(x)$ 로 선택한다.

제가 수학과가 아니고 통계 쪽에 약하다 보니 여러 사이트에 있는걸 제가 이해한 방식대로 편집했네요. 여러 개 글 본 것 중에 좋다고 생각한 사이트입니다!

참조 : <https://wooono.tistory.com/102>

## [ML] 베이지안 최적화 (Bayesian Optimizat...

wooono.tistory.com  
Hyperparameter Optimization이란, 학습을 수행하기 위해 사전에 설정해야 하는 값인 hyperparameter(하이퍼파라미터)의 최적값을 탐색하는 문제를 ...

참조 : <https://blog.naver.com/PostView.nhn?blogId=dpfkdl&logNo=221678800067&from=search&redirect=Log&widgetTypeCall=true&directAccess=false>

## Bayesian Optimization

blog.naver.com  
이 글은 Bayesian Optimization에 대한 깊은 수학적 이해보다는 Bayesian Optimization의 작동원리를 ...

참조 : <https://shinminyong.tistory.com/37>

# 하이퍼 파라미터 Bayesian Optimize

shinminyoung.tistory.com

1. 들어가며 베이시안 최적화(Bayesian Optimize)를 통한 하이퍼 파라미터 최적화 방법을 소개해드리겠습니다. 요즘 대회들을 많이 참가하면...

'Data scientist > Machine Learning' 카테고리의 다른 글

[ML/DL] 베이시안 최적화(Bayesian Optimization)란?

[ML/DL] 회귀(Regression)의 정의와 구현

[ML/DL] 군집화의 정의와 종류 및 구현

[ML/DL] XGboost의 정의와 구현 및 hyper parameter 설정

[ML/DL] 앙상블 학습 (Ensemble Learning): 3.Boosting(부스팅)이란?

[ML/DL] 앙상블 학습 (Ensemble Learning): 2. Voting(보팅)이란?



나아무늘보

혼자 끄적끄적하는 블로그 입니다.