



Chapter 9

Hyper-Parameter Tuning with Cross-Validation

CONTENTS

1. Motivation
2. Grid Search Cross-Validation
3. Randomized Search Cross-Validation
4. Scoring and Hyper-Parameter Tuning



01.

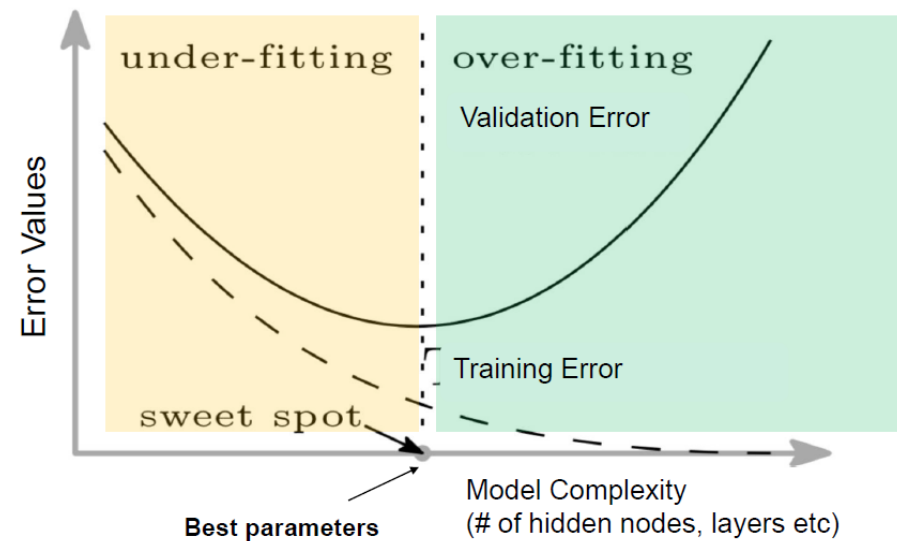
Motivation

Motivation

- 왜 하이퍼 파라미터 튜닝인가
 - 계량경제 모델과 기계학습 모델의 목적과 철학
 - MVUE
 - Minimize overall errors
- Uncertainty
 - **Aleatoric Uncertainty**
 - 데이터 생성과정에서 발생하는 무작위성.
 - **Epistemic Uncertainty**
 - 정확한 모델의 파라미터를 모르기 때문에 발생하는 불확실성
 - **Out-of-Distribution Uncertainty**
 - 데이터가 원래 훈련 데이터 영역의 범위를 벗어난 경우

Motivation

- 하이퍼 파라미터 튜닝
 - 적절히 이뤄지지 못할 경우 과적합 되기 쉬움
 - 금융 데이터에서는 편향보다는 분산, 즉 과소적합보다는 과적합이 더 큰 이슈
- 종류
 - Grid Search Cross-Validation
 - Randomized Search Cross-Validation



Conclusion

- 결론

Approach	ML	DL	Manual/ Auto	Cost	Space expl.	History	Parallel/ Distributed
Babysitting	👍	👎	🐼	💰	Low	✅	No
Grid	👍	👎	💻	💰 💰 💰 💰	High	👴	Yes
Random	👍	👍	💻	💰 💰 💰	Medium	👴	Yes
Bayes SMBO	👍	👍	💻	💰 💰	Medium - Driven	✅	*

02.

Grid Search Cross-Validation

Grid Search Cross-Validation

- 정의
 - 완전 탐색을 통해 CV 성능 최대화 하이퍼 파라미터 조합 확인
- 특징
 - 기저 데이터 구조를 잘 모를 경우 가장 먼저 취할 수 있는 방법
- 구현
 - Scikit-learn의 GridSearchCV

03.

Randomized Search Cross-Validation

Randomized Search Cross-Validation

정의

- 매개변수를 분포로부터 추출

장점

- 차원에 상관 없이 검색하고자 하는 조합의 개수 통제 가능
- 효율적인 계산량
 - Grid Search의 경우 너무 큰 계산량 요구

특징

- 로그-유니폼 분포 샘플링

- 수식

$$F[x] = \begin{cases} \frac{\log[x] - \log[a]}{\log[b] - \log[a]} & a \leq x \leq b \text{에 대해서} \\ 0 & x < a \text{에 대해서} \\ 1 & x > b \text{에 대해서} \end{cases}$$

- 대부분의 하이퍼파라미터는 음이 아닌 값을 받음
- 함수의 반응이 선형이 아닌 매개변수의 경우 0.01부터 1 사이의 증가와 1부터 100까지의 증가가 비슷한 효과

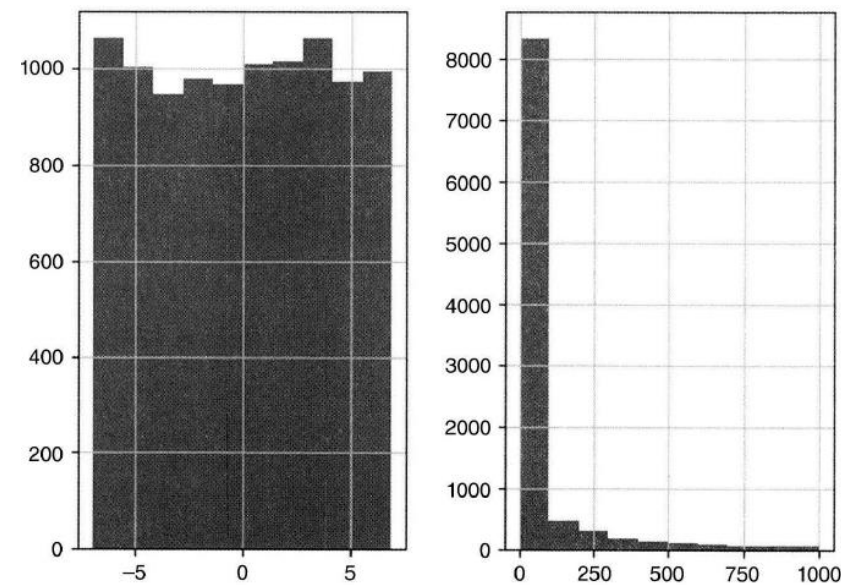
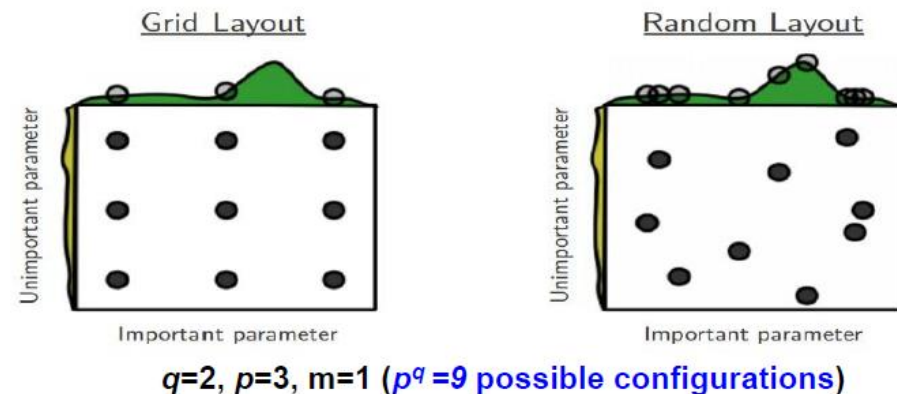


그림 9-1 logUniform_gen 클래스를 테스트한 결과

04.

SCORING AND HYPER-PARAMETER TUNING

SCORING AND HYPER-PARAMETER TUNING

- 기타 고려 요소
 - Scoring
 - 음의 레이블링이 많은 경우(메타-레이블 응용)
 - f1
 - 정밀도와 재현율 사용
 - Accuracy 혹은 neg_log_loss에 비해 유의
 - 모든 경우에 대해 동일한 정도로 관심이 있을 경우
 - Accuracy, neg_log_loss
 - neg_log_loss(cross-entropy-loss)
 - $L[Y, P] = -\log[\text{Prob}[Y[P]]] = -N^{-1} \sum \sum y_{n,k} \log[p_{n,k}]$
 - $p_{n,k}$: 레이블 k에 대한 예측 n과 연계된 확률
 - Y는 K중 하나로 이진표시행렬

SCORING AND HYPER-PARAMETER TUNING

- 기타 고려 요소
 - Neg_Log_Loss
 - 정확도는 높은 확률의 실수와 낮은 확률의 성공이 상쇄될 수 있음
 - 반면 Log Loss는 예측에 대한 확률 고려
 - 표본 가중값을 적용해 분류기 점수화

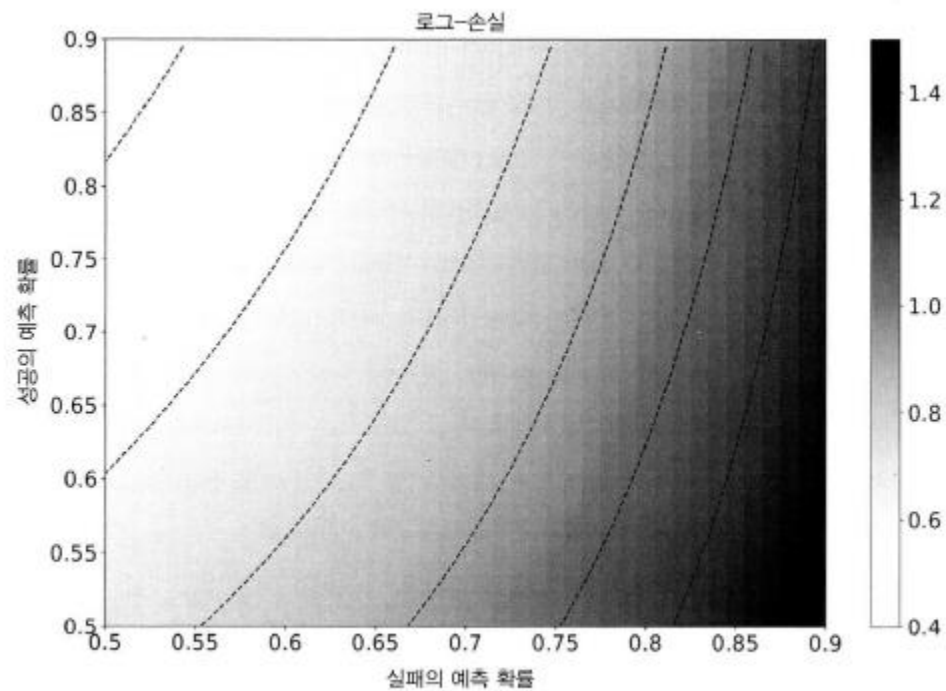


그림 9-2 성공과 실패의 예측 확률의 함수로 나타낸 로그-손실

05.

Conclusion

A low-angle, upward-looking photograph of several modern skyscrapers reaching towards a bright blue sky with scattered white clouds. A white commercial airplane is visible in the center of the frame, flying upwards. The perspective creates a sense of height and grandeur.

Thank you