

# Data Science(Machine Learning, Deep Learning)

- 소스 깃 참고

## <정규화 - Regularization>

Overfitting - 과적합(Overfitting)은 머신 러닝 및 딥 러닝 모델에서 주요한 문제 중 하나로, 모델이 학습 데이터에 너무 맞춰져서 새로운 데이터에 대한 일반화 능력이 떨어지는 현상

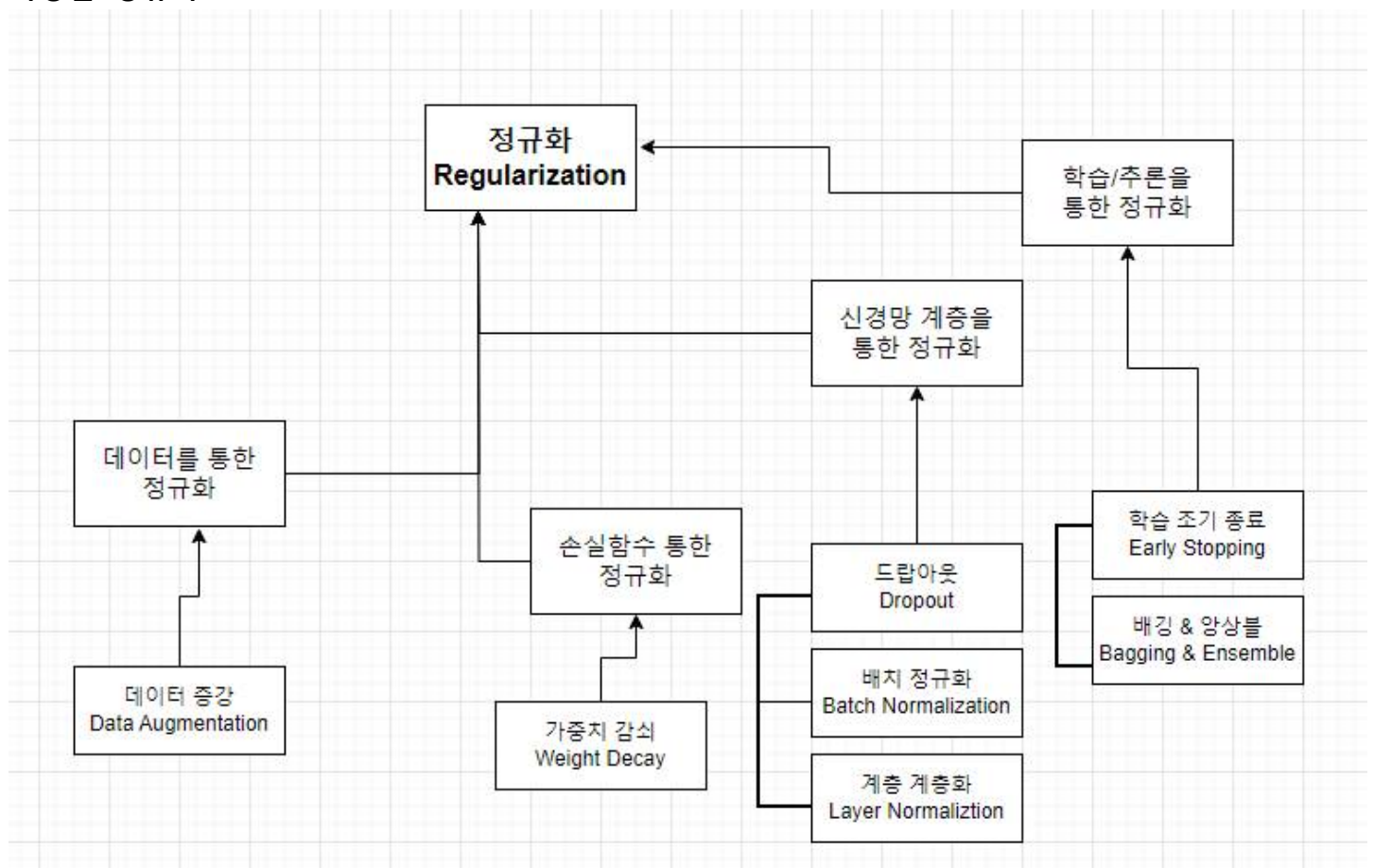
| 범주       | 학습 데이터셋 | 검증 데이터셋 | 테스트 데이터셋 |
|----------|---------|---------|----------|
| 가중치 파라미터 | 결정      | 검증      | 검증       |
| 하이퍼 파라미터 |         | 결정      | 검증       |
| 알고리즘     |         |         | 결정       |

학습 데이터셋, 검증 데이터셋, 테스트셋의 역할 비교

[정규화(Regularization)의 의미]

정규화(Regularization)는 Overfitting을 낮추고 모델이 학습 데이터로부터 적절한 특징(Feature)들을 학습하여 일반화 오차를 낮추는 기법

다양한 정규화



Early Stop(학습 조기 종료)

Early Stop(학습 조기 종료)은 딥러닝 모델이 오버피팅 되기 전에 학습을 종료하는 기법

정규화시 model.train()을 추가해서 정규화를 넣어줘야 함.

그래야 Dropout 이나 Batch를 함.

머신 러닝 학습 전에 미리 사용

model.eval()은 모델을 평가 모드로 전환하는 함수입니다.  
Loss 확인 후 결과 확인시 사용

정규화 적용과 미적용  
미적용 : 97%, 적용 : 98%  
정규화 = 모델 개선