

Part. 06

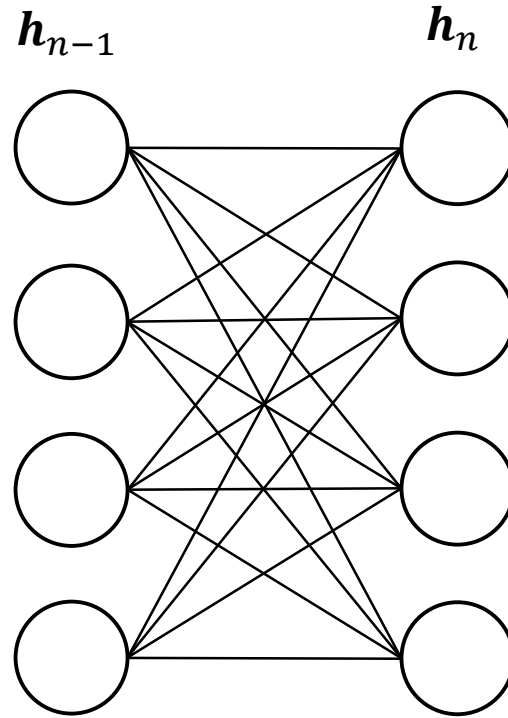
Generalization

| Drop-Out

FASTCAMPUS
ONLINE

강사. 신제용

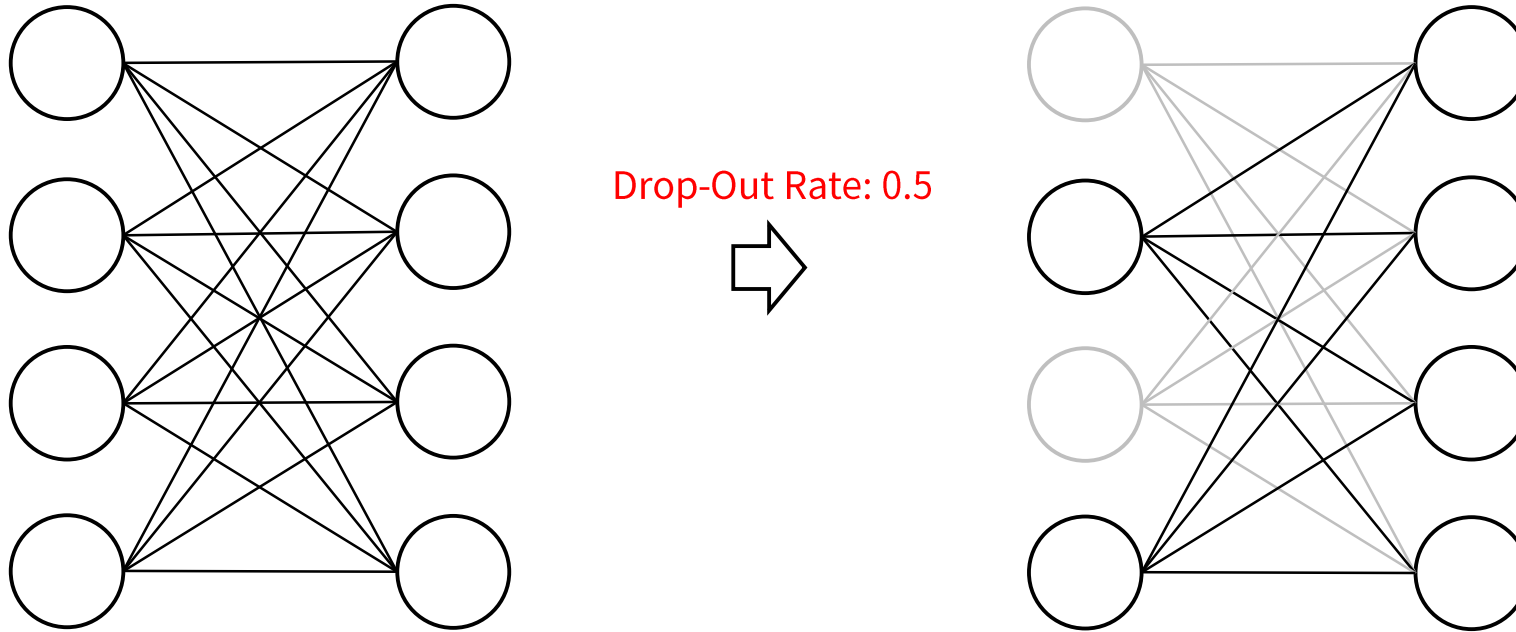
I Fully-Connected Layer



$$h_n = a(W_n h_{n-1} + b_n)$$

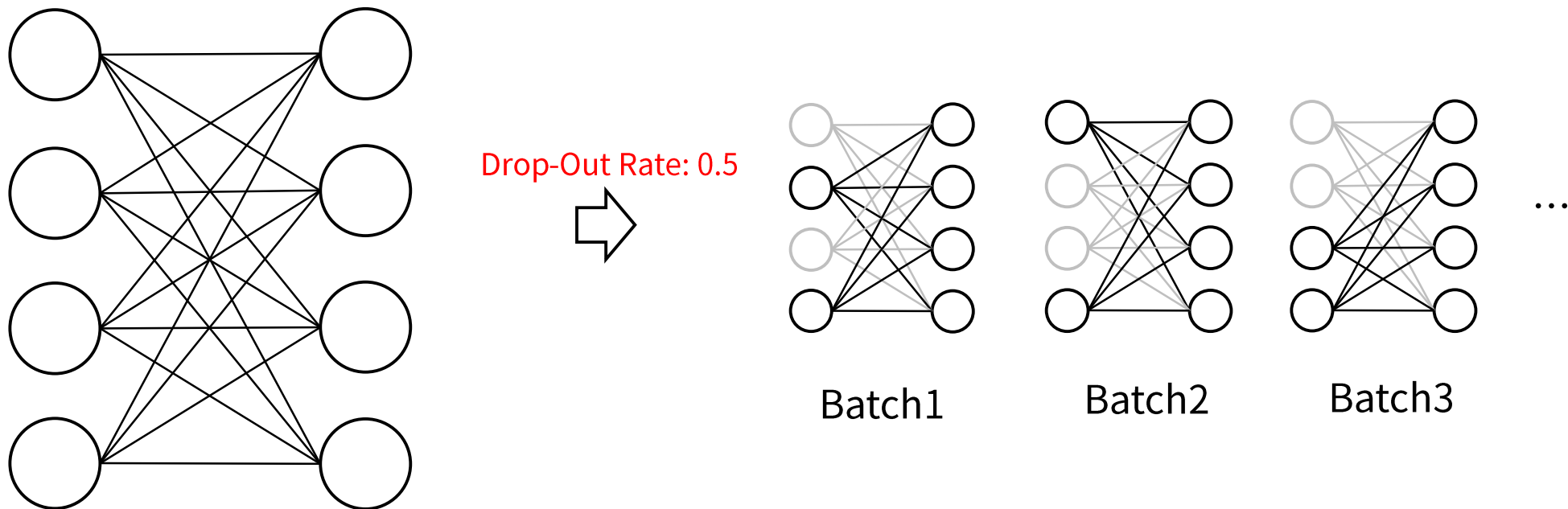
이제는 익숙하게 알고 있을, 전결합 계층을 다시 한번 떠올려 보자.

I Drop-Out



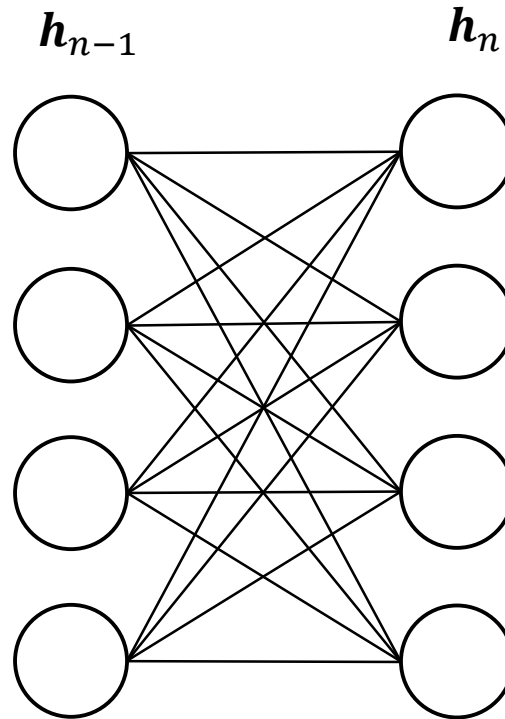
Drop-Out Rate (0 ~ 1)의 확률로 뉴런을 제거하는 기법을 Drop-Out이라고 한다.

I 학습 단계에서의 Drop-Out



Mini-배치 학습을 할 때, 각 배치별로 매번 새로운 Drop-Out을 적용하여 학습한다.

I 추론 단계에서의 Drop-Out

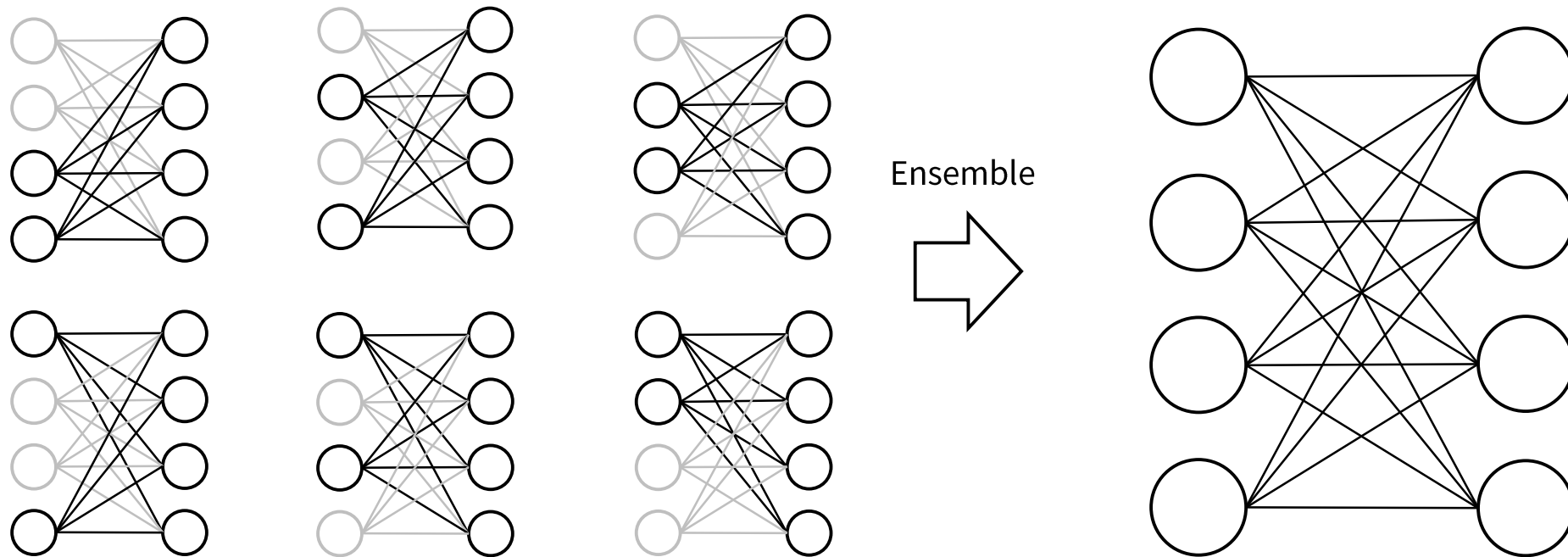


$$h_n = a(\alpha W_n h_{n-1} + b_n)$$

α : Drop-Out Rate

추론 단계에서는 모든 뉴런을 함께 사용한다. 단, 이 때 Scaling을 적용한다.

I Drop-Out의 의미



일부의 특징 (Drop-Out의 특정 Realization)만을 사용해서도 좋은 결과를 낼 수 있다.

각각의 가능성에 대해 좋은 결과를 내도록 학습할 경우, 이를 평균하면 (Ensemble) 균형있는 결과를 얻는다.

즉, 어느 특정한 Feature에 집중하여 Overfitting 되는 결과를 막아준다.