

# Teacher-Student Framework



## Teacher Model

**크기:** 대형 (수백만~수십억 파라미터)

**성능:** 최고 정확도

**용도:** 지식 제공

**예시:** ResNet-152, BERT-Large



## Student Model

**크기:** 소형 (10~100배 작음)

**성능:** Teacher에 근접

**용도:** 실제 배포

**예시:** MobileNet, DistilBERT

Knowledge Transfer



## 전달되는 지식



**Soft Targets**  
(확률 분포)

**Feature Maps**  
(중간층 출력)

**Attention Maps**  
(주목 영역)