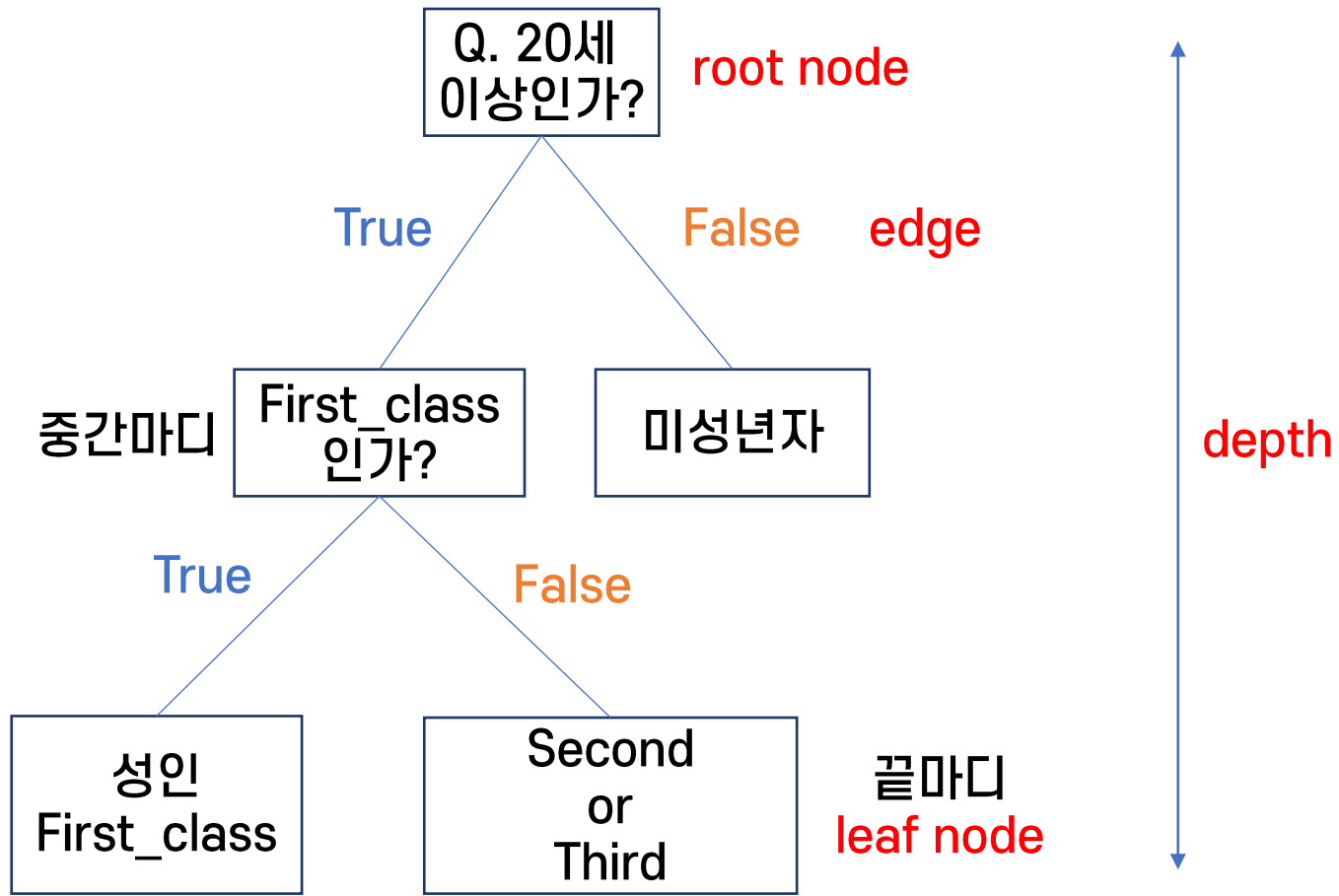


**의사결정나무**

**Decision Tree**

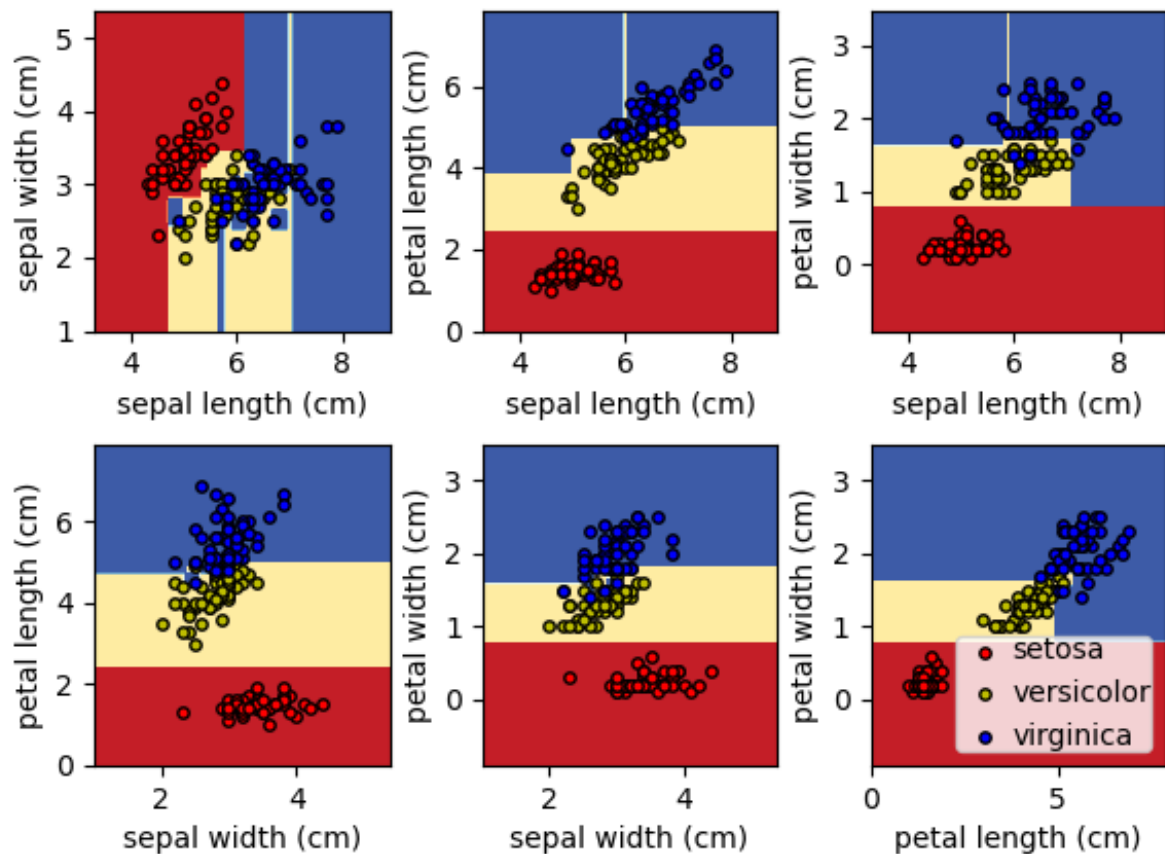
# 의사결정나무 개념



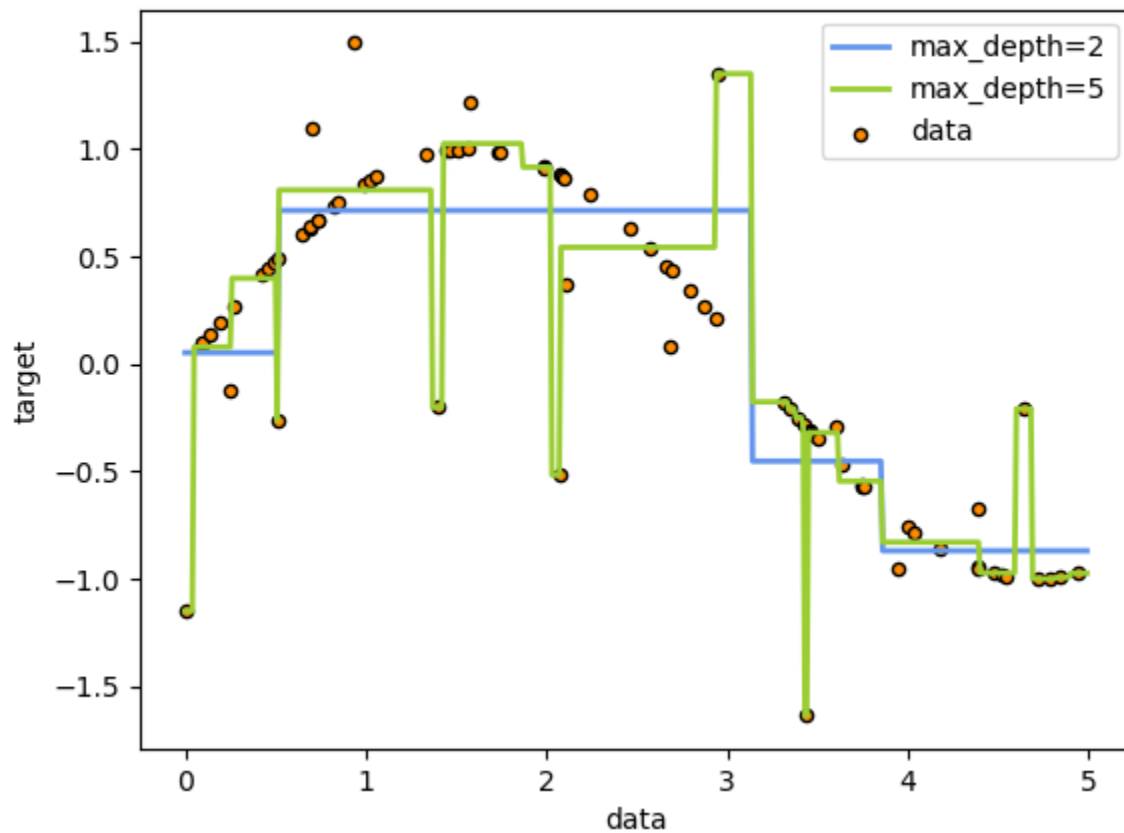
- Root node : 처음 시작하는 마디
- Leaf node : 끝 마디
- Edge : 연결 선
- Depth : 깊이 (최장으로 연결된 질문의 수)

# 의사결정나무 개념

Decision surface of a decision tree using paired features



Decision Tree Regression



# 의사결정나무 개념

---

- 장점

- 1) 데이터 스케일에 영향을 받지 않는다.
- 2) 특히, 데이터 내에 이진 특성과 연속적 특성이 혼합되어 있어도 무관함
- 3) 결정트리 모델이 어떻게 훈련되었는지 경로로 해석이 가능하다. 즉, 설명력이 좋다.

- 단점

- 1) 가지치기를 하더라도 과대적합이 되는 경향이 있다.
- 2) 모델의 일반화 성능이 좋지 않다.

- 매개변수 튜닝

- 1) max\_depth  
최대 가지치기 수
- 2) min\_samples\_split  
내부 노드를 분할하는데 필요한 최소 샘플 수
- 3) min\_samples\_leaf  
리프 노드에 있어야 하는 최소 샘플 수