

## ✧ KNN의 공식

- $y = \operatorname{argmax}_v \sum_{D_x} I(v = y_i)$
- KNN의 공식 x를 중심으로 하는 데이터 집합  $D_x$ 에서 가장 많은 Level을 가지고 있는 v값을 찾음

## ✧ Weighted - KNN의 공식

- $y = \operatorname{argmax}_v \sum_{D_x} w_i I(v = y_i) , w_i = \frac{1}{d(x, x_i)^2}$
- Unknown level에서 거리의 제곱의 역수를 가중치로 하여 거리가 멀수록 가중치를 떨어뜨리도록 모델을 설계함

## ✧ KNN의 종류

- KNN Classifier와 KNN Regressor
- KNN Classifier : 범주형 데이터 예측
- KNN Regressor : 연속값 예측
- target값에 따라 사용하는 모델이 달라짐
- 옵션은 Weighted의 가능 여부에 따라 4가지로 분류됨

## ✧ 데이터 정규화

- 데이터 속성들을 하나의 scale로 통일해야 함
- 가장 많이 쓰이는 방법은 통계학에서 쓰이는 z정규화 방식임

- $z_i = \frac{x_i - \operatorname{avg}(x)}{\sigma}$ ,  $\sigma$ 는 표준편차  $\operatorname{avg}(x)$ 는 x평균