

## 8. ML\_DL

### 8.1 Logistic regression

#### 1. Binary Classification

- 분류 알고리즘 중 정확도가 높은 알고리즘
- 알고리즘의 활용 예시
  - 1) Email 판별 : Spam인지 아닌지 판별(Spam이면 1, 아니면 0)
  - 2) 주식의 주가가 오를지 떨어질지 판별(오르면 1, 떨어지면 0)
  - 3) 신용카드 사용시 도난카드인지 아닌지 판별(도난카드 1, 정상카드 0)

#### 2. Machine Learning이 0과 1 사이의 값으로 예측

#### 3. 통상적으로 0.5 이상이면 1로 간주, 0.5 미만이면 0으로 간주

### 8.2 learning rate

- 여러 개의 값으로 Test
- Cost function의 값을 변화에 따라 적절한 learning rate를 찾아야 함
- 데이터와 실행 환경에 따라 learning rate는 다르기 때문에 경험적으로 찾아야 함
- 초기값은 0.01 정도로 설정 후 값을 수정하며 적절한 값을 찾아야 함

### 8.3 Overfitting

- Overfitting이란 Training data set에는 적합 하지만 실제 데이터에는 적용이 잘 안되는 경우
- Overfitting의 해결방안
  - 1) 많은 training data를 이용
  - 2) feature의 개수를 줄임(중복 및 불필요한 feature를 생략)
  - 3) Regularization
- weight 값이 클수록 그래프의 커브가 커지고 overfitting이 발생되므로 너무 큰 값이 되지 않도록 조정