핀테크 이론

1. 정리

1. 머신러닝의 분류

지도학습(supervised): label을 부여하는 학습 방식(= 개별 독립변수에 대한 종속변수를 개별적으로 부여하는 방식)

->주로 분류, 회귀, 시각/음성 등의 감지와 인지 모델에 사용

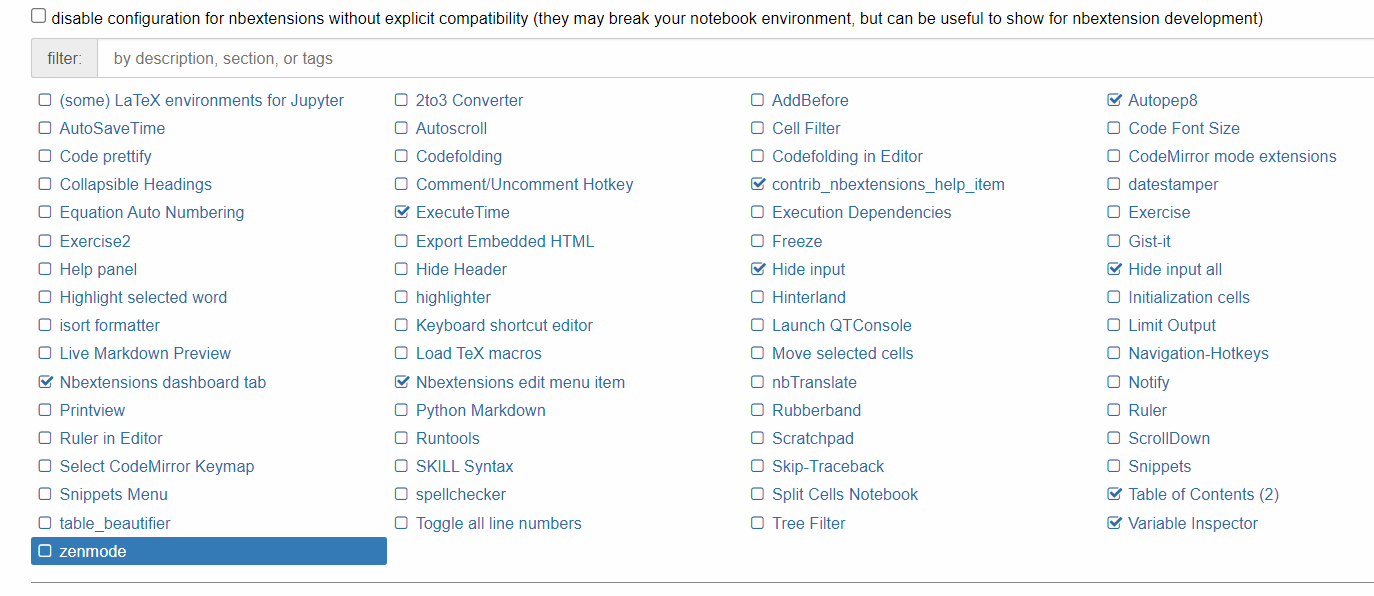
비지도학습(Unsupervised): label을 부여하지 않는 학습 방식

->군집화와 차원 축소 등에 사용

강화학습: 과정에서 penalty와 보상을 부여하는 방식

->최적 요건 도출 등에 사용

주피터 nbextensions 추가사항



-퍼셉트론(perceptron)

-의미: 인간의 뉴런 신경망을 본뜬 네트워크 구조

-기계학습 = 함수근사

만들어진 함수 근사 = 모델

-레이블

y=f(x)에서 y변수를 레이블이라고 한다. x값은 독립변수, y는 종속변수라고 취급한다.

여기서 f()는 알고리즘이 된다. 이 알고리즘을 만들면서 가장 최적의 알고리즘을 모델이라고 한다.

여기서 x값에 대한 실제 y값과 알고리즘의 y값의 차이를 에러라고 한다.

-회귀란? (regression)

-> 지도학습의 종류 중 한가지이다.

-> 임의의 알고리즘을 먼저 만들고 그것에 대한 에러 값을 구하면서 최적의 모델을 만드는 과정

독립변수에 대한 종속변수의 값에 따라 알고리즘을 만드는 것이 아니라 임의의 알고리즘에 대해서 실제 값과 알고리즘 결과값 사이의 차이인 에러 값을 통해서 최적의 모델을 만드는 것이다.

-분류(classification)

->입력을 두개 이상의 레이블(유형)으로 분할하는 것(ex:스팸분류 등)

->해당 모델 학습에 대해서 레이블을 제공해야한다.

-> 분류 유형의 숫자에 따라서 이중분류, 다중 분류 등으로 나뉠 수 있다.

-> 분류 알고리즘 예시: kN, SVM,의사결정트리, DNN

-비지도 학습

입력 값에 대해서 사람이 기준을 던져주면 스스로 분류하는 것을 의미한다.

이때, 이 분류에 대해서 지도학습처럼 어떤 의미를 알지 못하고 특징에 따라 데이터를 분류만 할 뿐이고 이 이후 과정에서는 인간의 개입이 필요하게 된다.

2. 궁금증 정리사항