Softmax Regression (Multinomial Logistic Regression)

Logistic Regression은 binary classification 이라고도 부른다.

데이터를 1과0 두가지 그룹으로 나누기위해 사용하는 모델이다

softmax는 데이터를 2개이상의 그룹으로 나누기위해 binary classification을 확장한 모델이다

hardmax: 통계에서 가장 큰 값을 찾는 개념

softmax: 통계에서 새로운 조건으로 가장 큰 값을 찾는 개념

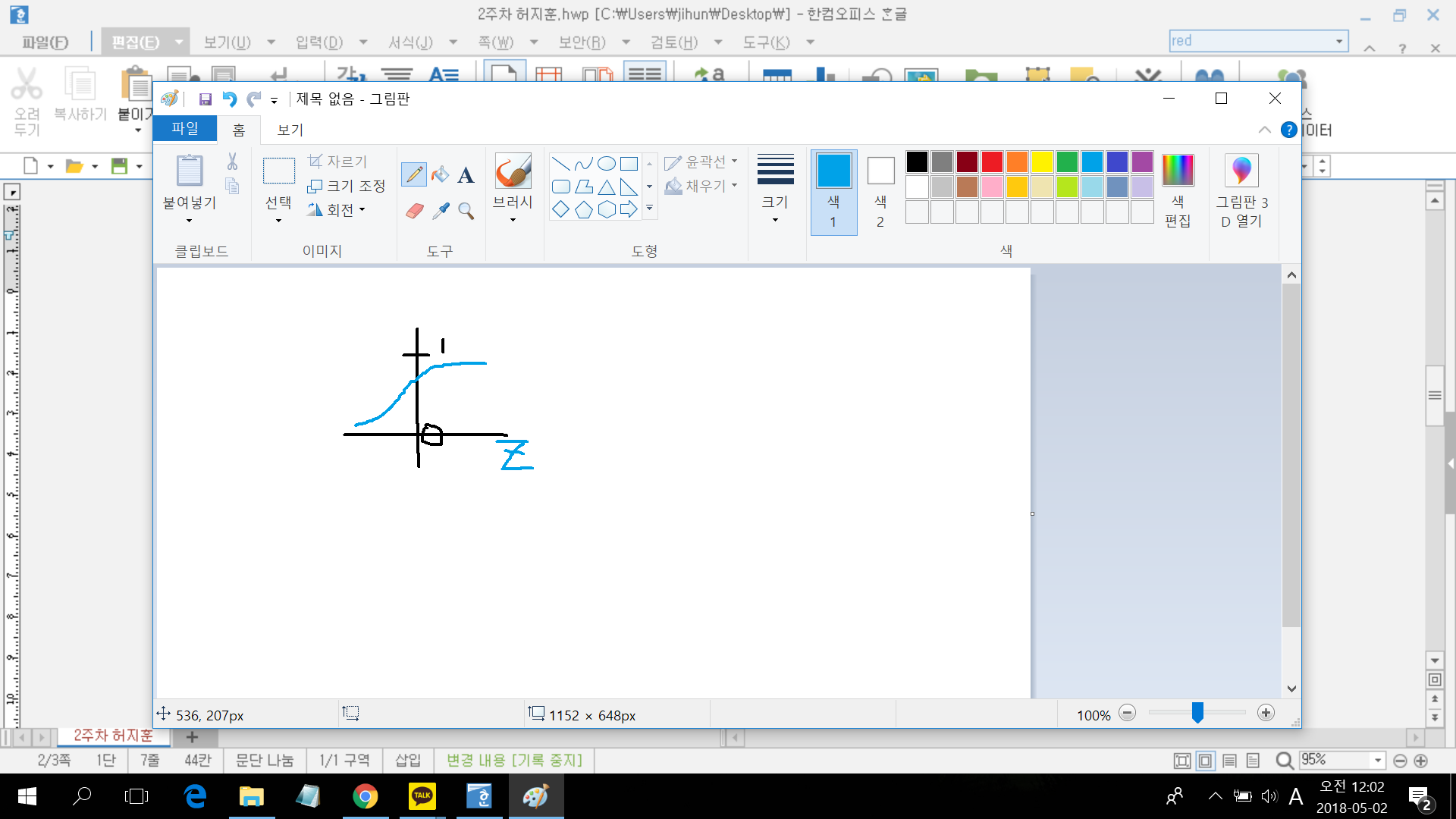
H(x)=Wx 형태로 일반적으로 Linear로 출발을했다

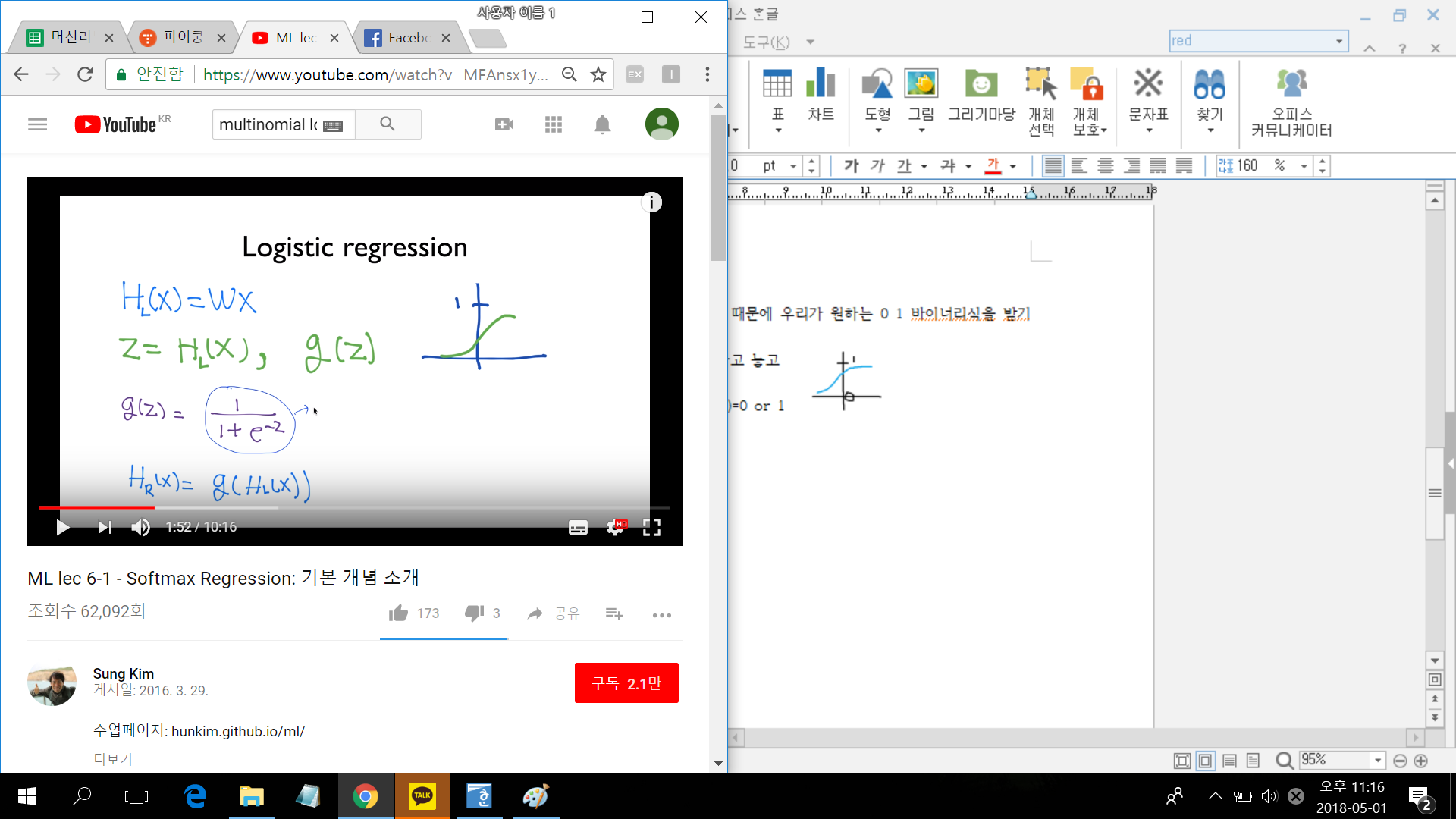
이것의 단점은 리턴의값이 실수, 양수, 음수이기 때문에 우리가 원하는 0 1 바이너리식을 받기 힘들다

그래서 우리가 생각한방법이 H(x)를 하나의 Z라고 놓고

Z=H(x),

g라는 함수 g(z)를 사용해서 예측 값을0이나 1 혹은 그 사이 값으로 만들어준다. g(z)=between 0 and 1

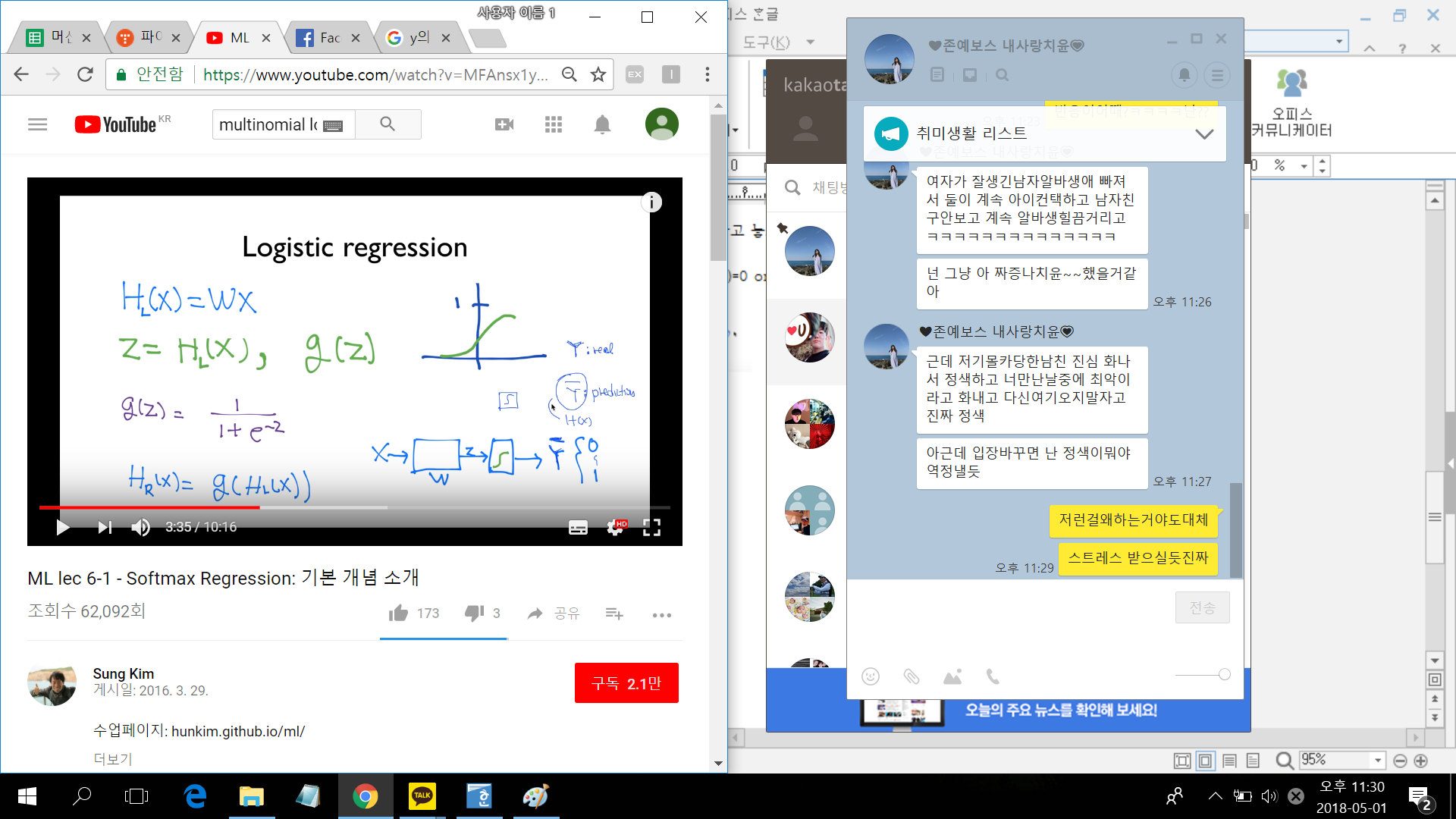
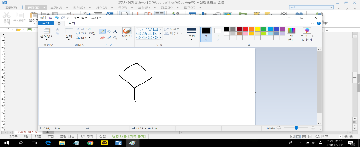




그래서 사람들이 연구해서 만들어낸 시그모이드(로지스틱) 함수이다

->1과 0사이로 압축 시킬수있게 해주는 역할

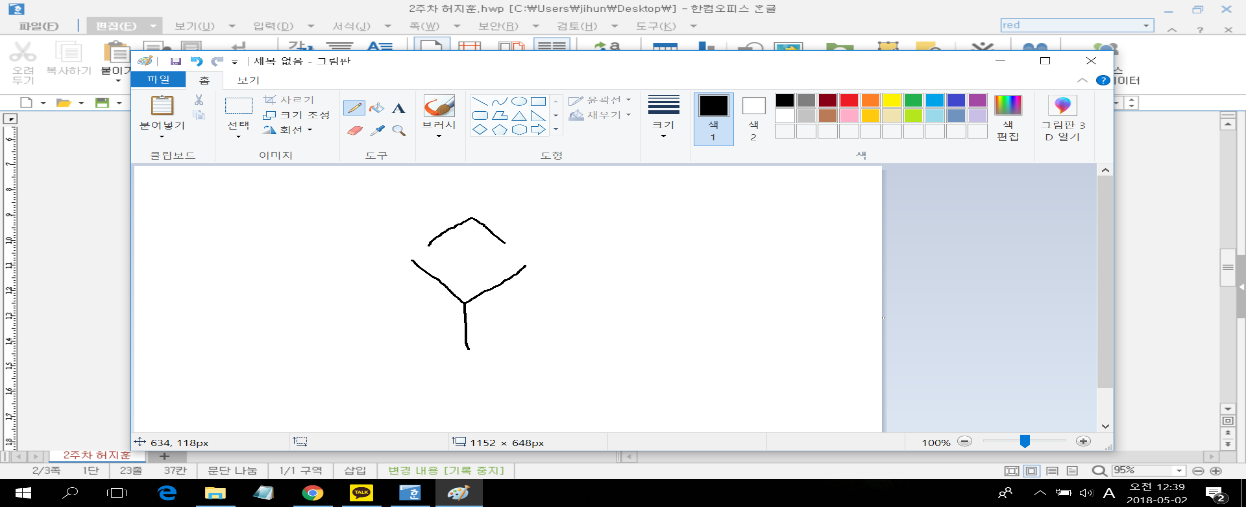
X라는 입력

W유닛으로 계산을 한다 나오는 값이 Z값 (Z=H(x))

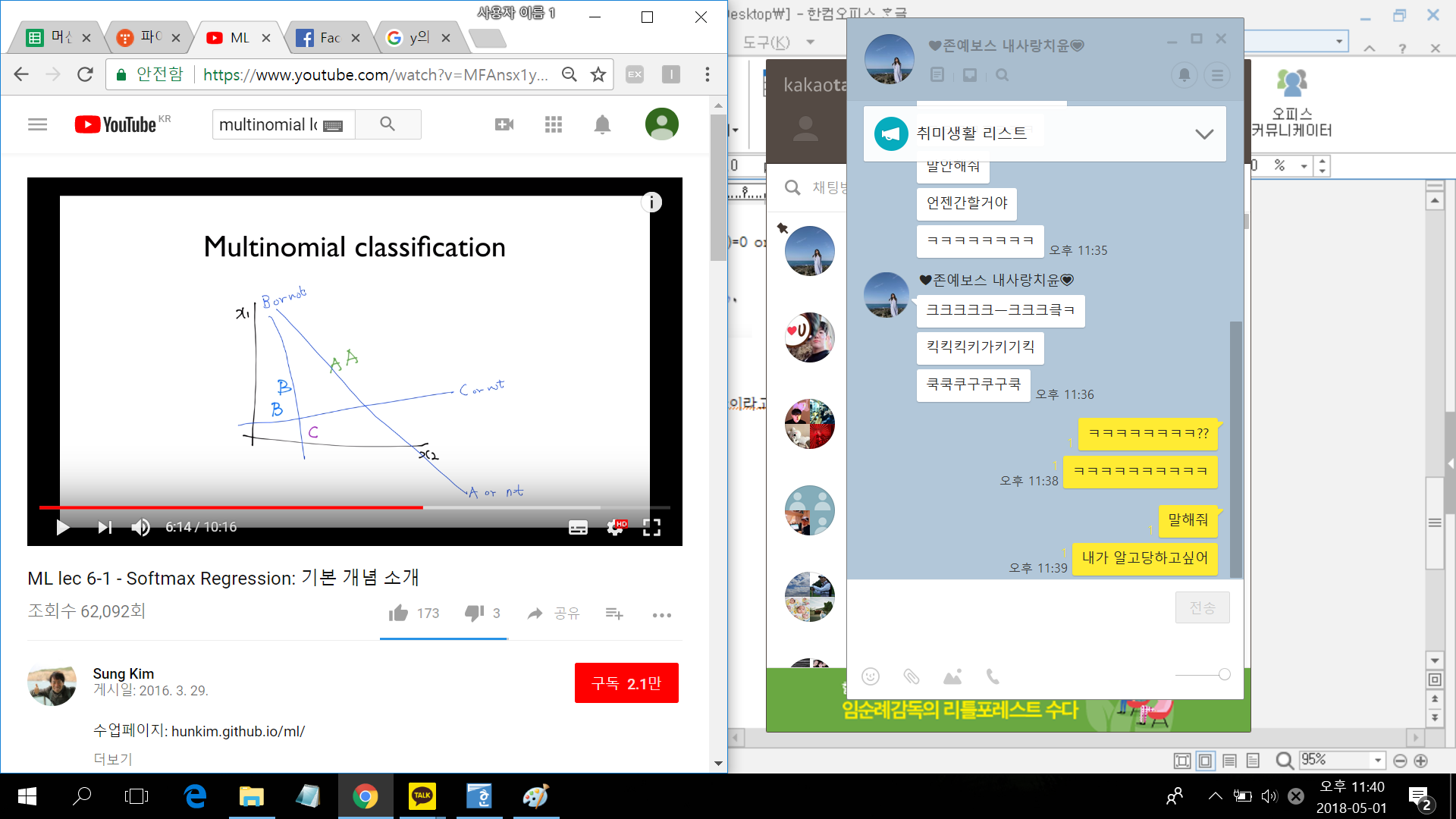
이 값을 시그모이드함수에 통과시킨다.

그래서 나온 결과가 Y의 hat이라고 부르고 이 값은 0~1사이,

x에 대응하는 realdata를 Y라고 부른다.



Y의 hat

Multinomial Logistic Regression

여러개의 클래스들이 있다

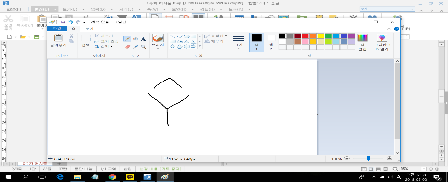
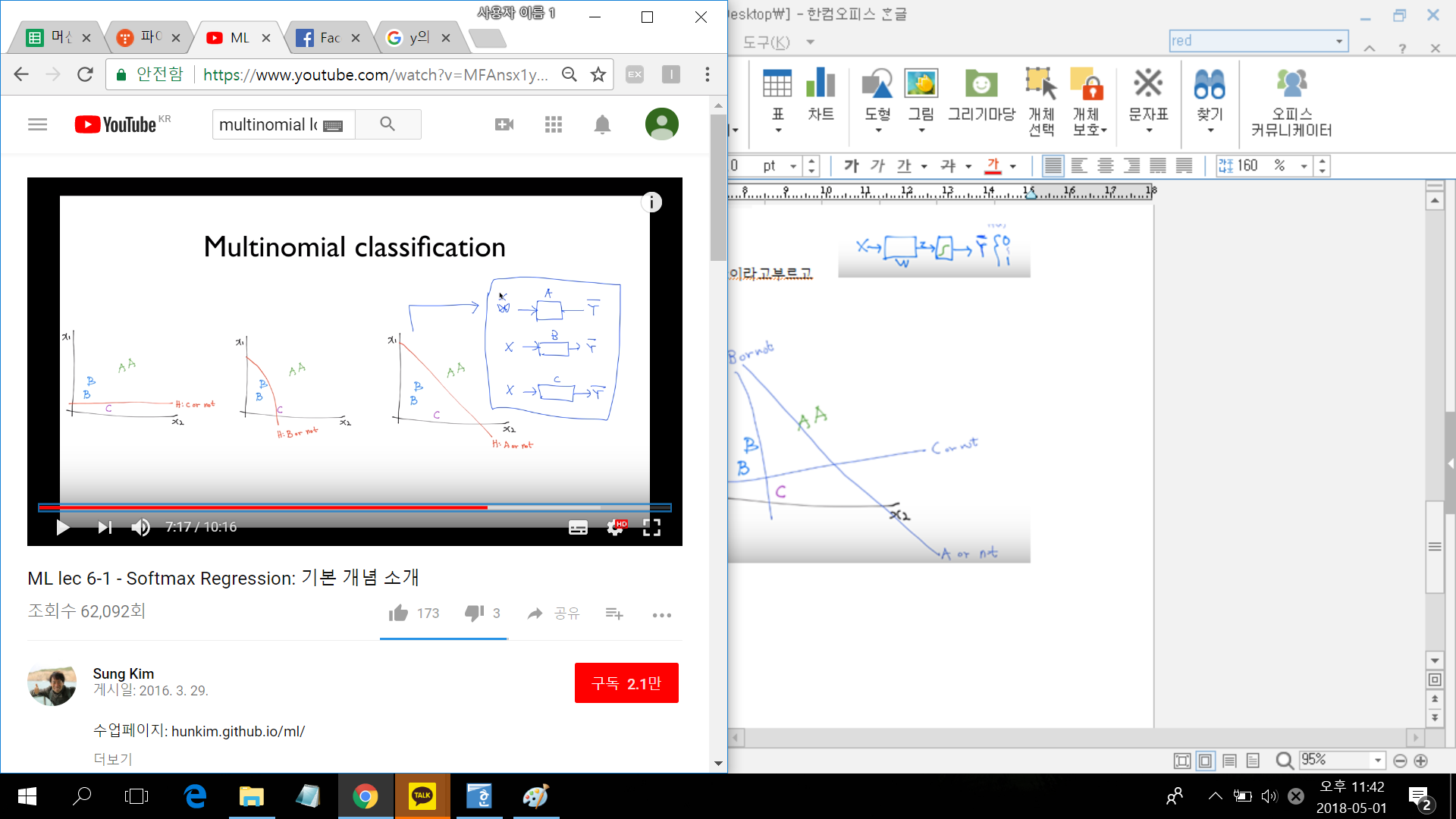
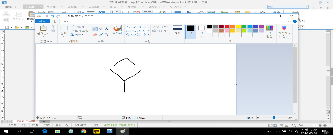
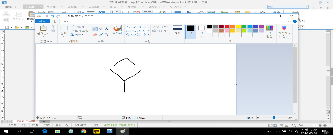
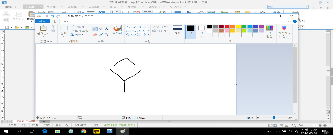
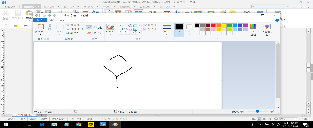
A, B, C의 3개의 그룹이 있고, 오른쪽 그림에서 binary classification을 사용해서 나누었다.

각각의 선은

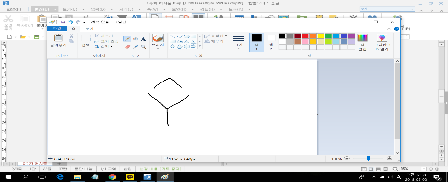
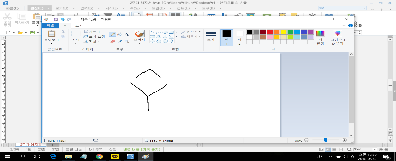
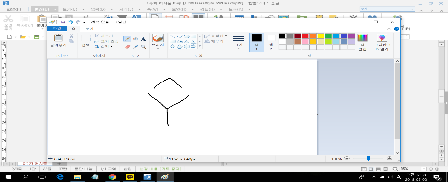
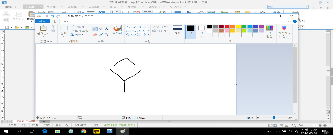
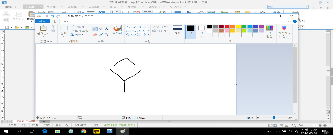
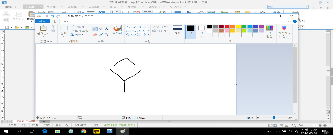
1)A냐 아니냐

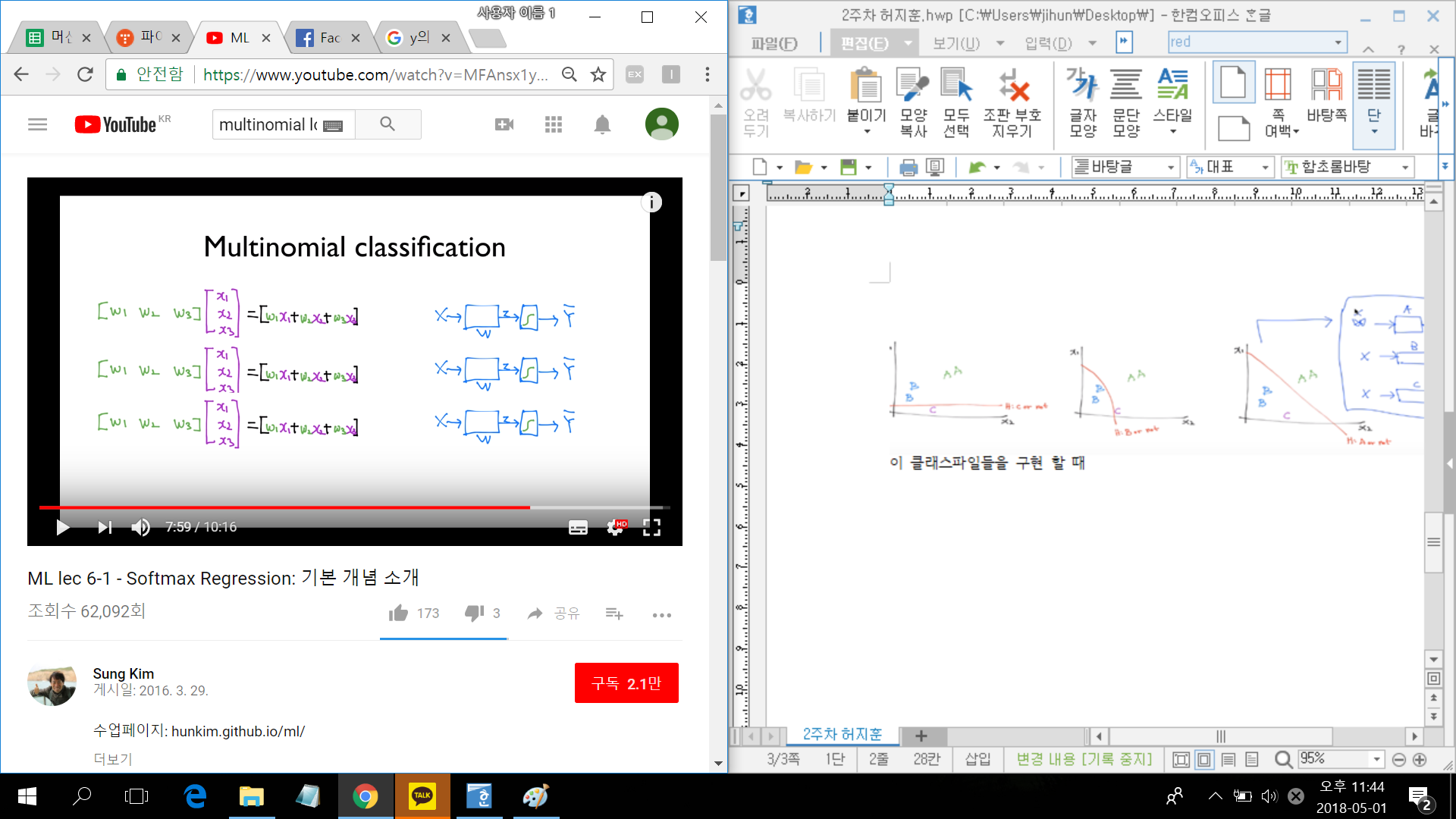
2)B냐 아니냐

3)C냐 아니냐로 나눌 수있다.

이 클래스파일들을 구현 할 때 각각 A인지 아닌지 B인지 아닌지 C인지 아닌지 독립된 클래스 파일 형태로 구현이 가능하다

행렬 계산한 값이 우리가 원하던 H(x)의 값이다. H(x)=Wx





독립적으로 계산하면 복잡하고 구현하기에도 복잡하다. 그래서 W를 축 늘린 형태로 나타낼 수 있다.

YA의 헷이 HA(x)값이 나온다.

YB의 헷이 HB(x)값이 나온다.

YC의 헷이 HC(x)값이 나온다.

이러한 값들을 한번에 계산이 가능하고 독립된 클래시피케이션들처럼 동작을 하게된다.

이렇게 나온값이 실수가 나오기 때문에 각각 시그모이드 함수를 써야한다.

함수 사용후 나온 값들의 합은 1이 되어야 한다.

