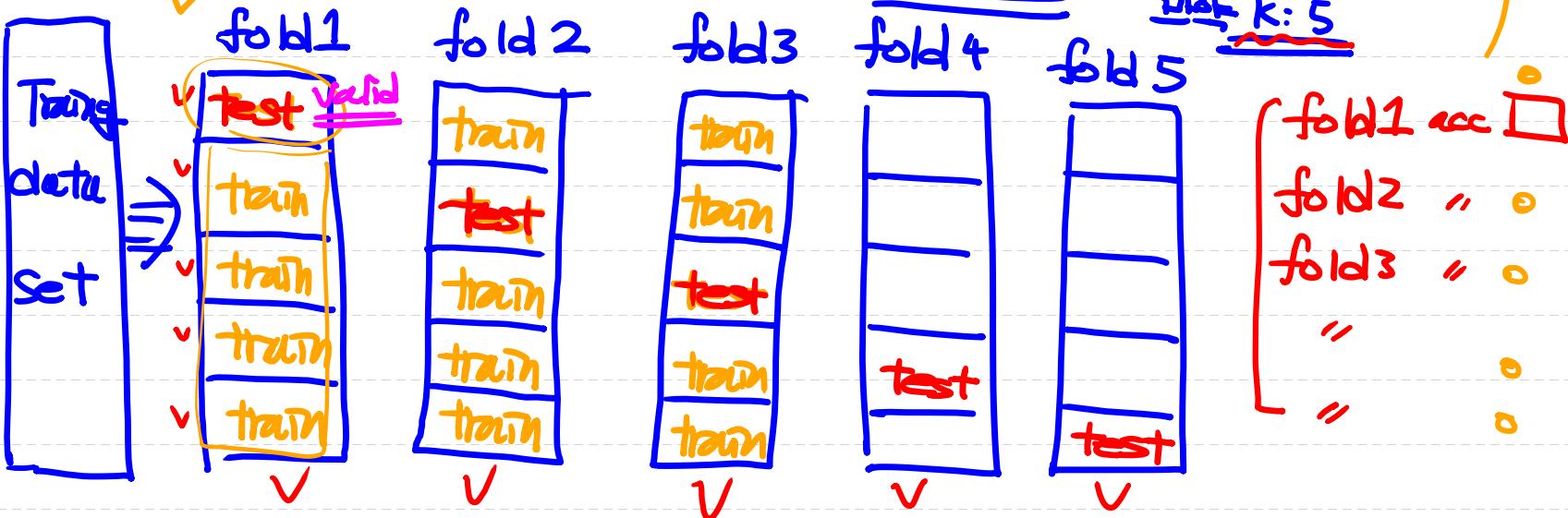


- Evaluation 할 때, 데이터가 충분하지면 !! (Training, Validation, Testing)

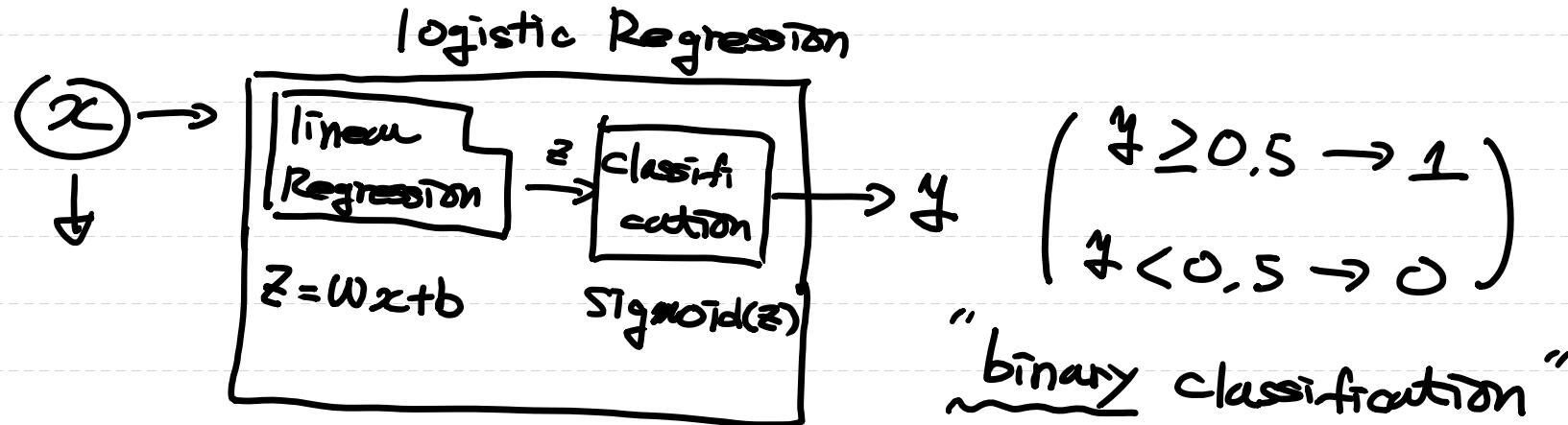
↓ 데이터량이 적으면 ?? (Underfitting)

epoch의 개수 늘리면 Overfitting

↓ Cross Validation (CV) k-fold k: 5



• 10/5(월)



→ logistic regression은 뉴런을 반복시 선형회귀를 적용함.

(Accuracy (정밀도))

• 성능평가 지표 (Metric)

→ 성능평가를 하는 방법 → 정답 Vs 모델이 예측한 것에 대한 정답 비율.

Confusion Matrix		실제 정답(+)	
		TRUE	FALSE
Model 결과	TRUE	TRUE POSITIVE	FALSE POSITIVE
	FALSE	FALSE NEGATIVE	TRUE NEGATIVE

- * Precision (정밀도) → 우리의 Model이 TRUE로 봉죽한 것 중에 정답으로 TRUE인 것에 대한 비율
- Recall (재현율)
- Accuracy (정확도)

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

• Recall : 실제 TRUE인 것 중에 우리의 Model이 TRUE로 예측한 비율.

$$\checkmark \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \Rightarrow [\begin{matrix} \text{Sensitivity (믹스도) } \\ \text{Hit Rate} \end{matrix}]$$

→ 왜 precision과 recall을 구별해 두었지, precision과 recall 중에
어떻게 더 중요하니?

✓ A라는 회사가 있어요.. 어떤 서비스의 기준을 개발. 사용을 업계별으면 그通り
고객이 A라는 회사를 찾을 확률을 기준.

✓ like.

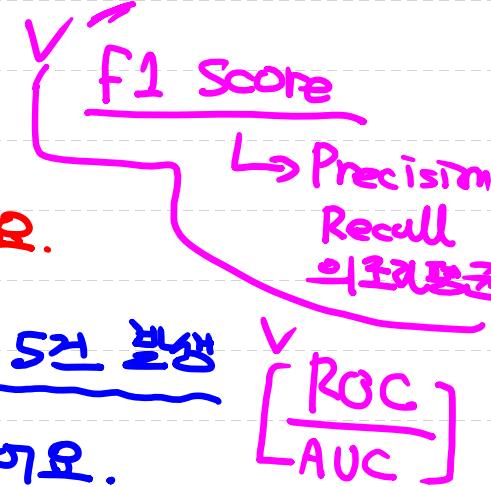
→ 이 기준을 이용하면 사용자 고객이 99.9% 찾을 수 있어요?

" " " " "
60% 찾을 수 있어요?
T_{clike}

- 흔치지 가중을 하고 해당 기준을 바라보자!!

→ 고양이 검출 기준은 초대로 오설출을 하지 않아요.

$\begin{cases} A \rightarrow \text{검출율 } 99.9\%, \text{하지만 오설출이 } 1200 \text{ 건 발생} \\ B \rightarrow \text{검출율 } 60\%, \text{하지만 오설출이 거의 없어요.} \end{cases}$



어디에 흔적화노트에 따라서 어느 기준이 더 좋지 그 내용이 달라져요!

기준 \rightarrow $\begin{cases} \text{검출율} \Rightarrow \text{Recall} \text{ (재현율)} \\ \text{정확도} \Rightarrow \text{Precision} \text{ (정밀도)} \end{cases}$

$$\rightarrow \text{Accuracy} \text{ (정확도)} \rightarrow \frac{TP + TN}{TP + FN + FP + TN} \Rightarrow \frac{\text{적외선}}{\text{증거지표}}$$

- ✓ Learning rate ! $\Rightarrow \underline{w, b}$

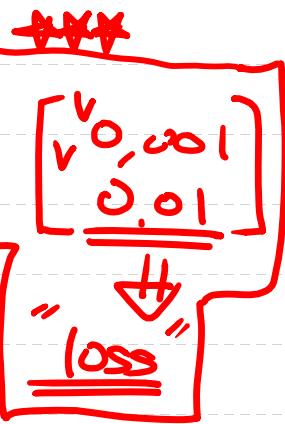
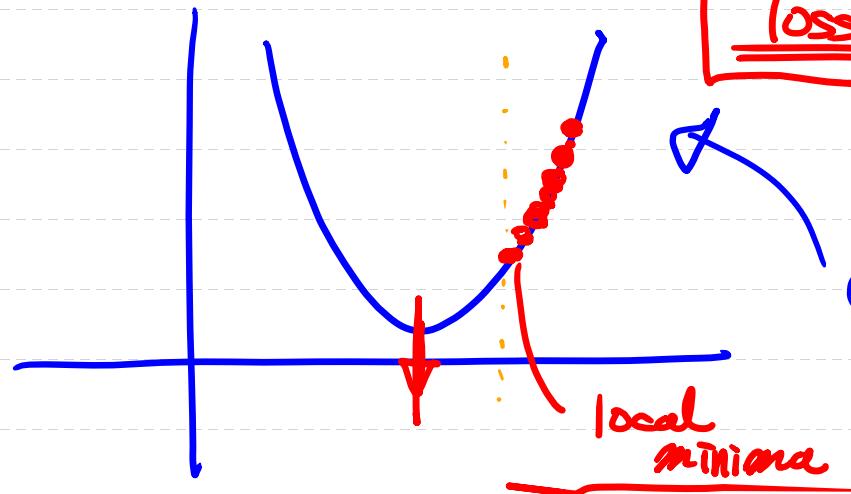
$$w = w - \alpha \cdot \frac{\partial E(w, b)}{\partial w}$$

- ✓ Normalization !

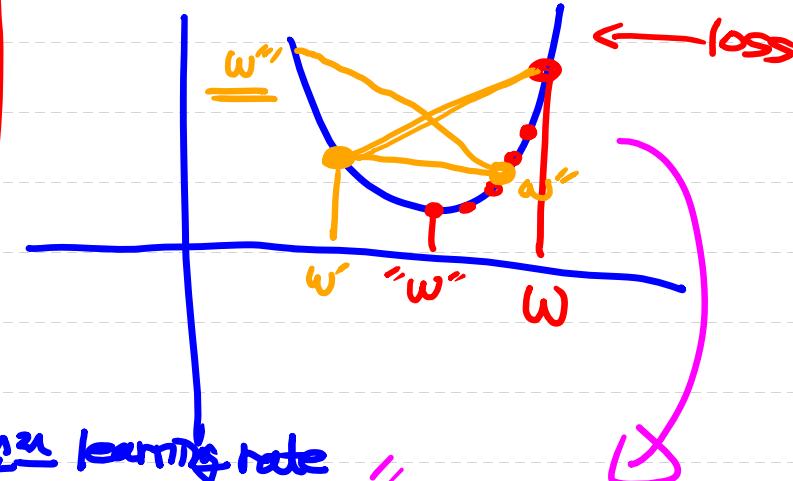
- ✓ Overfitting !

- Evaluation !

\hookrightarrow k-Fold
cross validation !



① 적절한 learning rate 보다 무리의 α
값이 크면,



적절한 learning rate,
② 무리의 α 값이 보다 작을 때
"Overshooting"

• Overfitting 은 해결해야 하는 문제 !!

• ① 많은 양의 Training data. ↗ 가끔 기분이 되는 조건인데 우리가 program 속으로 해결이 불가능!

✓ • ② feature의 개수를 줄여야 해요 !!

(독립변수) → 중복변수와 연관이 없는 독립변수들 제거. (선택과제 제작)
컬럼

③ Weight의 값을 의미적으로 조절! w 의 값이 커질 때 Curve를 그려보면 Overfitting의 여지가 많아져요.

"Regularization" →

④ Deep learning \Rightarrow "Dropout"

⑤ 학습률의 구조 줄여요! \rightarrow epoch 구조
직접하게 조작

전체
✓ 학습데이터를 이용해서
1번 학습 \rightarrow "1 epoch"

• Normalization (정규화)

* Min-Max Normalization

설명: 0 ~ 1 사이로 모든 feature를 scaling

단점: 이상치가 있으면 scaling이 이상!!

Standardization (표준화) → z-score Normalization

설명: 이상치가 상대적으로 포함!!

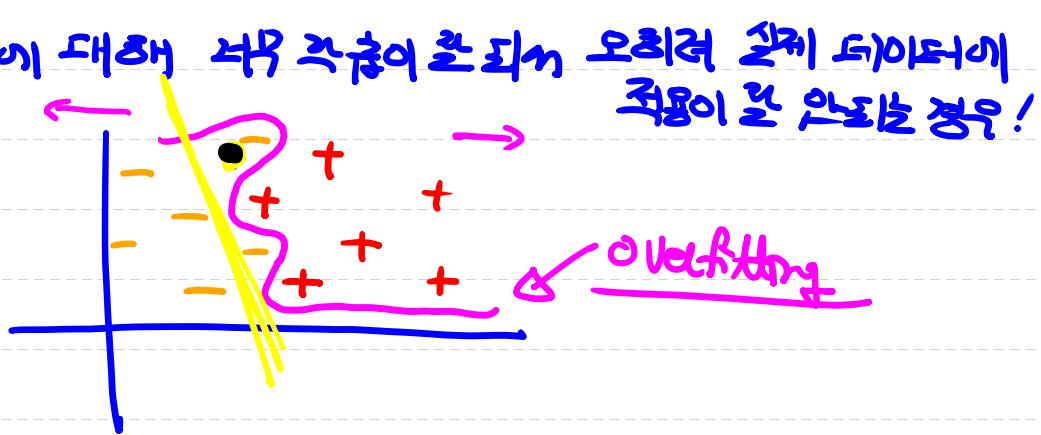
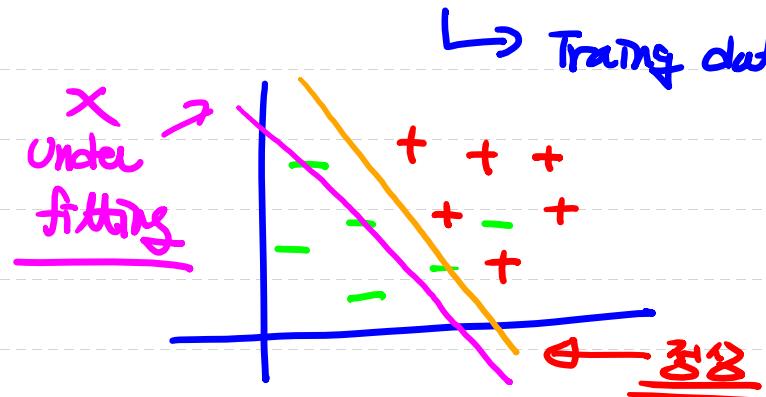
단점: 모든 feature들이 대체 동일 척도로

scaling을 못해요!!

• Underfitting → 너무 대충 학습해서 예측이

잘 이루어지지 않음 경우.

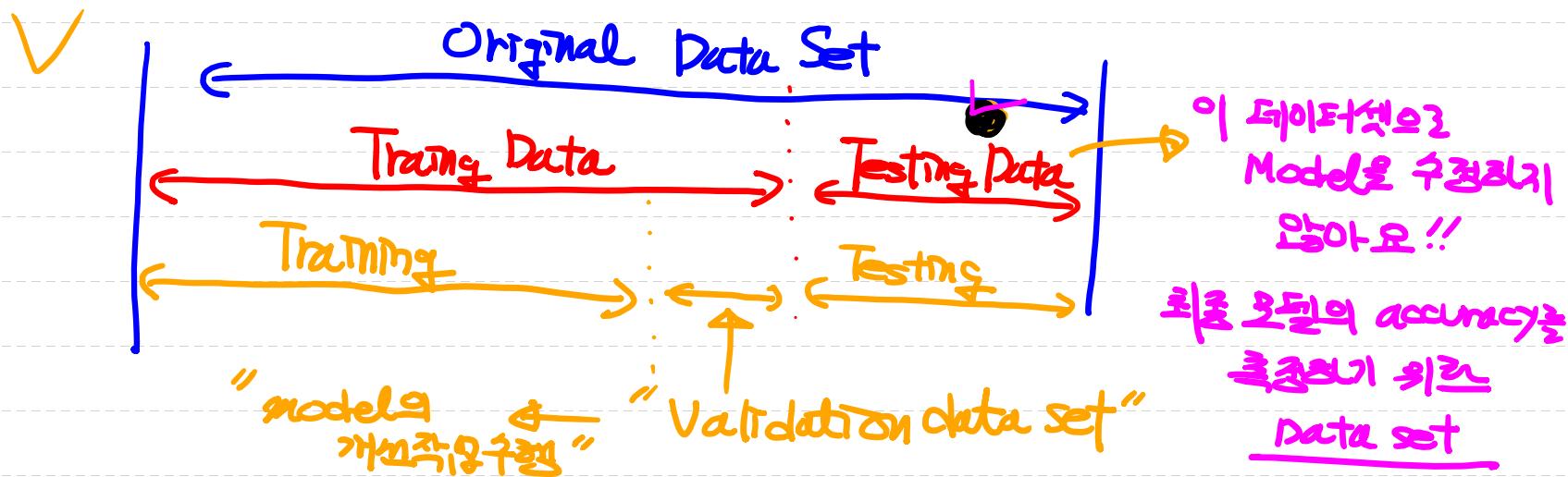
Overfitting ⇒ ↗



- Evaluation(평가) → 우리가 가지고 있는 Training Data Set을
어떻게 이용해서 성능평가를 할 것인가?
성능평가방법 → $\begin{array}{l} \cdot \text{Precision} \\ \cdot \text{Recall} \\ \star \cdot \text{Accuracy} \end{array}$

성능이 아주 좋게
나와요!!

- 우리가 피해야 하는 평가방법: Training Data Set으로 훈련하는
Training Data Set으로 테스트하는 X.



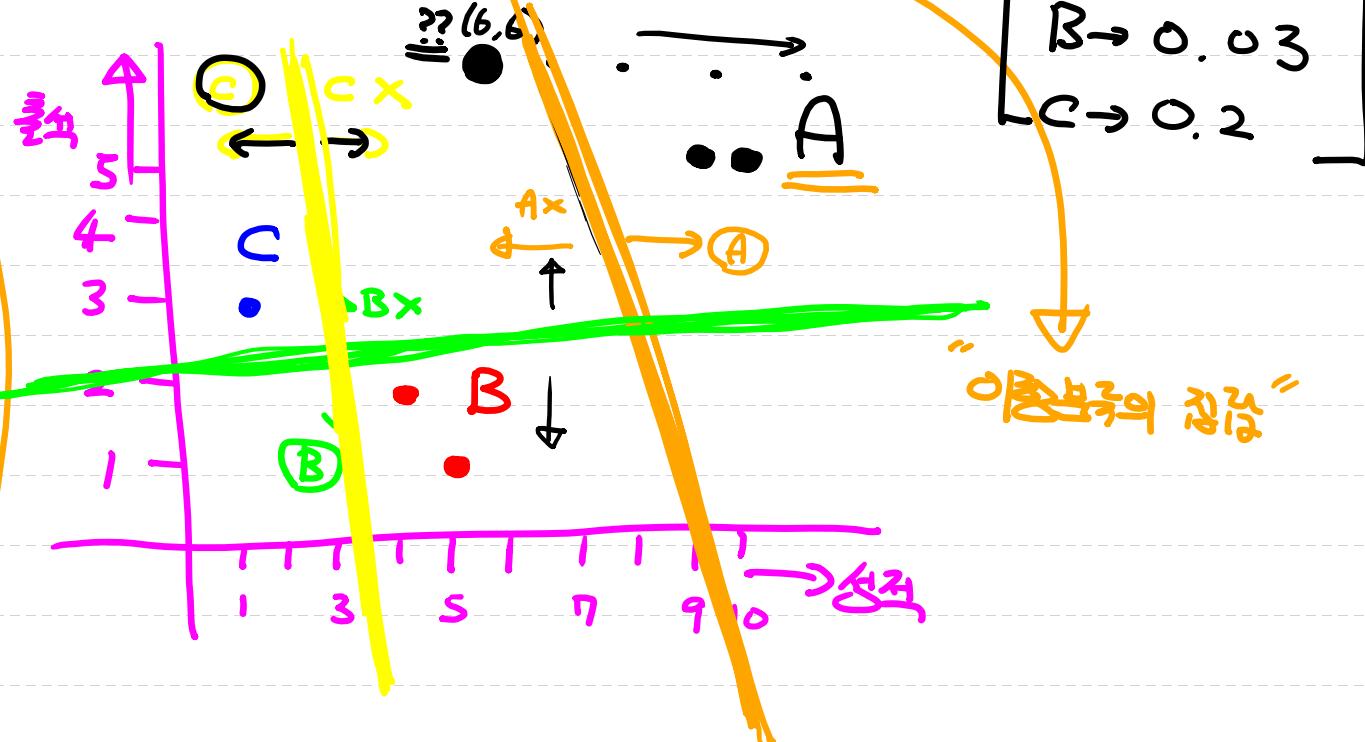
• Logistic Regression (Binary classification) \rightarrow 이항분류

\Rightarrow 둘 중 하나의 분류를 예측 (0 or 1, Pass or fail ...)



\rightarrow Multinomial Classification (다중분류)

성적	학적	Grade
x_1	x_2	t
10	5	A
9	5	A
5	1	B
4	2	B
1	3	C



"Logistic"

✓

✓

Linear \Rightarrow Sigmoid

✓ A

✓ * input data

A

$$\begin{pmatrix} x_{01} & x_{02} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \\ \vdots & \end{pmatrix}$$

$$\cdot \begin{pmatrix} w_{A1} \\ w_{A2} \end{pmatrix} + b_A =$$

$$\begin{pmatrix} x_{11}w_{A1} + x_{12}w_{A2} + b_A \\ x_{21}w_{A1} + x_{22}w_{A2} + b_A \\ x_{31}w_{A1} + x_{32}w_{A2} + b_A \\ \vdots \\ x_{n1}w_{A1} + x_{n2}w_{A2} + b_A \end{pmatrix} \xrightarrow{\text{Sigmoid}}$$

B

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ \vdots & \end{pmatrix}$$

$$\cdot \begin{pmatrix} w_{B1} \\ w_{B2} \end{pmatrix} + b_B =$$

$$\begin{pmatrix} x_{11}w_{B1} + x_{12}w_{B2} + b_B \\ x_{21}w_{B1} + x_{22}w_{B2} + b_B \\ x_{31}w_{B1} + x_{32}w_{B2} + b_B \\ \vdots \\ x_{n1}w_{B1} + x_{n2}w_{B2} + b_B \end{pmatrix}$$

C

"

$$\cdot \begin{pmatrix} w_{C1} \\ w_{C2} \end{pmatrix} + b_C =$$

$$\begin{pmatrix} \vdots \\ \vdots \end{pmatrix}$$

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix}$$

$$\cdot \begin{pmatrix} w_{A1} & w_{B1} & w_{C1} \\ w_{A2} & w_{B2} & w_{C2} \end{pmatrix}$$

$$+ (b_A \ b_B \ b_C)$$

One-hot
One coding

$$(x_{11}w_{A1} + x_{12}w_{A2} + b_A, x_{11}w_{B1} + x_{12}w_{B2} + b_B, x_{11}w_C + x_{12}w_{C2} + b_C)$$

softmax

Sigmoid

