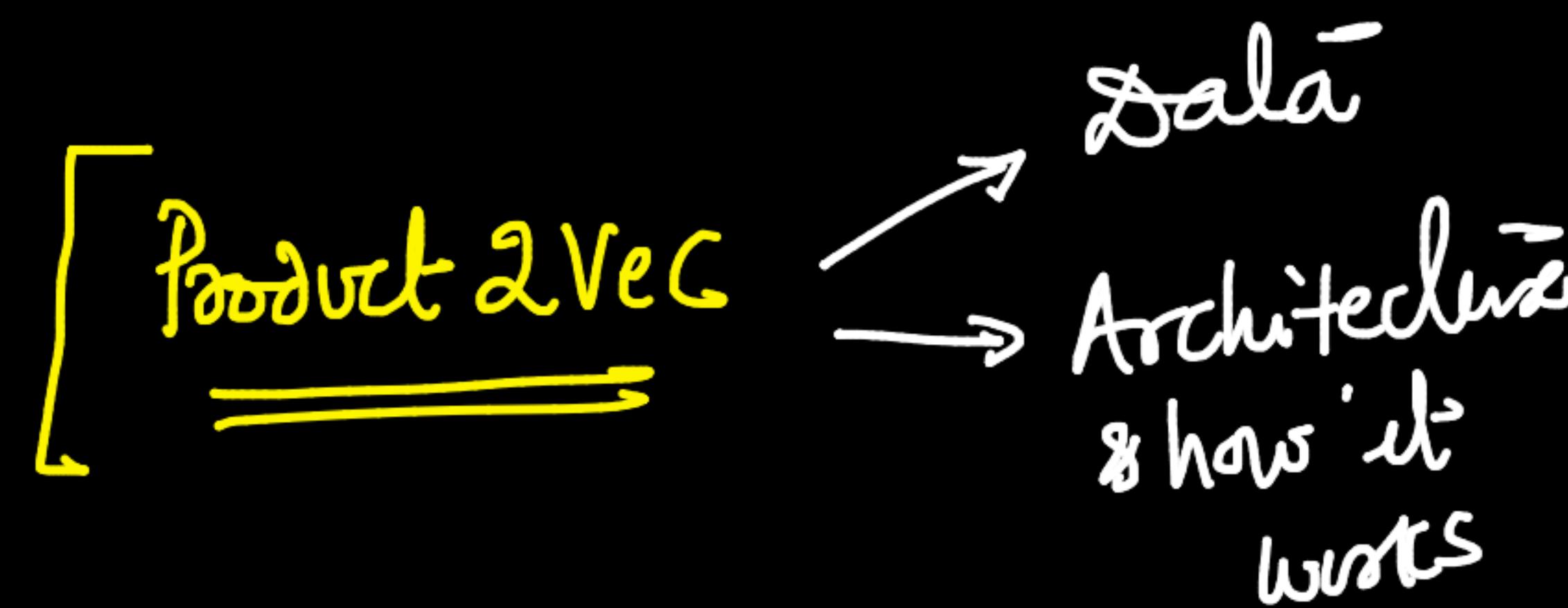
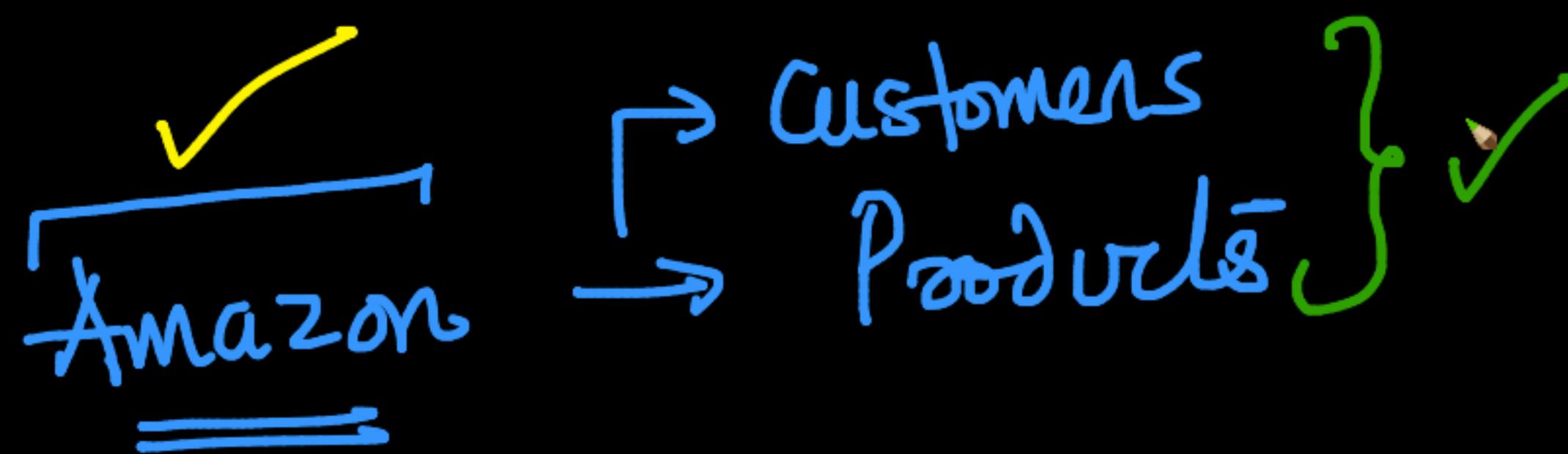


Q1



Scenarios-based

Tech:- MLPs,
W2Vec
AE
:

obj:- Similar products should be closer

a

Product-description (text)



Word2Vec [avg/tf-idf]

↳ not very well

Word2Vec

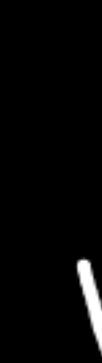
- text corpus
- context & focus words

Alt: sub-word embedding / fasttext



Products in a sequence:

$\{ \overbrace{P_1, P_2, P_3, \dots}^{\text{time stamps}} \}$



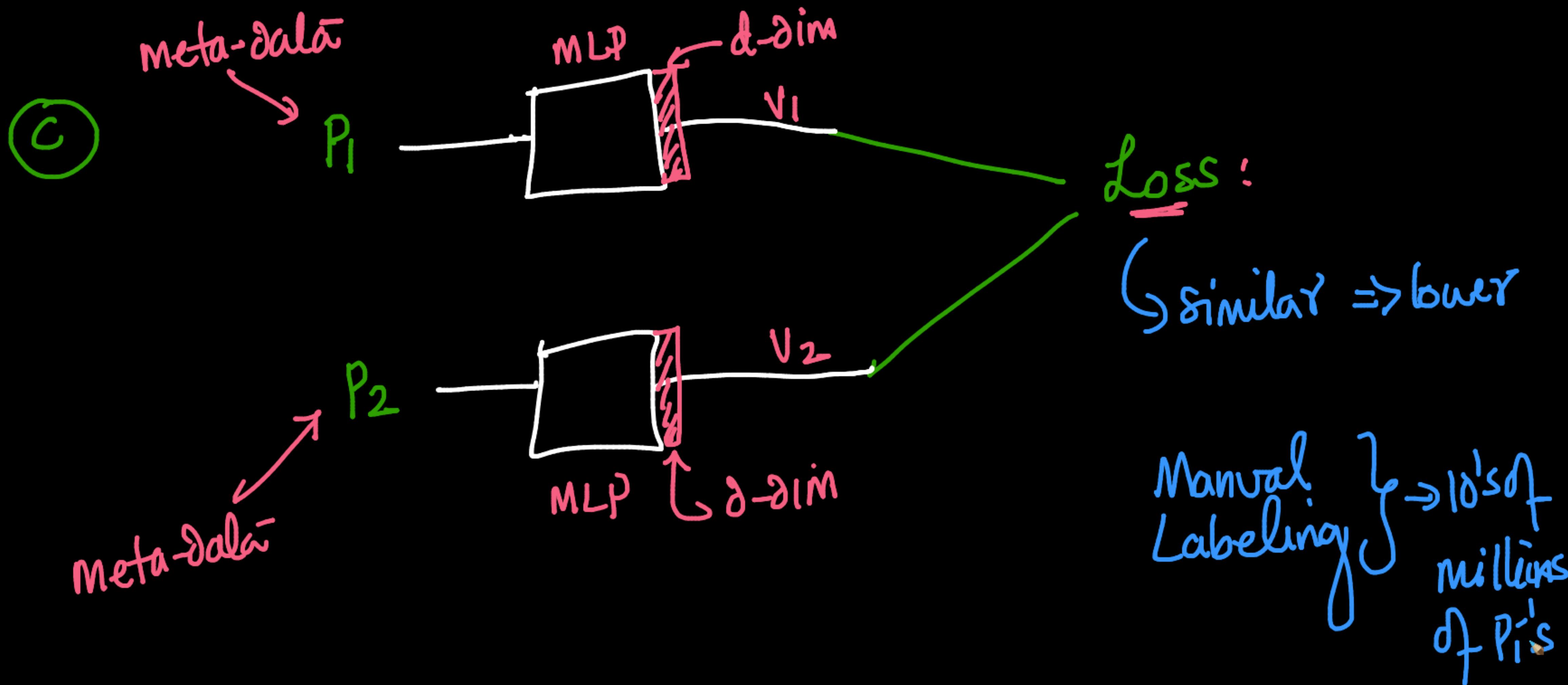
Similar to W2Vec

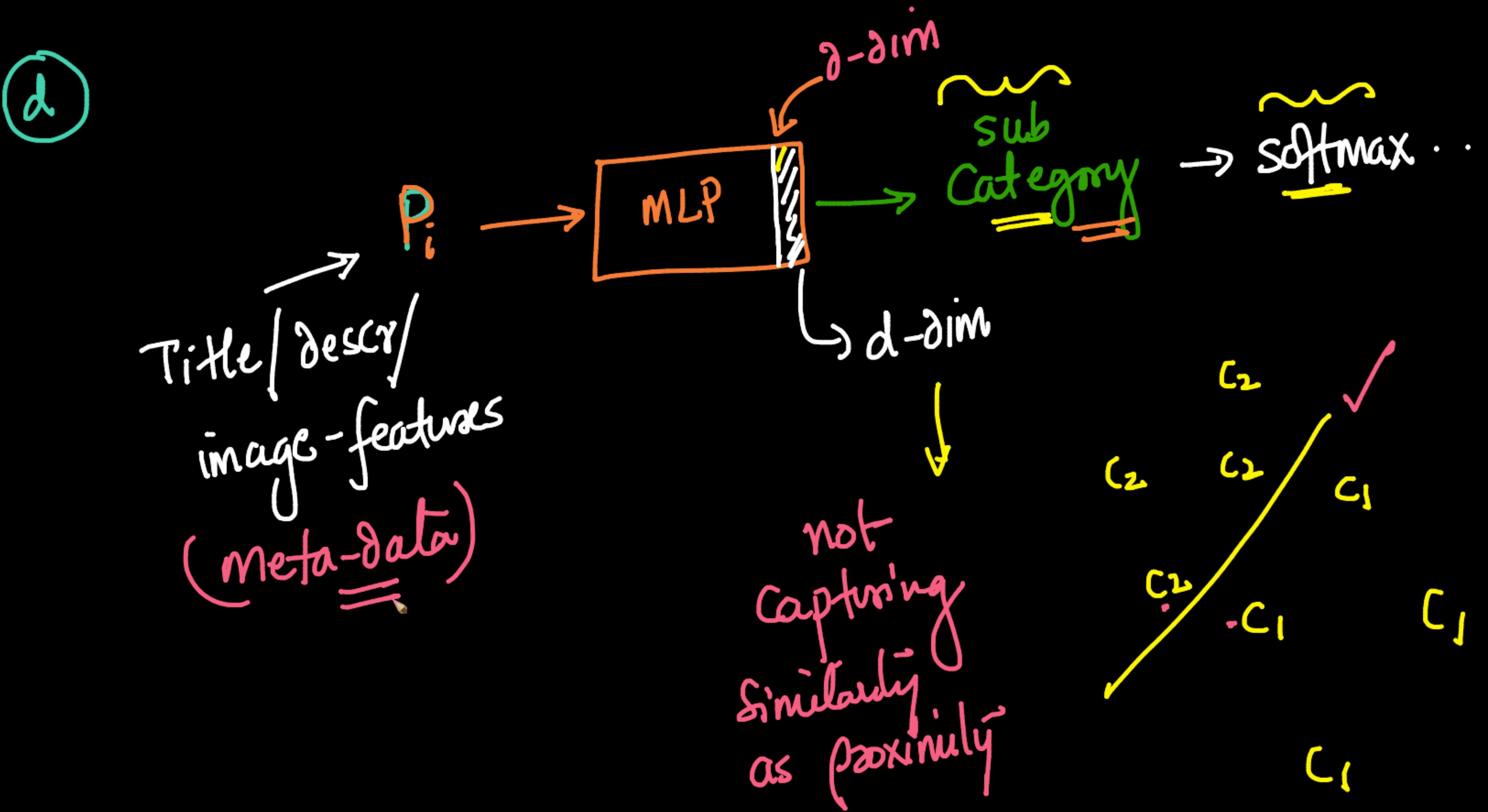
millions of these sequences

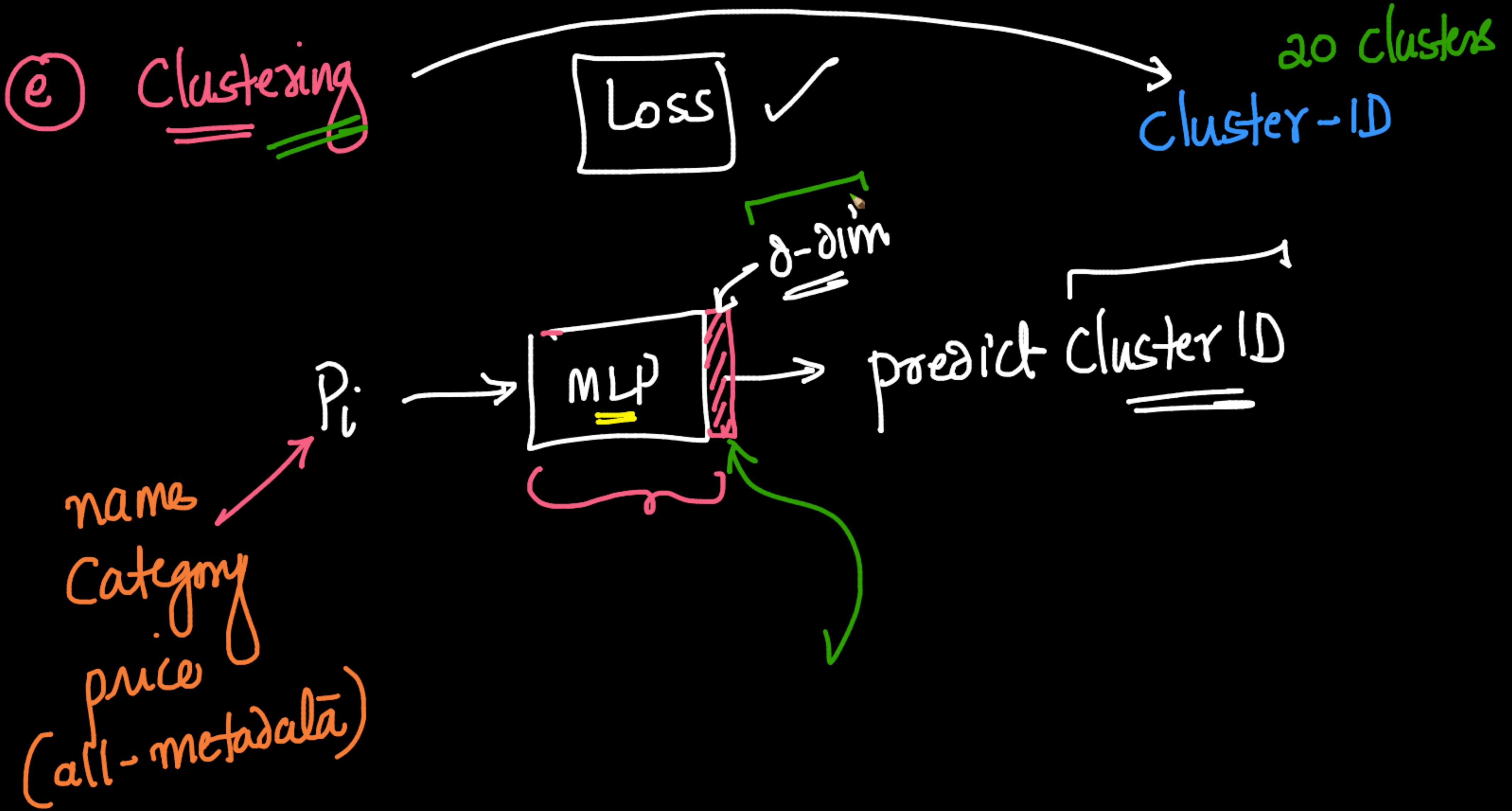
- ↳ sequences for family purchases
- ↳ not very reliable

$\{ \text{NOT } \underline{\text{sequence}}: \overbrace{\text{single-cart}}^{\text{Set}} \}$

↳ Hyp: each cart consists of similar items







f

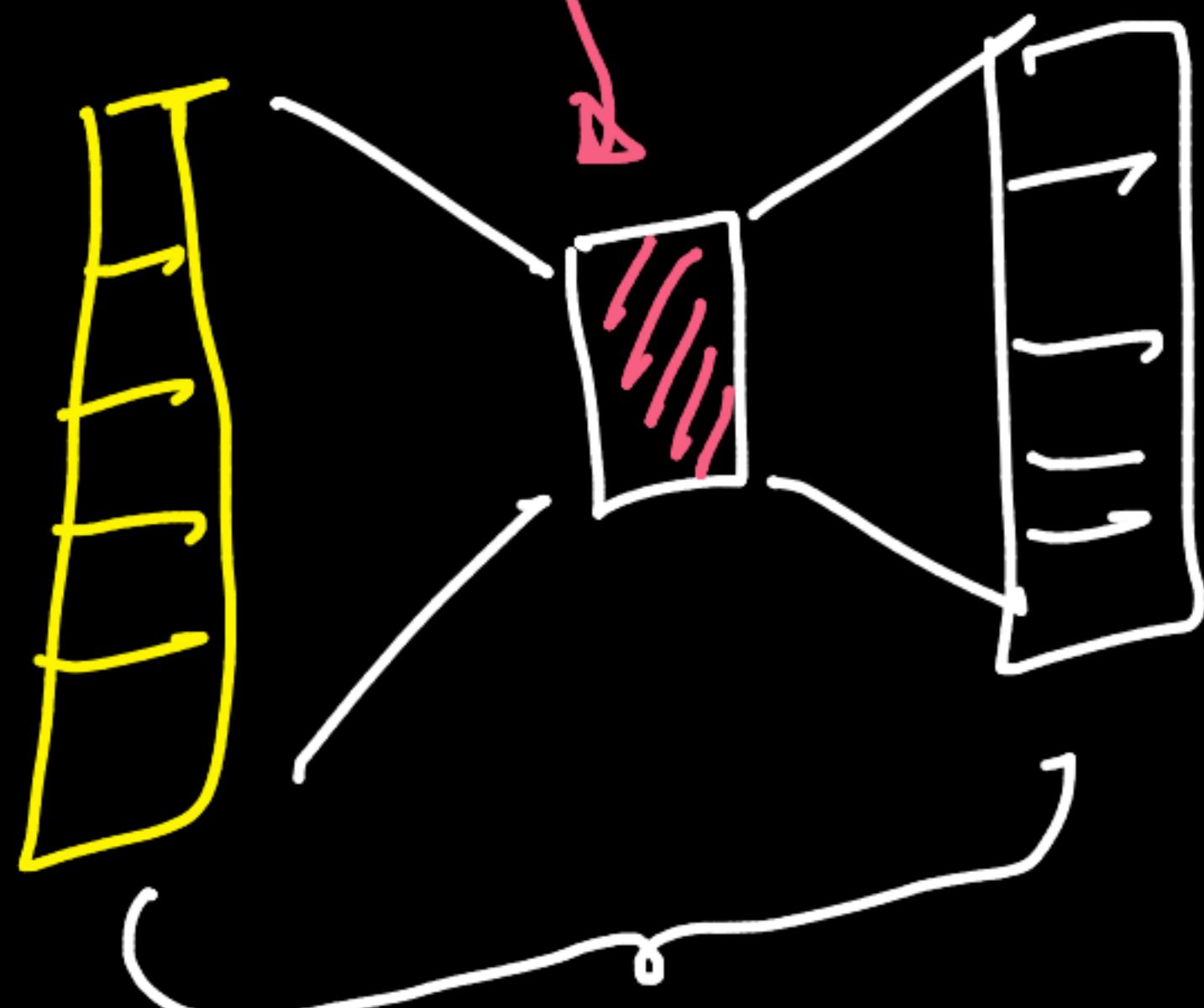
like sent2Vec

red polo t-shirt

Product-desc

↓
 $n \times m$ Matrix

word co-occurrence



(with one-layer)
AE

SVD:

$n \times d$ Matrix

⑨

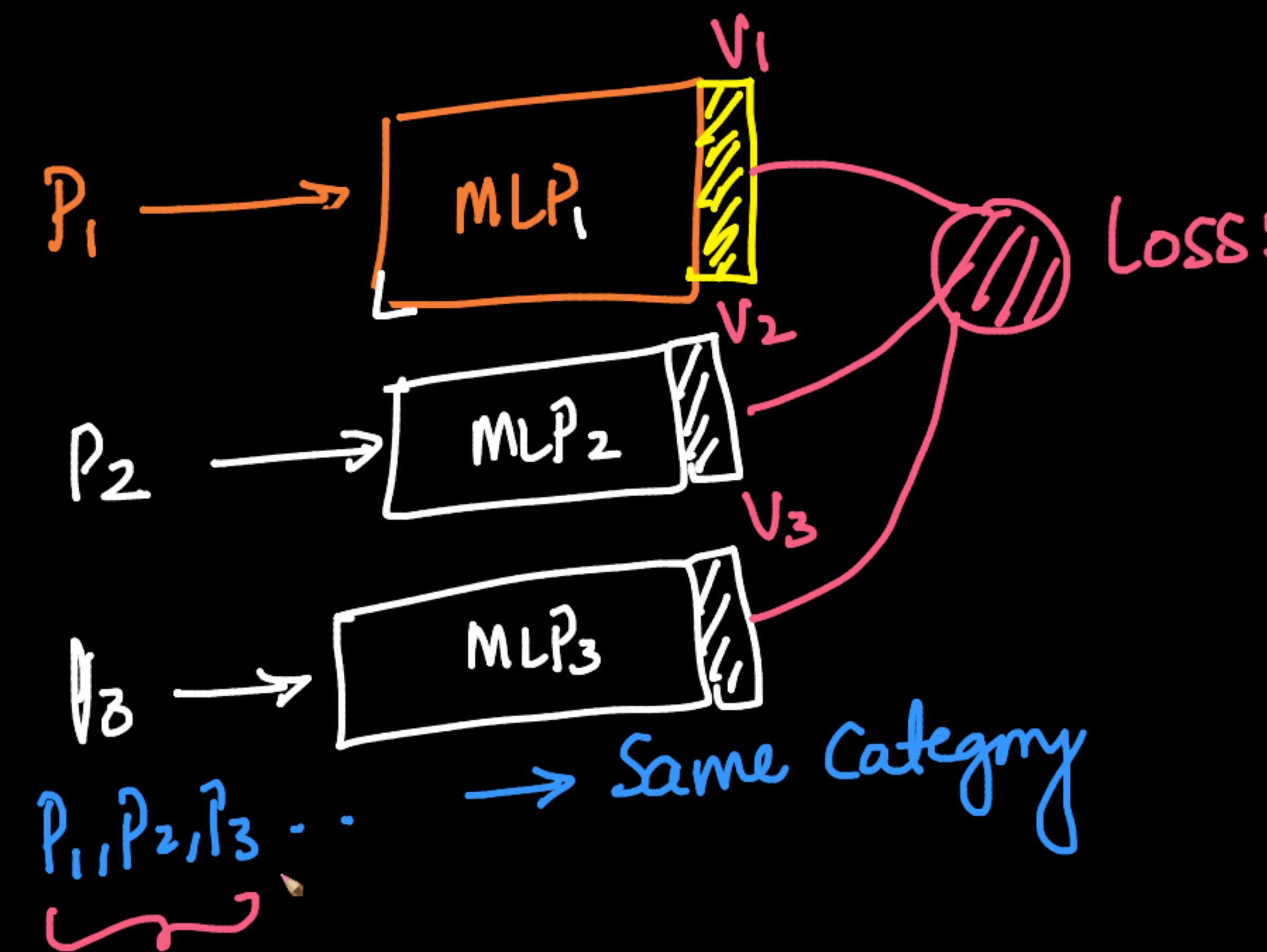
Sequence of products $\rightarrow \{ \text{NN in the metadata space} \}$

c_1 \downarrow *focus word*
 P_i $\underline{P_{10}}$ P_{12} P_{14}

\downarrow
[W2V-architecture] $\xrightarrow{\text{d-dim}}$
 P_{10} $\xrightarrow{\text{dense}}$

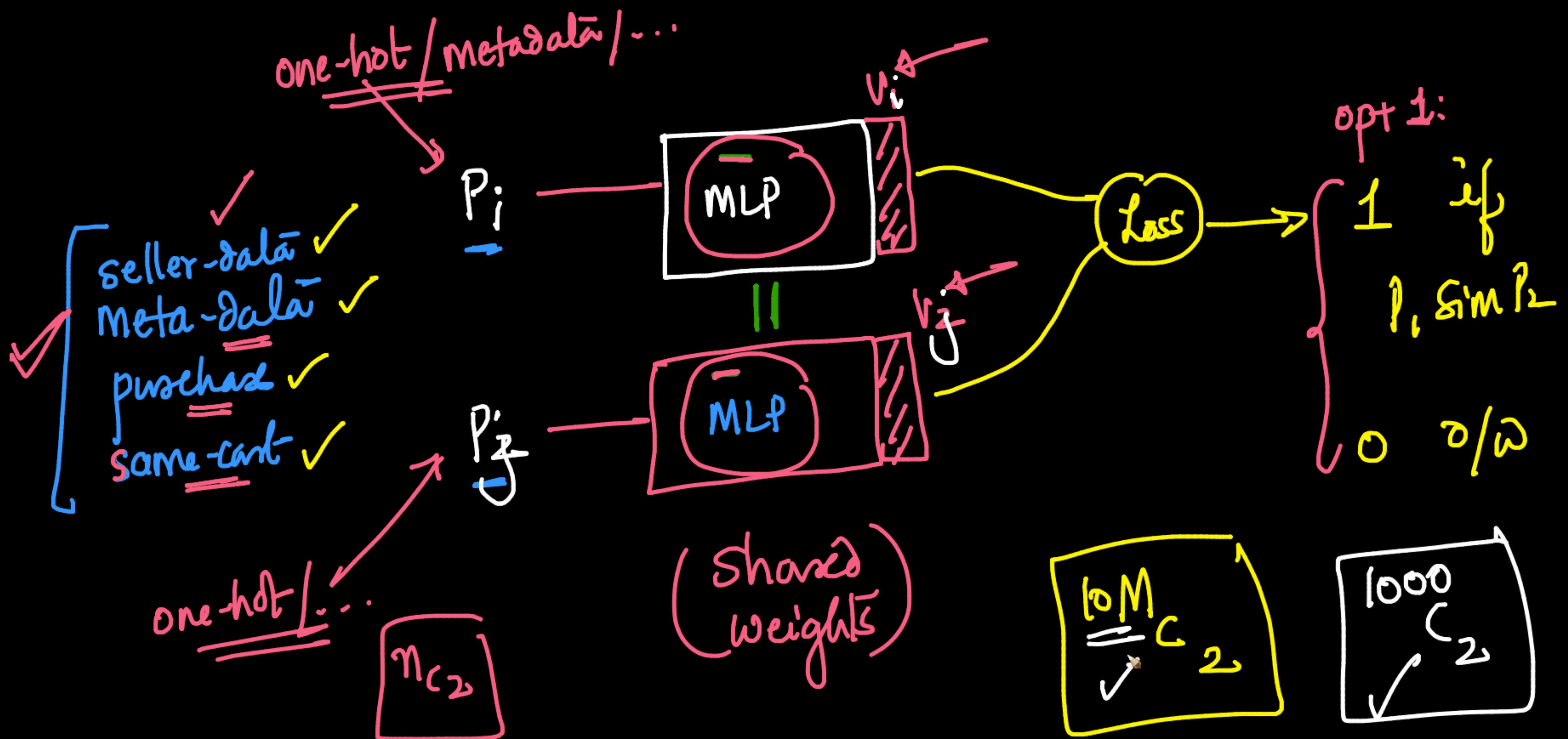
small-issue!

P_i, P_j
[$d(v_i, v_j)$,
 $\approx d(P_m, P_n)$]



$$\min \left[(V_1 - V_2) + (V_2 - V_3) + (V_1 - V_3) \right]$$

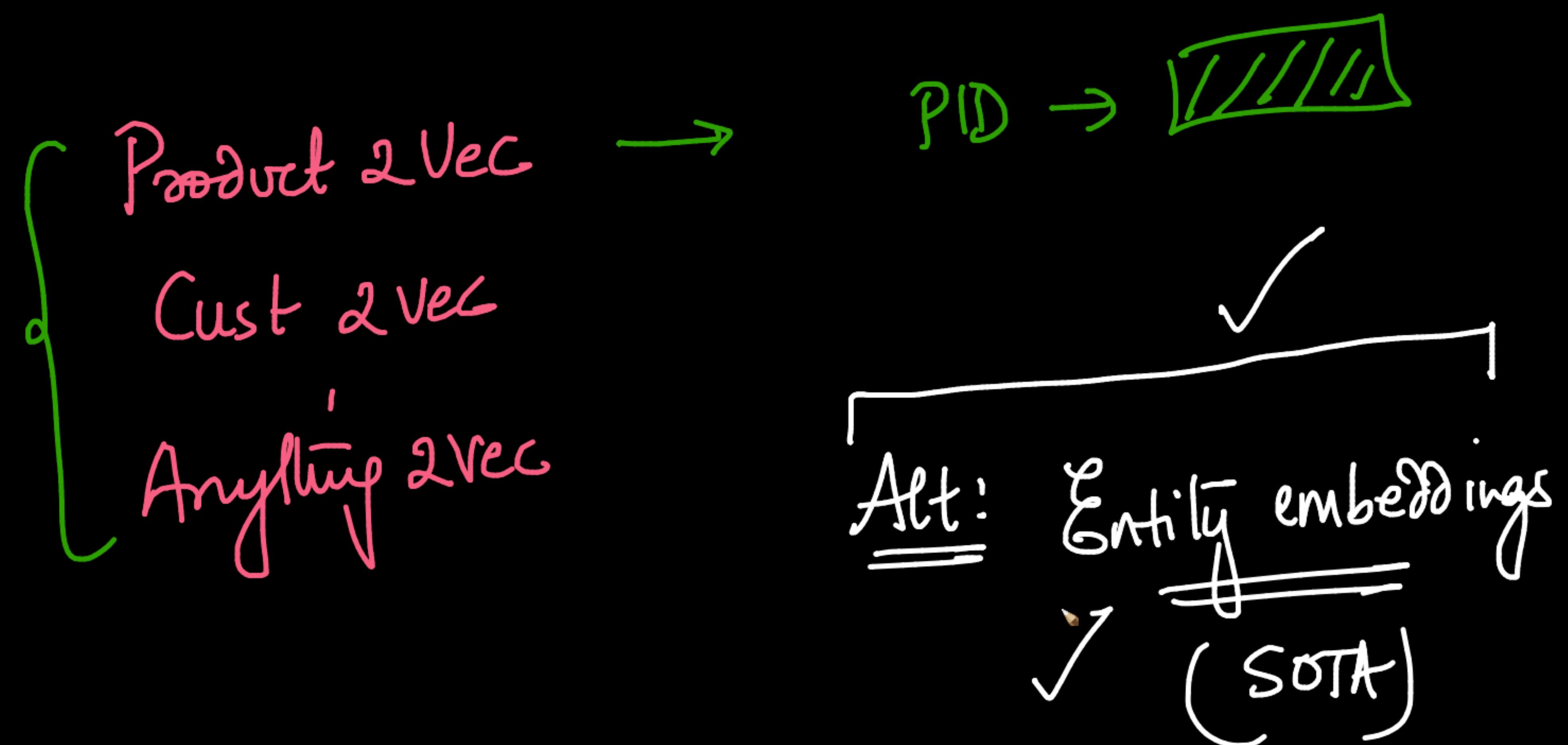


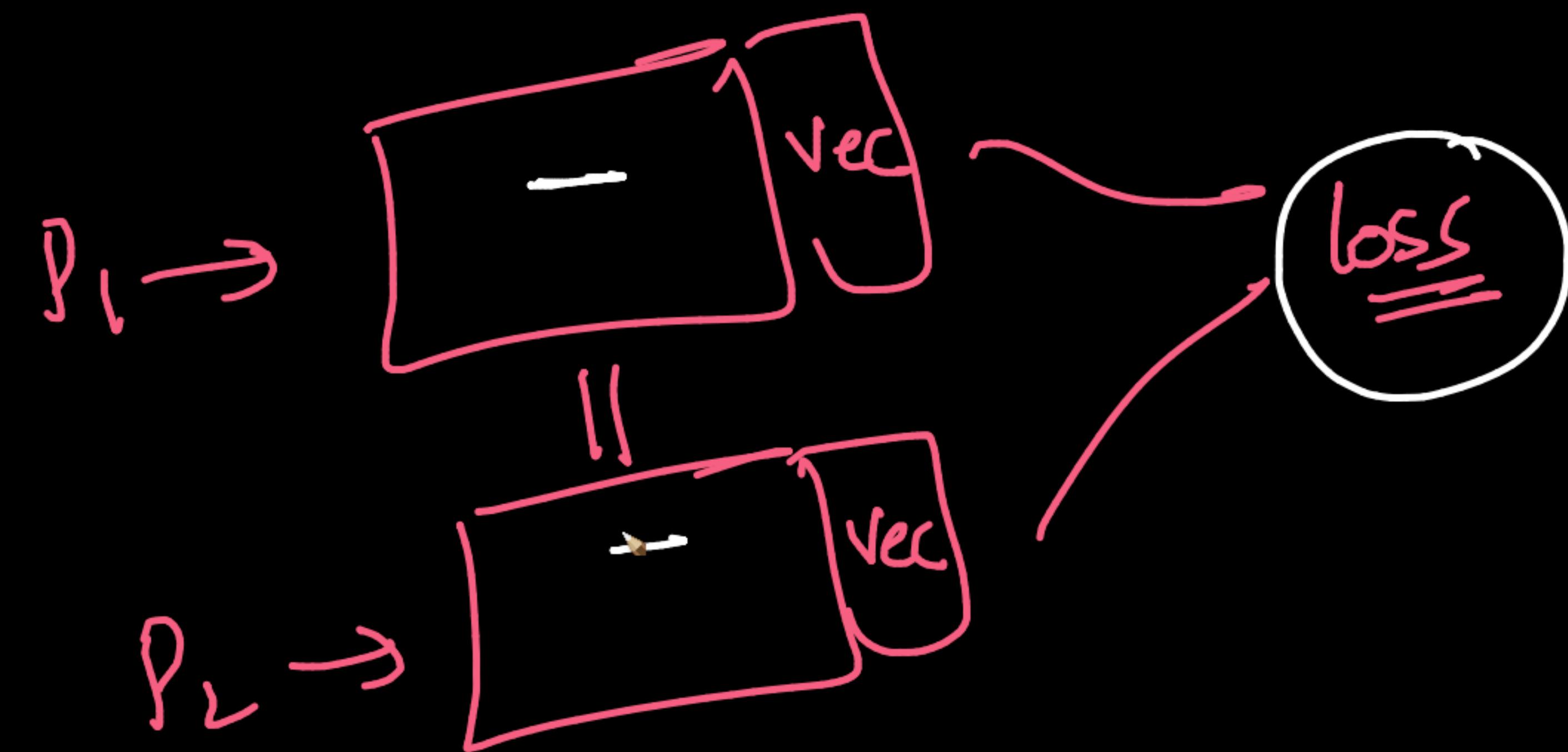


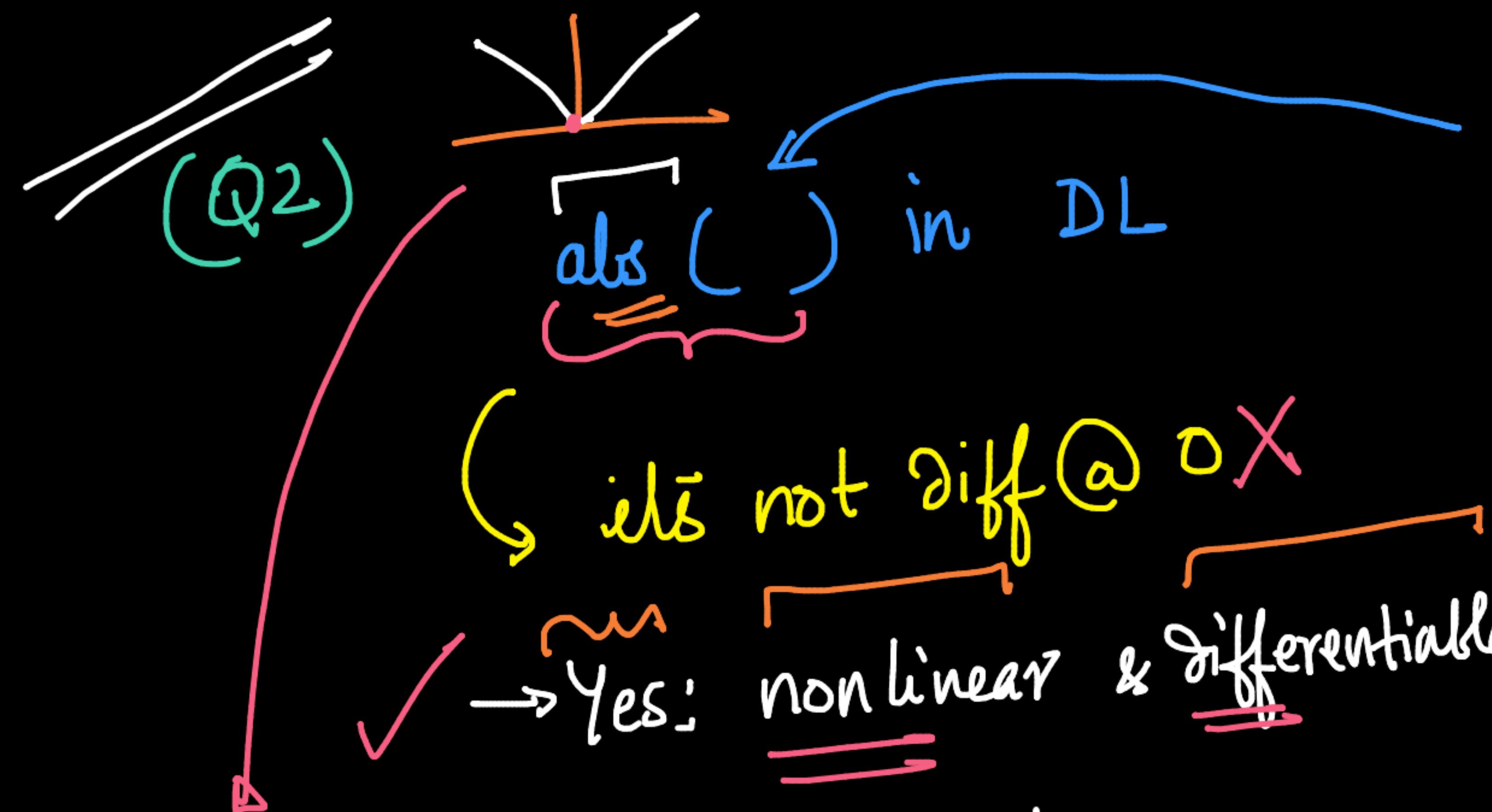
popular choice

Option 2

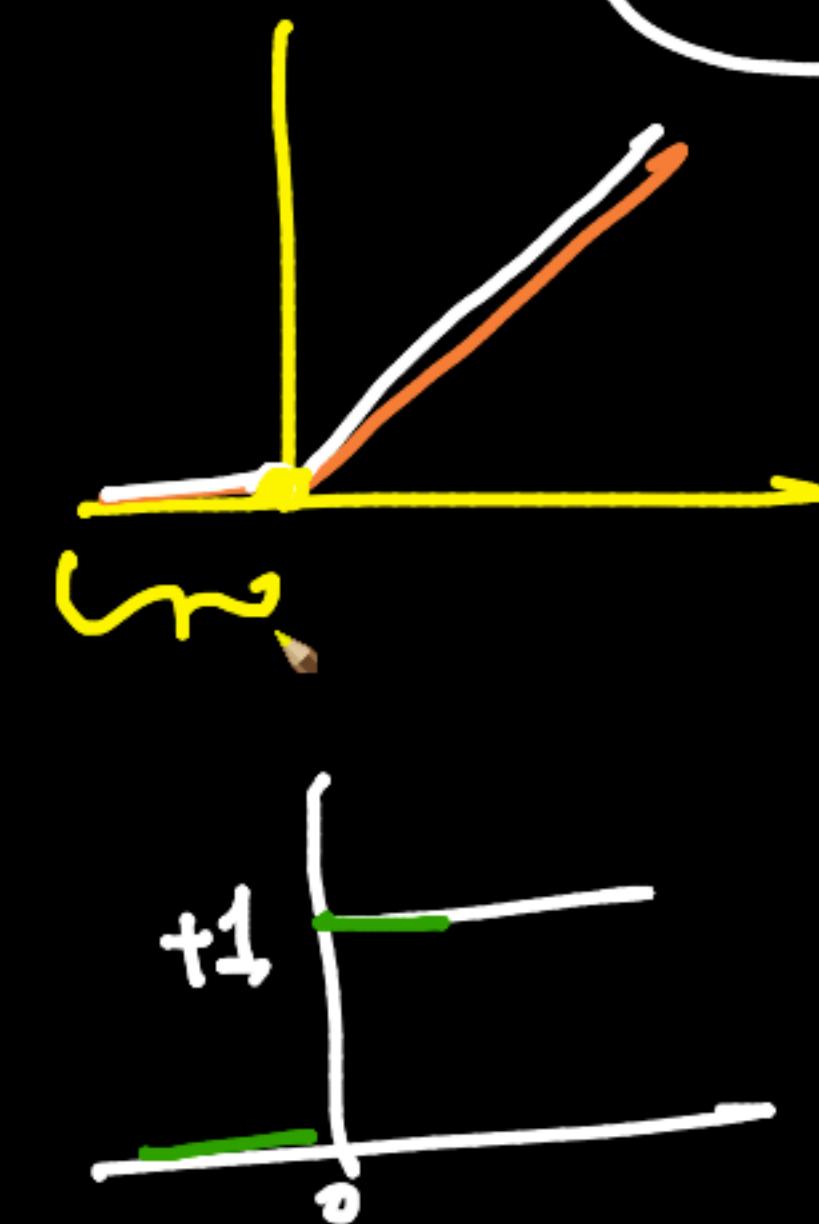
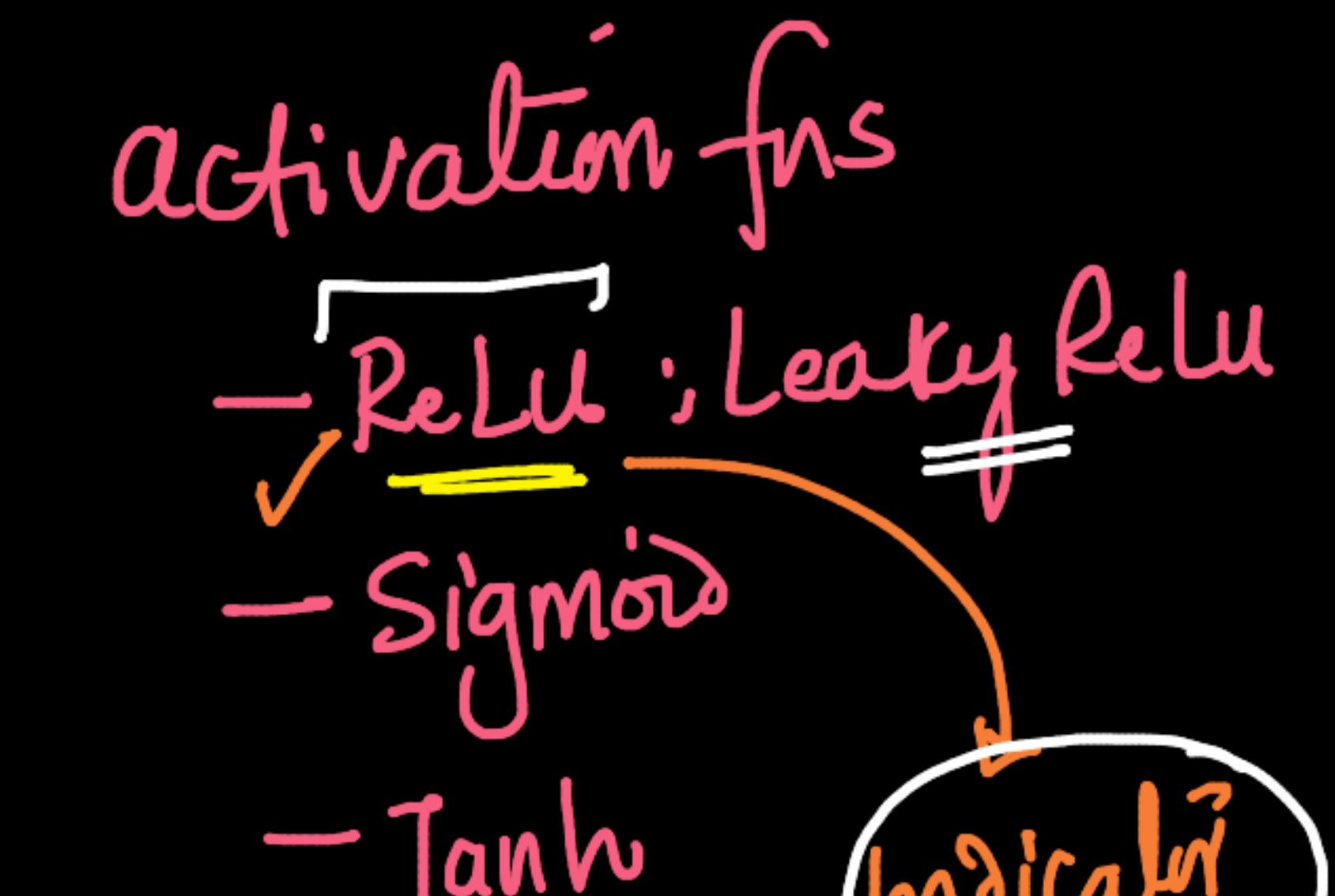
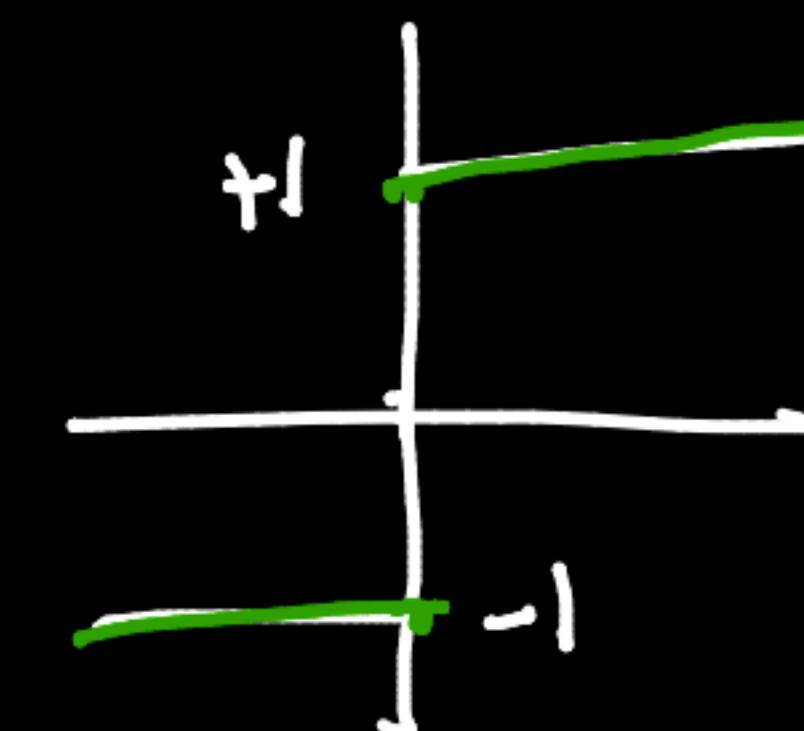
$\left\{ \begin{array}{l} \text{if } P_1 \leq P_2 \xrightarrow{\checkmark} \min d(v_1, v_2) \\ \text{else} \xrightarrow{} \min \sqrt{d(v_1, v_2)} \end{array} \right.$



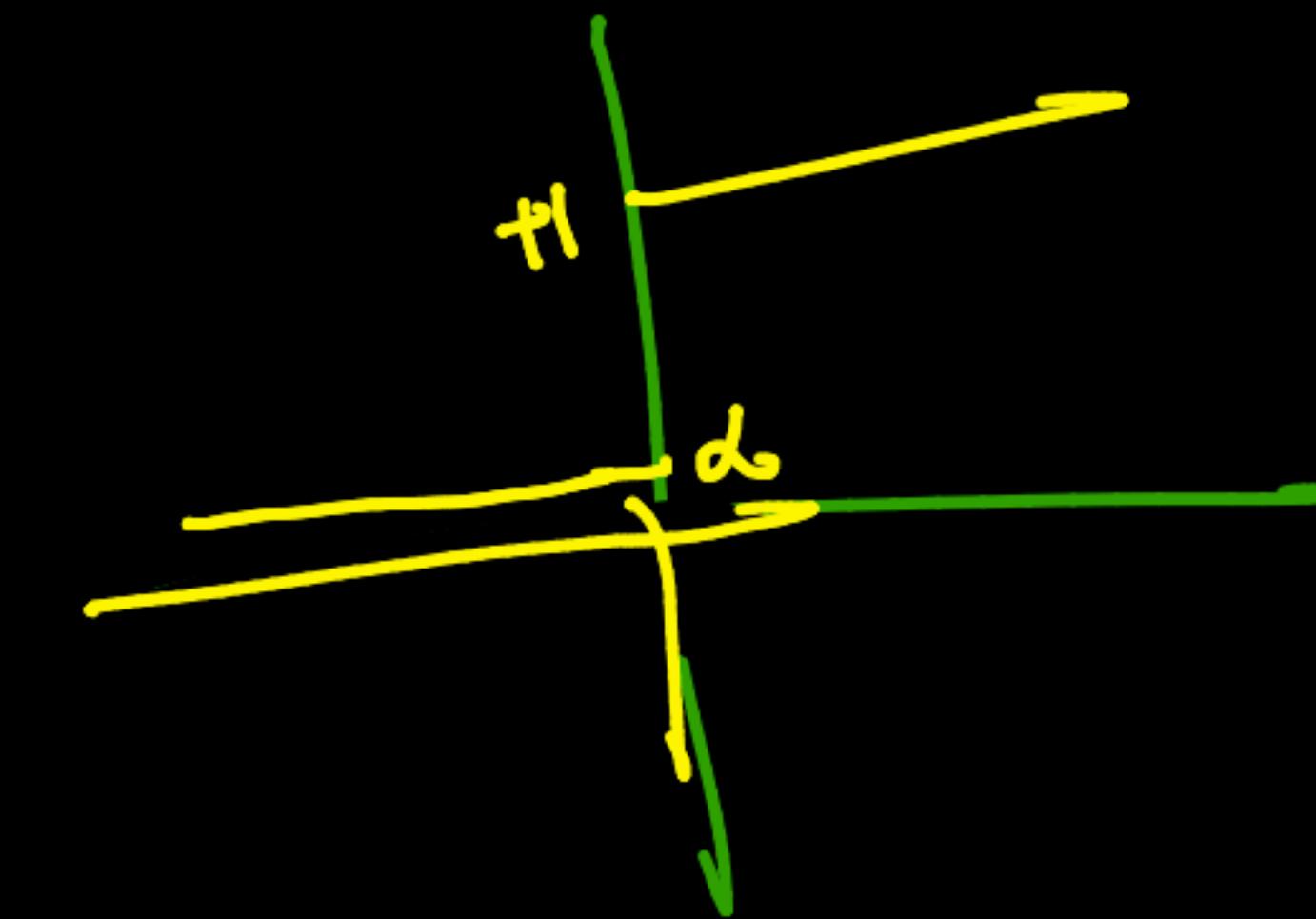
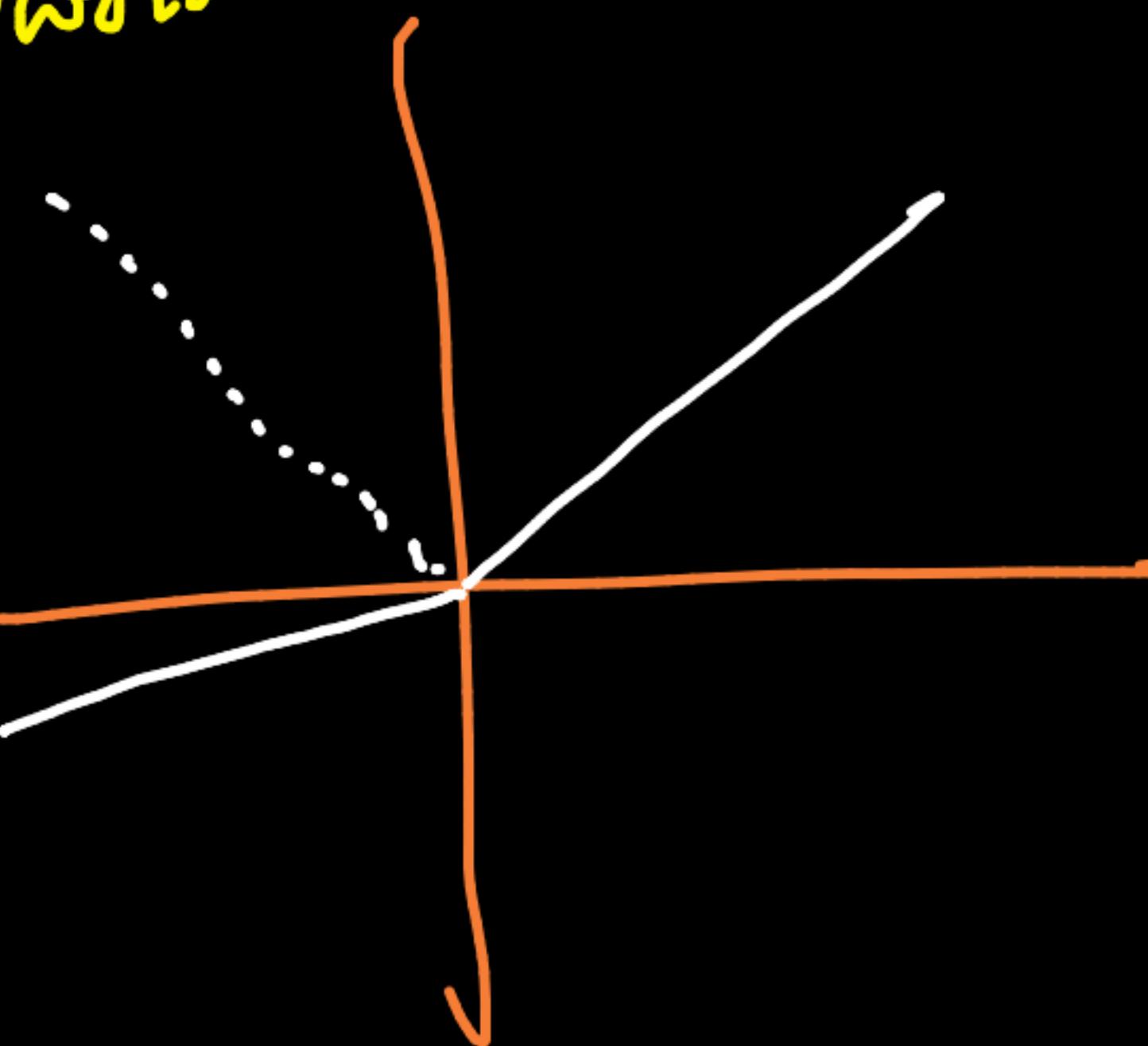
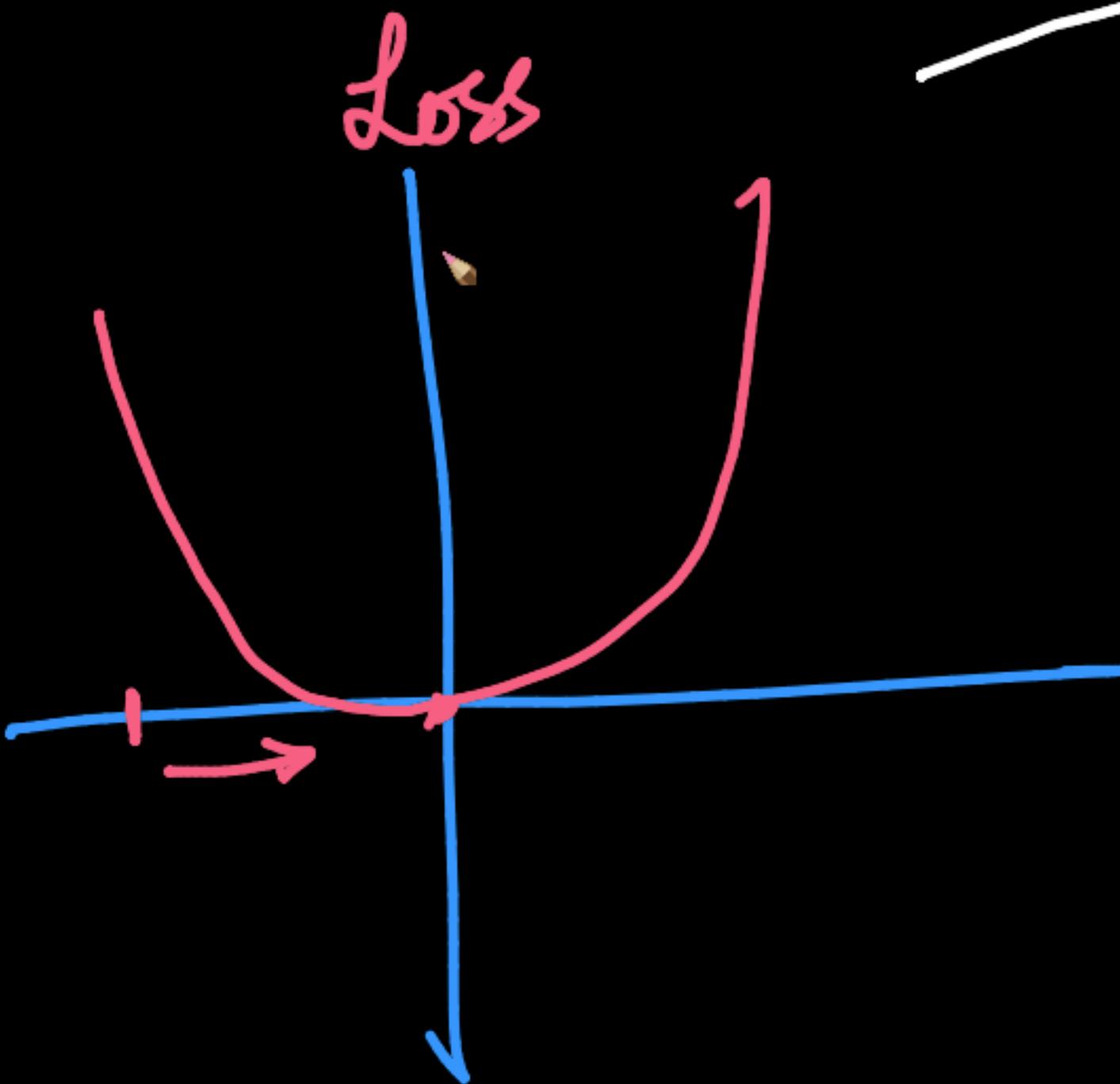


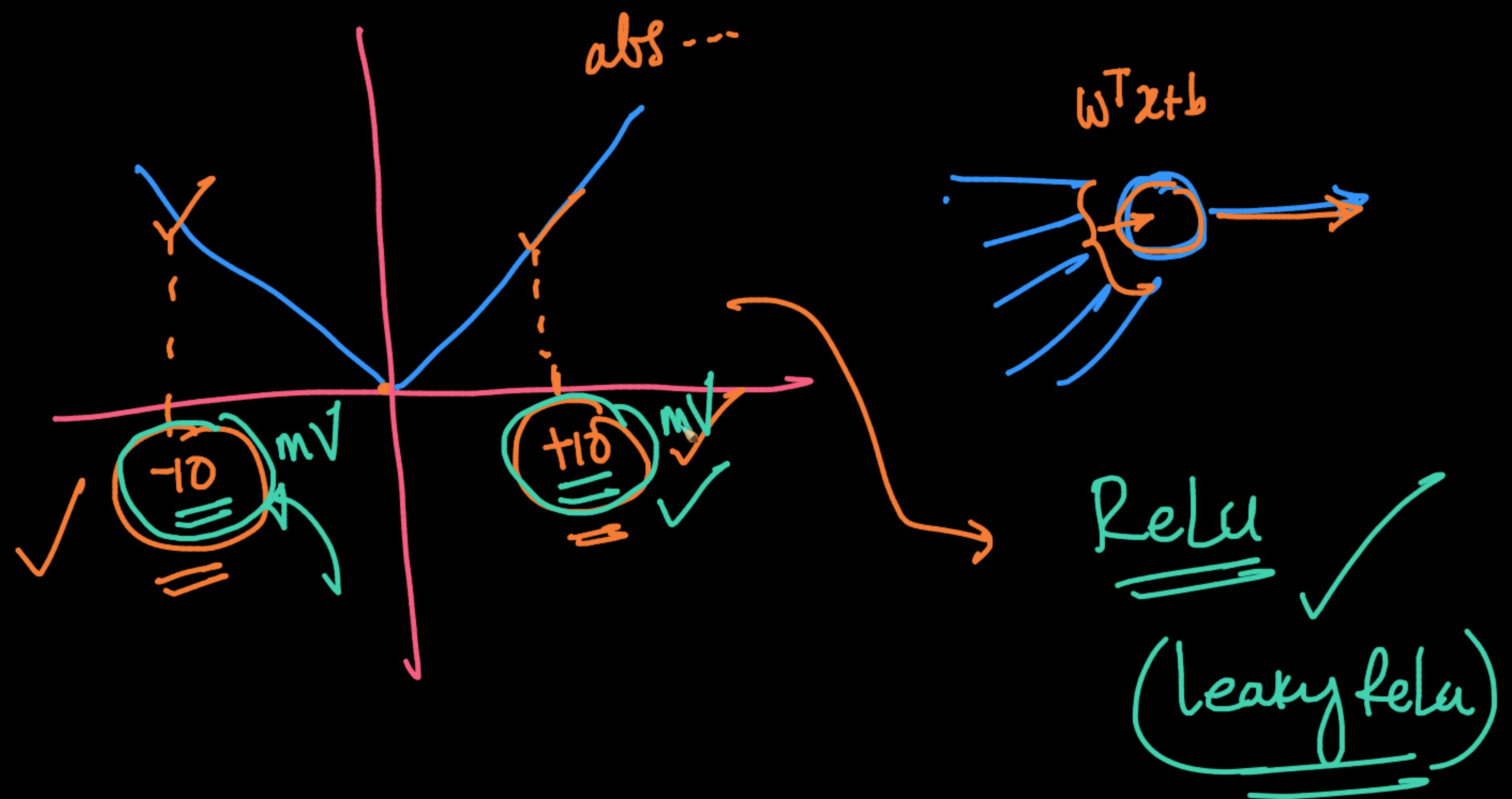


\times exploding gradient
Vanishing



$$W_{\text{new}} = W_{\text{old}} - \eta \sqrt{\frac{\partial L}{\partial w}} \Big|_{W_{\text{old}}}$$





Q3

d-dim

GD



$W_{10} \in \mathbb{R}^d$

$\nabla f = 0$

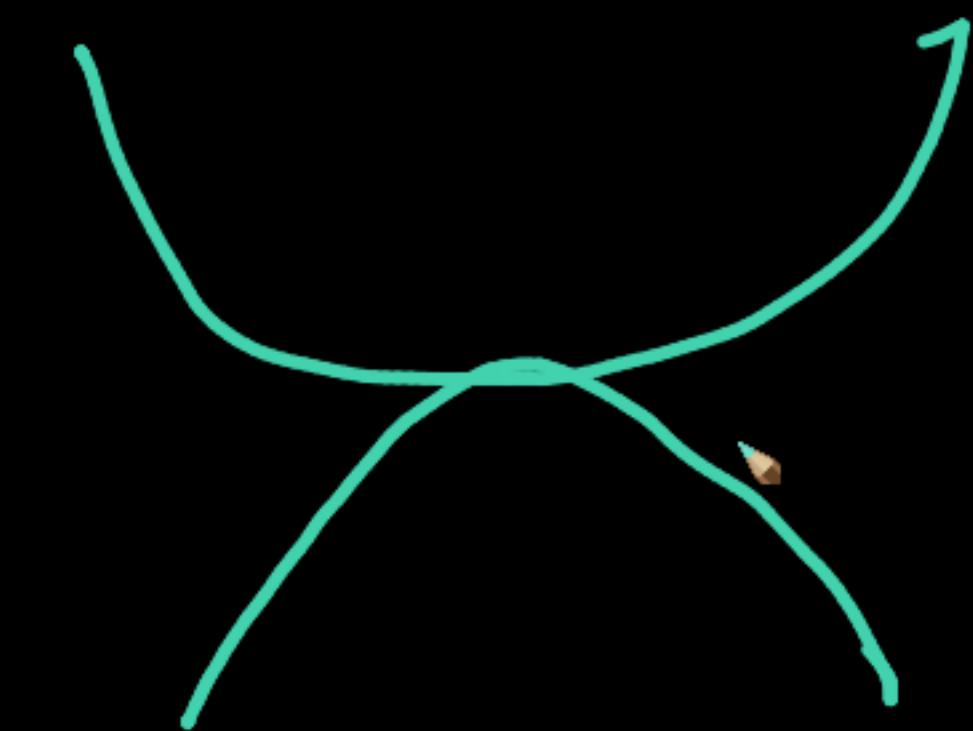
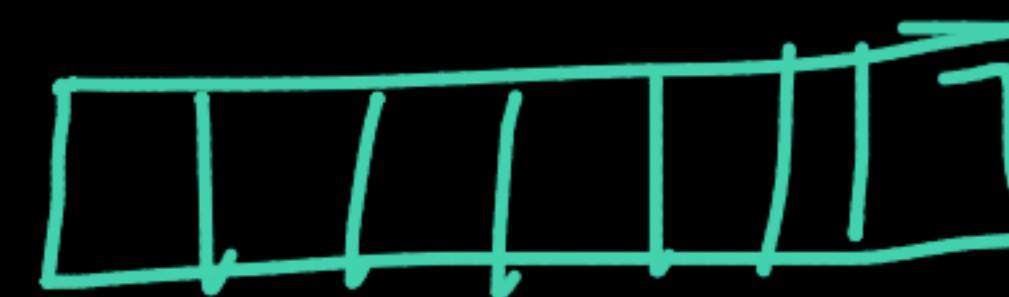
Codim 0

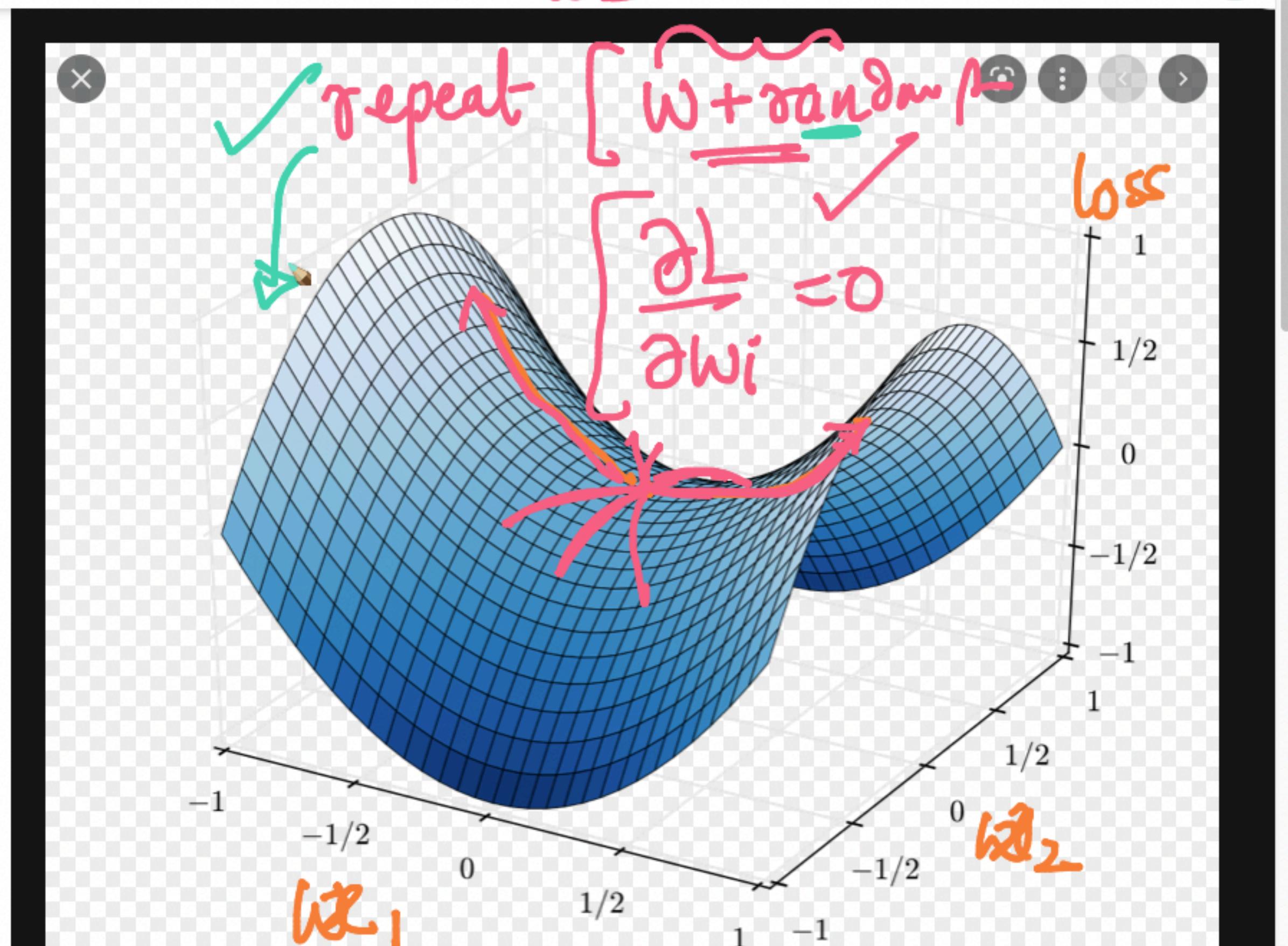
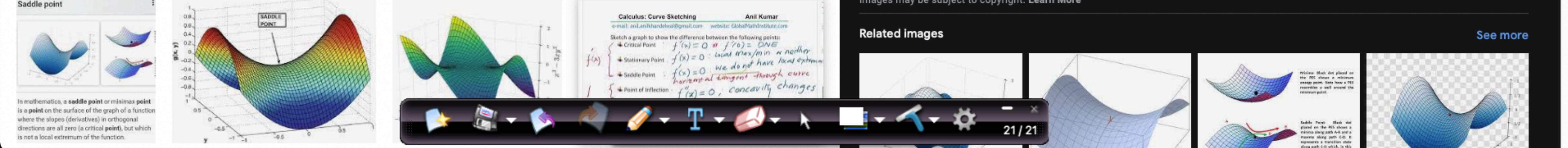
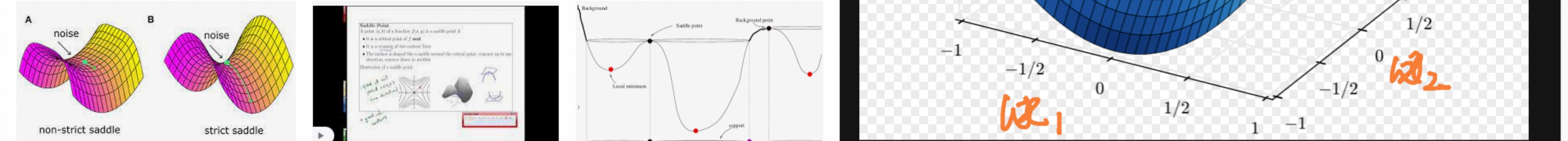
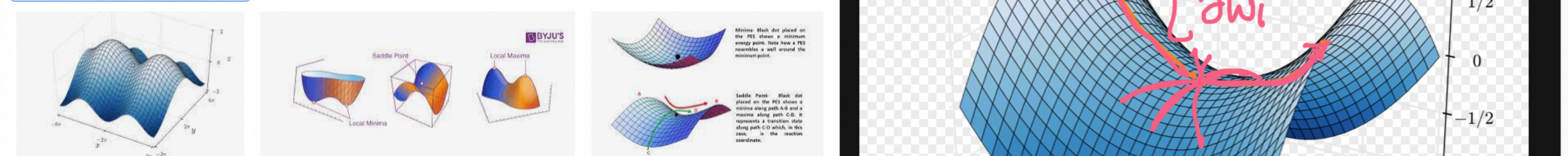
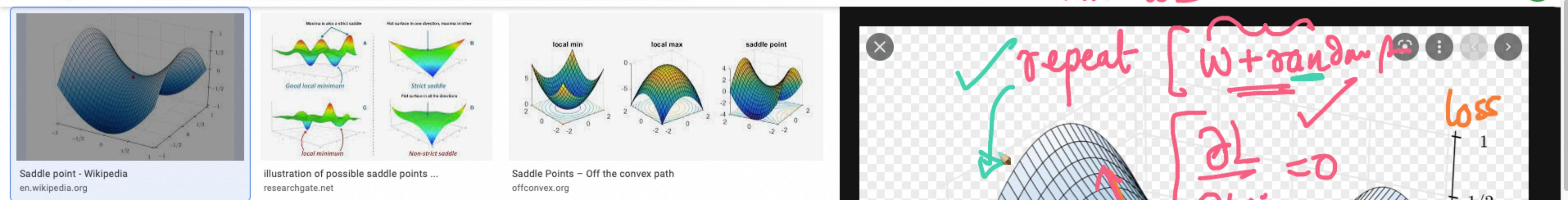
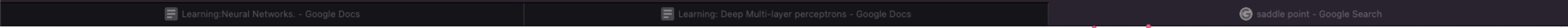
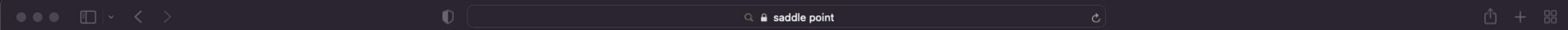
Saddle point
or not



2nd -derivative test

expensive



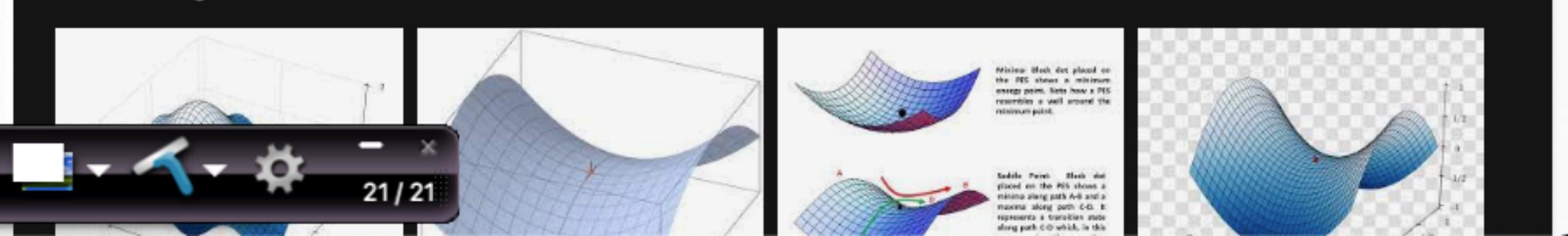


W Wikipedia

Saddle point - Wikipedia

Images may be subject to copyright. Learn More

Related images



Visit

See more

Learning:Neural Networks. - Google Docs | Learning: Deep Multi-layer perceptrons - Google Docs | saddle point - Google Search

Google saddle point

Saddle point - Wikipedia en.wikipedia.org

illustration of possible saddle points ... researchgate.net

Saddle Points – Off the convex path offconvex.org

Saddle point - Wikipedia en.wikipedia.org

Saddle Points in Multivariable Function byjus.com

File:Minima and Saddle Point.png ... en.wikipedia.org

Noise Helps Optimization Escape From ... frontiers.org

Definition of grey-level blobs from ... researchgate.net

W Wikipedia Saddle point - Wikipedia

Images may be subject to copyright. Learn More

Related images

See more

Calculus: Curve Sketching Anil Kumar e-mail: anil.khandelwal@gmail.com website: GlobalMathInstitute.com

Sketch a graph to show the difference between the following points:

- Critical Point: $f'(x) = 0$ or $f'(x)$ is DNE
- Stationary Point: $f'(x) = 0$: local max/min or neither
- Saddle Point: $f'(x) = 0$; we do not have local extrema horizontal tangent through curve
- Point of Inflection: $f''(x) = 0$; concavity changes

In mathematics, a saddle point or minimax point is a point on the surface of the graph of a function where the slopes (derivatives) in orthogonal directions are all zero (a critical point), but which is not a local extremum of the function.

22 / 22