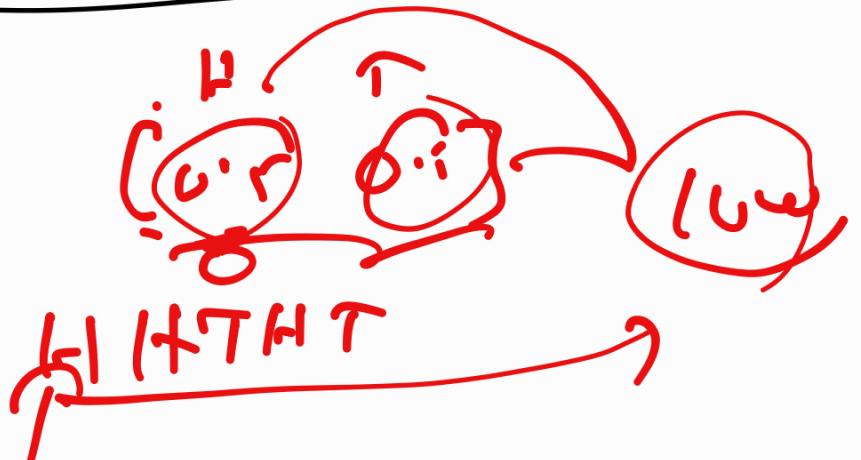
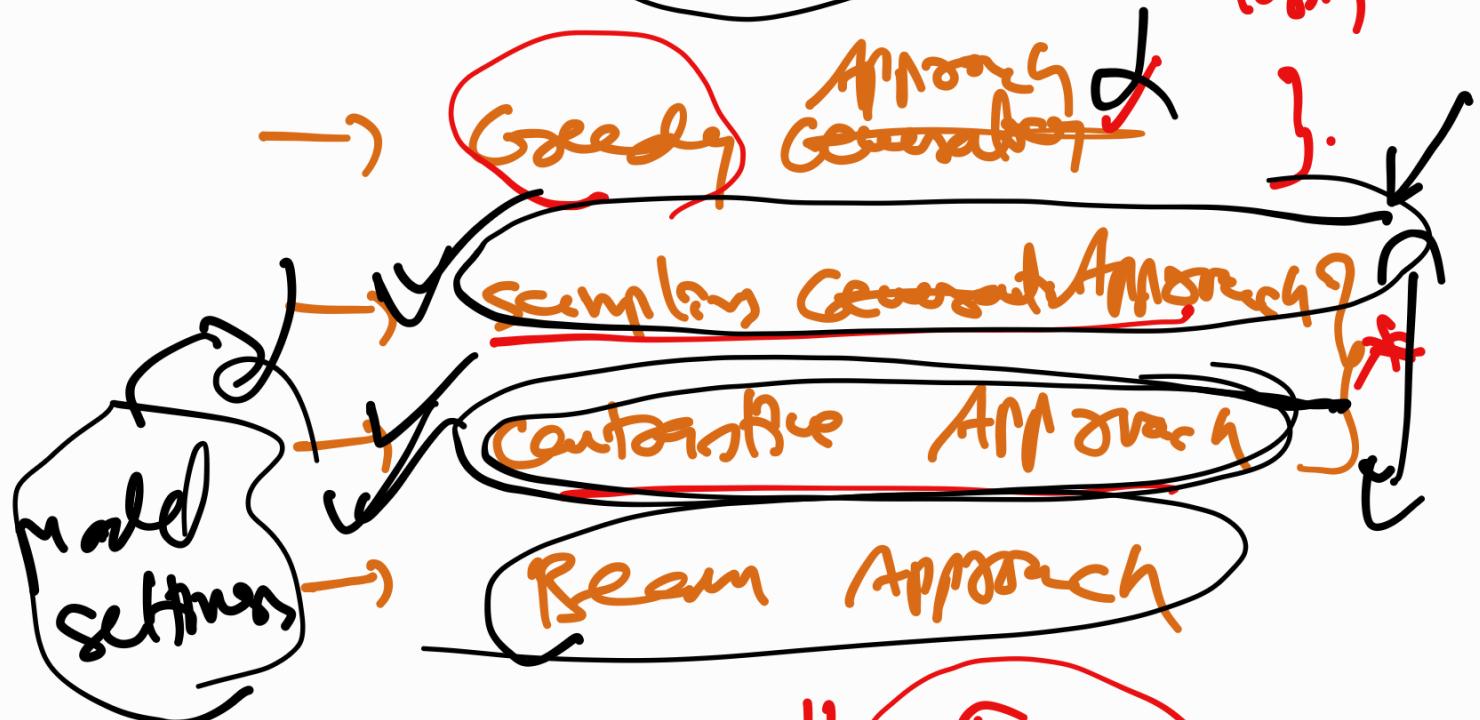
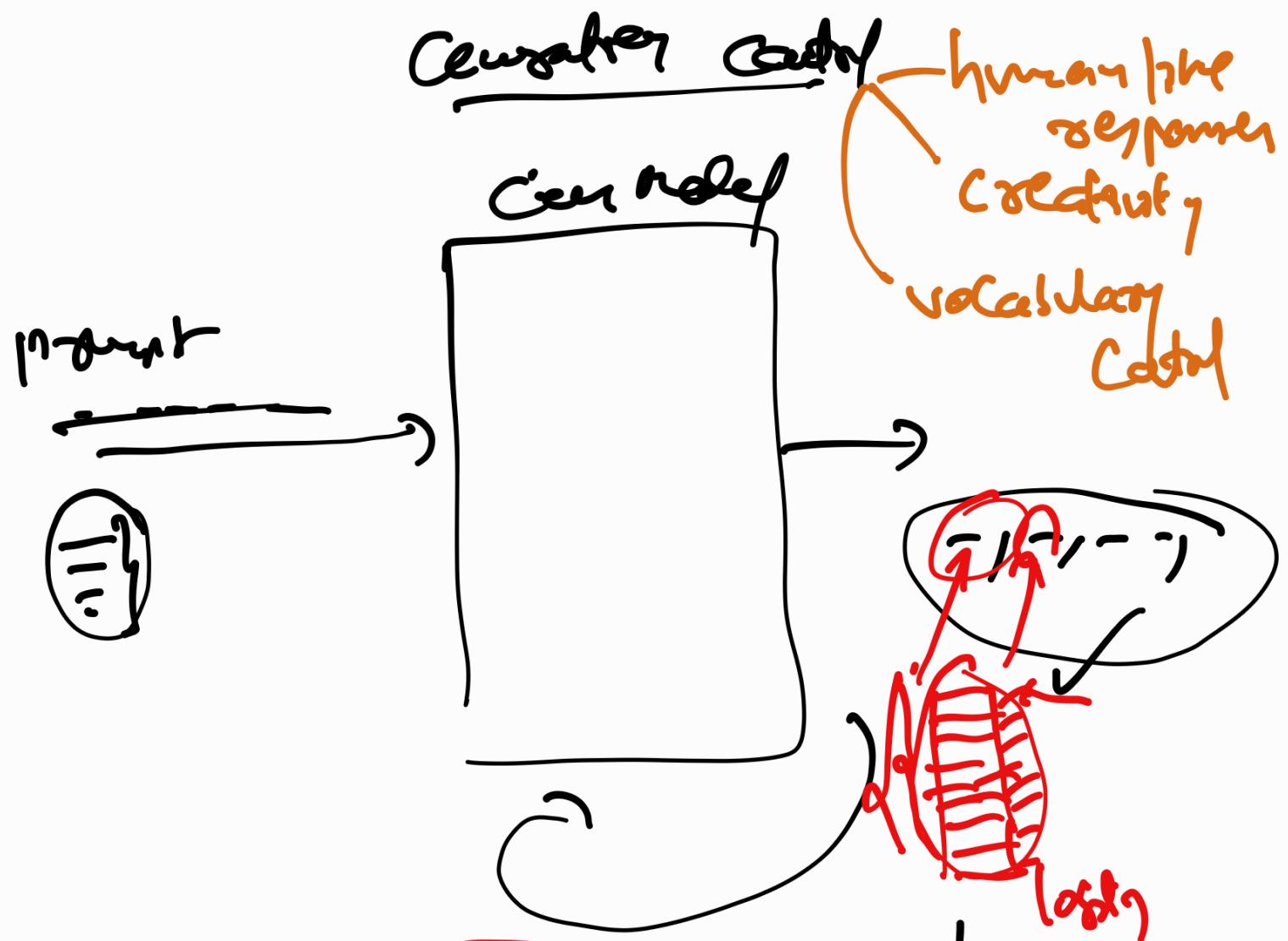




- NN Board,
 - Python Board,
 - NN for predictive modeling
 - NN learning
 - NLP model (sequential)
 - ↳ Harry Potter r-finder
 - ↳ control about grammar
 - GPT2
 - ↳ E =
 - GPT2 - hf
- Complex version
- euler. deGolr
decolor w/o
recolor w/o



Sampling Approach

- 1) Introduce Decoding / random sampling
- 2) ~~no. of tokens filtering~~
- 3) modify lost
 - type (K)
 - type (P) (nucleus)
 - temperature

try-k → pick the ~~format k-~~
tokens with
current k possibilities/
tries

tip-p → pick the smallest no. of tokens who sum
 $\hat{t} \quad t.$ scorer = p

A hand-drawn graph in red ink on white paper. A horizontal line represents time, with tick marks labeled '01 - 3ar' (likely referring to 1970), 'Ca' (likely referring to 1980), and '1990'. Above the line, a series of vertical red strokes of decreasing height represent a value that starts high at the first tick mark and gradually declines towards the end of the line.

$$\text{score} = p$$

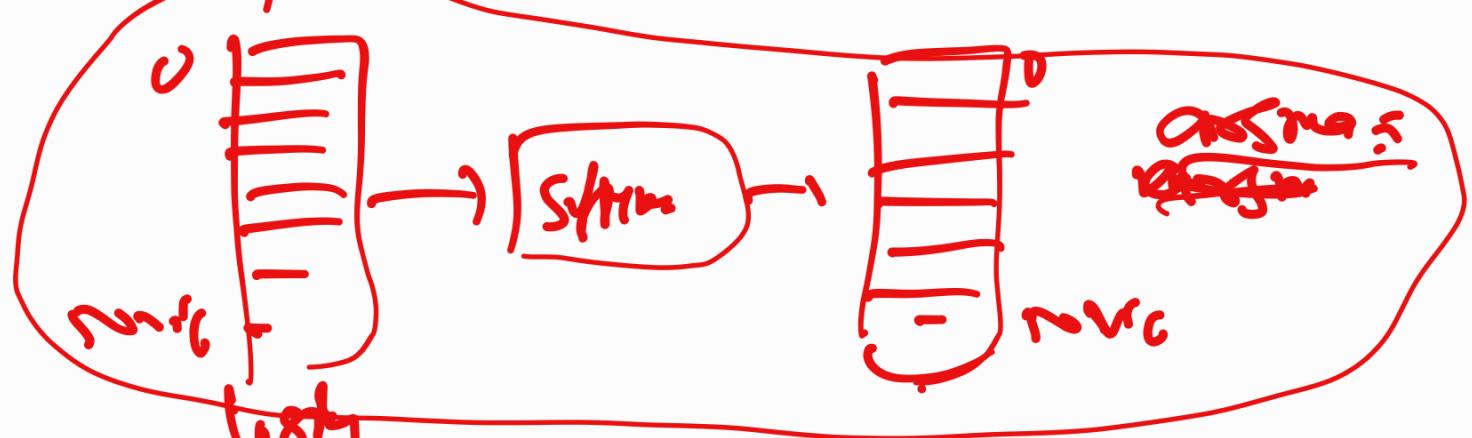
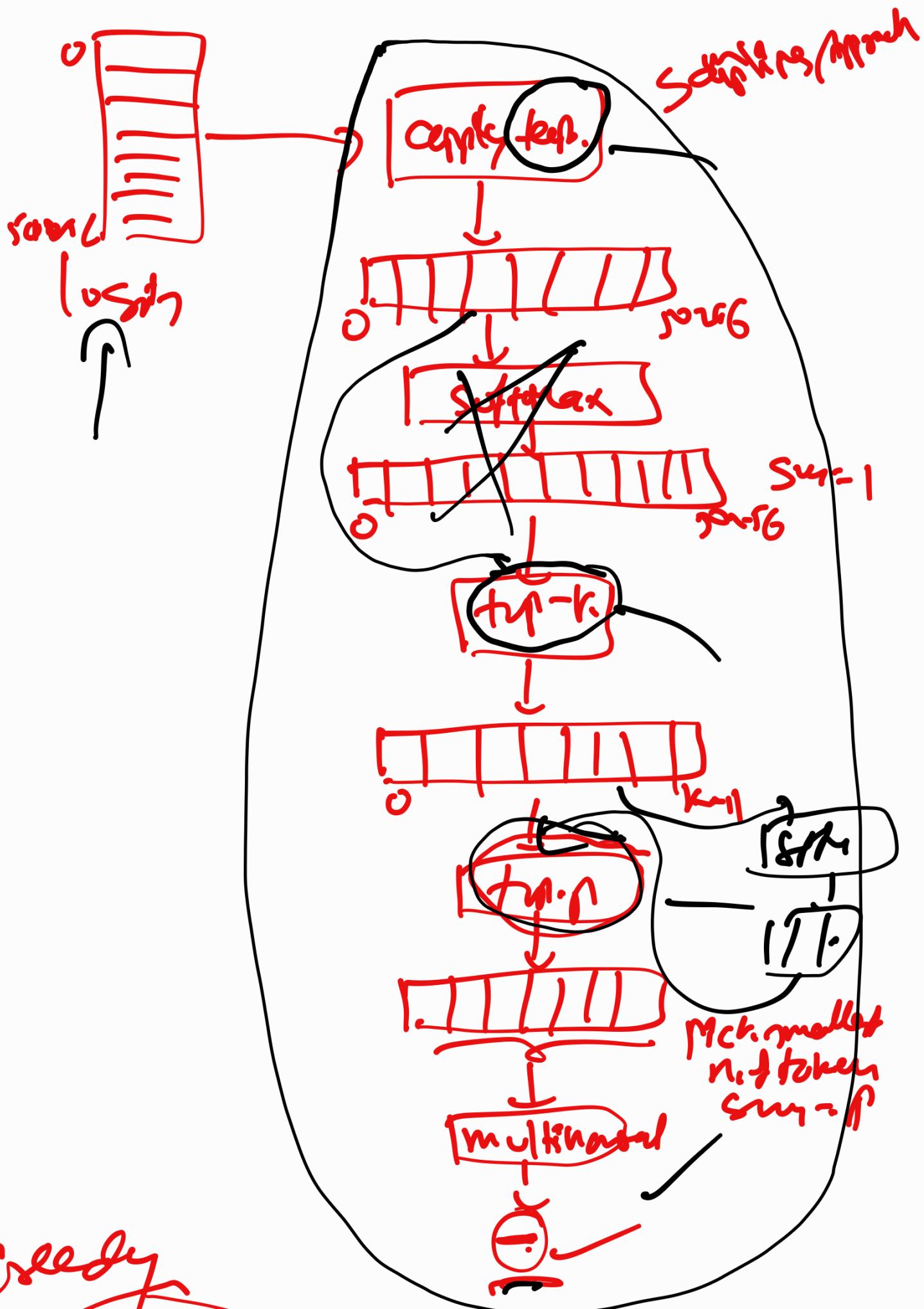
A hand-drawn diagram in red ink illustrating a surgical procedure. The top part shows a vertical tube labeled 'Ust.' (likely urethra) with an arrow pointing to the right labeled 'Sutura' (suture). To the right is another vertical tube labeled 'mos' (likely mosquito) with a curved arrow pointing upwards. The bottom part shows a vertical tube with two circular structures labeled 'bot' (likely balloon) attached to it. An arrow points from this tube to the right, labeled 'Sutura' (suture), which then points to a vertical tube labeled 'mos' (likely mosquito).

4 T
c/c off

$$\begin{array}{l} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array}$$

$$\frac{e^{\alpha}}{e^{\alpha}-1}$$

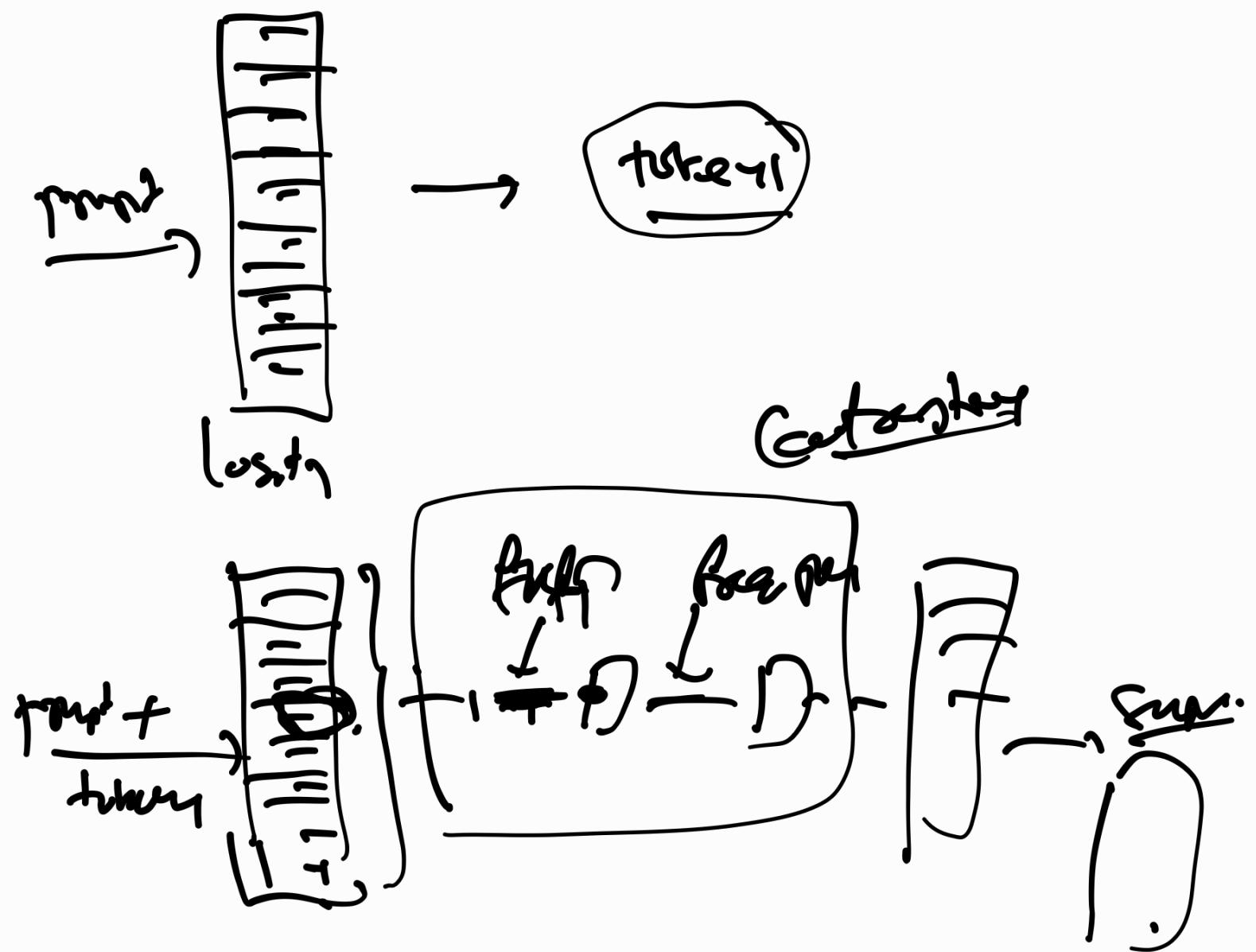
C → S
210
100%
100%



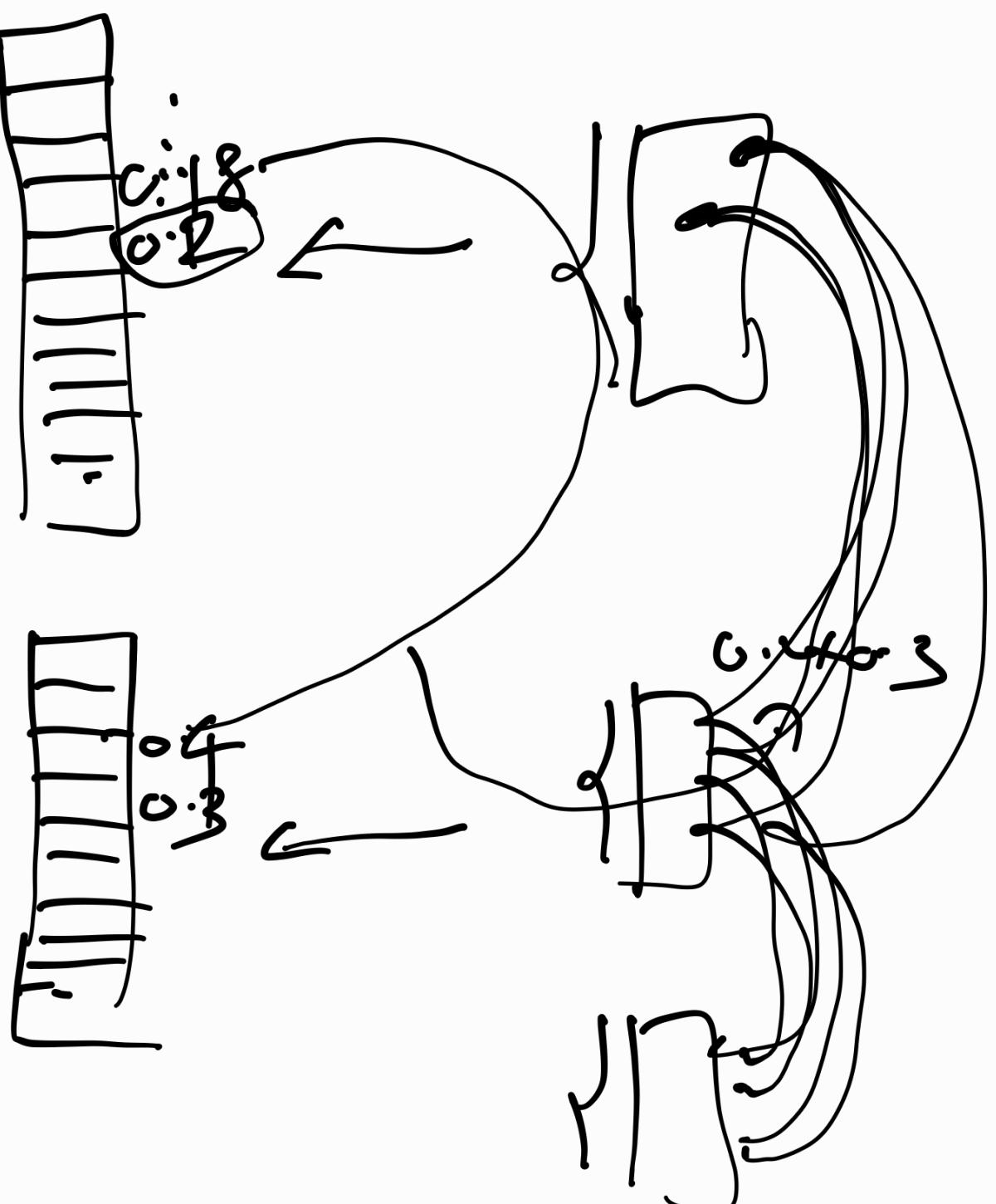
content-aware Approach:

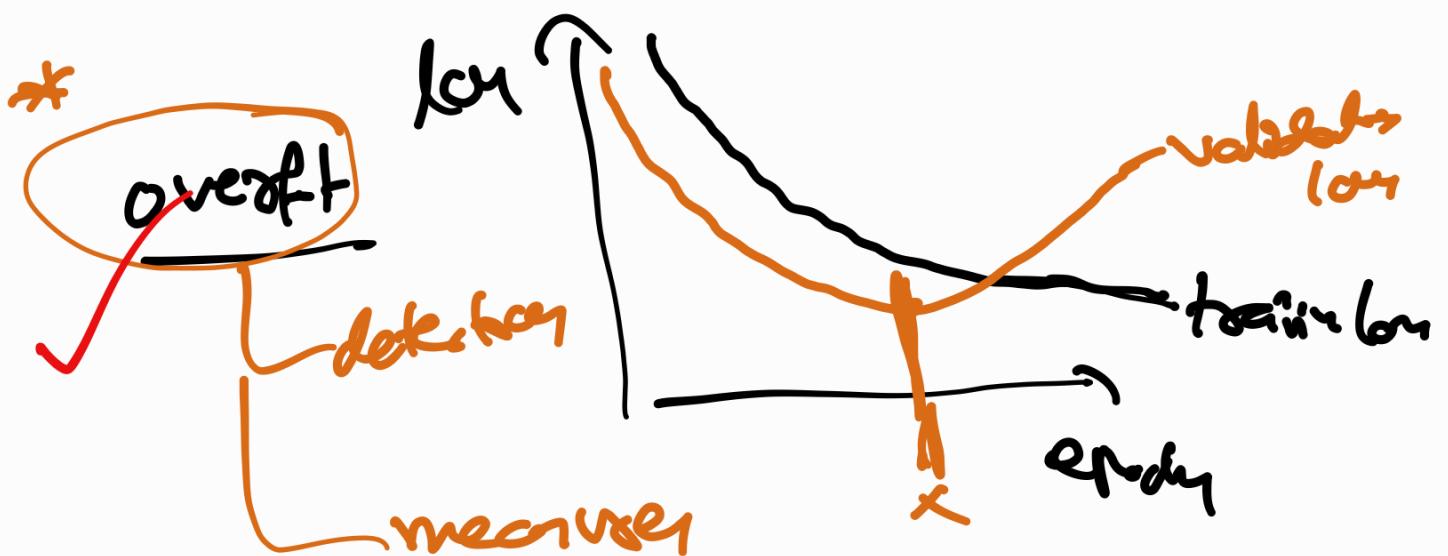
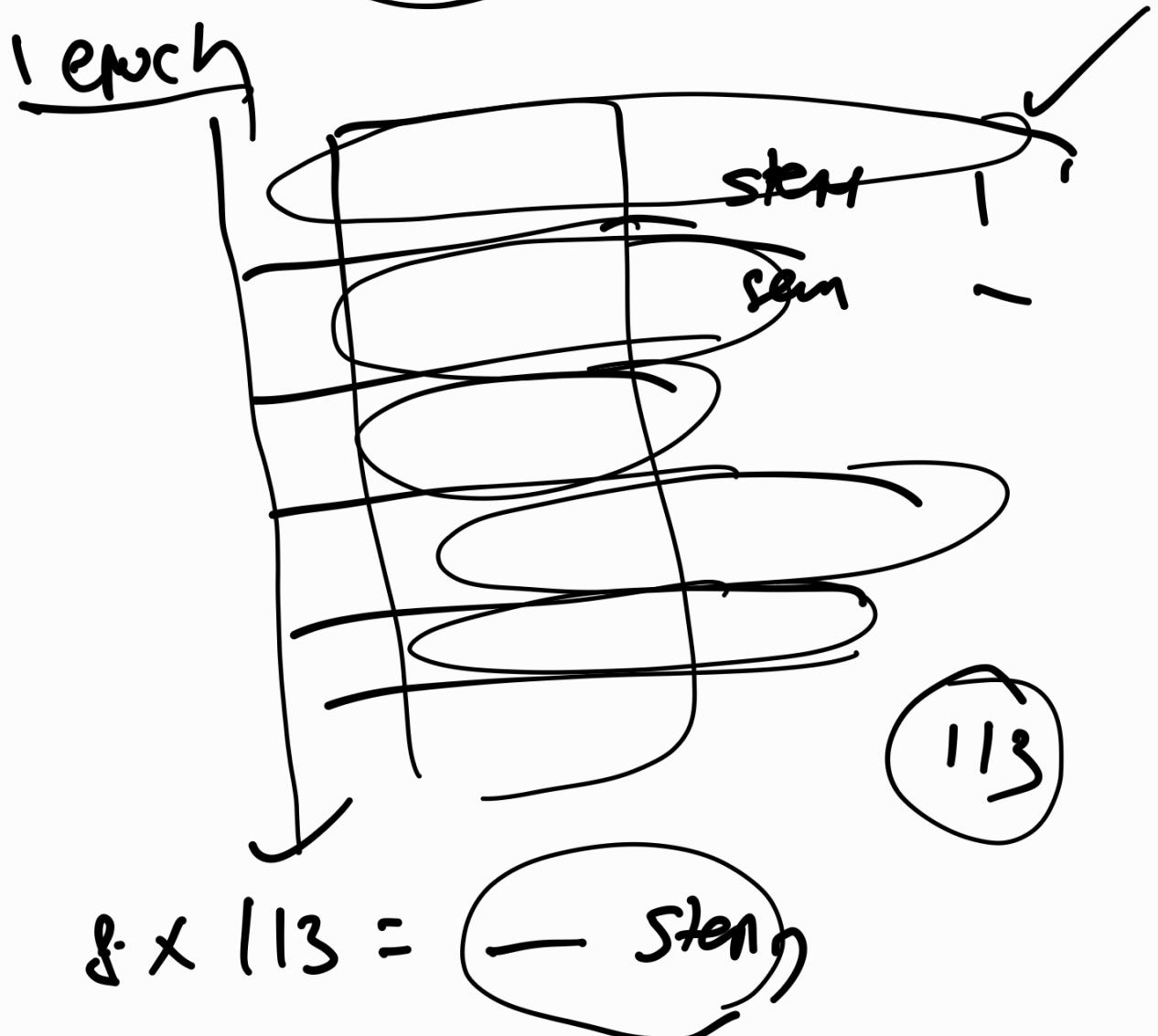
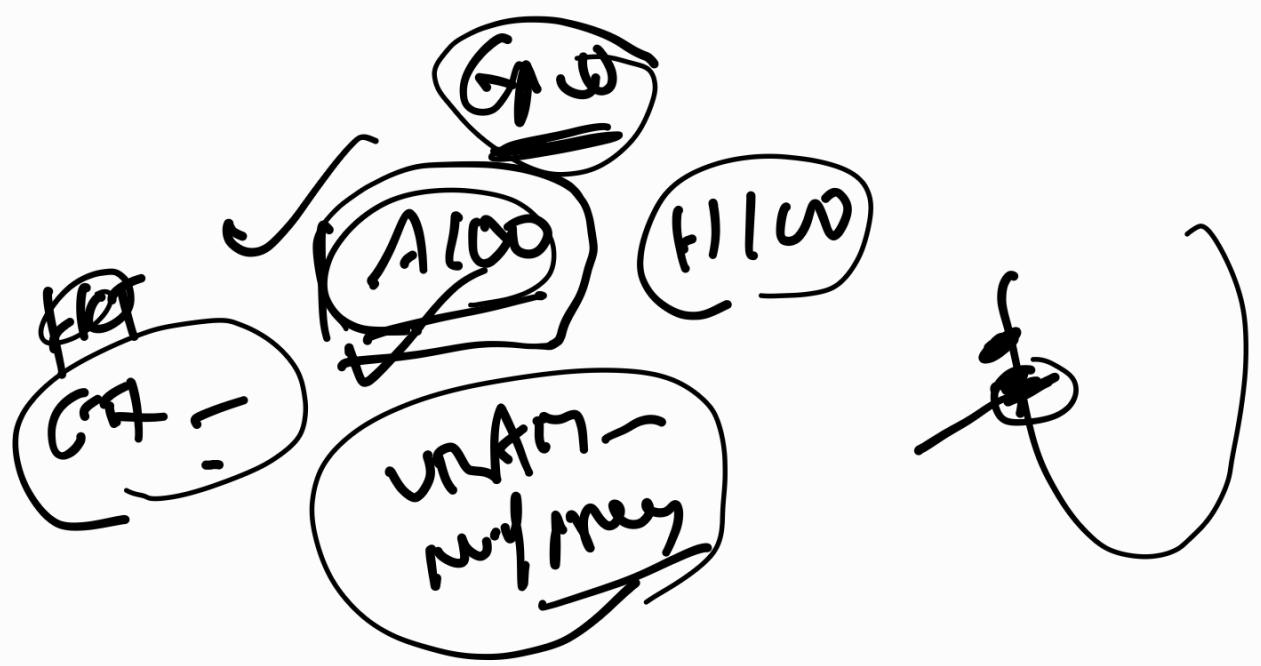
Penalize seq. history

Fixed penalty
~~dynamic seq. penalty~~

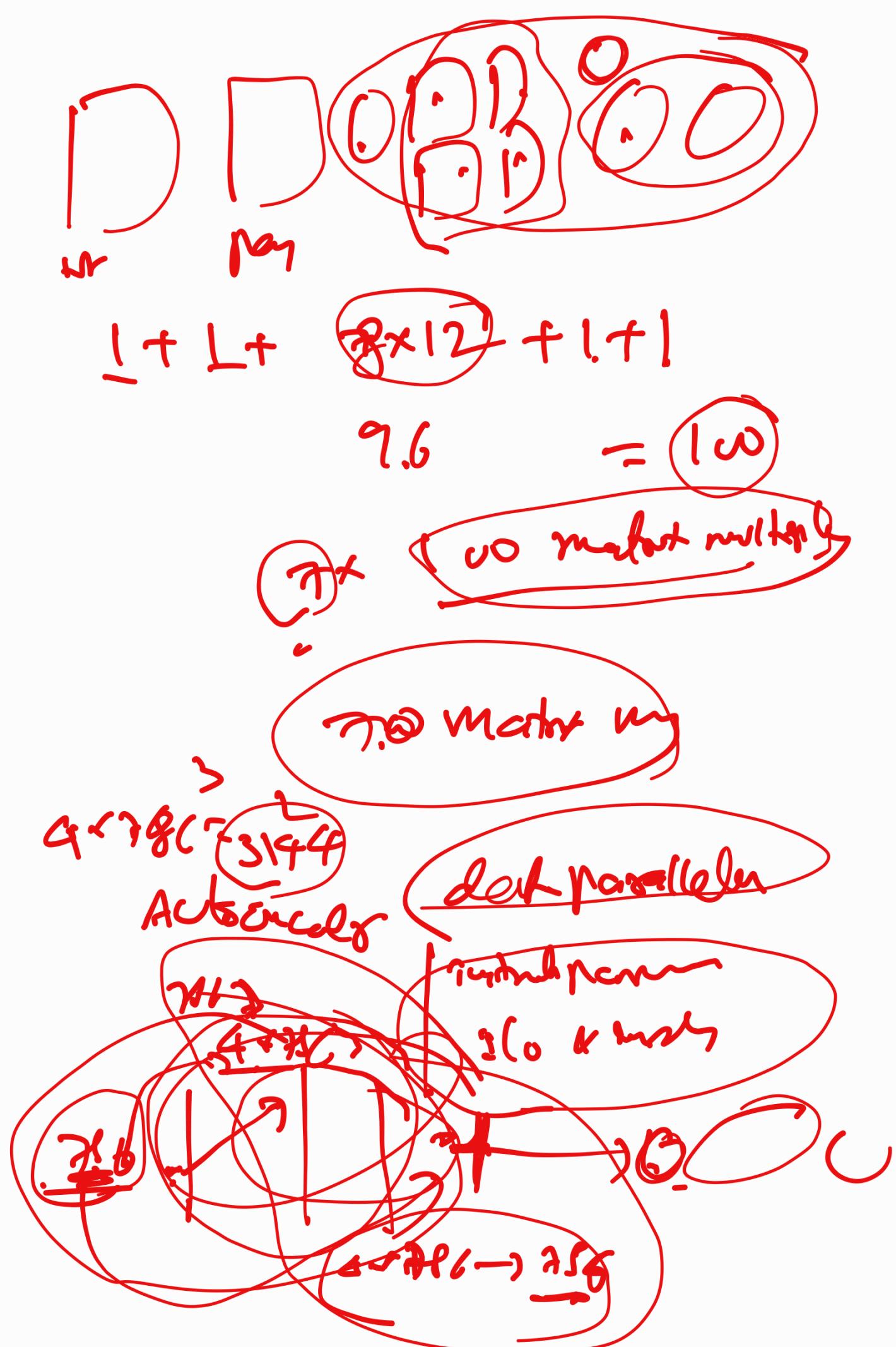
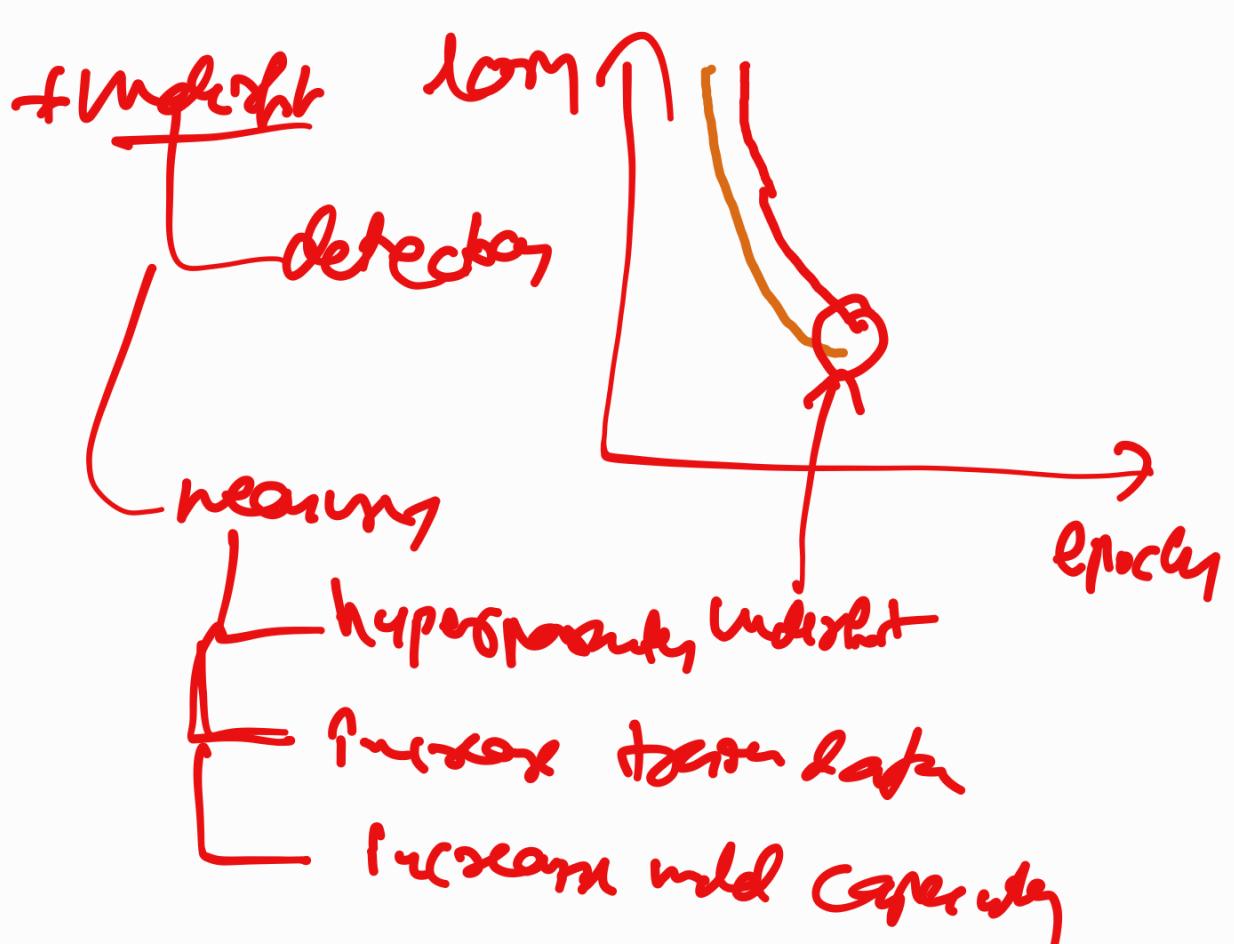


Beam Approach:





- ① increase train loss
- ② increase model capacity
- ③ regularization.
- ④ hyperparameter adjustment
 - batchesize 100



$$\frac{\partial L}{\partial \omega_0} =$$

The diagram shows four circles, each containing a mathematical expression. The first circle contains $\frac{\partial L}{\partial \omega_0}$. The second circle contains $\frac{\partial L}{\partial \omega_1}$. The third circle contains $\frac{\partial L}{\partial \omega_2}$. The fourth circle contains $\frac{\partial L}{\partial \omega_3}$. A red arrow points from the first circle to the second circle.