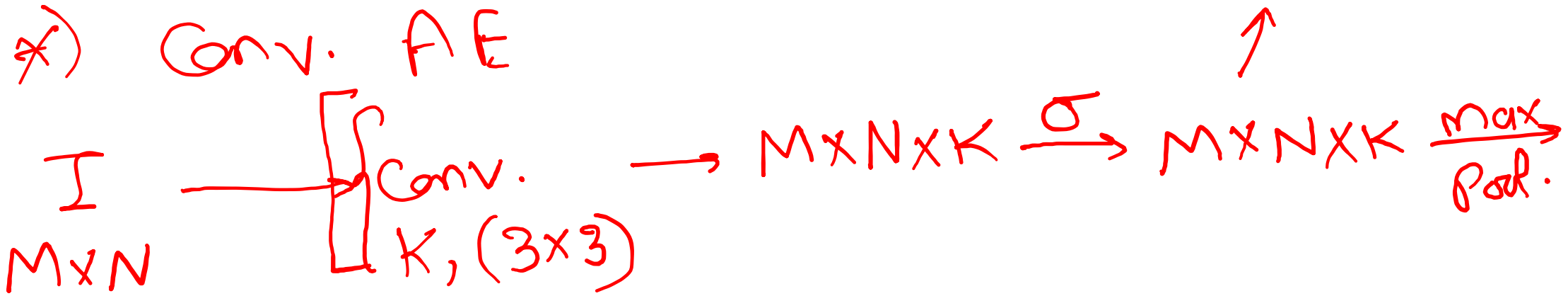
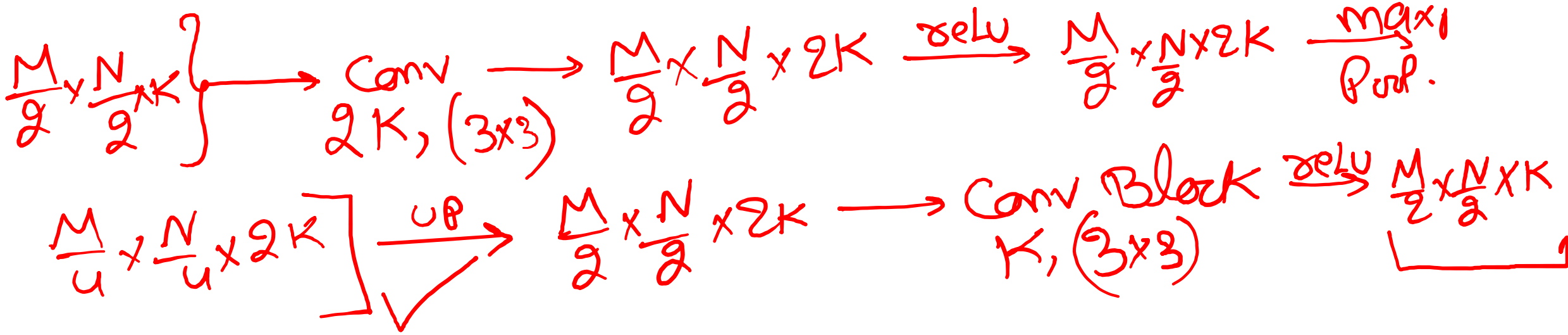
Lecture 21

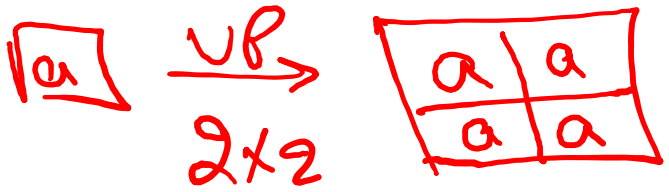
6 Dropout / Batch Normalization. Single conv. Blocks.

*) Conv. AE





[.] encoding process. Up-Sampling → NN



$M \times N \xrightarrow{K, 3\times 3}$

$M \times N \xrightarrow{3\times 3 \text{ conv.}} M \times N$

$\vdots$

$M \times N \xrightarrow[3\times 3]{K^{th}} M \times N$

$\Big\} K$

$\frac{M}{2} \times \frac{N}{2} \times K \xrightarrow[2\times 2]{UP} M \times N \times K \xrightarrow[1,(3\times 3)]{Conv.}$

$\underbrace{M \times N \times 1}$

decoded image $\hat{I}$

MSE $\equiv$

a batch of B Sampled

$$\sum_{i=1}^{B} \left\| (I_i - \hat{I}_i) \right\|_2^2$$

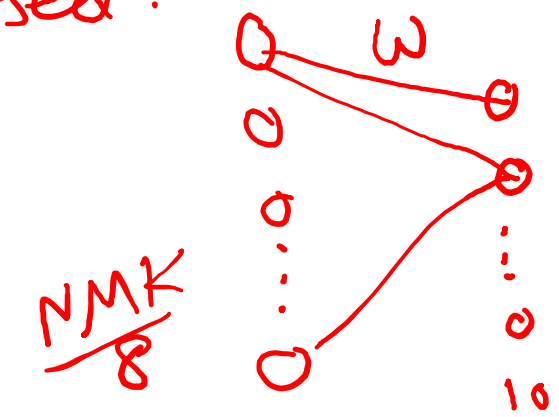
$I_{M \times N} \longrightarrow$ encodes $\longrightarrow \frac{M}{4} \times \frac{N}{8_4} \times 2K \xrightarrow{\text{flatten}} \frac{MNK}{8} \times 1$

$I \longrightarrow$ encodes $\longrightarrow$ flatten $\longrightarrow \frac{MNK}{8} \times 1$  activation. Softmax.

FC layers $\rightarrow$ 'C' nodes.

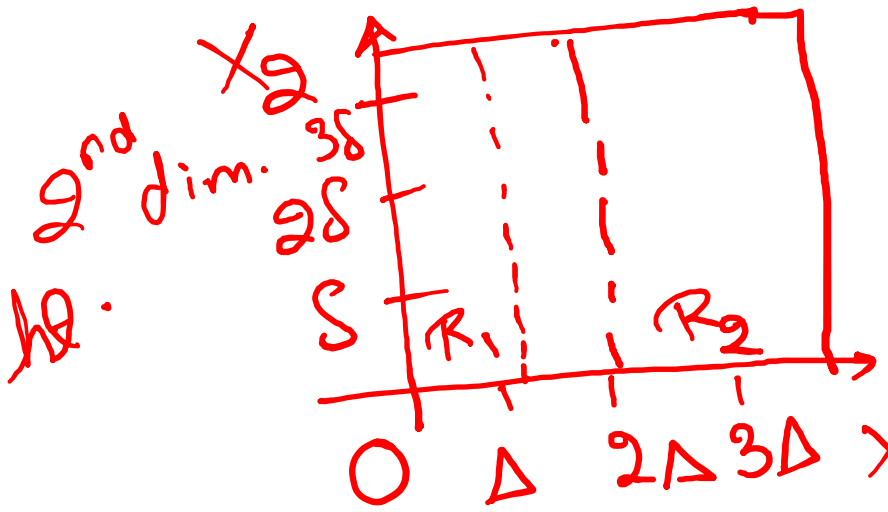
classification task - 'C' classes.

loss $\rightarrow$ multiclass cross entropy



Decision trees.

*) 2-D input



① (age, wt.) → wt

(1, 2) → 3

(2, 2) → 4

(5, 7) → ?

$$\sum_{i \in R_1} \left(y_i - \hat{y}_{R_1}\right)^2 = \varepsilon_{\Delta_1}$$

$$\sum_{i \in R_2} \left(y_i - \hat{y}_{R_2}\right)^2 = \varepsilon_{\Delta_2}$$

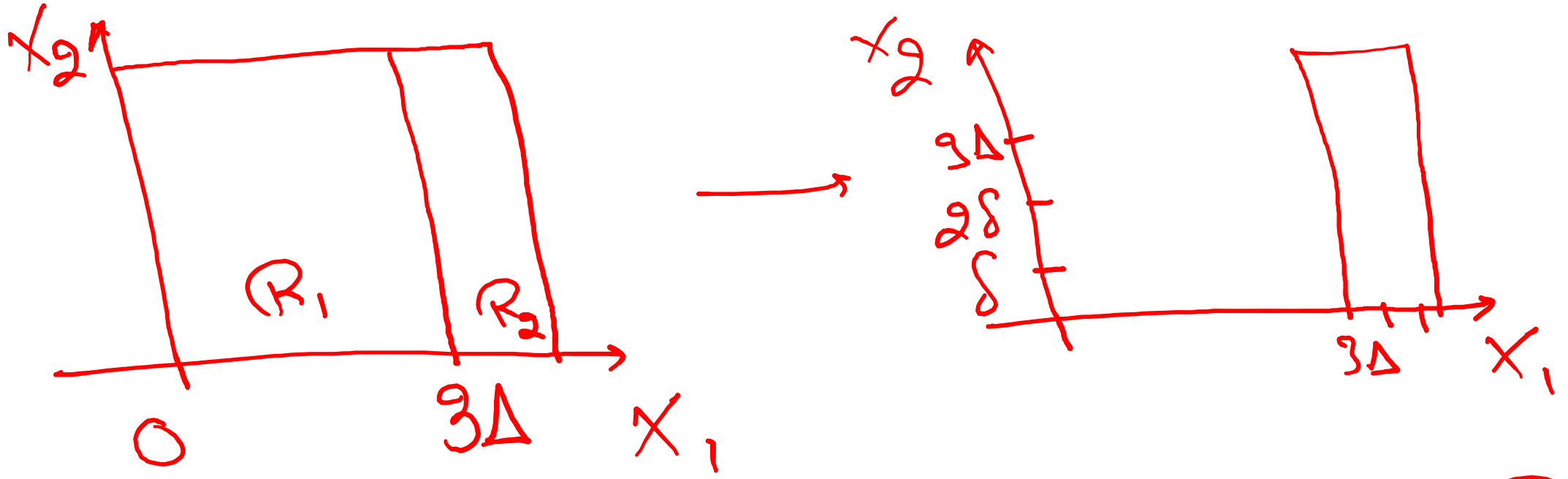
Total error ε_Δ: $\varepsilon_{\Delta_1} + \varepsilon_{\Delta_2}$

$\varepsilon_{2\Delta}$, $\varepsilon_{3\Delta}$

cut across X_2 at δ, 2δ, 3δ

E_δ, $E_{2\delta}$, $E_{3\delta}$

$\min\{ \varepsilon_\Delta, \varepsilon_{2\Delta}, \underbrace{\varepsilon_{3\Delta}}_{\text{min.}}, E_\delta, E_{2\delta}, E_{3\delta} \}$



2nd split → (X_1, X_2) choose either R_1 or R_2

*) Decision tree - classification.

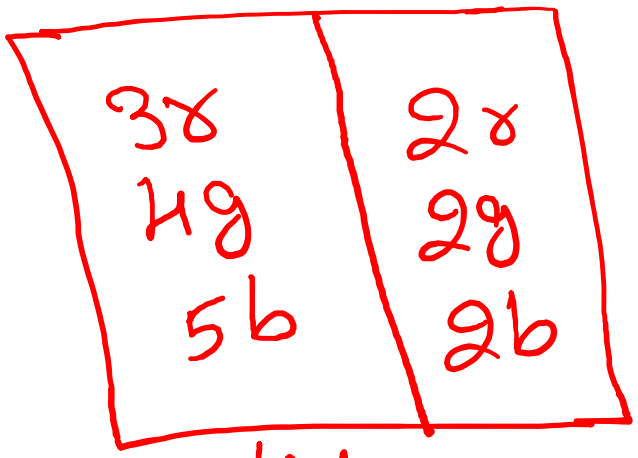
Loss → Gini index.

$$G_m = \sum_{k=1}^{K} p_{mk}(1 - p_{mk})$$

'm' denotes node.

K → " no. of classes.

p_{mk} → Prob. of class 'k' in node/region 'm'



m = 1 / left m = 2 / right.

3 - classes → γ, g, b

γ → k = 1

g → k = 2

b → k = 3

m = 1, k = 1, $P_{11} = \frac{3}{12}$, $P_{12} = 4/12$

$P_{13} = 5/12$

$P_{21} = P_{22} = P_{23} = 1/3$

G for m = 1

$G_1 = P_{11}(1 - P_{11}) + P_{12}(1 - P_{12}) + P_{13}(1 - P_{13})$

Total $G = \omega_1 G_1 + \omega_2 G_2$

$w_1 \rightarrow$ weight of region 1 $\rightarrow 2/3$

$w_2 \rightarrow$ weight " " 2 $\rightarrow 1/3$

$K = 2 \rightarrow$ binary code.

$$G = p_{m_1}(1-p_{m_1}) + p_{m_2}(1-p_{m_2})$$

$$= 2p_{m_1}(1-p_{m_1})$$