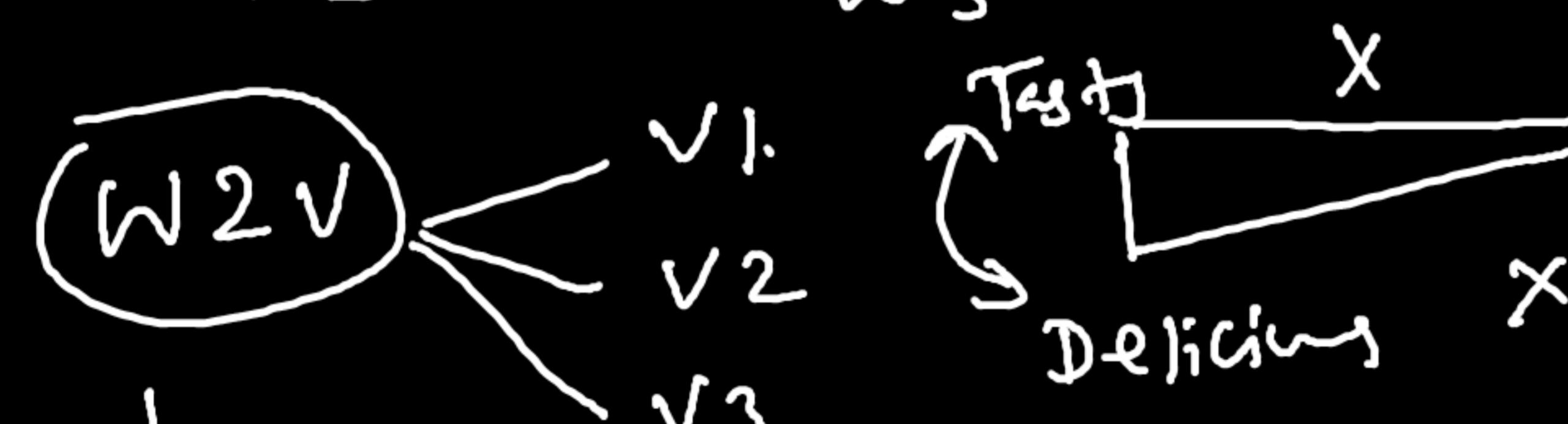


Word2Vec - Google - 2014

Text → vector

↳ Semantic Meaning

Tasty Delicious
 w_1 w_2



↳ Dense vector

↳ non-zero

→ State of the art Model (SOTA)

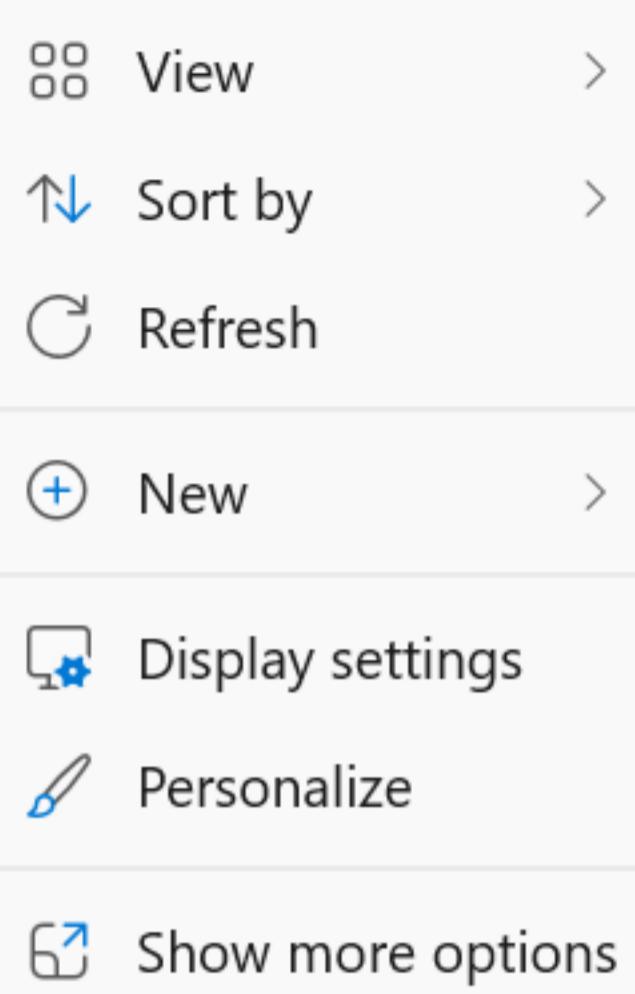
BOW / TF-IDF

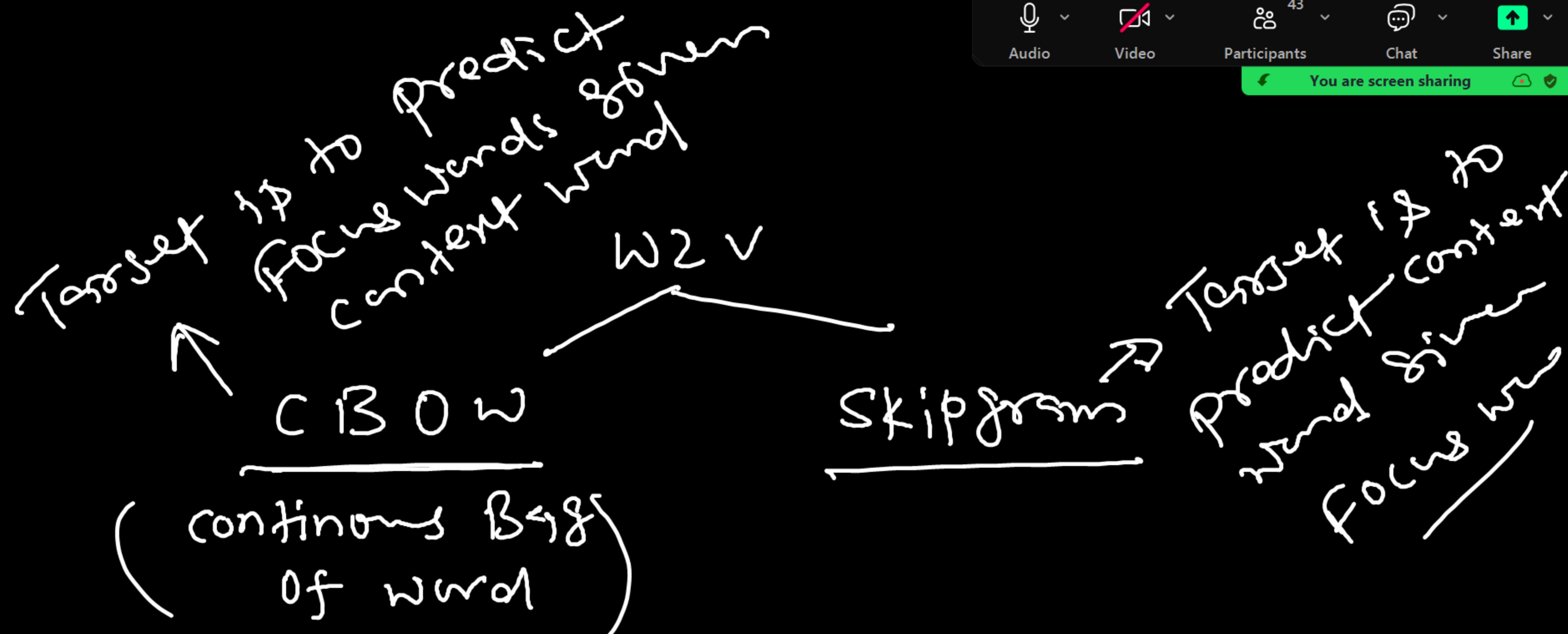
① w_1 & w_2 are similar then are also closer

② relationship

$v_{\text{Man}} - v_{\text{Woman}}$ || $v_{\text{King}} - v_{\text{Queen}}$
 $\text{Indigo} - \text{New}$ $\text{Deno} || \text{USA} - \text{WD}$.

- Semantic Mean
- Dense vector / matrix
- Matrix factorization
- Deep learning approach → Black Box





(continuous Bag of words)

C1 C2 = ... C3 Cf

Hi, my name is K-Sundaram

content words

Focus words

Algorithm / Optimizer

Hierarchical Softmax

Negative Sampling

