Word Embeddings: · Bepresent words as vectors of numbers. · Useful for creating simularity blus words. 1 · Captures sementic meaning Through vector sperations. T. C Cosine Similarly ---Measures similarity blu two verturs. · Formula:-A-B Cosine similarly = 3 11A11 × 11811 · Vectors w/ high cosine similarity point in the some Linetion Skip Gram Hodel: · Predict serrounding words (content) given a torget word. CE · Efficient for learning word embeddings. · Training involves manimizing the prob of content words given a target word. 10 3 Negative Sampling: 1 · Improves training efficiency by using logistic 4 regression model. -0 · Introduces negative examples (works not in cordint) to prevent trivial solutions.

1

	NAME
· Formula for the model's prediction:	THE SAME OF THE SA
	Bernelle
Prediction = o (Vingut · Vendent)	устранования
where o is the sigmoid fune.	
•	solony, make the
	resigned to contin
Training Process:	PRODUCTION OF THE PERSON NAMED IN
	nor appropri
· Initialize embeddings & content matrice Kandwuly.	acceptable)
· Use dot product & sigmoid to predict if	eservica
words are neighbors.	National Confession
	NUMBER OF
· Calculate exvor!	Microscopine
Error - Target - Prediction.	der and a second
- larger - president	necession and
. Adjust embeddings to minimize error.	manufacture (
· Hay with Emperatings to manifest error.	No. or head of
	goonumus
. Important Hyperparameter.	and town
-> Window Size: - Smaller (2-15) captores	and amount
words similarity, larger (15-50) captures	D. G.
relatedness.	No. of a second
	-
Number of regative sample: - Typical	4

5-20, 2-5 for large datasets. Analogy Operation: * Example: - "Kiny" - "Man" + "woman" = Queen. . Demonstrate vector with metic in embe ding