- ROUGE-N (**R**ecall-**O**riented **U**nderstudy for **G**isting **E**valuation), a metric within the broader ROUGE metric collection, is a vital metric in the field of natural language processing and text evaluation.

- It assesses the quality of a candidate text by measuring the overlap of n-grams between the candidate text and reference texts, and ranges between 0 and 1. A score of 0 indicates no overlap between candidate and reference texts, whereas a perfect score of 1 indicates perfect overlap.

- ROUGE-N provides insights into the ability of a system to capture essential content and linguistic nuances, making it an important and versatile tool used in many NLP workflows.

- It is a recall-based metric — a complement to the precision-based BLEU score.

- Formally, ROUGE-N is an n-gram recall between a candidate and a set of reference texts. That is, ROUGE-N calculates the number of overlapping n-grams between the generated and reference texts divided by the total number of n-grams in the reference texts.

- Mathematically, we define ROUGE-N as follows:

$$\text{ROUGE-N} = \frac{\sum_{S \in \text{Reference Texts}} \sum_{\text{n-gram} \in S} \text{Match(n-gram)}}{\sum_{S \in \text{Reference Texts}} \sum_{\text{n-gram} \in S} \text{Count(n-gram)}}$$

where Match(n-gram) is the maximum number of n-grams co-occuring in a candidate text and set of reference texts. It is clear that ROUGE-N is analogous to a recall-based measure, since the denominator of the equation is the sum of the number of n-grams on the reference-side.

- **Example:**

Assume we have the following candidate and reference texts:

Reference #1	`A fast brown dog jumps over a sleeping fox`
Reference #2	`A quick brown dog jumps over the fox`
Candidate	`The quick brown fox jumps over the lazy dog`

Splitting our candidate and reference texts into unigrams, we get the following:

Reference #1	[`A`, `fast`, `brown`, `dog`, `jumps`, `over`, `a`, `sleeping`, `fox`]
Reference #2	[`A`, `quick`, `brown`, `dog`, `jumps`, `over`, `the`, `fox`]
Candidate	[`The`, `quick`, `brown`, `fox`, `jumps`, `over`, `the`, `lazy`, `dog`]

Recall that our ROUGE-N formula is: $\frac{\#\text{ of overlapping n-grams}}{\#\text{ of unigrams in reference}}$

There are 5 overlapping unigrams in the first reference and 7 in the second reference, and 9 total unigrams in the first reference and 8 in the second. Thus our calculated ROUGE-1 score is $\frac{12}{17} \approx 0.706$