

What is the NER Model for Live Captioning?

GLOBAL ALLIANCE
OF SPEECH-TO-TEXT CAPTIONING...



The NER model is a method for determining the quality of live captions. Unlike Word Error Rate (WER), which is a simple measurement of words spoken versus how many of them are transcribed correctly, NER places emphasis on meaning and, more specifically, on idea units and how accurately they are captured in the captions.



How are NER scores calculated?

NER measures the number of edition and recognition errors deducted from the total number of words in live captions, including punctuation marks. The number is then divided by the total number of words in live captions and multiplied by 100. A score of 100 indicates that the content was captioned perfectly. The pass threshold in the NER model is 98%.



What are the types of errors?

[E] - Edition Errors

Edition errors are usually caused by a decision of the live captioner to edit or omit original spoken text, for example, because the original speech rate is too fast. These errors represent the loss of an idea unit (a piece of information).

-1.0

Critical Error: False Information
Editing/Paraphrasing error in the captions provides false but plausible information to the viewer

-0.5

Major Error: Loss of Main Point
Caption loses main point of an idea

-0.25

Minor Error: Loss of Detail
Caption keeps the main point but loses a detail of an idea

0.0

Correct Edition
Paraphrase captures full meaning.

[R] - Recognition Errors

Recognition errors are usually misrecognitions caused by mispronunciations/mishearings or by the technology used to produce the captions. These errors may be insertions, deletions or substitutions, and are calculated by comparing the respoken text to the original text

-1.0

Critical Error: False Information
Wrong word/phrase/punctuation error provides false but plausible information to the viewer

-0.5

Major Error: Nonsense Error
Wrong word/phrase/punctuation affects readability in a way that you cannot understand the idea

-0.25

Minor Error: Benign Error
Wrong word/phrase/punctuation affects readability but you can still understand the idea

How is NER scored?

Scores are calculated manually by professionals who have been trained in the NER model so as to provide consistent scoring. The scoring is accompanied by an assessment of the caption quality, which includes not only accuracy, but also delay, speed, positioning and technical issues amongst other aspects.

GLOBAL ALLIANCE
OF SPEECH-TO-TEXT CAPTIONING...



<https://speechtotextcaptioning.org/>