## Foreword

All the definitions and wordings are not self-written, but the definitions from the sources mentioned in each title.

## BLEU (Bilingual Evaluation Understudy) [1]

is a measurement for the difference between an automatic translation and a human referenced translation of the same starting situation.

### Evaluation process

The BLEU algorithm automatically translates consecutive words and compares them with the reference translation. It measures the fitting translations with different weights.

### How does it work?

The strength of BLEU is the correlation with human judgement. It is not looking on each word, but on the whole sentence written in the text, which helps to look at the context and meaning of the sentences and not just on the wording.

The score of a BLEU test depends on the complexity of the subject, how consistent test, training and optimizing data is and how much data is available for training. You can expect high BLEU scores, when your model is trained within a narrow domain and training data matches the test data.

### Used in

* Attention Is All You Need  
  arXiv:1706.03762v7 [cs.CL] 2 Aug 2023

## Encoder-Decoder-Modelle [2]

is a special type of architecture in ai and machine learning, which tries to understand transform and generate data. It is built with two main components.

### Encoder

The encoder is getting input data, for example a text to translate, and transforms the text into a static bunch of data, which are known as context vectors or internal representations. It abstracts the main features of the input values to extract the relevant data.

### Decoder

After the encoding process the decoder extracts the information of the encoded data and translates it to the requested language. The decoder learns based on the encoder information to generate the best result for the output.

The combination of encoder and decoder enables the model to resolve complex language translation, summaries and creation of humanlike language.

### Used in

* Attention Is All You Need  
  arXiv:1706.03762v7 [cs.CL] 2 Aug 2023

## Self-attention Layer [3]

is a mechanism to capture dependencies and relationships within input sequences. It allows the model to identify and weigh the importance of different parts of the input sequence by attending to itself.

### How it works

Self-attention works with transforming the input sequence into three **vectors**: query, key and value. These **vectors** are obtained through linear transformations [4] of the input

# Sources

**[1] Was ist eine BLEU-Bewertung?**  
<https://learn.microsoft.com/de-de/azure/ai-services/translator/custom-translator/concepts/bleu-score>  
(07.04.2025, 13:20)

**[2] Encoder-Decoder-Modelle-Anwendung**  
<https://we-make.ai/encoder-decoder-modelle-anwendungen/>  
(07.04.2025, 14:20)

**[3] Self-Attention Layer**  
<https://h2o.ai/wiki/self-attention/>  
(10.04.2025, 11:32)

**[4] Linear Transformation**  
<https://www.youtube.com/watch?v=On6wkamacRE>  
(10.04.2025, 12:39)