Introduction

Motivation

Supervised Learning has been the most fundamental technique for training Deep Learning Models since the development of the backpropagation algorithm. This is not surprising, as it provides a model with clear domain-specific learning signals, which is the most direct and efficient way of solving a problem or learning a task, respectively.

The development of artificial intelligence using labeled data, however, has its limitations. To teach Machine Learning models, particularly Deep Learning models, increasingly complex tasks, more labelled data is required. Naturally, generating millions of labeled examples becomes more difficult as the underlying tasks become more complex, making the development of such models more expensive and less feasible.

Because of this, Self-Supervised Learning has received an increased attention from the scientific community over the last few years. This is because Self-Supervised Learning does not rely on creating labeled data by hand, e.g. through human annotation, but receives it from the context of the data. To train a Large-Language Model (LLM) to understand written text, for example, words in a sentence are masked or deleted, respectively. The task of the model is then to predict those missing words.

Figure 1:

This has three advantages: Firstly, labels do not need to be created by hand, as it is easy to randomly mask words in a sentence and use them as the targets to predict during training. Secondly, because there are massive amounts of text available on the internet, a massive amount of training data can be generated. And lastly but most importantly, the model learns to write text that represents the world we live in. This becomes clear with the example seen in Figure ref{language_masking}. Here the model would have to predict the words “moon” and “bright” based on the context/words remaining after masking. In order to do so successfully, the model has to learn that only the moon shines at night, not the sun, and that if the moon shines, it is usually bright.

The aforementioned example illustrates an important characteristic of Self-Supervised Learning: It forces the model to learn common sense and the world that we humans live in.