

# Wie funktioniert eigentlich eine LM?

Albérique und Ruben L.

June 5, 2025

# 1 Einführung

Wir kennen sie alle und haben sie mit Sicherheit schon einmal benutzt: LM(Language Model)s. Kommerzielle Modelle, die wir normalerweise benutzen, sind sogar LLM(*Large* Language Model)s. Sie sind weitaus leistungstärker als kleine Modelle, sie funktionieren aber größtenteils gleich. Hier werden wir probieren, zu erklären, *wie*.

# 2 Grundlagen

Neuronale Netzwerke sollte uns allen ein Begriff sein. LMs besitzen auch solche Netzwerke, sie sind jedoch ein wenig abgewandelt: Bei einer LM heißen diese Netzwerke "Transformer". Sie funktionieren grundsätzlich gleich, nur werden die Netzwerke in Schichten eingebaut, dann heißen sie Transformer. Jedoch können die Neuronen nur mit Zahlen arbeiten, nicht mit Buchstaben. Wie geht man dagegen vor? Im Prinzip ist es sehr simpel: Das Netzwerk wandelt einzelne Buchstaben, Wörter oder Teile von Wörtern in Zahlen (Tokens) um. Diese werden in einem Vokabular gespeichert, damit immer die Buchstaben immer dem gleichen Token entsprechen.