

# Seminar Large Language Models - Transformer Architektur, SS 2024

3. April 2024

## Struktur des Seminars

### 1 Einführung

#### 1.1 Deep Learning - Basics

Einfaches Modell, Konzept des Trainierens

#### 1.2 Sequences to Vectors

Wie wird Text in Vektorform repräsentiert?

### 2 Architektur

#### 2.1 Recurrent Neural Network

Architektur zur Verarbeitung von sequentiellen Daten

#### 2.2 Transformer Architektur

Verarbeitung sequentieller Daten (SotA)

##### 2.2.1 Überblick

Bestandteile von Transformer-basierten Modellen

##### 2.2.2 Transformer Architektur - Attention Mechanismus

Kern einer Transformer-basierten Architektur

##### 2.2.3 Transformer Architektur - Blöcke mit Attention Heads

Transformer Modell verstehen

#### **2.2.4 Transformer Architektur - Encoder - Decoder**

Bestandteile verschiedener Modelle und deren Funktion

### **3 Training**

Erstellen/Trainieren eines Transformer Modells

#### **3.1 Training Task - Masked Language Modeling**

#### **3.2 Training Task - Next Token Prediction**

#### **3.3 Transformer - Optimierung und Fehler (Loss)**

### **4 Anwendungen**

#### **4.1 Generative Large Language Models**

#### **4.2 Biologie - DNABert**

#### **4.3 Finanzwesen - FinGPT**

### **5 Literatur**