

### **1. Działanie rekurencyjnych sieci neuronowych:**

Rekurencyjne sieci neuronowe (RNN) są modelem uczenia maszynowego, który operuje na sekwencjach danych, uwzględniając poprzednie dane z sekwencji w analizie kolejnych elementów.

Każda warstwa RNN przyjmuje sekwencję danych i wyznacza nowy stan wewnętrzny, który jest używany do analizy kolejnego elementu sekwencji.

### **2. Działanie mechanizmu uwagi:**

Mechanizm uwagi pozwala sieciom neuronowym wybierać, na których elementach sekwencji się skupić podczas przetwarzania.

Umożliwia to modelom koncentrację na ważnych częściach danych i efektywne przetwarzanie długich sekwencji.

### **3. Statyczne osadzenia słów:**

Statyczne osadzenia słów to reprezentacje słów jako wektory liczbowe w przestrzeni o niskiej wymiarowości.

Przekształcają słowa na wektory, zachowując pewne semantyczne zależności między nimi.

### **4. Model językowy, w tym znakowy model językowy:**

Model językowy jest modelem probabilistycznym, który przewiduje prawdopodobieństwo wystąpienia kolejnego słowa w oparciu o kontekst sekwencji słów.

Znakowy model językowy operuje na poziomie pojedynczych znaków zamiast na poziomie słów.

### **5. Pre-trening w kontekście modeli językowych:**

Pre-trening to proces wstępnego uczenia modelu na dużym zbiorze danych, który może być później dostosowany do konkretnego zadania poprzez proces fine-tuningu.

### **6. Wsteczna propagacja błędu w czasie:**

Wsteczna propagacja błędu w czasie (Backpropagation Through Time - BPTT) to algorytm propagacji błędów w rekurencyjnych sieciach neuronowych.

Pozwala on na obliczenie gradientu funkcji kosztu w czasie, co jest niezbędne do uczenia modeli przetwarzających sekwencje.

### **7. Długa pamięć krótkotrwała (LSTM):**

LSTM (Long Short-Term Memory) to specjalny typ rekurencyjnej sieci neuronowej, który może zachowywać informacje przez dłuższy czas.

Wykorzystuje mechanizm bramek, aby decydować, które informacje są istotne i mają zostać zachowane w pamięci.

### **8. Architektura modeli: GPT, BERT:**

GPT (Generative Pre-trained Transformer) oraz BERT (Bidirectional Encoder Representations from Transformers) to modele językowe oparte na architekturze transformer.

Są to duże modele, które osiągnęły doskonałe wyniki w zadaniach związanych z przetwarzaniem języka naturalnego.

## **9. Duże modele językowe (LLMs):**

Duże modele językowe (Large Language Models) to modele językowe o dużej pojemności, które są w stanie generować wysokiej jakości teksty, odpowiadać na pytania i wykonywać inne zadania związane z językiem naturalnym.