

Start time 10/25/25 12:10:32

Completion time 10/25/25 13:17:36

Die Frage	Antwort des Teilnehmers	Kommentar des Autors
Das Transformer model besitzt "positional awareness". Daten die ins Model hineingefüttert werden, laufen also zuerst durch keinen Positional Encoder. Das übernimmt das Modell selbst.	False	correct
Beschreibe die Logik hinter Scaled Dot-Product Attention, wie es im Transformer model verwendet wird. (Keinen Aufsatz, 2-3 Sätze genügen)	Es wird dabei das Dot-Product von zwei Vektoren berechnet, um deren Ähnlichkeit zu bestimmen. Je ähnlicher die verglichenen Vektoren, desto mehr Aufmerksamkeit bekommen diese.	correct
Erkläre kurz welche Aufgabe das Input Embedding in einer Transformer Architektur hat. (1-2 Sätze)	Beim Input Embedding werden die Eingaben in das Modell tokenisiert und einer Position zugewiesen, welche dann für das Positional Awareness eine Rolle spielt. Das gehört noch zum "Encoding" der Eingabewerte.	False, das input embedding ist dafür verantwortlich den Vektor aus Tokens in einen Vektor von höheren dimension (d_{model}) umzuwandeln.
Die Outputlayer eines FNN für Regression, welches auf die Rechnungen aus dem Projekt (Neural Predictive Calculator) trainiert wurde, spuckt gleich die resultierende Dezimalzahl raus.	Nein! Das ist falsch, die Outputlayer spuckt einen hochdimensionalen Vektor aus, welcher zuerst noch verarbeitet werden muss.	False, output der output layer, welche aus einem einzigen neuron bei regressions modellen besteht ist eine dezimalzahl, die die das model berechnet hat.
Anhand des Beispiels von dem Projekt (Neural Predictive Calculator), erkläre was overfitting ist.	<p>Overfitting führt dazu, dass das Model beginnt, die Resultate von den Eingaben "auswendig" zu lernen. Dabei lernt es aber nicht, selbst auf die Lösung zu kommen.</p> <p>Wenn ich also Inputs eingib, welche nicht beim Training verwendet wurden, gibt mir das Model eine falsche Antwort mit hohem Verlust (Loss)</p>	correct
Neuronale Netze designed für Regression geben als Output immer eine Dezimalzahl raus. Obwohl das Model eigentlich mit Ganzen Zahlen als Lösung trainiert wurde ($1+2+3 = 6$). Das kann einwenig counterintuitive sein, erkläre den Grund für die Dezimal zahlen im Output.(2-3 Sätze)	Die Inputs sind zwar immer ganze Zahlen, jedoch werden diese nach dem Tokenisieren mit Gewichtungen (W) und Biases (b) verrechnet, welche auch Dezimalstellen haben können. Das beeinflusst dann, was der Output sein soll.	correct
Soweit habe ich die Feed-forward Neural Networks (FNNs) verstanden. (1: gar nichts, 10: Ich habe alles verstanden was mir beim Gespräch unterrichtet wurde.)	8	FNNs gut verstanden
Soweit habe ich transformer Modelle verstanden.(1: gar nichts, 10: Ich habe alles verstanden was mir beim Gespräch unterrichtet wurde.)	6	Die Transformer sind etwas schwieriger, beim Gespräch habe ich ebenso den Eindruck erhalten, dass Marton mit den Transformern mehr mühe hatte. War zu erwarten, sie sind komplizierter aufgebaut.
Wie stark ist die Schuld deinem Gesprächspartner(mir) zu geben, für Themen/Bereiche die du nicht verstanden hast? (Im Vergleich dazu, wenn jetzt eine professionelle Lehrperson meine Stelle eingenommen)	3	Entspricht einem guten Wert. Die Resultate des Gesprächs sind gebrauchbar und verallgemeinbar.
Ich fand die konkreten Beispiele anhand des Projekts hilfreich um ein Allgemeines Verständnis über das Thema zu ergreifen.	9	(Outlier) Die Beispiele, haben gleich wie mir als ich das Thema gelernt habe, Marton auch stark geholfen.
Das erklären anhand eines Models für Regression hat mir das Verstehen von neuronalen Netzen vereinfacht	7	Eine durchschnittliche Zahl, um 2 kleiner als die Beispiele. Man kann somit sagen, dass Regressions modelle einen kleineren unterschied ausgemacht haben.

Falls Regressions modelle die generelle Funktionsweise von bestimmten Neuronalen Netzen für dich vereinfacht haben, warum?	Ich konnte bestehendes Wissen verwenden und dieses an das neue Thema anknüpfen.;Im Gegensatz zu den meisten anderen Modellen ist der Input in die behandelten Regressions Modelle eine Zahl, genau so wie der Output. Das Umwandeln von Wörtern (und Ähnlichem) zu Vektoren hat (oder hätte) mich verwirrt.;	Das bestätigt, dass das einführen ins Thema mithilfe von Regressionsmodellen von Vorteil ist. Der erste Teil ist von Marton selbst geschrieben, er meint hier den Übergang von linearer Optimierung zum Regressionsmodell.
Welche Themen/Bereiche hast du besser verstanden?	Am besten erklärt fand ich die Behandlung von Tokens/Vektoren und wie diese verarbeitet werden. Dies war mir bisher tatsächlich noch nicht ganz klar. Das gilt für die FNNs als auch Transformers.	Das ist eben von Regressionsmodellen ermöglicht worden, weil diese mehr mit Zahlen arbeiten und die Tokenizer Verschlüsselung somit einfacher ist.
Welche Themen/Bereiche blieben unklar?	Ich habe eigentlich alle Unklarheiten geklärt. Ich wurde auch öfters gefragt, ob ich noch offene Fragen habe. Natürlich würde ich noch die Details im Selbststudium behandeln müssen, wenn das jetzt ein Schulfach wäre. Es könnte höchstens noch das Trainieren vom Modell ein wenig mehr erläutert werden. (z.B. wie werden der Bias und die Gewichtung beim Lernen angepasst und weshalb?)	Ja, wie Backpropagation funktioniert habe ich nicht erwähnt. Und die genauere Funktion von Optimizern wurde auch nicht besprochen. Das liegt vor allem daran, dass ich in diesem Projekt mich auf die Architektur der Modelle konzentriert habe und mich somit also mit diesen Themen nicht beschäftigt habe.
Wie fandest du den Gesprächspartner (mich)? Habe ich vielleicht zu schnell oder zu langsam gesprochen? Oftes stockern und suche nach Wörtern hat dich gestört? Alles Feedback willkommen!	Ich habe das Gespräch mit dir als angenehm und informativ empfunden. Bei Unklarheiten konnte ich immer Fragen stellen oder eigene Beispiele einbringen, damit ich die Verbindungen zu bestehendem Wissen machen konnte. Das einzige, was manchmal gestört hat, war die gelegentliche Annahme, dass ich bereits mit einem Thema vertraut war. Als ich aber nachgefragt habe, konnte dies auch in einer angemessenen Tiefe erklärt werden. (z.B. das Dot-Produkt von Vektoren)	Ja, das ist wahr, ich muss da aufpassen. Gut, dass er sich gemeldet hat und mich rechtzeitig stoppen konnte und es ihm erklären. PS. Der Bezug zu bereits bestehendem Wissen scheint wichtiger zu sein als ich dachte.
Genereller Eindruck zum Gespräch und sonstige Bemerkungen	Die Präsentation gab mir den Eindruck, dass du dich gut mit diesem Thema auseinandergesetzt hast. Bei jeder Frage konntest du mir eigentlich selbstständig eine Antwort geben, auch ohne Recherche oder googeln zu müssen. Das zeigte mir, dass du das behandelte Thema auch wirklich verstanden und verinnerlicht hast.	Das freut mich. Danke schön, sehr nett.