**eeq-qdde-agt (2023-03-17 12:11 GMT-4) - Transcript**

# **Attendees**

Ali Mirzaei, Amir Feizpour, Ammar Khan, Debjani Mukherjee, Dikshya Mohanty, Karim Khayrat, Karim Khayrat's Presentation, Kua Chen, Marzieh Zare, mohammed fahad, Nikhil Varghese, Percy Chen, Suhas Pai, You Cheng, Yujing Yang

# **Transcript**

*This editable transcript was computer generated and might contain errors. People can also change the text after it was created.*

Karim Khayrat: So yeah, I'm gonna talk about two things today. I'm going to talk about Self-instructs next and but first, we'll start with any bars. Which is a paper on. On few short learning for information retrieval tasks. I'll it imparts us for inquisitive. Parrots for search, I guess like,

Karim Khayrat: Yeah, in in the sense that parrot, repeat you know what you say. So we'll see if the paper matches the name. So yeah, what's a was a challenge that paper aims to solve is that you guys know it's expensive to gather data for the train retrieval models in particular. It's expensive to gather human data. And and the solution that they presented the few short learning approach for generating, synthetic training data sets. So in summary, this is the model overview.

Karim Khayrat: a Given a document, let's say, you know, a paragraph or some piece of tech And a prefix, which is basically the prompt right? That consists of N, pairs of questions and relevant documents. So you go let's say to charge GPT you give it documents and some questions really to that document and give it, let's say three pairs. That's what you use in the paper there. Randomly chosen from a data sets from the Marco Ms. Marco data set. And that creates basically a prefix. In the prompt. And then you use a User much better, language model. And what you're trying to find tune To generate a question. That's relevant to document, that's the input, the new document that you want to create questions for.

Karim Khayrat: so, we will come to an example after this, but, but basically, you if you set n equal to three you can possibly generate thousands if not millions of training examples, From a random document collection. and after you generate, after the step where you generating so many synthetic questions to your documents,

Karim Khayrat: Not all of them, of course, are going to be of equal quality, right? So what they do is they filter some top k. In this case, the filtered They generated 100,000 and then you filter the top 10,000. Based on the log probability of the generated output. So you know, in a transformer model every every output has some probability. So, based on the lot, the largest log probability of the output and the they took the filter them and they took the top 10,000 examples.

Karim Khayrat: So this is basically model overview. You have a few short, input three examples for documents D. This is Document D. But they give it. And the examples are not one in the in this image. So you have some or they're shown behind, no, actually, they're not going in this image, the way they presented it. So, you have the document, we don't know a lot about the effects of caffeine during pregnancy on you or your baby. So it's best to limit the amount you get each day. This could be like from a web in the website or or any other website, for example, and you're asking the language model To give a question on this document, right? So you input this to a language model G And it recreates some questions like What are the effects of caffeine during pregnancy and there's a probability of this question being generated, right? And so you connect these generated question document pairs

Karim Khayrat: and you select the top key based on the probability. Of the generative question. And then you get some training pairs and then you input these training pairs to fine tune a re-ranker, right? So what I read anchor does is it takes the question and the document and it scores from zero to one in the paper you say it's a classification model. So I assume the use same structures classification model to to basically classify whether this question and document match are there relevant or not. So this is the relevant SEA score. so, relevancy, given of given relevancy, given the document and question, Okay.

00:05:00

Karim Khayrat: So I guess there's a question.

Amir Feizpour: Yeah, I can ask a question. So that the three examples are they just?

Karim Khayrat: Yeah.

Amir Feizpour: Like what you're showing is just a very bottom of The prompt I assume that there is instruction and…

Karim Khayrat: Yes.

Amir Feizpour: there are three examples, and then what you're showing in blue is what came after.

Karim Khayrat: Yes, exactly. Yeah.

Amir Feizpour: Okay. and, What's the setup for the language model G? So, is it producing only one question is producing several questions. The talk about that at all.

Karim Khayrat: Um, let me I I believe you you can, you can produce as many questions for document. As you want. I'm not I don't remember exactly the details of of what they did. But yeah in principally you you generate multiple questions? Because also the prompt that the sample, three examples, right? Those changes. So you can have Different. Different questions based on the prompt based on the three examples that you give it previously and…

Amir Feizpour: Okay, are the examples?

Karim Khayrat: you sample three different ones each time.

Amir Feizpour: Are the three examples? Chosen? Dynamically or are those like heart condition.

Karim Khayrat: I don't think they're fixed. I think there's children dynamically. I think the fixed number that they talk about is three. Like, three examples that that would make a more sense like the number three effects. But the examples themselves are are sampled.

Amir Feizpour: Place and…

Karim Khayrat: And they are sampled from them from the Marco data set.

Amir Feizpour: I assume it's not.

Amir Feizpour: I see, are they randomly sampled or like, is there a retrieval happening there?

Karim Khayrat: No, I believe they're randomly sample because there's what will there are no reason to retrieve it on something particular at this point. So there are generating question document pairs, right? So there's no retrievals you have a document randomly sampled And you just,

Karim Khayrat: of course, like they talk about things about, let's say, if you, if you do domain specific, Domain specific sampling for documents, right? But they don't, they don't support that in their conclusion, that you should. You should go that route,…

Amir Feizpour: Presentation, and…

Karim Khayrat: you know?

Amir Feizpour: the re-ranker. What's the purpose of it? I didn't quite follow. There is that to choose the question. So, the selection of the top K has happened already. So reruncle is being.

Karim Khayrat: Right. Thank you. Yes exactly. They're fine tuning the the data set I guess this is this is, for example.

Karim Khayrat: Let me see. Variation. Here you go. You the randomly a hundred thousand documents. From from using this process. Like the The Q comma DPs. and, Degenerate. Yeah, here you go. The generate only one question per document. That that answer your question. But in principle, you can generate as many questions as you want. So it's not like

Karim Khayrat: Part of this. This. Yeah, it's not like Important. I don't think it's a very important detail but here they generate one question for document. Using the Gpt's query and the language model. And then the Select 10,000 pairs. You'll find you. And then some detail I missed out Skipton, is that, yeah, you are printing a classifier, so they need negative examples. So, what they do is they just use B, m, twenty five on on the document. At it. So they have a question, right? So the they can just query a document using bm25. I think the select like 100 documents so there and…

Nikhil Varghese: You.

Karim Khayrat: then they select one random one of them. So it's like a bad match. Let's say.

Karim Khayrat: I ran them random sample and and they use that as a negative sample.

Amir Feizpour: So what is the classifier you mentioned?

Karim Khayrat: The classifier is there is the rear anchor. The classifier, the reactor takes the document and a query. And it it outputs whether they're relevant or not.

00:10:00

Amir Feizpour: Your audio is cutting out.

Karim Khayrat: but,

Amir Feizpour: We cannot hear you can.

Karim Khayrat: Can you hear me now? Hello. Yes.

Amir Feizpour: Yes, yes.

Karim Khayrat: Oh yeah.

Karim Khayrat: Yeah, so so, where did I cut off? It did I cover the negative examples? Part?

Amir Feizpour: a yeah, like the like you moved to the next slide so here

Karim Khayrat: Yeah, so so this is just an example of Of a prompt. If you if you you have like a document, the relevant query, which is generated by gp3 and and example, two an example, three,

Karim Khayrat: Are you? This is this is the prompt. This is a question.

Karim Khayrat: this is an example of the few short input that's given to To the to the query.

Amir Feizpour: Thank you.

Karim Khayrat: Yeah. So GPT 3 now takes this. and,

Amir Feizpour: You.

Nikhil Varghese: so, what's Like I icon quite understand why,…

Karim Khayrat: Butter.

Nikhil Varghese: what's that? Good question. And a bad question.

Karim Khayrat: Okay. What the?

Nikhil Varghese: It's the line doesn't seem very far.

Karim Khayrat: Yeah. So these are two methods in which they prompted gpt3. So the left side is is the standard one that you expect, but then they tried having a good question. And a bad question means a good question, is a good question. And a bad question is some is one. That's not very relevant to the document. and,

Karim Khayrat: They try to. Have and then charge the PT. Just generates the good question but it's more like It, it knows what the bad question and what's a good question. So it can differentiate between generating a good question and a bad question but ideas to generate higher quality questions. Let's say using this kind of prompt.

Karim Khayrat: and,

Karim Khayrat: Yeah.

Amir Feizpour: You want to thank you. So,

Karim Khayrat: And and they will get to the to the detail to the to the results later like whether this method got got anything improved, anything or not. But yeah,…

Amir Feizpour: so,

Karim Khayrat: that that was the idea between behind it.

Nikhil Varghese: Yeah, don't matter.

Amir Feizpour: so I, I

Nikhil Varghese: Seems The method seems reasonable. but, the questions over here, maybe the model is both models are not too bad. So it's hard to really say. One is a clear winner.

Amir Feizpour: so I I'm a little So the left side are the questions from Barcode Edison.

Amir Feizpour: If you're saying anything, we cannot hear you.

Karim Khayrat: Yeah, I don't. yeah, so so they have yes, they have

Karim Khayrat: No, they don't they don't have the start with.

Karim Khayrat: With. Human examples. Let me pull up the paper just to make sure that I got this right.

Amir Feizpour: Right. Because You know, I'm also trying to wrap my head around. What is a good question and what is about question?

Karim Khayrat: Okay.

Nikhil Varghese: It feels more like. Being more specific I guess. Versus a more vague question.

Amir Feizpour: Yeah. Actually, you know, or another way that seems. To me, that might be the definition of good question. Like A lot of what they're saying. Bad questions. They have like single word answers, but the quote unquote. Good questions have probably long format answers. That might also be, you know, the difference.

Nikhil Varghese: You.

Amir Feizpour: well, my question is Where did these questions come from? Is it from the market answer? For

00:15:00

Karim Khayrat: Where did the good and bad questions come from, right?

Amir Feizpour: You know, that's a

Karim Khayrat: Okay, so yeah. I got yeah I got I got the part where the, the talk about the bad questions. For for the second prompt template called Guided by Bad Questions gbq. Which is the, which is illustrated on the right? The on that. On this, on the current slide.

Karim Khayrat: It's it's similar to the vanilla template, but we encourage the language model, contextual aware questions, then the ones from Ms, Marco for that, we use that Marco questions as examples of bad questions and manually. It once as good questions. So rather than finding the answer and part of the input document, the full context of the document will contribute to the answer. With this prompt the model generator good and a bad question for each document. You

Karim Khayrat: Because the gpq prompt by design generates questions. There are different from MS. Marker ones. We use a vanilla prompt to generate questions that are you? Oh yeah.

Karim Khayrat: As they use Ms Markle. Yes.

Amir Feizpour: Uh, Karim. Is there a way to fix your audio problem? Because you're cutting out a lot and it's barely understandable.

Karim Khayrat: maybe I'll just call then on On Ali,…

Amir Feizpour: Don't you? I'm guessing that it is your headphone not your connection.

Karim Khayrat: I use my phone.

Amir Feizpour: Is there another headphone you can use or you? Maybe talk about that?

Karim Khayrat: Yeah.

Percy Chen: I think you are charities for.

Amir Feizpour: Smart, was that?

Karim Khayrat: Yeah.

Percy Chen: Are you a trail? Is his phone?

Karim Khayrat: Yeah. Sorry about that, and joining with my phone. Now, so, That should.

Amir Feizpour: You know, whatever you

Amir Feizpour: You. Okay, yeah. Keep going.

Karim Khayrat: Okay, cool. Yeah. So

Karim Khayrat: all for, for, Question Madison.

Amir Feizpour: You. Yeah, okay. So to summarize, then they're using the questions from the market data set as bad questions. And they are essentially writing. The good questions themselves, just trying to make them more complicated.

Karim Khayrat: That exact. so, in this case, Supervised. Because really re The. I guess. This. You don't they don't need a lot of do. It's like you need a lot of examples. Was. just, Each time.

Karim Khayrat: and,

Karim Khayrat: Yeah, so this is this is the result.

Karim Khayrat: Have. On the market data sets.

Karim Khayrat: On the track ideas and listen data sets. And with different metrics. so,

Karim Khayrat: for, for A unsupervised version with.

00:20:00

Karim Khayrat: Without without the

Karim Khayrat: They have.

Karim Khayrat: State-of. This time. So of course, this paper was was published Quite some time ago. Of aliens.

Karim Khayrat: and, And they have. Or they use them the mono T5.

Karim Khayrat: Different parameters to 20 million parameters and 3 billion parameters. And they compared that. To. Davinci Curie and other from, from the Openai.

Karim Khayrat: opening a you provide the API for and, Yeah, their metrics one of them, this case.

Amir Feizpour: Books. Sorry,…

Karim Khayrat: An achievement meth. Evil meth.

Amir Feizpour: can I Can ask a big question, I'm not sure what I'm looking at. So,

Karim Khayrat: You.

Karim Khayrat: oh, Yes. So On like the Mon. Over here is the River. Already. so the the inference, the inference stage for the retrieval is first Actually. A bag of. It was with BM. Life.

Karim Khayrat: it's I related to 100 documents, but I'm not. remember Ex, Remember.

Karim Khayrat: the Mon I Dec. One. and, Mod. Equal number of, Examples in their frontier.

Karim Khayrat: Published the State.

Karim Khayrat: Which is.

Karim Khayrat: You. and,

Karim Khayrat: Since generating. Degrees. and, The. The.

Karim Khayrat: It for the courage. just, Of. Is from.

Karim Khayrat: and, The Great Document.

Karim Khayrat: Again. is instead of

Karim Khayrat: Was.

Karim Khayrat: And the score of Monotypically So it's,…

Amir Feizpour: You carry,…

Karim Khayrat: Monotypically directly. it's…

Amir Feizpour: sorry. Hold on. Hold on.

Karim Khayrat: Yeah.

Amir Feizpour: It's impossible to. hear what you were saying, there is still

Karim Khayrat: okay, so I I left,…

Amir Feizpour: oh, That's much better,…

Karim Khayrat: hello, can you

Amir Feizpour: I don't know what changed, but yeah, that's much better.

Karim Khayrat: I changed, I changed from my Wi-Fi to to 3G. Hopefully, that fixes things. Sorry.

Amir Feizpour: Yeah, your your home Wi-Fi sucks. You need to change that.

Karim Khayrat: If it's been I'm sorry but yeah, don't don't, don't go to bed. Don't go to Rogers, don't go to anything here. Okay so sorry version, two of the language model instead of the opening eye, Curie model the use and gptj for for generating the synthetic query. So the they move to an open source model and,

Karim Khayrat: The GPG. Pro is prompted with three examples from Ms, Marco? and,

Karim Khayrat: Instead of filtering. By by probability, they filtered by mono, T5 model. Until? let's say the standard one just to

Karim Khayrat: get a score from it. And of course, this is more expensive. This step would probably be more expensive than than filtering by probability. But at the end of the day, if you, if you have the hardware to run this fine tuning, this step should be relatively be cheap. And since you're you're only doing this once to generate the synthetic data and then a mono T5 3B is fine-tuned on the synthetic data and here they they compare. You know. At the mono T5. Trained on Marco alone. Trained on Marco plus in parts V1. And sorry, fine tunes and on Marco plus in in parts version one, which is the previous paper.

00:25:00

Karim Khayrat: And the version 2, which is the updated one. And the good thing is that the scores don't change by much. So an average, it's 0.539 versus 0.545. And in this step, the

Karim Khayrat: In this step, of course they they used everything is open source so they didn't rely on the open AI API at all over here. but, but I would criticize this part is because in version 2, They use this query document pair filtering, not based on the probability, but based on on filtering with the model itself.

Karim Khayrat: But version one didn't do that. They just looked at about probability. So probably there were they would get a better score if they did that and version one and, and kept the open AI API as this. I'm just speculating here but it's not like a fair comparison, of course, like if you want to compare open source versus close source, This step I think is is quite important. So maybe this this step using gptj brought down the scores down but this step fixed it. Bit by providing higher quality filtering.

Amir Feizpour: So, again, like I don't understand this table, the rows are different data sets.

Karim Khayrat: Yeah, the the every rose, a different dataset.

Amir Feizpour: Okay. So how do I read this? So they, they did a retrieval with Bm25 then fine-tuned with Marco. And that's a performance. They're getting after that.

Karim Khayrat: This is BM 25. Pure beam 25. And this is the

Karim Khayrat: ndcj of the top 10. I think maybe Suhas is better to explain this metric. That they're, they're reporting.

Amir Feizpour: Ation. Okay. So the metric is a separate question.

Karim Khayrat: If? Yeah.

Amir Feizpour: What I'm curious about is What's the difference between? Like, what is the column that that is called Marco?

Karim Khayrat: Yeah, so mono t53b fine-tuned on on the Marco data set.

Amir Feizpour: Okay, so, fine tuning on Marco and then tried on all of these data sets. These are the numbers. And then in courses added,…

Karim Khayrat: Yes.

Amir Feizpour: I see.

Karim Khayrat: You. So on, on the average metric. You know, improves.

Karim Khayrat: You know. Marginally. So

Amir Feizpour: but, Yeah, it looks like there is no advantage in adding imports and imports too. Is there

Karim Khayrat: But, I'm sorry. So I'm, I'm not familiar with how this small of a difference makes in practice. but, Yeah, there's

Karim Khayrat: that doesn't seem that way.

Amir Feizpour: Okay.

Karim Khayrat: Yeah, so so yeah. Does anyone else have any questions?

Karim Khayrat: Otherwise, you can move on to the next paper.

Karim Khayrat: Okay, so I'm gonna stop sharing this tab. I'm going to start another one with with Self-instruct.

Karim Khayrat: Yeah.

Karim Khayrat: Okay, so so this I'll give a top level overview. there's, so, for self-instruct, the goal is to generate And instruction data sets, synthetic instruction data, sets not for a particular. Instruction like the previous one was a particular instruction. Like, you, you given a document generate the question, right? Imagine that you want to generalize this. Given. Any task generate the prompt for that task so it's a lot more general. And basically what we've seen before with the document question is a subset of So, how does this work? So initially, It.

00:30:00

Karim Khayrat: The data the the input to the to the process is 175 seat tasks which is like a reasonable number of seats. And every task it has one instruction and with one instruction, of course, and one instance per task, for example, you can say sort these numbers In ascending order. So that's the instruction. And then the instance per task, is the input and the output. So the input would be the array of numbers. You want to sort and the output would be the sort of and some and you'll see later that you know some tasks don't have an input, you just have an output.

Karim Khayrat: Like let's say right the letter to whatever or they can also be stated in us in in a four in a rigid form of you Know, instruction input output. Like right, a letter according to the following instructions and then the input would be You know through the school. And then the output will be the letter to the school. So you can they talk about this in the paper and detail and the the allow for both kinds of instructions. Let's say for more robustness So now you have a task tool. So this task pool consists of the tasks and then instruction and the instance per task. And this task tool is going to grow, so it starts with 175 but the aim is to grow it to bootstrap the model to generate more and more of this task pool. Okay, so the task pool, you give it to the language model.

Karim Khayrat: And it first step is instruction generation. So first step is to just generate instruction, not the input to not the output you instruction. So here's what the prompt looks like. So, come up with a series of tasks. That one. Instruction for existing task their student instruction class. So and so on. And then they ask the model, the the model to keep on generating and they stop at 16 tasks, but I guess they stop based on the limit for the output at that time. But yeah, you can potentially give it a larger problem and and more tasks. But again he there's also question of, you know, the cost of doing this with opening the I API. So yeah, so this instruction from so it's very cool. In a sense. Now you're just generating the instructions. You're not generating any input or

Karim Khayrat: The next step. Will will skip over this classification. Task Identification parts Later. I'll explain why? If, if you're interested but basically you can look at the bottom task input outputs over here and the instruction here is, for example, give me a quote from a famous person on this topic. And you have that, you have the instance for the for the staff. So you have the input and output. And you're asking the model. To generate both the input and…

Karim Khayrat's Presentation: You.

Karim Khayrat: the output for the given instruction. So this is this is an interesting prompt. So, instant generation come up with examples for the following tasks. Try to generate multiple examples of possible in the task doesn't require additional input. You can generate the output directly. This is the, This is the prompt start of the prompt and i, The problem wouldn't have fit in the slides on breaking it down. So, this is like the top of the prompt example task, which exercises are best for you, reducing belly fat at home, and this that doesn't require any input. It's it's already, you know, gives the output directly But these are all good suggestions lying leg raises like in and out. Planks syrups pipeline sector.

Karim Khayrat: And then there's another task convert, 85, Fahrenheit to Celsius, and then there's another task sort, the given list sendingly, and over here, the problem. It's given in the prompts, you know, an example of generating two examples. Instead of one example. So here's example one, here's the input is output. Example, 2 input output.

Karim Khayrat: and here's another task, turn down a job offer by sending an email to a recruiter and here's the letter and then finally, The right before you give this prompt to GPT, you ask you the task and you input that you input one of the instructions. That you generated using the previous prompt over here. and then, so basically your bootstrapping, everything over here, your generating suctions, and then from the instructions that you generated your generating, these input output pairs, Yeah, and yeah. So do you do you have any questions till this point?

00:35:00

Amir Feizpour: And you repeat everything you said again, I'm kidding.

Karim Khayrat: Yeah, sure. So

Amir Feizpour: It. Can you do summarize it because it was a little longer. So in case anybody lost the overall picture

Karim Khayrat: so, in summary, You you first step you generate instructions. You ask Gpt? Or any large language model to generate instructions. For you given a set of example, instructions. And then after you collect these instructions, you ask Chat GPT to. To generate the input and the output for these instructions.

Karim Khayrat: Right. And that way, you generate a lot of synthetic. Data. That. Overall includes. both the instruction and the tasks for those instructions for sorry, the input on the output for those instructions.

Amir Feizpour: And the goal of this process is to create a data set for fine-tuning.

Karim Khayrat: Yes. So goal is of this process is to fine-tune A model that can then be used like, charge GPT. Like where you where you ask it for instructions. You give it some instructions and it follows your instructions.

Amir Feizpour: I see. So the goal is to create an instruction. Fine-tuning this

Karim Khayrat: Exactly. And now the filtering step. They try to filter to decrease the overlap between the similarity between the generated tasks. So you have a task, if it's very similar, let's say in the data set, they try to increase diversity of the data set. By filtering out. Filtering out tasks that are similar to each other.

Karim Khayrat: And yeah, so that's that's it at the top level, I think it's it's It's a more general data synthesis approach than the document question retrieval. Of course, you have different Aims in each one. But what I've, what I like about this one is that it actually

Karim Khayrat: Um it's it actually can generate the way they set things up together over here. You can actually generate, you know, And unlimited amount of training data by using another model. So don't, this is also something that's interesting is like if you have open AI now. You know, those everything is closed source for them, but you can query open. Yeah, let's say you have an infinite amount of budget. You can acquire you open AI and get it, you know, a killer data set and then train your own model. And you can get a similar performance to open AIs, you know, close source solution. So just because a company closed sources, Their, their data, but exposes the models via an API. You can still get, you know, use their model to to boost an existing model.

Karim Khayrat: and, I'll show a Web page now that that they use this approach to improve. Using charge gdpt, but in just trying some some some results. So what they did over here, they did a human human based score. So they're, they're human experts, which which involves five of the authors of the paper. And ranked the the outputs of the model as a B C or D. So A is correct and satisfying response. B is acceptable response with minor imperfection see response to the instructions but has significant errors and these completely and invalid or irrelevant response. So this is with vanilla gpt3. and,

00:40:00

Karim Khayrat: gps3 plus t0 training, I forgot what this these This three. Will correspond to.

Karim Khayrat: And then there's the gpt3 self and self-instruct, which is what the paper, what this process that I showed here does. So You get a significant improvement? In performance. and they say it's comparable to, you know, stock GPT one, two, and three,

Karim Khayrat: And yeah. So so that's that's where the the that's where the left of at that point of the paper, But then,

Karim Khayrat: Let me.

Karim Khayrat: Self.

Karim Khayrat: f\*\*\*\*\*.

Karim Khayrat: You first?

Karim Khayrat: okay, so that's where the left of the paper, but everything since then, there was a recent Recent development. I will share another tab like

Karim Khayrat: so,

Karim Khayrat: yeah, so there was a recent development called Alpaca So they, they introduced alpaca 7 billion Model. 7 billion. Parameter model fine-tune using the llama 7B model. Which is open source. On 52k instructions, following demonstrations and these 52 instructions were generated using the self-instruct process.

Karim Khayrat: And they claimed that alpaca behaves qualitatively similar to opening eyes, text DaVinci, DaVinci 003 while being surprisingly small and easy and cheap to reproduce. So they they use a budget of less than $600. To, to do this. And they have a similar similar. Let me zoom in over here. So they have a similar, you know, similar process, the use text DaVinci and 175 Self-estruct CT tasks.

Karim Khayrat: And then they generated 52k instruction following examples. And some support advice fine-tuning of the meta. Model of the metas llama 7 billion model to get alpaca 7 billion.

Karim Khayrat: Well, and The evaluation sorry about difficult, it's all it's it's a blog post. So there's not a lot of technical details and they say that they did some things differently. But here's the Web demo. Okay. And yeah, so they claim that this is as good as as DaVinci 0, 0 3. But from my, from My Um, for my experience, I query that with some something that I know about like, for example, I know about books on Food Dynamics, right? So I asked you to recommend books on introductory food dynamics and just gave me like books that I would never read but charge GPT gave me actually books that I have read in, in my academic year. So so

Karim Khayrat: So, and maybe their claim is, it's similar in. In standard. You know standard time more standard, or more widespread, more common tasks like writing a letter or doing or others, or summarizing a document or getting questions from document and so forth. But there's, there's still a limitation because you still have a relatively small model and it would inherit everything that the original lama 7 billion model has. So if so that's that's something that I noticed thing on it and yeah encourage you guys to play around with it. but yeah, like this is this, I think it's an exciting development in the sense that if you can, if you can, if you have a budget of $600, you can potentially, you know,

Karim Khayrat: create something that that similar to charge GPT. Fine tunes on your own documents. That's also, you know, and and the recipe is quite intuitive, you know, you just generate instructions and then documents that. Oh yeah.

00:45:00

Karim Khayrat: So running out of time, I guess. So you have any questions? I can go back to the slides.

Amir Feizpour: I'm good. Anybody else questions?

Amir Feizpour: Okay, if not something. So everyone. Thanks Karim for presentation. We will see you all next week. If you haven't decided what paper you will look at, but come back for

Meeting ended after 00:45:36 👋