

LLM Recommender

Transcribed on August 5, 2025 at 7:52 pm by Minutes AI

Speaker 1 (00:00)

In mechanical engineering.

Speaker 2 (00:02)

Okay.

And Elena.

Speaker 3 (00:06)

Yeah.

Welcome everyone.

So I'm Anna, I work as a senior portfolio researcher for DWS.

I've been doing that for the past eight years and I'm also currently doing a part time PhD with Philip at UCL, also in computer science focusing on AI applied to financial market analysis.

Speaker 2 (00:25)

Do you want to say a little bit about dws?

Speaker 3 (00:28)

Oh yeah, dws.

I think we rebranded ourselves quite a bit in the past.

So it's the former Deutsche Asset and Wealth Management.

So we are the asset management part or arm of Deutsche Bank.

We are kind of a separate entity and are a multi asset asset manager, meaning that we cover all investment products from the liquid businesses from e.g.

eTFs, equities, fixed income to illiquid investment products like private credit.

And we have also multi asset team.

So so we cover basically the full range of investment products for retail and institutional clients and specifically on the department I'm working in we manage or basically the client relationships or the products for institutional clients.

Everything from like sovereign wealth funds, pension, pension funds, insurance companies.

So basically that spectrum and that's where we also kind market our quantitative strategies for those kind of clients.

Speaker 2 (01:28)

Good.

So we think there are well over 20 different LLM models ranging from you know, sort of the standard sort of foundation models like chat, GTP to deep models for you know, multi, sorry multi, multi models, you know, like genetic AI and mixture of experts.

There's also a number of deep models specializing on things like concepts etc and then there are an increasing number of specialist models for you know, education and what, and what we, what we're interested in doing.

Because most people know hardly any, you know, most people who want to use LLMs know hardly anything about them.

And so is it possible to build a recommender system that if you tell them which financial application you're interested in, it can then suggest which LLM model might work best for you.

That's the basis.

The other part of the, of the project is because there are LLMs coming out all the time, particularly for specific domains.

We'd like to build a sort of horizon scanning system that identifies, you know, new LLMs as they emerge.

Yes.

Is that we all agree on.

Speaker 3 (03:06)

Yes, that's what we're doing.

Speaker 2 (03:07)

Good.

So Sarah, do you want to start by saying what you think the structure of this project if you break it down into components?

Speaker 4 (03:19)

Yeah.

So I don't know if you read the files I sent to you, but what we're thinking to do is divide our project into three modules.

So the first one's going to be like kind of a radar that's going to track new models and they're going to be like classified.

The second module would be like a recommender, so we'd build like recommendation system and the third one would be the explainer which should basically output a full dashboard of information about the recommended module, the recommended model, why we recommended it.

And maybe we're thinking to like also provide like advice like prompt engineering or rag setups for the specific task so we can reduce maybe like hallucinations and improve like basically make sure our users get reliable results.

Speaker 2 (04:25)

Good, thank you.

So now go to the Indian students.

Do you have any questions for Sarah about the project?

No.

Was that a no?

Speaker 5 (04:39)

Actually we already had a meet with Sarah earlier discussing the whole project and everything.

So whatever doubts we had, we already kind of pitched them in front of.

Speaker 4 (04:51)

Yeah, I told you to prepare questions if you have had any.

But if not, like I have my own questions prepared as well as taking yours into consideration from our last meeting.

Speaker 5 (05:02)

Yeah, let's begin with that then.

Speaker 4 (05:04)

I really want to.

Would love to hear like Anna's thoughts on the project and like if you see this going somewhere.

Speaker 3 (05:12)

Yeah, so yeah, I read through the brief and the presentation that you sent, looked over it and also thought about basically from a practitioner's perspective.

Right.

Because we analyze or have in touch with a lot of startups and companies in that space that basically provide also solutions and we develop our own LLM solutions in house and thinking about basically what kind of criteria or what are things we are looking at.

And that goes a bit more detail right when you start building that up.

But I feel also kind of the cost aspect of having, for example local LLMs.

So that's kind of where you can probably also dive deeper into that aspect because then when you productionize it, right, Costs is always something where everyone tries to reduce costs.

However, if the larger LLMs are the more general ones, they become larger and larger.

So even the input and the output kind of tokens cost you even more.

So for some use cases, as for example you also mentioned in, in your brief with sentiment analysis you have smaller models like bert.

Where also in a lot of like quant conferences you often hear kind of the term like Occam's razor that you want to, when you want to solve a problem, like for example sentiment analysis, that you would try to have the smallest model possible in order to solve that kind of problem.

So you would rather go for example with a BERT based transformers model to do sentiment analysis rather than for Example chatgpt or like a GPT or LLM based model because these are more general tasks.

And so yeah, I mean it's more like aspects but I think it's a great idea because often what I see when also we start developing AI applications or source them out is the question of which LLM to use, how much is there actually a difference between LLMs and a lot of companies have for example a preference because I don't know, we use the gcp, the Google Cloud platform, so we would have more free access to Gemini.

So to some extent our internal systems provide access to Gemini, but we also have for example the LLAMA models and then it's often a question of how much is actually the module improvement of using one or the other.

So I think it's basically a good project especially for corporations or people to kind of get started with developing AI applications and provide basically the groundwork.

I think the only question is right, what kind of characteristics you gonna source or have in a more unique way collected in order to build the recommender system?

Because I think that's basically where you can actually make a difference.

What I've seen on the market and one thing when I read through your brief was I think, I mean potentially you could also think about because you said specifically for example for finance you want to provide a recommender system and since a lot of there's some research in that space where you have financial corporations or also academic institutions in the finance space applying LLMs and also explaining in detail in their academic papers why they've chosen one LLM over the other.

I mean sometimes it's cost aspect, sometimes it's just they thought to try that out and I feel like maybe you could also collect that kind of information and then analyze it and use it as additional characteristics, for example for your recommender system.

So these were just like general things when I read through that would be quite, quite helpful.

And of course the aspect that you're going to provide like kind of like a library of what's out there, what's general, what's finance specific, I think would be helpful.

And then of course drawing down what you can do because right, it's not just recommending you can use, I don't know, Claude and not Gemini.

You mentioned rack systems or prompt engineering techniques.

So the more detail you provide of how such an architecture could look like, I think having a tool that actually explains it in simple words, it's something where I would also see for example also my colleagues that are not that quantitative or that familiar in that space is something where I would think there is value in it.

But of course it's some things trying out what works and what doesn't.

But I think it's a good idea and I think definitely helpful because I haven't seen such a specific tool, especially for specific aspects, be it finance or other things on the market or some startups.

Speaker 4 (09:56)

Yeah, that's really great input.

Thank you so much.

Yeah, so we are thinking like what we.

Because like when I was having the meeting, first meeting with my teammates is that the question that arise the most is how do we make our system trustworthy?

Like, how do we get our users to trust us?

So we're thinking maybe we could start off by looking at like benchmarks or like leaderboards.

There's like we find like some little boards out there, but like we want to take everything into considerations, especially like scientific papers.

Yeah.

For that also I wanted to ask.

Speaker 3 (10:38)

You, like maybe just one point because that reminds me of.

Right.

Because the whole AI, generative AI kind of research area is currently emerging.

So I mean, it's still in development in the sense of there are no strict guidelines, as you said, they are like benchmarks or like leaderboards, but every leaderboard tells you a slightly different story.

So exactly what you said.

If you combine all of them that you find you could provide people with kind of a better understanding where they excel and where they don't.

Same goes for prompt engineering techniques.

Like everyone has a different name maybe for the.

For the same thing.

So by kind of mentioning that concept is similar to that and kind of taking the definitions and all that and kind of having like a library for that, I think that definitely adds value to such a system.

Yeah, Sorry, I didn't want to interrupt you.

Just something came to your mind.

Speaker 4 (11:31)

Yeah, I was just.

Thank you for that.

I was just gonna ask you, like, what kind of.

Because we haven't looked deeper at financial tasks that might be solved by AI.

So I really wanted to ask you, like, what do you think could benefit from this sort of recommended tool?

Like, what kind of task would our recommended tool be able to help with?

Speaker 3 (12:00)

Yeah, I mean, a lot of experiments that have been done in the beginning when ChatGPT came out was like sentiment analysis.

But personally, the plain vanilla sentiment, where you just look at positive tonality and negative tonality is not really something I would say that's what you should use an LLM for.

Because it's a probabilistic language generation model.

It's not naturally kind of done to do sentiment analysis.

However, you can create kind of like custom data with LLMs.

So that's also one thing that I explore in my research.

I create with an LLM like a synthetic data set based on my human rules.

And then I create kind of my custom, I mean, you can call it sentiment or stock impact score, whatever.

So things like generating synthetic data, a lot of people use it for sentiment.

So you could still have it in your system because you would just recommend people would still do it and want to know if I do sentiment analysis, what's the best approach?

But then you would probably go down the road of you'd rather use a BERT model than an LLM.

But yeah, basically that then you have a lot of things about when you have tables that you can describe, for example, what the numbers in a specific table mean.

Other use cases in the quant space that I've seen is LLM to SQL.

So where you kind of have like no code environment.

So someone that for example is not familiar with SQL or has data in the background and he's like, okay, I want to compute for all my securities of the S&P 500, the earnings per share figures for the past 12 years.

And then I don't know, I also want to square them.

And then it would basically transform it in a SQL query and it would retrieve it from your database.

So tasks like that I've seen also applications of LLMs.

Otherwise what we also have, I mean, a lot of reporting tasks, but that goes into that direction of how you extract, for example, information from tables or how you combine different data sets.

For example, if you would write a fund report of your performance of your fund of the last month, you would have all the figures, what drove the performance with securities, which sectors.

And then you would take an LLM, take also like a new source and combine all that data to kind of generate the final report, for example.

But then you have a lot of connection points.

So it's not just the LLM, it's also which kind of data source you're going to use.

All that thinking about that, I mean otherwise, I mean a lot of like back office functions like compliance or like risk controls, we have kind of use cases across the company, which I don't know, it's not my specialty, so I'm not that excited about that.

But still very, very valid ones.

I think one thing that I've also seen is, I mean translation obviously is a big thing, but also things like when you have like specific regulation policies within a company, for example, you have to be careful whether you say I mean at the beginning of the Ukrainian war was like do you say conflict or war?

Because it's a very political sensitive topic.

So in our company we had specific rules that you how you're gonna kind of to the outside world communicate these kind of topics when you explain for example performance.

And so you would basically you always had a human component where a human needs to check once you mention that whether it's the correct way.

But you could do that for example with an LLM based system.

So it's more like controls that you could also automate with LLMs and then of course text generation.

So when you write, when you write reports or client pitches and one thing, I mean it often write the leverage comes when you connect it to some database and have like a rack based system to kind of analyze all your documents and reduce the hallucination problem.

So the hallucination problem that you mentioned, I mean that's especially important in the financial industry where we are so highly regulated.

So therefore you can see that a lot of financial organizations are a bit slow in implementing full AI systems.

Because first of all we are regulated.

We can't really.

We always need the human in the loop component.

But everyone is very cautious that right.

If you have a wrong number that that could be a significant kind of financial impact on your company.

And one thing is, for example you have, I mean that's very specific example that I've applied it for.

And some of my colleague you have things called requests for proposals.

So a lot of clients send us questionnaires about our strategies to kind of tell them what our strategy is doing.

So there are specific questions which are similar from client to client.

In the past 10 years, there's not huge changes, but some and we have like a database of all the historic questions with the answers.

And so we are using basically an LLM based system to retrieve, to retreat to retrieve the answer that fits the new question with the historic ones and then write basically the answer before refine it and things like that.

I mean it's very broad.

Right.

Because finance is a big industry.

I can only say what we do for investment management firms and what I've seen around.

But yeah, it's kind of like in a nutshell what I would say some areas which are very promising.

Speaker 2 (17:17)

Would you reckon I would start off with a, say an area like sentiment just to you know, get going?

Speaker 3 (17:23)

Yeah, I think sentiment is probably the most straightforward one that I've seen at least in the technical quant community, people using LLMs.

But yeah, or as Sarah Mentioned like a rack based system where you just like I have, I don't know, an annual report of a company and you want to extract information from there in a very factual, correct manner.

So these kind of two are the most broad one or like the most I've seen.

Speaker 2 (17:58)

Questions from anyone.

Speaker 4 (18:00)

Yeah, I have.

Okay, go on, go on.

Speaker 1 (18:04)

Okay, so.

Speaker 5 (18:07)

When we are talking about in the context of making an LLM recommender system in finance, what do you think is like the hardest part to get right?

Like we have to get right back in the system.

Like talking about tracking, advising, explaining or let's say recommending.

And as Sarah mentioned before, how would you recommend like the user's trust part?

How would you recommend measuring the trust or let's say usefulness of the model recommendation?

Speaker 3 (18:36)

Yeah, I mean for the build up phase I would say because basically the difficulty is right of actually creating the characteristics that gonna build your recommender engine.

And so I feel that's where you, you can excel of combining different sources, updating them in real time or like on a monthly basis.

So being up to date, having a broad set of sources and looking at statistics where others may not have done it.

Like as I mentioned, you can take academic reports or news articles and kind of count which model, how has it been mentioned and what's the traction on that model.

You could also track kind of Reddit or social media platforms of what the traction there is.

So by using alternative data, incorporating or transforming it into some form of characteristics for your recommender system, I feel that's where you can make a difference in the recommender engine.

And then in the end, I wouldn't say because right you first of all building that up is going to take some time and then the overload valuation of how helpful it is, I mean that's more like a trial thing.

You can give it to, to either financial professionals or to students in finance departments that use LLMs for some of their work and then kind of collect feedback.

Because at the end of the day, I mean you don't necessarily need thousands of data points of feedback.

You could do it in a small scale.

However, you could also collect feedback systematically.

Like the same way as the Reinfor enforcement is done for ChatGPT, right?

When, when you do, when you test ChatGPT for free, they kind of use you as kind of a test away.

It gives you like two solutions and it asks you which one do you like more.

And in a similar way you could also systematically do your evaluation if you have like sent this thing for example to, to.

To A lot of people and see what kind of their feedback is.

But I think small scale feedback of I don't know, 20, 30 people should be at the beginning enough to kind of nudge you in the right direction.

Speaker 1 (20:47)

So I wanted to ask something.

You said that one of the ways that we could use LLMs was to generate data, right?

Yeah.

Can you specify for what are we exactly generating data of?

Like what are we generating data off?

Speaker 3 (21:08)

Yeah, for example, I mean one example I can say, which I'm looking more specifically into it is as I mentioned with sentiment analysis, sentiment looks the tonality of positive and negative words.

But if you read for example a statement, I don't know, Apple is opening up a new factory in, in, in China, for example, everyone will have a specific article about it.

Like an analyst, fundamental analysts would put this maybe into context and would be like, okay, there is like a wide scale of factories being opened up by Apple across Asia and there's expected to be significant growth opportunities.

And therefore if I would rank it, how important it is, I don't know, would be eight or something.

And so you could basically the same way as for example a human being would analyze text, you could generate score.

Because generating scores or numerical values is difficult with LLMs because naturally it's not something they've done or seen.

So the way I look at it, like if you think about Internet, what's out there?

Like there is not that much out there about custom data where someone created scores or put their viewpoints in in order to train on it.

And so you can code basically human rules or fundamental rules of experts and then generate synthetic data points and then have thousands of sources.

But of course, for example, this one would for example use news articles and then you would generate kind of a custom score in order to run for example another model or have that kind of tagging model for how a fundamental analyst for example would look like or would do this kind of work.

That's kind of the direction of.

But there are also a lot of other use cases where I've seen people trying to generate synthetic data for like, for, for like stock markets.

But again it's numerical values.

It's, it's quite difficult and it's like some use cases I see it's more people trying things out than it's something you would actually put into practice.

Speaker 1 (23:18)

Yeah, I think I agree on the fact that numerical data like suppose I give charge any LLM a prompt and it gives me.

Suppose I give two LLMs the same prompt and they generate two different datas how do I consider which one is better?

So like that is one thing that I feel is very weird.

And that was something Sarah, we discussed with Sarah as well.

Like if suppose I have, I ask a recommendation, I ask an LLM a certain recommendation and it gives me two of them.

Like suppose two different LLMs give me two recommendations.

How do I compare them and say that which one is better?

I think that was one of the.

Speaker 3 (24:00)

At the end of the day, right, the recommender system is never going to be perfect.

But you need to find approximations of how you can recommend something with a good conscience.

And in this case if of course for that specific task there is no leaderboard if you say okay, you want to for example, create synthetic data.

But LLMs have been used for synthetic data generation.

So that's why I mentioned at the beginning when I read the brief I was thinking of academic research because academic research or people that publish papers are the first ones that actually put this into tests and have some findings.

Plus people that write blogs about these kind of things.

So online social media data is probably something which going to help you the most in order to make an approximate decision of what LLM is more helpful.

And some LLMs or like the Claude ones, for example, they're very natural sounding, they're very great of writing text.

However, more recently for example also Gemini, I've seen that they kind of picking that up.

But also Gemini is being used for more for example analytical tasks.

And then if you read blogs or analyze blogs and see that you can exactly see that trend that people also have similar findings.

And so by breaking down the tasks that you're having, for example synthetic data, generation, finance and then maybe another keyword, let's say you work with keywords, then you could build up your recommender system of those topics that are.

Then the data behind is collected through alternative data that you can for example find online.

And that would probably be I would say currently the only approach because most of the tasks that people write would try to use the recommended system either have been tried out a couple of times by someone or could be also very niche.

But like a very niche task is always part of like a bigger group.

So you could have some kind of taxonomy that you go through.

But then yeah, depends how specifically you're going to create this recommender system.

Speaker 1 (26:10)

Yeah, and I think when, when we are searching for some data that has previously been created, like you mentioned, of course academic papers, they are the most reliable and as we go down Reddit and I would say vlogs, they to some extent become.

Don't you think they become unreliable?

Because can't anyone just write stuff on Reddit?

And like, how do you, if we are making an LLM to consider that considers those into account as well, like, won't it be.

Speaker 3 (26:47)

Yeah, there you need, of course, when you have things like Reddit, you need like consensus.

So basically you wouldn't rely on basically one person saying, okay, I use that LLM for synthetic data generation.

But if 50 people on Reddit said, oh, I prefer Gemini to Claude to generate synthetic data or some topic relating to synthetic data, then you could say kind of the consensus on the social media goes towards that model.

Plus you select it for maybe academic papers or also publications like Medium or something.

You could also track the followers of blogs or things like that.

I mean, it's just, you could go very broad.

Of course you have to select one thing.

But in terms of like, for example, Reddit it is, right, I wouldn't rely on one person, but if there is some form of scale, it would be fine.

But something is still better than nothing.

So imagine a world where nobody has ever used LLMs for synthetic data generation, but 10 people on Reddit did it.

You could still, if you have the data, put it in your recommender system and say there is some evidence that that model is good for synthetic data generation for these kind of tasks outside of finance.

So having in your recommender system the disclaimer of, I'm recommending I don't know, Gemini over Claude, because that you could still have a certainty level and say with, I don't know, 40% certainty or the certainty is not high so that the user still understands, okay, there's some evidence, but I may still need to dig a bit deeper into that.

So.

But the more context you provide to the user, the more helpful it is.

I mean, you always have to compare it, right?

The person that's gonna use your tool, I mean, he's either gonna go onto Google, put it in there and just do his search for a couple of hours for this task, what he's gonna do or is going to type it into your tool and then you have to provide basically additional marginal value to what this person would otherwise do.

So yeah, that's why, I mean like it's never going to be perfect, but you can, the more context and the more information you provide to the person that's going to use your tool, the, yeah, the more, the it's still some value and it's going to be helpful.

Speaker 1 (29:06)

Yeah, that, that makes sense.

And also the AI models that are mostly used in finance.

There is one of the metric that it uses is and that's called intelligence.

And again that again relies on a lot of benchmarks and two or three of them actually involve indirect human interaction.

Like I haven't exact, I don't exactly know the details behind it, but it was just one of the sites that we were looking at like so there.

One of the, one of the metrics actually uses humans indirect human interaction for judging how correct particular generated data is other than I guess Reddit and stuff or other Google searches.

So maybe we can also consider that.

Speaker 3 (30:06)

Yeah, I think it's helpful but also always right consider when there's like human interaction are they actually domain experts or are these kind of like random people who have kind of judged it for financial tasks?

So I feel with those kind of benchmarks often it's good to read but once you go deeper sometimes you find errors of.

Yeah, they were like I don't know either.

Any finance field would be nice but also people without a finance background who judged financial tasks for example.

But if we're generally speaking if they were selected.

Well a human feedback loop is always more, always helpful in the evaluation step than not having that in.

So yeah, you could basically have your own weighing scheme.

So if you say the benchmarks that you're collecting had some human component, you're going to weigh them higher for the final score.

So you could also implicitly do that so that not all the benchmarks gonna weight it equally at the end.

Speaker 1 (31:14)

Yeah, thanks Sarah.

That's it.

Thanks Anna.

That's it from the side.

Speaker 2 (31:23)

Any other questions from anyone?

Speaker 4 (31:25)

Yeah, I have a final question for Anna.

So is there a way that you currently assess whether an LLM or an AI model is reliable enough to use within your company or organizations?

Are there any internal practices, evaluation methods or checks that you use to ensure things like data privacy, low hallucination risk risks or even like efficiency of these models?

I just wanted to ask that.

Speaker 3 (32:01)

Yeah, I mean a lot of questions in one but basically in terms of LLM usage, I mean it's one driven by costs and also by security.

As I mentioned in our internal cloud we currently have Gemini and the Llama models and Deutsche bank specific LLM which generally fine tuned Llama model.

So that's more for security reasons why we have that.

However, we also work with external parties which provide for example document extraction AI software.

Like it's a rack based system where they provide us the opportunity to compare different LLMs but it's more kind of like from a user perspective, a user runs through it and then it provides feedback.

So the thing with a lot of AI applications and the problem we're having is that you don't have a systematic analysis of which LLM is good or bad.

It's more like the portfolio manager, like the style of this one a bit more than the other one.

But it's very kind of subjective.

I would sense a bit small scale feedback for my purposes as a quant researcher.

So that's kind of how we've been going along.

And then the problem of hallucination that you mentioned, I mean that's very, very important.

So we kind of prefer of taking longer to develop AI systems and making sure that they're factually correct.

And so we, we rely a lot on like the Rack infrastructures in, in between and having basically linking the information specifically to the output so that for example the portfolio manager that reads the final text can specifically track back where this coming from.

So specific for numerical values which we then don't change and just take them in.

So that's something what we do.

So you have a basic structure of your text and then you just pull the specific numeric values from your database without having them as part of your LLM system.

So you just insert them into the text where you need them.

And so yeah, but it's very important and mostly like rag rather than you would do it with the LLM itself because their hallucination risk is high.

Specifically when you want specific numeric values.

Speaker 1 (34:24)

Yeah, yeah, you've got me.

Speaker 2 (34:28)

Sorry sir, you had a question?

Speaker 4 (34:30)

No, no, no, I said that's it from my side.

Speaker 2 (34:32)

Right.

Speaker 5 (34:35)

I had one last question.

So as far as I am aware, when we are talking about the finance team, usually AI models usually lack experience explainability.

And as far as I have researched, like this is the base problem and we are trying to fill the gap and this is the gap.

So how critical, like with the surge of explainable AI and everything.

So how critical do you think explainability is for getting LLMs approved?

For using like regulated workflows or anything like that?

Speaker 3 (35:09)

I mean the regulatory landscape is a bit complicated in the sense, I mean, thing is, as long as you have a human in the loop component, which everyone does, as far as I know, it is basically fine of using LLMs and generative AI models because you can make sure that someone is checking it in the end.

But fully automated processes and agentic AI is something which is being tested a lot across the spectrum.

However, due to also the European regulation, financial institutions are not allowed to apply them in production processes.

But specifically, basically for the thing that you are developing, because you develop a recommender system for LLMs, I feel the explainability aspect is not as important because.

Right.

I mean the LLM is kind of the product that you want to recommend, but your system still needs to be transparent in the sense of what kind of data you're using, what kind of characteristics, what kind of weighing schemes you apply.

So by like providing people explainability of the AI model for your recommender system, I think that that's enough.

And then basically yeah, explainability is important.

But as far as I see it, like the beginning of when LLMs came out people were like yeah, it's a black box, we can't use it.

But nowadays I feel it's more seen as kind of like a computer.

Like nobody understands fully how a computer works but they use it on a day to day basis.

And people have, have come more accustomed to what an LLM does and what kind of output you can expect from it.

And so explainability is still important.

But I feel in the financial industry people stopped screaming of I need a deterministic model that's going to give me the same answer for every question.

They understood that it's a realistic model that going to give you probably similar answers but not the same and we have to use it with caution.

And then implementing rack based systems in the middle ground and having deterministic models in between to make your content less prone to hallucinations.

That's the trend that that's been going on for, for the past basically year.

Speaker 5 (37:22)

I had one more question, more of like a generic doubt.

So we discussed like three types of let's say module, the Radar one, advisor and Explainer one.

So let's say if you were deploying this internally at your firm, which module would you prioritize building out first?

I want to know your take on this tool.

Speaker 3 (37:43)

Yeah, I mean as also mentioned, I feel the more you provide additional value then what anybody would be able to do with a quick Google search or doing research for a day is added value.

And the more you fine tune or make your characteristics in determining which LLM is good for that specific use case, I feel the more value it will have and the more useful it is as an end user.

And then that's more for example how I as a quant and technical person would use your system.

And then if I think about colleagues who use LLMs but are maybe not as technically advanced then for them of course the output layer that's going to explain them which one is better or what kind of setup or infrastructure set up.

For example, you need a rack system or you recommend that for sentiment you use that and that's a more granular level for a non technical person.

You give in the output.

I feel that's also something where you can financial professionals would like to see but of course financial professionals always split do people understand LLMs and know what they are or people only using as kind of like a tool and then see from there.

So I feel both are important.

But I feel for your project computer science and what you want to build, the foundation is definitely going to be the recommender system with the specific characteristics and then you always can create a nice dashboard in the end and provide information.

But I feel that's probably going to be the core focus and what's going to take you probably the most time and also to make a difference of what's currently already out there.

Speaker 5 (39:37)

Thank you, that was really helpful.

That's it from my story.

Speaker 2 (39:40)

Good.

Are there any other questions from any of the other people?

A new hub, please.

Still there.

Speaker 5 (40:00)

Do you have any questions.

Speaker 2 (40:05)

Rehan?

Speaker 1 (40:08)

No sir, I'm actually.

Speaker 2 (40:10)

Okay, so over to you now to build something team.

Anyway, thank you very much to Anna Helena for you know, giving up your valuable time to give advice and maybe we can reconvene in three weeks time to see your system.

Okay, thank you very much everybody for joining us.

You'll brief the other students who weren't able to join, will you?

Yeah.

Okay.

Thank you very much everybody.

Goodbye.

Goodbye.

Speaker 3 (40:48)

Thank you.