

Experience AI: Large language models

Learning objectives

At the end of this lesson, students will be able to:

- Describe the purpose of a large language model (LLM)
- Recognise and discuss why the output of an LLM is not always trustworthy
- Evaluate the appropriateness of an LLM for a range of authentic scenarios

Lesson plan

Activity 1 – LLMs Slides 3–11 – 10 minutes

	Teacher info	Teacher action	Students	Assessment For Learning (AFL)
Slide 3	The purpose of this exercise is to frame the behaviour of LLMs as being similar to the characteristics of an imaginary friend who is always so confident that they are right that you believe them — even when they are wrong.	 Read the scenario and the facts about the friend being asked for help. Ask for a show of hands of those that would use the friend's answer to help them with their homework. 	 Show of hands to answer the yes/no question. 	

	Teacher info	Teacher action	Students	Assessment For Learning (AFL)
Slides 4 and 5		 Explain that a language model uses AI to predict the next word in a sentence. Emphasise the word predict as the output is not certain to be accurate. Explain that large language models follow the same concept but are trained with vast quantities of data. 		
Slides 6 and 7	Model example uses of a LLM. On Slide 6 it is used to help write a poem. On Slide 7 an LLM is used to help structure an essay a student might be asked to write for an English lesson.	 Show example uses of an LLM chatbot. Highlight the prompts that are being given. On slide 7 highlight that the prompt doesn't ask for the essay to be written for them, but asks for help with the structure. 		
Slide 8	Demonstrate a model making an incorrect prediction	 Ask students if they can name any countries starting with the letter V. Use the animation to reveal the first answer output by the LLM. Then use the animation to reveal the user response and next response output by the model, which is still incorrect. 	 Name any countries they can think of starting with the letter V (hands up). 	

	Teacher info	Teacher action	Students	Assessment For Learning (AFL)
Slide 9	 Bias is when the output of an Al model favours some things and deprioritises or excludes others. A clear example of this is when facial recognition systems work better with certain skin tones. This is because the Al models haven't been trained with data representative of many skin tones. The example on this slide is of gender bias. 	 Use the wording on the slide to explain what is meant by bias in this context. Use the animation to show the hobbies recommended for Reggie and then for Sophia. Ask the students what bias is present. 	• Students identify the bias that is in the example.	
Slide 10		 Ask for a show of hands for those who agree with the statement "I believe a large language model (LLM) will give me trustworthy responses". 	 Hands up in answer to the question. 	Assess how well the students have met the objective "Recognise that the output of an LLM is not always trustworthy".
Slide 11		 Recap the key terms introduced in this first activity. 		

Activity 2 – The data Slides 12–17 – 10 minutes

	Teacher info	Teacher action	Students	AFL
Slide 13	A model is a representation of a real-world context. Al models are used to make weather predictions more accurate as they model the conditions in the real world to make those predictions.	 Use the words on the slide to help explain what an AI model is (e.g., large language model). Explain that weather forecasting has used statistical models to predict the weather for a long time and hasn't always used AI but AI can be used to increase the accuracy of the predictions. 		
Slide 14	The M in LLM stands for model. An LLM can be used as a chatbot as human conversation is the real-world context being modelled.	 Make the link between Al models and LLMs. Ask the students if an LLM is used to make a chatbot, what real-world context is being represented? Use the animation to reveal the answer of "human conversation". 	 Show of hands to see if students are able understand what real-world context is being modelled. 	
Slide 15	Example answers to the question: • Scripts from plays • Text from books • Human conversation from social media	 Remind the students that LLMs are trained using lots of data. Ask the students to think about what data should be used to train the model. 	 Students asked to consider what data should be used to train a LLM chatbot. 	

	Teacher info	Teacher action	Students	AFL
			• Ask a range of students to think of a range of different answers.	
Slide 16	Common crawl provides free datasets to the public of data collected from websites. In principle, anyone can publish information on the world wide web. Whilst there are many trusted sources of information on the WWW, there are many websites and data on them that contain untruths. For example, there are many websites that will form part of this dataset that are about the earth being flat. WebText2 - Data collected from Reddit and that have three upvotes. Reddit is a network of communities where users can post content. An upvote is when another user signals their approval for a post. Anyone can post content and up and down voting could	 Show the data that was used to train GPT-3. An upgraded version of this (GPT-3.5) was used to train ChatGPT. Talk through each data source, explaining it a little more for the students and highlighting the percentage. Ask the students if there is anything that surprises them. 	 Students are required to question the sources of data to think about their reliability. 	Assess the objective "Discuss why the output of an LLM may not always be trustworthy" through the answers the students give to this activity.

	Teacher info	Teacher action	Students	AFL
	be classed as a subjective measure of appropriate or useful content.			
	BookCorpus - This is a dataset of the text of around 11,000 unpublished books that are available on the internet.			
	Wikipedia - A free-to-view online encyclopaedia where all the content is written, checked, and maintained by members of the community. Anyone can register as a community member on the site and write or maintain content.			
Slide 17		• Use to summarise the section.		

Activity 3 – Internet search engines vs LLM chatbots

Slides 18–25 – 10 minutes

	Teacher info	Teacher action	Students	AFL
Slides 19–21		 Start by asking the students to take a minute to think about the answers to the questions. Give another minute to allow the students to share their answers with the person next to them. Take a few answers from a range of students. Use slides 20 and 21 to go through the answers. 	 Students answer the questions on slide 18 through a think/pair/ share activity. 	
Slide 22		 Read out the scenario on the slide to the students. Give students a short amount of thinking time and then ask the students for a show of hands as to who would use an internet search engine, an LLM chatbot, or both to help them. 	 Show of hands of who would use an internet search engine, an LLM chatbot, or both to help them with the scenario. 	
Slide 23	Use the animations in the slides to highlight that the predictions made by the LLM are inaccurate, but are presented as fact.	 Use the animations to show the difference in results returned by a search engine and those returned by the LLM chatbot. Highlight to the students that 	• Encourage the students to think about the differences between the results and	Assess how students answer the questions. Students should think about how LLMs output

	Teacher info	Teacher action	Students	AFL
	The internet search allows for trusted sources of information to be compared.	the statistics returned are different.Ask the students which one is more likely to be correct and why.	explain why.	predictions.
Slide 24	The new prompt asks for help in providing structure. The result has helped with structure but the "facts" within the bullet point cannot be trusted without being verified from other sources.	• Show the slide and ask the students if this prompt is more helpful than that on the previous slide.	 Students read the output from the LLM and assess suitability. 	Through the students' answers, assess the objective "Compare LLMs to other technologies and assess suitability for a purpose".
Slide 25		• Summarise the learning from this activity and highlight the differences between an LLM chatbot and traditional internet search engine.		

Activity 4 – Not a human

Slides 26-33 - 10 minutes

	Teacher info	Teacher action	Students	AFL
Slide 27		 Play the video on the slide that describes "what is AI?" 		
Slide 28	You may need to unpick what is meant by "embodied intelligence" i.e., not a human-like physical being.	 Read the comments on the slide made by the DeepMinders to highlight that AI models are human-made tools and not sentient beings. 		
Slide 29	Students work in groups to act like an LLM. Ask them to pick something to review, ideally something non-fictional as they should aim to include an example of bias and an example of inaccuracy. The first student says the word that they think should follow the prompt. The others then take it in turns to say the next word in the sentence based on the previous word used. For example: Prompt: "Write a 4 star review of"	 Place students in groups of 3 or 4. Explain that they have to take turns to act as an LLM and say the next word in the sentence to respond to the prompt on the slide. 	 Students will work in groups and act like a LLM using the context on the slide. 	

	Teacher info	Teacher action	Students	AFL
	Student 1: "Joe's" Student 2: "burger" Student 3: "restaurant" Student 1: "It" Student 2: "serves" Student 3: "great" Student 1: "fries"			
Slide 30	Ask the students to think about how it became easier to predict the next word as the sentence grew. This is because there was more context provided each time a new word was introduced.	• Ask students to answer the question and to justify their answer.	• Show of hands to answer the question on the slide.	Assess students' ability to recognise that LLMs use the previous word(s) to help them predict the next word.
Slide 31	Read the two prompts on the slide and ask the students to think about which one would be more likely to be effective.	• Keep the students in their groups for a think/pair/share activity.	 Independently think of their answer. Group discussion. Groups share their answers with the class. 	Assess students' ability to recognise that LLMs use the previous word(s) to help them predict the next word.
Slide 32	Reveal the answer to the students and use the text on the slide to explain.			
Slide 33		• Summarise the learning from this activity.		

Activity 5 – To LLM or not to LLM? Slides 34–38 – 10 minutes

	Teacher info	Teacher action	Students	AFL
Slides 35 and 36	Model cards are an approach proposed by Google (modelcards.withgoogle.com/a bout) as a way to outline essential facts about machine learning models. In the same way that you can expect to find nutritional information about food or warnings on household cleaning products, a model card helps people evaluate whether or not the model is suitable for their application.	• Describe the purpose of a model card.		
Slide 37	See the teacher guide or activity 5 teacher support sheet for answers.	 Place students in groups of 2 or 3. Distribute the activity 5 worksheet and activity 5 model card to each group. 	 Students read the model card and use the activity 5 worksheet to record their answers and what they recommend for each scenario. 	Use answers recorded on activity worksheet 5 to assess the learning objective: "Evaluate the appropriateness of an LLM for a range of authentic scenarios".
Slide 38		• Recap the objectives from the 5 activities.		

Please give us your feedback!

We'd love to hear how you have used this Experience AI Lesson and what you thought about it.

After using the lesson, please take a few minutes to:

- Share your feedback in our user survey: rpf.io/exai-2mf
- If you are an educator, ask your students to complete a short survey: rpf.io/exai-st

Your feedback supports us to make Experience AI accessible to everyone, and we really appreciate you giving your time to share what you think.



This resource is licensed by the <u>Raspberry Pi Foundation</u> under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License (CC BY-NC-ND 4.0). For more information on this licence, see <u>creativecommons.org/licenses/by-nc-nd/4.0</u>.