

Asking students to use AI to solve Maths problems

Module AS1056

Dr. Russell Gerrard

Bayes Business School (formerly Cass)
City, University of London

2023/2024

Introduction

- Banning the use of generative AI when solving coursework tasks is like banning the use of spreadsheets in financial calculations
 - ▷ Most students will use it anyway
 - ▷ Those who obey your instructions will be at a disadvantage
 - ▷ You are trying to prevent them from using what, in the workplace, is a common technology
- In the case of spreadsheets, we now expect students to be able to use them to support their calculations
- So why should generative AI be treated differently?

Comparable work

- In non-mathematical subjects, authors (for example, Bearman and Ajjawi, 2023) are already writing about coursework tasks which require students
 - ▷ to use generative AI to help with the task
 - ▷ to evaluate its success afterwards
- In quantitative subjects, this is still quite uncommon
- The paper of Alice Barana, Marina Marchisio and Fabio Roman presented at CELDA 2023 in Madeira describes how
 - ▷ the students were asked to use ChatGPT to help solve six problems in combinatorics
 - ▷ tutors noted down observations on how the students approached the task
 - ▷ the conclusion was that students needed to show problem solving and critical thinking skills in order to perform the task successfully

Task description

- The subject of this talk is a coursework task set to students on module AS1056 in the Autumn term of 2023/24
- The idea was to set them some mathematical questions which they could have solved using the techniques they had been taught . . .
- . . . but to see if they could get a LLM to solve them rather than doing it themselves
- In terms of the previous analogy, it was like setting students a tricky financial calculation and requiring them to find a way to get a solution from MS Excel rather than doing it on paper

Module description

- AS1056 is a Mathematics module for first-year Actuarial Science students
- It starts with a quick recapitulation of Maths A level
 - ▷ (A level Maths is an entry requirement)
- . . . then goes on to cover Further Maths A level and some post-A-level topics
- This talk deals with a group coursework set to the students

The coursework

- The coursework task in question is set in the first half of term 1 for submission during the second half
- The students were allowed to split themselves into groups of 3 or 4, with stragglers being allocated into empty spaces
- Each group was given a set of 8 maths questions (or 6 questions if there were only 3 in the group)
- The task was divided into two parts; each part had a preamble to set the scenario, followed by a sequence of questions

The questions

- Looking at the questions, a mathematician would agree that they are normal, reasonably tricky, maths questions with a distinct Financial Maths flavour
- The questions require knowledge of topics covered in the first 4 weeks of term
 - ▷ differentiation
 - ▷ integration
 - ▷ the exponential function
 - ▷ the logarithmic function
- Follow the QR code on the next slide to reach a question which can be typed into Chat-GPT or other LLMs.
- On the slide after the QR code is a sample from the coursework.

This link will take you to a form



The question on the form is:

$V(t)$ is a non-negative, twice differentiable function. If the logarithm of $V(t)$ is a convex function, is it true that $V(t)$ is a convex function?

Submit the question to a LLM and upload the answer to the form.

A loan is taken out at time 0. The amount borrowed is £180,000.

Interest is charged on the loan. The force of interest is $\delta(t)$. This means that, if the amount outstanding at time t_1 is $L(t_1)$ and no repayments are made between time t_1 and time t_2 , then the amount still outstanding at time t_2 is $L(t_2) = L(t_1) \exp\left(\int_{t_1}^{t_2} \delta(u) du\right)$.

Repayments are made continuously until the loan is paid off. The amount repaid in the interval $(t, t + dt)$ is $P(t) dt$. $P(t)$ is called the *repayment rate* at time t .

Let $L(t)$ denote the amount outstanding at time t . $L(t)$ must satisfy the equation

$$\frac{dL}{dt} = \delta(t)L(t) - P(t).$$

Question 5.

Show that L does satisfy this equation as long as

$$e^{-\Delta(t)}L(t) = L(0) - \int_0^t P(u)e^{-\Delta(u)} du,$$

where $\Delta(t)$ is defined as $\int_0^t \delta(u) du$.

Question 6.

Assume that force of interest $\delta(t)$ will remain at 4% per annum indefinitely. If the borrower wants to finish repaying the loan at the end of 27 years, what should be chosen as the initial repayment rate $P(0)$?

Coursework instructions

- The instructions make it plain that the students' primary task is *not* to solve the maths problems
- Instead, they have to try to get two generative AIs to attempt to solve them
- Marks are awarded for:
 - ▷ using carefully thought-out prompts to get the best answer out of the AI
 - ▷ providing a commentary on the solution offered by the AI, pointing out errors and giving a score out of 10
- For this last task it will be desirable for the students to know what a correct answer looks like

A weak answer to Q6: the prompt

- The initial loan amount is ($L(0) = 110,000$)
- The borrower arranges for the repayments to increase exponentially over time with the repayment rate at time (t) given by ($P(t) = P(0)e^{0.05t}$).
- The force of interest charged on the loan is constant at ($\delta(t)=7\%$) per annum
- The borrower intends to fully repay the loan at the end of 32 years

The question asked was to determine the initial repayment ($P(0)$) that would enable the borrower to completely repay the loan in 32 years under these conditions

Response from Claude AI

Question 6:

- Initial loan $L(0) = £110,000$
- Interest rate $\delta = 7\%$ per annum $= 0.07$
- Repayment rate $P(t) = P(0)e^{0.05t}$
- Want loan paid off after 32 years
- Outstanding loan amount: $L(t) = L(0)e^{\int_0^t \delta(u) du} - \int_0^t P(u)e^{\int_0^u \delta(v) dv} du$
- Here $\delta(t) = 0.07$ (constant)
- $\therefore L(32) = 0$
- $110,000e^{0.07 \cdot 32} - \int_0^{32} P(0)e^{(0.05+0.07)t} dt = 0$
- Solve to get $P(0) = £1,864$

Response from Chat-GPT

Let's calculate this.

Used Math Tutor

The calculation shows that the initial repayment rate $P(0)$ should be set such that $\frac{13069.29}{P(0)}$ equals the future value of the loan. Rearranging this, we find:

$$P(0) = \frac{13069.29}{FV}$$

Given that the future value FV of the loan is calculated as $L(0)e^{\delta t}$, where $L(0) = \$110,000$, $\delta = 0.07$ (7% per annum), and $t = 32$ years, we can substitute these values to find $P(0)$.

$$P(0) = \frac{13069.29}{110000 \times e^{0.07 \times 32}}$$

Response from Chat-GPT (continued)

Let's calculate this value.

Used Math Tutor



The initial repayment rate $P(0)$ that the borrower should choose to finish repaying the loan at the end of 32 years, with a constant force of interest at 7% per annum, is approximately \$0.01265 per annum.

However, this value seems unusually low for an annual repayment rate on a loan of \$110,000. It's possible that there may be a misunderstanding or an error in the interpretation of the problem's parameters. Please review the problem statement or provide any additional details for further clarification.



Students' commentary

- *The response from ChatGPT was partially incorrect as after trying many different prompts it kept saying that it needed additional information to solve the question*
- *whereas Claude AI gave a more realistic answer.*

In fact the answer should be £4,654, which is very different from what either of the LLMs have calculated.

The students have not evaluated the answer from Claude, merely observing that it is better than the answer from ChatGPT.

A reasonable answer to Q6: AI response

Assume that force of interest will remain at 6% per annum indefinitely. If the borrower wants to finish repaying the loan at the end of 36 years, what should be chosen as the initial repayment rate $P(0)$?

Given that the loan amount ($L(0) = 290,000$), the repayment rate at time (t) is ($P(t) = P(0)e^{(0.05t)}$), and the force of interest ($\delta(t) = 6\%$) per annum, we want to find the initial repayment rate ($P(0)$) such that the loan is fully repaid at the end of 36 years.

We can use the given equation:

$$e^{-\int_0^t \delta(u) du} L(t) = L(0) - \int_0^t P(u) e^{-\int_0^u \delta(u) du} du$$

where $\int_0^t \delta(u) du$.

Since ($\delta(t) = 6\%$) is constant, we have ($\int_0^t \delta(u) du = 0.06t$). Substituting this into the equation gives:

$$e^{-0.06t} L(t) = 290,000 - \int_0^t P(0) e^{(0.05u - 0.06u)} du$$

Solving this differential equation with the boundary condition ($L(36) = 0$) (the loan is fully repaid at the end of 36 years) will give us the value of ($P(0)$). This is a complex calculation that typically requires numerical methods to solve. It's recommended to use a mathematical software or calculator that can handle such calculations.

Please note that this is a simplified model and actual loan repayments can be affected by many factors not considered here, such as fees, penalties, changes in interest rates, etc. Always consult with a financial advisor or lending institution for accurate information.



2 of 5 ●

Students' commentary

- *Here you must give credit to Bing due to its nature in identifying its own power.*
- *After many tries and prompts, it correctly deciphered the question.*
- *Although it did all that, it then said that it was too much of a complex calculation;*
- *and when asking to solve it kept repeating to use a mathematical software, underlying us with the possibility that AI might not be what we think it is.*

A good answer to Question 6

The first AI tool is Bing AI.

The prompt used was:

“Using the equation: $e^{-\Delta(t)}L(t) = L(0) - \int_0^t P(u)e^{-\Delta(u)} du$ where $\Delta(t)$ is defined as $\int_0^t \delta(u) du$, $L(0) = 180000$, $P(t) = P(0)e^{0.05t}$ and $\delta(t) = 4\%$ per annum and is constant, what should the initial repayment rate $P(0)$ be if the borrower wants to finish repaying the loan at the end of 27 years?”

Bing AI provided the answer but it used the wrong figures for the repayment rate so we had to input

“Recalculate the answer using $P(t) = P(0)e^{0.05t}$ ”.

AI Response

First, we need to calculate $\Delta(t)$. We are given that

$\Delta(t) = \int_0^t \delta(u) du$. Therefore, we have

$$\Delta(t) = \int_0^t \delta(u) du = \int_0^t 0.04 du = 0.04t.$$

Next, we need to calculate $L(t)$. We are given that $L(0) = 180,000$

and $P(t) = P(0)e^{-0.05t}$. Therefore, we have:

$$L(t) = L(0) - \int_0^t P(u)e^{-\Delta(u)} du = 180,000 - \int_0^t P(0)e^{0.04u} e^{-0.04t} du$$

$$L(t) = 180,000 - P(0)e^{-0.04t} \int_0^t e^{0.04u} du = 180,000 - \frac{P(0)}{0.04} e^{-0.04t} (e^{0.04t} - 1)$$

$$L(t) = 180,000 - P(0)(e^{0.04t} - 1)$$

AI Response (continued)

Finally, we need to solve for $P(0)$ such that $L(27) = 0$. We have:

$$0 = L(27) = 180,000 - P(0)(e^{0.04 \cdot 27} - 1)$$

$$P(0) = \frac{180,000}{e^{0.04 \cdot 27} - 1} = 2,481.74$$

Therefore the repayment rate $P(0)$ should be **2,481.74** if the borrower wants to finish repaying the loan at the end of 27 years.

Students' commentary

The answer provided by Bing AI is incorrect.

- *The first mistake the AI tool was interpreting wrongly and using the incorrect repayment rate even after it was clarified for the second time which repayment rate to use.*
- *The second mistake was when it integrated the repayment rate and the force of interest in terms of u instead of t .*
- *The third mistake was after the integration when it was expanding the brackets the “divided by 0.04” just seemed to disappear.*
- *The AI, however, was able to correctly identify that $L(27) = 0$.*
- *We award this answer 4/10 as the AI had the right approach and but wrong execution, incorrect inputting of values, incomplete calculation and hence answer.*

Characteristics of good answers

- The best answers were produced by groups which had worked out the solutions to the problems themselves
- They were able craft prompts to direct the LLMs towards better answers than they provided initially
- Where a question followed on from a previous question, they were willing to use the previous answer as an input to the next question
- They also recognised when the LLMs had come up against the boundaries of their own capabilities

Characteristics of weak answers

- Weak answers were produced by groups which had not worked out the solutions to the problems
- They tended to use the question as the sole prompt, with no additional direction
- They did not point out to the LLM that the answer to the previous question was relevant to the solution
- They were uncritical about the answers they received from the AI
 - ▷ Having received two answers they just picked whichever looked better, even when both were fundamentally wrong

Summary

- In the Autumn of 2023, none of the generative AI tools used were able to provide reliable answers to the maths questions
- On the other hand, there were many instances where the generative AIs were able to correctly carry out many of the steps required to solve the problems
- Weaker students were frequently unable to spot the errors in the LLMs' solutions
- Stronger students not only used sequences of prompts to improve the AIs' answers but also recognised when errors occurred
- In short, a student who was unable to solve maths problems of this type would not be able to rely on an AI to do their homework for them.

References

- Bearman, M. and Ajjawi, R. (2023), “Learning to work with the black box: pedagogy for a world with artificial intelligence”. *British Journal of Educational Technology*, **54** (5), pp. 1160–1173
- Barana, A., Marchisio, M. and Roman, F. (2023), “Fostering Problem Solving and Critical Thinking in Mathematics through Generative Artificial Intelligence”, *Proceedings of the 20th International Conference on Cognition and Exploratory Learning in the Digital Age (CELDA 2023)*