Sequence

Finished Survey - 6 people

Visualization Usefulness

I think that I would like to use the sequence Visualization Frequently.

Strongly Disagree [0%] Disagree [14%] Neutral [43%] Agree [43%] Strongly Agree [0%]

I found the sequence Visualization unnecessarily complex.

Strongly Disagree [0%] Disagree [29%] Neutral [29%] Agree [43%] Strongly Agree [0%]

I thought the visualization was easy to read and use.

Strongly Disagree [0%] Disagree [43%] Neutral [29%] Agree [29%] Strongly Agree [0%]

I thought the visualization was very cumbersome to read and use.

Strongly Disagree [0%] Disagree [43%] Neutral [29%] Agree [29%] Strongly Agree [0%]

I thought the visualization was very useful for explaining the model's behavior.

Strongly Disagree [0%] Disagree [14%] Neutral [29%] Agree [43%] Strongly Agree [14%]

Model Performance

What is your perception of the reliability of the model in generating code?

Highly Reliable [29%] Not Reliable [43%] Impossible to tell [29%]

Would you rely on the model to assist you in source code generation tasks?

Yes [43%] No [29%] Impossible to tell [29%]

Open Questions

What information from the visualization did you find useful in explaining the model's predictions?

8. What information from the visualization did you find useful in explaining the model's predictions?

The clear differences based on color between tokens were helpful.

I found it useful to see the probabilities linked to each word.

The relative probability of the predicted tokens made the task of explaining why certain keywords were generated. It also helped to debug why the model might be making a mistake when it picks a keyword with a relatively low probability score.

probability of next tokens.

The model helps find out the most likely token, which potentially guides the rest of the sequence.

It was helpful to see how confident the model was in a particular token. This would be helpful in identifying areas where the developer should spend extra time analyzing the resulting code.

Being able to see the probability of each token and seeing the tokens in order was useful, as it gives insight into how much confidence the model has for each token.

What information from the visualization did you find useful in explaining the model's predictions?

9. What information from the visualization did you find useful in explaining the model's predictions?

This appears to be the same question as #8. Maybe this was intended to be "not useful". If so, I found the actual values to be confusing and not extremely helpful.

I found it useful to see the probabilities linked to each word.

Same as above.

probability of next tokens.

The model helps find out the most likely token, which potentially guides the rest of the sequence.
(This is the same question as 8)

The probabilities.

duplicate?

What other information (if any) would you like to see in the visualization?

10. What other information (if any) would you like to see in the visualization?

A method to better differentiate which data is useful rather than just the raw numbers.

I would like to see more about the relationships among words.

The list of keywords, and their probability, from which the model picked the current keyword.

different branches, second highest probability.

As I mentioned previously in one of the open-ended responses, it might be useful to see the second-highest probabilistic token and its subsequent sequences.

Potential branching of the second or third most likely token to be in that spot. Would be interesting if the developer could toggle this feature based on a threshold. If probability is lower than X threshold, then show additional tokens.

I would like to see the tokens with the second and third highest probability that did not get chosen by the model for each token that was chosen. Additionally, it would be useful to see some probability metrics or the average for the whole code snippet.

What elements of the visualization did you like most?

11. What elements of the visualization did you like most?

Color indications
The probability linked to each word.
The ease with which we could locate the probability of any given symbol/keyword.
probability.
Color coding - the concept is useful.
Color coding and the showing of the probability.
I liked that all of the probabilities were shown for each token and that the tokens were shown in a sequential order.

What elements of the visualization did you like least?

12. What elements of the visualization did you like least?

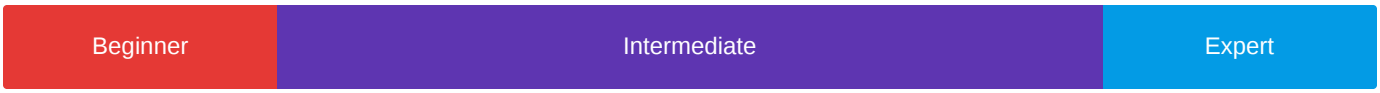
Raw values
Symbols and punctuation marks.
For a long line of code, it sometimes was hard to follow where exactly I was looking at, especially in case of repeating symbols in the line of code.
space information might not be so useful
Interestingly, choice of colors in the Color coding. Although I'd like to see the color mapping from red to green. Blue to red has less contrast, and was not obvious from the model.
The linear structure, python is built with whitespace as syntax. It would have been nice to show the visualization with that same structure.
The probabilities being shown in bubbles with lines being drawn to their tokens made the visualization a bit chaotic. I would have preferred if it were a bit more organized and easier to read

Demographics

Q2 - Background

Researcher (Academics) [30%, 3] Researcher (Industrial) [0%, 0] Applied Scientist [0%, 0]
Software Engineer [10%, 1] Machine Learning Engineer [0%, 0] Project Manager [0%, 0] IT Manager [0%, 0]
Devops Engineer [0%, 0] Other [0%, 0] Student [60%, 6]

Q3 - 3. What is your level of expertise in python?



Beginner [20%, 2] Intermediate [60%, 6] Expert [20%, 2]

Q9 - 5. Are you familiar with Abstract Syntax Tree (AST) representations?

No [20%, 2] Yes [80%, 8]

Q5 - 6. Do you have a background in machine learning? Formal (college classes, degree, certification) or informal (self-learning or other training)?

Yes, Formal [30%, 3] Yes, Informal [10%, 1] Yes, Formal and Informal [60%, 6] No [0%, 0]

Q6 - 7. Have you used AI-assisted tools (e.g., Copilot or ChatGPT) or Large Language Models (e.g., CODEX or GPT) to aid in your programming?

No [40%, 4] Yes [60%, 6]