

CHATGPT AS A SOFTWARE ENGINEER

ABE GIZAW

HENRY KRZYZEWSKI

STRUCTURE OF THE TALK

ABSTRACT

I

CHATGPT 3.5

II

CHATGPT 4

III

CONCLUSION

IV

Abstract

This presentation dives into the performance of ChatGPT versions 3.5 and 4.0 in the context of software development and will achieve the following...

- break down the strengths and weaknesses of these AI models
- Discuss both versions' new functionalities and improvements
- address potential challenges, such as the accuracy of the generated code

CHATGPT 3.5

HENRY KRZYZEWSKI

How ChatGPT 3.5 exceeds

- A focus on Reinforcement Learning from Human Feedback (RLHF)
- Designed to generate prompts within policies based on human values

RLHF is a semi-guided practice that includes hands-on training and is purposed to remove toxic responses and promote factual ones through a rewards-based system.

Step 1 - Collect demonstration data and train a supervised policy

Desired output is provided by a "Labeler"

The response is then used to fine-tune the parameters of GPT 3.5 via this supervised learning

Step 2 - Training a reward model

A prompt is provided but the language model is allowed to generate a set of outputs

The outputs are reviewed by the labeler and ranked from best to worst

These rankings are used to train a reward model

Step 3 - Updating policy using proximal policy optimization

One last prompt is supplied to the model

The model generates a response and we will use it as the input to our trained reward model

The output from the reward model will be used to generate a new reward for that prompt

The main purpose of RLHF is to...

- Filter responses that do not effectively answer the prompt or are negative
- Improve the quality of the response but also the perceived supportiveness
- Basically, make a more "Human" and more "sympathetic" chatbot

Room to improve for 3.5

- The probabilistic nature of ChatGPT
- Does not guarantee that the answers themselves are correct and can result in nonsensical answers
- No consciousness or means of deep thought/understanding

With no direct path for any one coding problem existing, the response is made entirely from relative associations and therefore is often prone to being non-functional.

These nonsensical responses, also known as hallucinations, have the potential to confidently provide false responses and even justify them when questioned

ChatGPT 3.5 and simple code generation

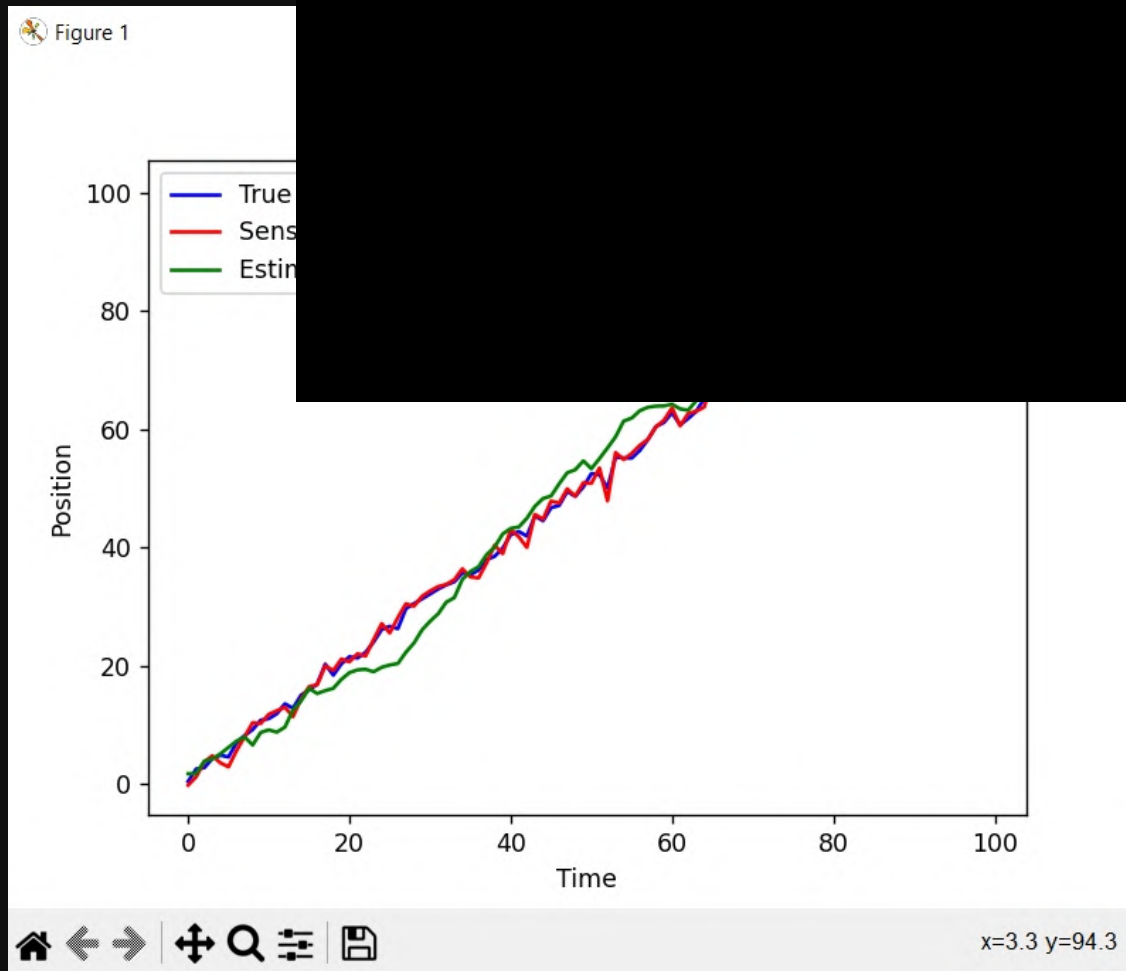
ChatGPT 3.5 was given the same prompt for implementing a binary search algorithm three different times and supplied three different implementations, all of which were functional

- Proficient in producing functional code in the case of less complex software
- There are infinite ways to solve a coding problem
- In the case of simple concepts with few interworking parts, ChatGPT 3.5 shows promise

If code has to integrate into a larger pre-existing piece of software it is more likely that ChatGPT 3.5 will not be able to process this as it would need to have context of all the current code to attempt to properly integrate the new portion.

This is one limitation we can identify immediately since it's not realistic to give all of the software content as a prompt, so we will move forward to identify the bounds of simple to complex more individual coding assignments.

ChatGPT 3.5 and complex code generation



based on sensor readings. I was pleasantly surprised when ChatGPT 3.5 gave me this response...

ChatGPT 3.5 gave me a fully fleshed-out example of a working particle filter. Particle filters are complex, but have plenty of documentation and online examples from GIT repositories.

odel, we can get

ementation of a
cting location


```
PS C:\Users\krzyzehj\CSSE413\Cutting Edge Research Project\ChatGPT 3.5 testing> & "C:/Program Files/Python39/python.exe" "c:/Users/krzyzehj/CSSE413/Cutting Edge Research Project/ChatGPT 3.5 testing/imageGuesser.py"
2024-04-11 00:55:25.349889: I tensorflow/core/util/port.cc:113] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from different computation orders. To turn them off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
2024-04-11 00:55:26.616013: I tensorflow/core/util/port.cc:113] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from different computation orders. To turn them off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
2024-04-11 00:55:32.225347: I tensorflow/core/platform/cpu_feature_guard.cc:210] This TensorFlow binary is optimized to use available CPU instructions in performance-critical operations.
To enable the following instructions: AVX2 FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
1/1 ██████████ 1s 1s/step
Traceback (most recent call last):
  File "c:\Users\krzyzehj\CSSE413\Cutting Edge Research Project\ChatGPT 3.5 testing\imageGuesser.py", line 61, in <module>
    main()
  File "c:\Users\krzyzehj\CSSE413\Cutting Edge Research Project\ChatGPT 3.5 testing\imageGuesser.py", line 54, in main
    top_classes = tf.keras.applications.imagenet_utils.decode_predictions(features, top=5)[0]
  File "C:\Users\krzyzehj\AppData\Roaming\Python\Python39\site-packages\keras\src\applications\imagenet_utils.py", line 136, in decode_predictions
    raise ValueError(
ValueError: `decode_predictions` expects a batch of predictions (i.e. a 2D array of shape (samples, 1000)). Received array with shape: (1, 1280)
PS C:\Users\krzyzehj\CSSE413\Cutting Edge Research Project\ChatGPT 3.5 testing>
```

- Ineffective with recent libraries and general concepts with less existing documentation
- ChatGPT 3.5 does not have the sufficient data to pull from and thus cannot train to properly
- The determining factor in ChatGPT 3.5s ability to perform as a software engineer exists primarily in the maturity of the content online which the model itself is training upon

How should ChatGPT 3.5 be used for software development?

You
Teach me about



ChatGPT

NumPy (Numerical Python) is a fundamental package for scientific computing in Python. It provides support for large, multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays. NumPy is widely used in various fields such as machine learning, data science, engineering, and scientific research due to its efficiency and versatility.

Here are some key features and concepts of NumPy:

1. Arrays in NumPy

At the core of NumPy is the `ndarray` (n-dimensional array) object, which represents a multi-dimensional array of elements of the same type. These arrays can be of one dimension (vectors), two dimensions (matrices), or more.



It is

working parts,

Does not have the capacity to consistently provide functional examples of new software with less mature online documentation

Extremely useful for learning about an unfamiliar library, new language, or any concept that requires explanation without impending detail

In cases of mature concepts with plenty of online documentation, ChatGPT 3.5 appears to be proficient in coding more complex software if it has access to ample relevant data to be train on

CHATGPT 4

ABE GIZAW



How is ChatGPT 4
different from 3.5?



1 Expanded understanding of current technologies

2 Improved context retention and comprehension

3 Advanced code creation abilities

4 Ability to take larger inputs and able to read images

CHAT PROMPT

“CREATE AN AVL TREE CLASS IN JAVA, FULLY BUILT OUT WITH INSERT, DELETE, ROTATIONS, GET, AND AN INORDER-ITERATOR”


```

class AVLNode {
    15 usages
    int key, height;
    30 usages
    AVLNode left, right;

    1 usage
    public AVLNode(int item) {
        key = item;
        height = 1;
    }
}

class AVLTree3 {
    16 usages
    AVLNode root;

    14 usages
    int height(AVLNode node) {

```

```

public class AVLTree<T extends Comparable<T>> implements Iterable<T> {
    26 usages
    private class Node {
        14 usages
        T value;
        8 usages
        int height;
        31 usages
        Node left, right;

        1 usage
        Node(T value) {
            this.value = value;
            this.height = 1; // New node is initially added at leaf
        }
    }

    6 usages
    private Node root;

    // Get the height of the node
    14 usages
    private int height(Node N) {
        if (N == null) return 0;
        return N.height;
    }
}

```



Some main differences included:

- **3.5 created nodes that specifically required integers, while 4.0 uses Java Generics (T extends Comparable<T>)**
- **4.0 Added a contains method, which is useful for tree traversal, but wasn't specifically asked for.**
- **4.0 Provided comments and Documentation while 3.5 Did not**
- **4.0 encapsulated the AVL tree functionality within the class, exposing only the necessary public methods**
3.5 Made everything default/public.
- **4.0 provides a lazy iterator. It iterates through the tree using a stack to keep track of the nodes.**
3.5 used a recursive approach to in-order traversal, which does not support lazy iteration

OTHER GPT4 BENEFITS FOR SW DEV

- Enhanced Language Understanding
- Integrated Data Processing
- Cross-Domain Knowledge
- Document Handling Capabilities
- Advanced Programming

GPT STORE

Dis

Q Sea

Top Picks

Featured

Curated top picks from this week



Adobe Express

Stand out with Adobe Express. Quickly and easily make impactful social posts, images, videos, flyers...

By adobe.com



Code Copilot

Code Smarter, Build Faster—With the Expertise of a 10x Programmer by Your Side.

By promptspellsmith.com



Social Butterfly-AI 2.0

AI Assistant for content development and social media strategy. Nuanced, guided...

By digitalmogul.co



Software Architect GPT

Builds new software architecture documents by understanding user requirements and design...

By V B Wickramasinghe

Trending

Most popular GPTs by our community

1



image generator

A GPT specialized in generating and refining images with a mix of professional and friendly tone.image generator

By NAIF J ALOTAIBI

2



Write For Me

Write tailored, engaging content with a focus on quality, relevance and precise word count.

By puzzle.today



**GPT 4 IS
NOT
PERFECT**



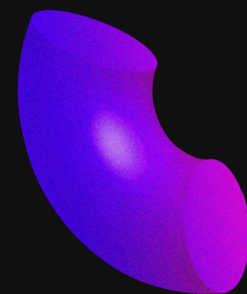
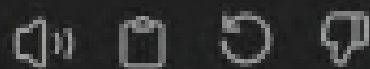
You

From where do you grab the information or data when I ask you to write any specific code?



ChatGPT

I generate responses based on a vast dataset of text from books, websites, and other sources that I was trained on up until my last update in April 2023. This training involved analyzing and understanding patterns, syntax, semantics, and the logic behind coding in various programming languages, as well as problem-solving strategies. I don't "grab" information or data from external sources in real-time. Instead, I use the knowledge embedded in me during training to generate code and explanations. If you have a coding question or a problem, I use this internal knowledge to try to provide a helpful response.



DISADVANTAGES OF GPT4

- Difficult to debug
- Lack of Contextual Understanding
- Incomplete Knowledge and Narrow Training Data

CHAT PROMPT

“I NEED A METHOD THAT CAN TAKE IN AN ASM METHOD NODE SIGNATURE OR DESC AND AN ASM METHOD NODE AND RETURN A METHOD DECLARATION IN UML FORMAT.”

Code:

```
public static String methodNodeToUML(MethodNode methodNode) {
    String methodName = methodNode.name;
    String descriptor = methodNode.desc;
    String signature = methodNode.signature;

    // Use signature if not null; otherwise, use descriptor
    String typeDescriptor = signature != null ? signature : descriptor;

    // Extract return type and parameter types
    Type returnType = Type.getReturnType(typeDescriptor);
    Type[] argumentTypes = Type.getArgumentTypes(typeDescriptor);

    // Convert types to UML-friendly format
    String umlReturnType = convertTypeToUML(returnType);
    String umlArguments = convertArgumentsToUML(argumentTypes);

    // Format and return UML declaration
    return umlReturnType + " " + methodName + "(" + umlArguments + ")";
}
```

```
void singleObject(java.lang.String)
void multipleObjects(java.lang.String, java.lang.Integer, java.lang.Double)
void mixPrimitiveAndObject(int, java.lang.String, Object)
Exception in thread "main" java.lang.IllegalArgumentException Create breakpoint Explain with AI ↗
    at org.objectweb.asm.Type.getTypeInternal(Type.java:443)
    at org.objectweb.asm.Type.getArgumentTypes(Type.java:331)
    at domain.diagramconverter.ClassMethodsUMLConverter.getMethodInfo(ClassMethodsUMLConverter.java:74)
    at domain.diagramconverter.ClassMethodsUMLConverter.convert(ClassMethodsUMLConverter.java:33)
    at domain.diagramconverter.ConvertASMToUML.generateDiagramByNode(ConvertASMToUML.java:40)
    at domain.diagramconverter.ConvertASMToUML.generateDiagramByPackage(ConvertASMToUML.java:61)
    at domain.diagramconverter.ConvertASMToUML.generateDiagram(ConvertASMToUML.java:40)
```

Result

- The code provided didn't function  correctly.

```
void singleObject(java.lang.String)
void multipleObjects(java.lang.String, java.lang.Integer, java.lang.Double)
void mixPrimitiveAndObject(int, java.lang.String, Object)
```

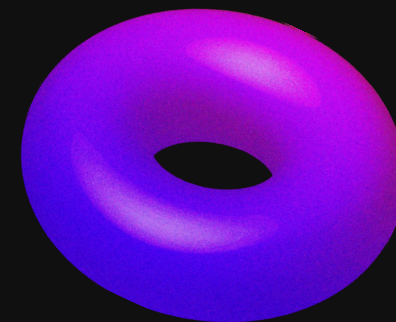
- Furthermore, the method constructors it generated, shown above the error message, clearly deviate from the standard UML format
- They lack parameter names, place the return type incorrectly, and omit access modifiers.

Conclusion:

Both ChatGPT 3.5 and 4.0 have a place in software development.

ChatGPT is good for straightforward coding tasks or automating repetitive parts of coding.

For complex tasks, though, they're best used as an auxiliary tool, not the main one you rely on.





Want to make a presentation like this one?

Start with a fully customizable template, create a beautiful deck in minutes, then easily share it with anyone.

[Create a presentation \(It's free\)](#)