

AI code generation report

Roblox prime numbers:

Oleh Basystyi

Anna Stasyshyn

Artur Rudish

Anton Valihurskyi

Maksym Zhuk

April 2024

1 Introduction

In this little survey our team have researched AI's capability to solve CodeWars tasks of different difficulty (primarily 7, 6, 5, 4 kyu) and different CodeWars categories:

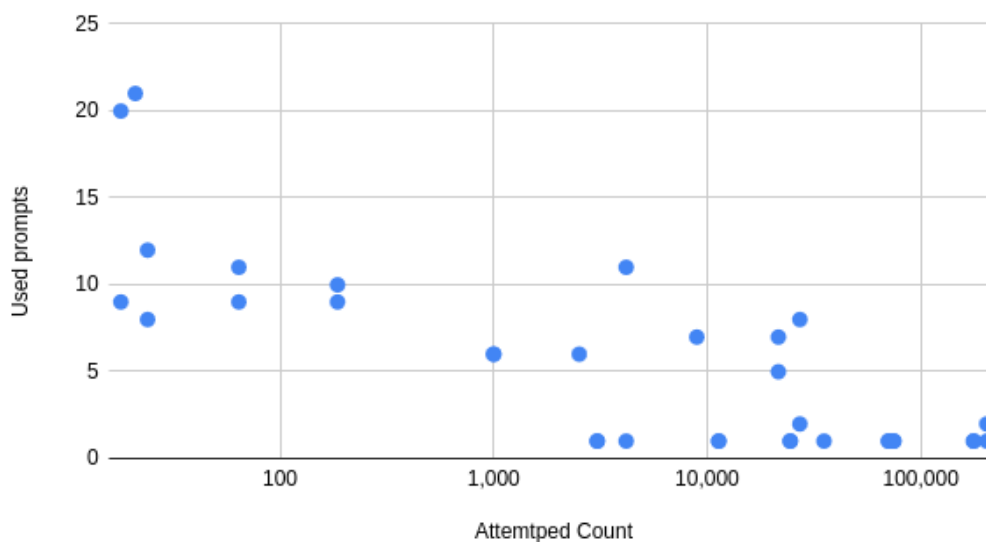
- Algorithms
- Strings
- Arrays
- Mathematics
- Linear Algebra
- Dynamic Programming
- Data Structures

As AIs we used *GPT-3.5*, *GPT-4*, *Claude* and *Gemini* and compared their performance on same tasks. Also we researched two ways of writing prompts to solve problems like this: templated and manual that must be used in different cases

2 Experimenting

To carefully carry out experiments we have set up [google sheets table](#) to log data about AI performance on solving particular tasks. In total we analyzed 18 problems and gained that kind of data:

Used prompts vs. Attempted Count



It shows the relation between number of people that attempted to solve a particular problem and number of prompts that we have used to make AI solve it. As can be seen from chart the general rule is that tasks attempted by smaller amount of people were more difficult to solve by AI. And actually we can see a breakpoint around region of 1000 attempts. Due to this we have used different techniques to get problem solution from AI