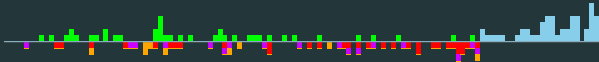


G: Garbage In, Garbage Out

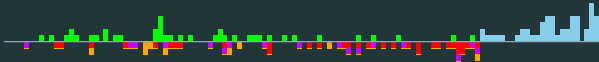
Problem author: Jeroen Op de Beek



Problem: Determine whether a text is human-made or LLM-generated.

G: Garbage In, Garbage Out

Problem author: Jeroen Op de Beek

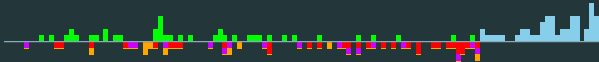


Problem: Determine whether a text is human-made or LLM-generated.

Human: The text is some concatenation of words from a list.

G: Garbage In, Garbage Out

Problem author: Jeroen Op de Beek



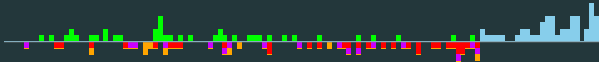
Problem: Determine whether a text is human-made or LLM-generated.

Human: The text is some concatenation of words from a list.

LLM: The text is randomly generated.

G: Garbage In, Garbage Out

Problem author: Jeroen Op de Beek



Problem: Determine whether a text is human-made or LLM-generated.

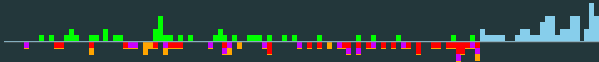
Human: The text is some concatenation of words from a list.

LLM: The text is randomly generated.

DP Solution: The text is human-made if the first 6, 7, 8, 9, or 10 letters occur in the word list, and the remainder of the text is also human-made, recursively.

G: Garbage In, Garbage Out

Problem author: Jeroen Op de Beek



Problem: Determine whether a text is human-made or LLM-generated.

Human: The text is some concatenation of words from a list.

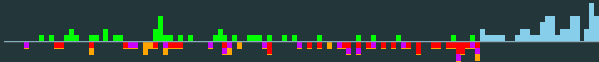
LLM: The text is randomly generated.

DP Solution: The text is human-made if the first 6, 7, 8, 9, or 10 letters occur in the word list, and the remainder of the text is also human-made, recursively.

Better: Only check if any word is a prefix: fails with probability at most $5000/26^6 < 0.002\%$ per case.

G: Garbage In, Garbage Out

Problem author: Jeroen Op de Beek

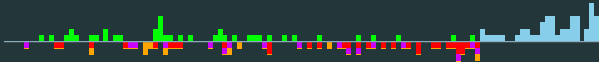


Problem: Determine whether a text is human-made or LLM-generated.

Observation: In a random text, nearly all length-6 substrings (6-mers) are different, since $26^6 \approx 300\,000\,000 \gg 300\,000$.

G: Garbage In, Garbage Out

Problem author: Jeroen Op de Beek



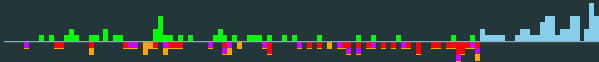
Problem: Determine whether a text is human-made or LLM-generated.

Observation: In a random text, nearly all length-6 substrings (6-mers) are different, since $26^6 \approx 300\,000\,000 \gg 300\,000$.

Solution: The text is LLM-generated if all 6-mers occur at most twice, and human otherwise.

G: Garbage In, Garbage Out

Problem author: Jeroen Op de Beek



Problem: Determine whether a text is human-made or LLM-generated.

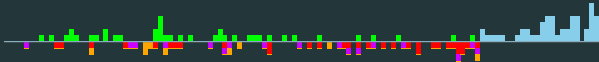
Observation: In a random text, nearly all length-6 substrings (6-mers) are different, since $26^6 \approx 300\,000\,000 \gg 300\,000$.

Solution: The text is LLM-generated if all 6-mers occur at most twice, and human otherwise.

Alternative: Count the number of distinct 6-mers instead.

G: Garbage In, Garbage Out

Problem author: Jeroen Op de Beek



Problem: Determine whether a text is human-made or LLM-generated.

Observation: In a random text, nearly all length-6 substrings (6-mers) are different, since $26^6 \approx 300\,000\,000 \gg 300\,000$.

Solution: The text is LLM-generated if all 6-mers occur at most twice, and human otherwise.

Alternative: Count the number of distinct 6-mers instead.

Statistics: 146 submissions, 39 accepted, 49 unknown