

Новый конкурс по программированию раскрыл слабые результаты ИИ

Недавно завершился первый этап нового «ИИ-челленджа» по программированию — K Prize, организованного Databricks и сооснователем Perplexity Энди Конвински. Победителем стал бразильский инженер Эдуардо Роча де Андраде, который получил \$ 50 000. Но удивило то, что он правильно решил лишь 7,5% заданий.

Конвински отметил, что целью конкурса было создать действительно сложный тест, чтобы проверить реальные возможности моделей. В отличие от других испытаний, K Prize использует новые данные с GitHub и не допускает подготовки моделей на конкретных задачах, что делает тест честным, но сложным.

Конвински даже обещал миллион долларов тому, кто создаст открытую модель с точностью выше 90% в этом teste. Он подчёркнул, что пока ИИ далеки от обещанных ролей врачей или юристов, а даже в программировании модели пока показывают низкий результат.