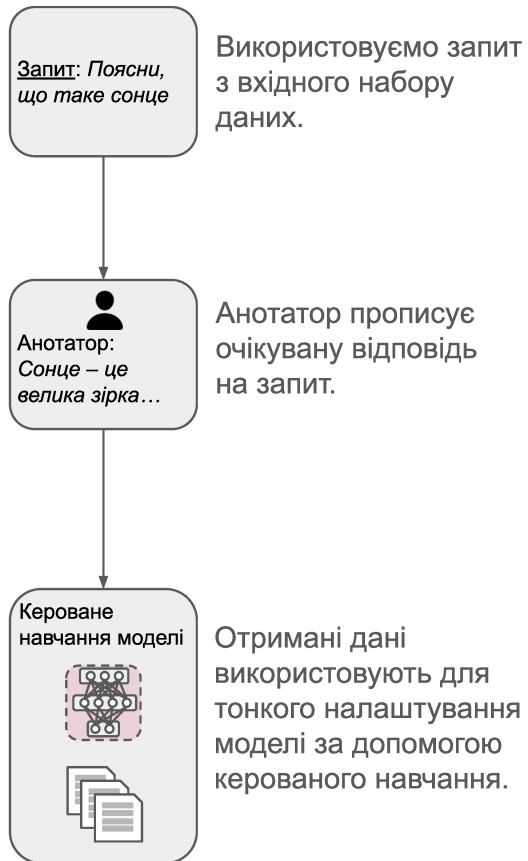


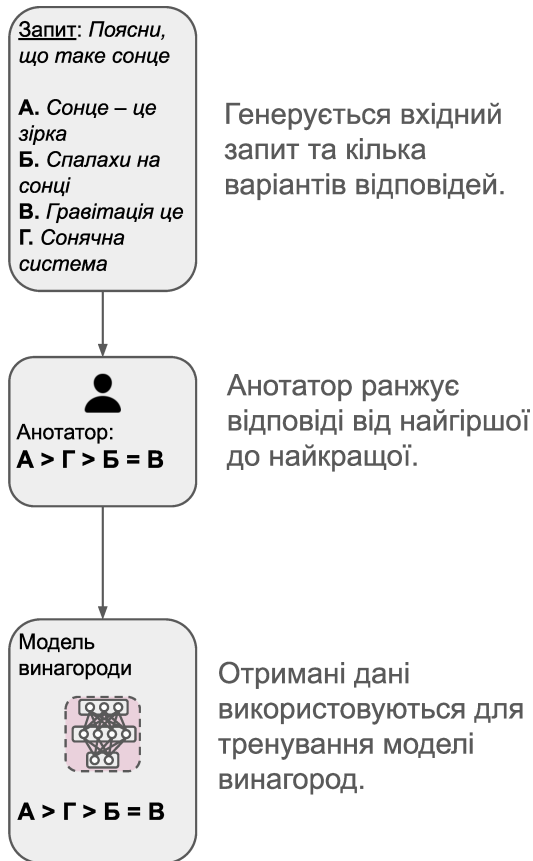
Крок 1

Збір даних для демонстрації та кероване навчання стратегії моделі



Крок 2

Збір даних для порівняння та навчання моделі винагород



Крок 3

Оптимізація стратегії моделі проти моделі винагород за допомогою навчання з підкріпленням

