โพสต์ทูเดย์
Post Today
Circulation: 320,000
Ad Rate: 900

Section: ธุรกิจตลาด/-
วันที่: พฤหัสบดี 24 พฤษภาคม 2561
ปีที่: 16 ฉบับที่: 5586 หน้า: B3(ล่างขวา)
Col.Inch: 31.67 Ad Value: 28,503 PRValue (x3): 85,509 คลิป: ขาว-ดำ
คอลัมน์: AI in trend: AI กับเกมนับเลข

AI กับเกมนับเลข

AI in trend

วันนี้ผมมาชวนลับสมองกับเกมนับเลขครับ เกมนี้น่าสนใจในแง่คิดแบบเอไอ (AI) เล่นกัน 2 คน เริ่มนับจาก 0 แต่ละคนบวกเพิ่ม 1 2 หรือ 3 สลับกันไปเรื่อยๆ ใครถึง 20 ก่อนชนะ สมมติว่าเราเริ่มก่อน ถ้าเราเลือกบวก 1 ค่าใหม่ก็จะเป็น 1 สมมติว่าเพื่อนของเราเลือกบวก 1 เหมือนกัน ค่าใหม่ก็จะเป็น 2 นั่นคือเมื่อการเล่นวนมาถึงเราใหม่อีกรอบ เราสามารถมองได้ว่าเราอยู่ใน “State” ที่ 2 ถ้าเพื่อนของเราเลือกบวก 2 เมื่อการเล่นวนมาถึงเราใหม่เราจะอยู่ใน State 3 และถ้าเพื่อนของเราเลือกบวก 3 เมื่อการเล่นวนมาถึงเราใหม่เราจะอยู่ใน State 4

จะเห็นว่า “Action” เดียวที่เราเลือกสามารถนำไปสู่ State ใหม่ที่ต่างกันออกไปได้ เทคนิคเอไอที่เหมาะกับงานแบบนี้คือ “การเรียนแบบเสริมกำลัง” หรือ “Reinforcement Learning (RL)” โจทย์ของ RL ต่างจากโจทย์เอไอทั่วไปที่เป็นเพียงการถาม-ตอบ สำหรับโจทย์ RL ระบบเอไอต้องตัดสินใจต่อเนื่องและหลายๆ ครั้งเราจะรู้ว่าการตัดสินใจดีหรือไม่ในตอนจบเกมเท่านั้น เทคนิค RL นี้เหมาะกับพวกเกมต่างๆ เช่น AlphaGO เป็นต้น

แนวความคิดหลักของ RL คือ เราจะพยายามประเมินค่า “ประโยชน์ (Utility)” ของแต่ละ State โดยวัดจาก “รางวัล (Reward)” ที่จะได้ในอนาคต หากเราเริ่มเดินจาก State เหล่านี้ไป จากนั้นเราเลือกแอ็กชั่นที่นำไปสู่ State ใหม่ที่มีประโยชน์สูงสุด สำหรับเกมนี้แผนการเดินที่ได้ เช่น ถ้าอยู่ State 19 ให้บวก 1 ถ้า 17 ให้บวก 3 ถ้า 14 ให้บวก 2... เป็นต้น

ถ้าเราวิเคราะห์โจทย์ของเกมนี้ดีๆ เราจะสามารถสร้างแผนการเดินนี้ได้เช่นกันโดยไม่ต้องใช้ RL จากวิธีเล่น เรารู้ว่าทั้งเราและเพื่อนนับได้แค่ 1 2 3 ดังนั้น หากมีใครอยู่ที่ State 16 เขาก็จะนับได้แค่ 17 18 19 ซึ่งไม่ถึง 20 และเมื่อเขานับ 17 18 19 เราก็จะสามารถนับต่อได้ถึง 20 และชนะไป นั่นแปลว่า State 16 นั้น เราไม่ควรอยู่ เพราะถ้าเราอยู่ไม่ว่าจะนับอย่างไร เพื่อนเราจะนับต่อได้ถึง 20 เสมอ มองอีกมุมหนึ่งคือเราควรจะนับให้ถึง 16 ก่อนอีกฝ่าย เพราะพอถึงตาเขา เขาจะได้อยู่ใน State 16

ด้วยวิธีคิดแบบเดียวกันเราจะได้ว่า เราควรจะนับให้ถึง State ที่หาร 4 ลงตัวก่อนอีกฝ่ายเสมอ แผนการเดินนี้ตรงกับผลที่ได้จาก RL นั่นคือ เอไอสามารถสกัด “องค์ความรู้ (Knowledge)” จากเกมเพื่อให้ได้คำตอบว่าต้องเล่นอย่างไรเพื่อให้ชนะ หรือคำตอบสำหรับคำถาม “How” แต่เอไอและแผนการเล่นนั้นไม่สามารถ “อธิบาย” ได้ว่า “ทำไม (Why)” ต่อให้เราเปลี่ยนไปใช้ Deep Reinforcement Learning ผลก็ยังเป็นเช่นเดียวกัน

จากเกมข้างบนจริงๆ แล้วยังมีข้อสรุปอีกอย่างที่สำคัญมาก นั่นคือ “ถ้าไม่อยากแพ้ต้องไม่เริ่มนับคนแรก” ข้อสรุปนี้นับว่าเป็น “ปัญญา (Wisdom)” ที่ได้จากการวิเคราะห์ปัญหานี้ ในปัจจุบันเรายังไม่มีกรอบการทำงานที่สกัดเอาข้อสรุปนี้ขึ้นมาได้เอง

จริงๆ แล้วถ้าเรารู้ล่วงหน้าว่า นี่คือสิ่งที่เราต้องการพบเราสามารถวิเคราะห์ Log จากการเล่นเพื่อให้ได้ข้อสรุปนี้เช่นกัน แต่ถ้าเรารู้ล่วงหน้า ข้อสรุปนี้ก็คงไม่น่าสนใจ และถ้าเราสามารถสกัดจนได้ข้อสรุปนี้มาแต่ขาดความเข้าใจว่าทำไม การจะนำมันไปใช้ก็คงยาก เพราะเราไม่รู้ว่ามันมาได้อย่างไร ■