

ปกิณกะ

คิดต่างเรื่อง deep learning

อนุมงคล ศิริเวทิน

ผมได้อ่านบทความ “เรื่องของ deep learning” โดย ชิตชนก เหลือสินทรัพย์ และสมชัย บวรกิตติ ลงพิมพ์ในพุทธชินราชเวชสารฉบับเดือนพฤษภาคม - สิงหาคม พ.ศ. ๒๕๖๑ เห็นว่าบทความนั้นควรจะมีความชัดเจนว่าเขียนมาเพื่อให้ใครอ่าน เท่าที่เป็นอยู่ไม่ชัดเจน เช่น ถ้าจะให้ให้นักคอมพิวเตอร์ทั่วไปอ่านก็จะมีข้อมูลพื้นฐานที่น่าจะรู้กันอยู่แล้ว ถ้าจะให้ให้นักวิชาการทั่วไปอ่าน ก็มีข้อมูลที่คนส่วนมากไม่เข้าใจ เช่น โครงข่ายนี้ควรมีกี่ชั้น แต่ละชั้นควรมีจำนวนเซลล์ประสาทกี่เซลล์

ในการสร้างความฉลาดให้เครื่องคอมพิวเตอร์เพื่อให้สามารถแก้ปัญหาที่ซับซ้อน ทำไมจึงพูดถึงยุคแรก ไม่จำเป็นทำให้คนอ่านสับสนเปล่าๆ ที่ว่าการพัฒนาปัญญาประดิษฐ์ให้เครื่องคอมพิวเตอร์ได้ ทั้งนี้เนื่องจากการคำนวณเชิงตรรกะใช้ค่าเพียง ๒ ค่าคือค่าจริงและค่าเท็จ (ค่าจริงมักแทนด้วย ๑ และค่าเท็จแทนด้วย ๐) นั่นไม่จริงเพราะสามารถใช้สูตรที่ปรับค่าให้อยู่ระหว่างศูนย์กับหนึ่งได้ หรือจะใช้ค่าที่มีหลายมิติก็ได้

และที่ว่าส่วนปัญหามันจำเป็นต้องเกี่ยวข้องกับค่าอื่นนอกจากค่า ๑ และ ๐ เช่น สีใบหน้า ขนาดโครงหน้า ตำแหน่งปาก ดังนั้นการวิจัยเพื่อสร้างปัญญาประดิษฐ์จึงเน้นการวิจัยให้เข้าใจถึงการทำงานของเซลล์ประสาทในสมอง คิดว่าไม่ใช่เหตุผล น่าจะเป็น “จึงเลือกใช้” หรือ “จึงพยายามเลียนแบบ” มากกว่า “ให้เข้าใจ” และจำลองการเรียนรู้ของเซลล์ประสาทโดยใช้บริบททางคณิตศาสตร์ซึ่งประกอบด้วยตัวแปรที่แทนค่าความต้านทานที่จุดเชื่อมต่อ (synapse) และตัวแปรที่แทนค่าของข้อมูลที่สอน เซลล์ประสาทเรียนรู้ข้อมูลโดยการปรับค่าความต้านทานที่จุดเชื่อมต่อจากแบบจำลองทางคณิตศาสตร์ของเซลล์ประสาท ซึ่งตีความได้ว่าเซลล์ประสาททำหน้าที่แบ่งข้อมูลที่ให้สอนเซลล์ประสาทเป็นสองกลุ่ม ข้อมูลนี้ก็คือคุณสมบัติที่สกัดมาจากสิ่งที่ต้องการให้เซลล์ประสาทรู้จำ เช่น ถ้าต้องการให้เซลล์ประสาทรู้จำหน้าคน รูปสมบัติของใบหน้าที่สกัดออกมาเพื่อเป็นข้อมูลที่ให้สอนเซลล์ประสาทอาจเป็น ความกว้าง ความยาวของหน้า สีของหน้า ขนาดตา ขนาดปาก เซลล์ประสาทเพียง

เซลล์เดียวไม่สามารถเรียนรู้กลุ่มข้อมูลที่หลากหลายและซับซ้อนได้ เช่น ข้อมูลของหน้าคนหลายคน เพื่อแก้ปัญหานี้จึงจำเป็นต้องใช้หลายเซลล์ประสาทมาเชื่อมต่อเป็นโครงข่ายหลายชั้น สิ่งที่น่าสนใจก็คือโครงข่ายนี้ควรมีกี่ชั้น แต่ละชั้นควรมีจำนวนเซลล์ประสาทกี่เซลล์ (จะต้องมีความชัดเจนว่าคำว่า “ควรมี” ในที่นี้ ผู้ใช้หรือ deep learning เป็นผู้กำหนด) เพื่อให้สามารถเรียนรู้และให้ผลการรู้จำที่ถูกต้องได้สูงสุด วิธีการหาค่าความต้านทานที่จุดต่อเชื่อมของทุกเซลล์ประสาทในโครงข่ายที่เหมาะสมสำหรับการรู้จำของแต่ละปัญหามีหลายวิธี (ประเด็นนี้ผู้อ่านที่ไม่ใช่ นักคอมพิวเตอร์ที่รู้เรื่อง neural net อยู่แล้วคงจะไม่เข้าใจ) บางวิธีก็ให้ผลน่าพอใจ แต่วิธีที่ไม่ดีตามที่ต้องการ

ปัญหาสำคัญของการประยุกต์โครงข่ายเซลล์ประสาทก็คือ มีสมบัติอะไรบ้างที่ต้องสกัดออกมาจากสิ่งที่ต้องการให้โครงข่ายประสาทรู้จำ และโครงข่ายเซลล์ประสาทควรมีโครงสร้างอย่างไร ประมาณ พ.ศ. ๒๕๒๓ ศาสตราจารย์ Kunihiko Fukushima ได้เสนอโครงข่ายเซลล์ประสาท Neocognitron เพื่อใช้รู้จำตัวพิมพ์เลขอาราบิกและตัวพิมพ์อักษรภาษาอังกฤษ โดยเอาจำนวนจุดภาพจากขนาดต่างๆ กันที่เป็นได้ทั้งหมดจากตำแหน่งเดียวและตำแหน่งต่างๆ ที่ใกล้เคียงกันในภาพผสมกันมา เป้าหมายของการใช้จำนวนจุดภาพก็คือต้องการให้โครงข่ายเซลล์ประสาทตรวจสอบว่ามีรูปแบบการจัดตัวของกลุ่มจุดภาพตามที่กำหนดหรือไม่ เช่น ตัวอักษร A จะต้องประกอบด้วยชิ้นส่วนที่เกิดจากจุดภาพดังนี้

กลุ่มจุดภาพเหล่านี้ใช้แทนคุณสมบัติของภาพในการสอนโครงข่ายเซลล์ประสาท วิธีนี้ทำเพื่อแก้ปัญหาว่าควรจะสกัดคุณสมบัติอะไรบ้างจากภาพตัวเลข นอกจากนั้น นิโอคอกนิตรอนยังแก้ปัญหาของโครงสร้างของโครงข่ายเซลล์ประสาทว่าควรเป็นอย่างไร โดยใช้วิธีง่ายๆ คือ ใส่จำนวนชั้นและจำนวนเซลล์ประสาทในแต่ละชั้นให้มากเท่าที่เครื่องคอมพิวเตอร์จะคำนวณได้ แนวคิดของศาสตราจารย์ฟูคิมิตสึ ได้ถูกนำไปขยายต่อโดยใช้เซลล์ประสาทที่ทำงานตามแบบจำลอง

คณิตศาสตร์ของ ศาสตราจารย์ McCulloch และ Pitts รวมทั้งใช้วิธีการปรับค่าความต้านทานที่จุดต่อเชื่อมแบบต่างๆ เช่นใช้วิธีปรับค่าแบบ Perceptron แบบ Backpropagation ซึ่งต่อมาก็ตั้งชื่อวิธีการนี้ว่า deep learning ซึ่งหมายถึงการเรียนรู้ของโครงข่ายเซลล์ประสาทที่ใช้จำนวนชั้นมากกว่า ๓ ชั้น (ชั้นขาเข้า ชั้นตรงกลาง และชั้นขาออก) และใช้จำนวนเซลล์ประสาทจำนวนมากในแต่ละชั้น

ข้อดีของ deep learning ก็คือ

๑. ผู้ใช้ไม่จำเป็นต้องสกัดสมบัติของสิ่งที่ต้องให้โครงข่ายเซลล์ประสาทรู้จำ เพราะ deep learning ใช้ทุกสมบัติที่เป็นไปได้ของสิ่งที่ต้องการเรียนรู้

๒. ผู้ใช้ไม่จำเป็นต้องรู้เรื่องการคำนวณภายในของโครงข่ายเซลล์ประสาทเพื่อกำหนดจำนวนชั้น จำนวนเซลล์ในแต่ละชั้น และการเลือกวิธีการปรับค่าความต้านทานของจุดเชื่อมต่อของทุกเซลล์ประสาทในโครงข่าย เพราะ deep learning ใช้วิธีลองโดยใช้จำนวนชั้นหลายชั้นและจำนวนเซลล์ประสาทในแต่ละชั้นมาก วิธีของ deep learning ให้ความสะดวกและเหมาะกับผู้ใช้ที่ไม่มีความรู้ในเรื่องการคำนวณเบื้องหลังของโครงข่ายเซลล์ประสาท ที่ว่าข้อเสียของ deep learning ก็คือ ระยะเวลาของการเรียนรู้นานมากเนื่องจากการใช้จำนวนชั้นและจำนวนเซลล์มากเกินความจำเป็น ควรขยายความคำว่า “จำเป็น” ในที่นี้ ปัญหานี้อาจแก้ได้โดยใช้

เครื่องคอมพิวเตอร์ความเร็วสูงที่มีหลายหน่วยประมวลผลกลางและหน่วยความจำขนาดใหญ่มาก แต่วิธีนี้นำไปสู่อีกปัญหา คือการใช้พลังงานไฟฟ้าจำนวนมากมหาศาลในการสอนโครงข่ายเซลล์ประสาท ไม่เหมาะที่จะนำโครงข่ายเซลล์ประสาทไปติดตั้งในหุ่นยนต์ที่ส่งทำภารกิจตามลำพังและไม่สามารถติดต่อกับศูนย์ประมวลผลขนาดใหญ่ และไม่เหมาะที่จะติดตั้งในโทรศัพท์มือถือที่มีขนาดของหน่วยความจำที่จำกัด มีจำนวนหน่วยประมวลผลกลางที่จำกัด และที่สำคัญที่สุดคือมีจำนวนพลังงานของแบตเตอรี่ที่จำกัด นอกจากประเด็นจุดอ่อนของ deep learning แล้ว วิธีการเรียนรู้แบบนี้ไม่สามารถใช้ได้กับการเรียนรู้ที่ข้อมูลไหลเข้ามาในกระบวนการเรียนรู้ตลอดเวลา ปริมาณข้อมูลจะเพิ่มขึ้นตามเวลาที่ผ่านไป ข้อมูลประเภทนี้เรียกว่าข้อมูลไหลหลัง (streaming data) เมื่อมีปริมาณข้อมูลเพิ่มขึ้นก็ต้องเก็บข้อมูลที่เข้ามาใหม่ ในที่สุดข้อมูลก็จะล้นหน่วยความจำของเครื่องคอมพิวเตอร์ นอกจากนั้นแล้ว เมื่อปริมาณข้อมูลเพิ่มขึ้นระยะเวลาการเรียนรู้ของโครงข่ายเซลล์ประสาท ก็ต้องเพิ่มขึ้นอย่างแน่นอนเพราะการปรับค่าความต้านทานที่จุดเชื่อมต่อของทุกเซลล์ในโครงข่ายต้องคำนวณจากข้อมูลทั้งข้อมูลที่เรียนไปแล้วกับข้อมูลที่เพิ่งเข้ามาใหม่ ตัวอย่างของข้อมูลประเภทไหลหลัง ได้แก่ข้อมูลที่เกิดในเครือข่ายสังคม และข้อมูลเกี่ยวกับเอกสารของลูกค้าธนาคาร

หมายเหตุ: ศัพท์บัญญัติของ deep learning น่าจะเป็นการเรียนรู้ของสมองกล ไม่ใช่การเรียนรู้ลึก