

สำหรับเผยแพร่โดยทันที

มหาวิทยาลัยสงขลานครินทร์ ประเทศไทยจับมือกับฮิตาชิ
พัฒนาเครื่องมือวิเคราะห์ความรู้สึกสำหรับภาษาไทย
เครื่องมือสามารถจำแนกคำพูดภาษาไทยเฉพาะแบบที่ใช้บนสื่อสังคมออนไลน์
โดยใช้เทคนิควิเคราะห์แบบอิงคลังศัพท์ซึ่งมีมากถึง 100 ล้านคำ

โตเกียว – 20 กุมภาพันธ์ พ.ศ. 2562 – มหาวิทยาลัยสงขลานครินทร์วิทยาเขตภูเก็ตร่วมมือกับบริษัท ฮิตาชิ จำกัด (TSE: 6501, “Hitachi”) ได้เริ่มโครงการวิจัยร่วมกันเพื่อการประมวลผลภาษาไทยธรรมชาติ¹

ในระยะแรกของการวิจัย

ทั้งสองฝ่ายได้ร่วมกันพัฒนาสร้างต้นแบบของเครื่องมือวิเคราะห์ความรู้สึกโดยใช้โปรแกรมปัญญาประดิษฐ์ AI (artificial intelligence)

ซึ่งสามารถจำแนกข้อความที่เขียนเป็นภาษาไทยเพื่อค้นหาความหมายที่สื่อถึงความรู้สึกเชิงบวก เชิงลบ หรือเป็นกลาง

เครื่องมือสามารถวิเคราะห์คำที่แสดงอารมณ์และความรู้สึกแบบอิงพจนานุกรมที่มีคำภาษาไทยมากถึง 100 ล้านคำโดยรวบรวมคำจากผู้ใช้สื่อสังคมออนไลน์

เป็นที่ทราบกันดีว่าการประมวลผลภาษาไทยยากกว่าภาษาอื่นๆ เพราะภาษาไทยมีภาษาพูดเฉพาะแบบ แต่เครื่องมือวิเคราะห์ความรู้สึกนี้ให้ความแม่นยำสูงเพราะสามารถรองรับภาษาพูดที่ใช้กันบนโซเชียลมีเดียเป็นจำนวนมาก

มหาวิทยาลัยสงขลานครินทร์และฮิตาชิจะนำเครื่องมือมาใช้จริงเพื่อวิเคราะห์คำที่คนไทยโพสต์แบบเรียลไทม์บนสื่อสังคมออนไลน์ไม่ว่าจะเป็น Facebook, Twitter และโซเชียลมีเดียอื่นๆ เพื่อจะประเมินประสิทธิภาพของโปรแกรมต้นแบบร่วมกัน โดยจะนำมาใช้งานจริงในเดือนเมษายนปี พ.ศ. 2562 เป็นส่วนหนึ่งของบริการวิเคราะห์ความรู้สึกของฮิตาชิ²

ในประเทศไทยการใช้เครื่องมือสื่อสารแบบเคลื่อนที่เติบโตอย่างก้าวกระโดดและมีอัตราการใช้งานของสื่อสังคมออนไลน์ทั่วประเทศ ประเทศไทยมีประชากรประมาณ 69 ล้านคน มีผู้ใช้ Facebook ประมาณ 50 ล้านคน

และผู้ใช้นี้ Twitter ประมาณ 12 ล้านคน^{*3}

ดังนั้นจึงมีความต้องการในการพัฒนาผลิตภัณฑ์และบริการให้ตรงตามต้องการของลูกค้าโดยวิเคราะห์จาก รูปแบบข้อความที่ผู้ใช้นี้สื่อสารบนสื่อสังคมออนไลน์

ซึ่งคนไทยมีการแสดงออกอารมณ์และความรู้สึกผ่านคำพูดมากมาย บ่อยครั้งที่ผู้ใช้นี้สื่อสารบนสื่อสังคมออนไลน์ใช้ตัวสะกดไม่ตรงตามหลักภาษาไทย มีการใช้เครื่องหมายแสดงเสียงที่ยาวขึ้น^{*4} สัญลักษณ์ที่ใช้แทนตัวเลขหรือเสียง การสร้างคำใหม่โดยเลียนเสียงจากธรรมชาติ และสัญลักษณ์ต่างๆ ที่แสดงสีหน้าและสื่ออารมณ์ ซึ่งยากที่จะประมวลผล ด้วยความท้าทายนี้เอง จึงจำเป็นต้องมีการจัดเก็บ รวบรวมข้อมูลมหาศาลก่อนการประมวลผลเพื่อแก้ปัญหาคำที่เขียนและสะกดคำที่ไม่เป็นไปตามหลักภาษาไทย

มหาวิทยาลัยสงขลานครินทร์ได้รวบรวมคำศัพท์ประมาณ 100 ล้านคำจากผู้ใช้นี้สื่อสารบนสื่อสังคมออนไลน์หลายรูปแบบเพื่อสร้างคลังศัพท์ภาษาไทยขนาดใหญ่ที่ใช้บอกความรู้สึกและ ให้ความแม่นยำสูงจึงถือได้ว่าเป็นสถาบันวิจัยชั้นนำแห่งหนึ่งที่ได้ประมวลผลคำภาษาไทย

ในเดือนตุลาคม ปี 2561 บริษัท ฮิตาชิ จำกัดเริ่มให้บริการวิเคราะห์ความรู้สึกโดยนำเทคโนโลยีมาใช้จำแนกและประเมินเสียงของลูกค้าโดยใช้แหล่งข้อมูลจากสื่อภาษาญี่ปุ่น บันทึกบทสนทนา และจากแหล่งอื่นๆ มาจัดทำเป็น 1,300 หัวข้อ ความรู้สึกและความคิดเห็น

เครื่องมือนี้ถือได้ว่าเป็นผลิตภัณฑ์ที่ร่วมกันสร้างครั้งแรกระหว่างมหาวิทยาลัยสงขลานครินทร์และบริษัท ฮิตาชิ จำกัด

โดยการผสมผสานผลวิจัยของมหาวิทยาลัยและความชำนาญด้านการออกแบบสถาปัตยกรรมระบบตลอดจน การใช้เทคโนโลยีประมวลผลผ่านระบบขนาดใหญ่ของฮิตาชิ ด้วยฟังก์ชัน hybrid noise-removal (ซึ่งผนวกเอาเทคนิคการเรียนรู้ของเครื่องและเทคโนโลยีอื่นๆ เข้าด้วยกัน)

และชุดคำศัพท์ที่รวบรวมคำไว้เป็นจำนวนมากทำให้เครื่องมือสามารถวิเคราะห์ข้อความภาษาไทยบนสื่อสังคม ออนไลน์ได้อย่างแม่นยำ ทั้งยังสามารถวิเคราะห์การแสดงความรู้สึกเฉพาะแบบได้มากมาย รวมทั้งการสะกดคำอีกด้วย

มหาวิทยาลัยสงขลานครินทร์และบริษัท

ฮิตาชิ

จำกัด

จะทำงานร่วมกันอย่างต่อเนื่องเพื่อให้เครื่องมือทำงานสมบูรณ์เมื่อมีการใช้ข้อมูลแบบเรียลไทม์

และจะพัฒนาการคัดกรองเพื่อวิเคราะห์ข้อมูลได้อย่างถูกต้อง (อาทิ การนำระบบการประเมินความรู้สึกทั้ง 7 ขั้นตอนมาใช้⁵) โดยตั้งเป้าที่จะเริ่มให้บริการจริงผ่านทางบริษัท ฮิตาชิ จำกัดในเดือนเมษายน ปี 2562

ทั้งสองฝ่ายจะยังเดินทางทำงานร่วมกันเพื่อปรับปรุงเครื่องมือการวิเคราะห์นี้

อาทิ

การเพิ่มคำศัพท์ตามข้อความและบริบทต่างๆ

1. การประมวลผลภาษาด้วยคอมพิวเตอร์ซึ่งโดยทั่วไปใช้เพื่อการสื่อสารในชีวิตประจำวัน (เช่นภาษาธรรมชาติ)
2. ข่าวประชาสัมพันธ์ (1 ตุลาคม 2018):
ฮิตาชิเปิดตัวบริการวิเคราะห์ความเชื่อมั่นเพื่อจำแนกและวิเคราะห์เสียงของลูกค้าในหัวข้อความรู้สึกและความตั้งใจประมาณ 1,300 ประเภทโดยใช้ AI (<http://www.hitachi.com/New/cnews/month/2018/10/181001.html>)
3. ข้อมูล ณ เดือนมกราคม 2562 ที่มา: StatCounter“ สถิติโซเชียลมีเดียประเทศไทย”
4. ประเภทของเครื่องหมายการออกเสียงกำกับอยู่เหนือตัวอักษรซึ่งโดยปกติจะใช้เพื่อทำเครื่องหมายสระเสียงยาว
5. การประเมินเจตนาขั้นตอนประกอบด้วย 3 ขั้นตอนบวก 3 ขั้นตอนลบและขั้นตอนที่เป็นกลางหนึ่งรวมเจตนาขั้นตอน

เกี่ยวกับการบริการวิเคราะห์ความรู้สึกของฮิตาชิ

เครื่องมือวิเคราะห์ความรู้สึกที่พัฒนาขึ้นในโอกาสนี้จะเป็นส่วนหนึ่งในการบริการวิเคราะห์ความรู้สึกของฮิตาชิ ภาคราชการไทยโดยจะเริ่มให้บริการในเดือนเมษายนปี 2562 ฮิตาชิมีศูนย์ Lumada ประจำภูมิภาคเอเชียตะวันออกเฉียงใต้ (ตั้งอยู่ในนิคมอุตสาหกรรมอมตะซิตี้ จังหวัดชลบุรี ประเทศไทย) ซึ่งเป็นศูนย์กลางประจำภูมิภาคที่ให้บริการนี้ บริษัท ฮิตาชิ เอเชีย (ประเทศไทย) จำกัด มีหน้าที่หลักในการขับเคลื่อนและเสนอบริการกับลูกค้าหลายกลุ่ม รวมทั้งบริษัทญี่ปุ่นที่มีสาขาในภูมิภาค และบริษัทคนไทยในอุตสาหกรรมยานยนต์ ภาคการผลิต โรงแรม โรงพยาบาล ธนาคาร และสถาบันภาครัฐ นอกจากนี้ฮิตาชิยังจะเสนอบริการเครื่องมือวิเคราะห์ความรู้สึกที่สนับสนุนภาษาอังกฤษและภาษาจีน และวางแผนที่จะสนับสนุนในอีกหลายภาษาภายใต้แผนการขยายธุรกิจของบริษัท

หากต้องการข้อมูลเพิ่มเติม กรุณาไปยัง

<https://www.hitachi.com/products/it/appsvdiv/service/sentiment-analysis/index.html>

หมายเหตุเรื่องเครื่องหมายการค้า

- Facebook เป็นเครื่องหมายการค้าหรือเครื่องหมายการค้าจดทะเบียนของ Facebook Inc. ในสหรัฐอเมริกาและประเทศอื่นๆ
- Twitter เป็นเครื่องหมายการค้าหรือเครื่องหมายการค้าจดทะเบียนของ Twitter Inc. ในสหรัฐอเมริกาและประเทศอื่นๆ
- ชื่อบริษัทและชื่อผลิตภัณฑ์อื่นเป็นเครื่องหมายการค้าจดทะเบียนหรือเครื่องหมายการค้าของเจ้าของนี้

หากต้องการข้อมูลเพิ่มเติมเกี่ยวกับเทคโนโลยี กรุณาไปยัง

<https://www.hitachi.com/products/it/appsvdiv/inquiry.html>

เกี่ยวกับบริษัท ฮิตาชิ จำกัด

บริษัท ฮิตาชิ จำกัด (TSE: 6501) มีสำนักงานใหญ่อยู่ในกรุงโตเกียว ประเทศญี่ปุ่น ให้บริการด้านนวัตกรรมเพื่อตอบโจทย์ความท้าทายทางสังคมโดยผนวกเทคโนโลยีด้านการผลิต

เทคโนโลยีสารสนเทศและผลิตภัณฑ์เข้าด้วยกัน บริษัทมีรายได้ของกลุ่มกิจการของปีงบประมาณ 2017 (สิ้นสุดวันที่ 31 มีนาคม ปี 2018) รวมทั้งสิ้น 9,368.6 พันล้านเยน (ประมาณ 88.4 พันล้านดอลลาร์สหรัฐ)

กลุ่มบริษัทฮิตาชิประกาศตัวเป็นพันธมิตรด้านนวัตกรรมกับ The Internet of Things และมีพนักงานประมาณ 307,000 คนทั่วโลก

ฮิตาชิมุ่งเน้นธุรกิจนวัตกรรมทางสังคมโดยใช้เทคโนโลยีดิจิทัลในหลายด้าน ได้แก่ ทั้งด้านไฟฟ้า/พลังงาน อุตสาหกรรม/การกระจายน้ำ การพัฒนาเมือง การเงิน/โครงสร้างทางสังคม/การดูแลสุขภาพ

หากต้องการข้อมูลเพิ่มเติมเกี่ยวกับฮิตาชิ กรุณาไปยัง <http://www.hitachi.com>