

Voice of Customer Analytics

จุดประสงค์ของ Assignment นี้คือ การทำ text analytics เพื่อวิเคราะห์ sentiment จาก social media โดยมีขั้นตอนดังต่อไปนี้

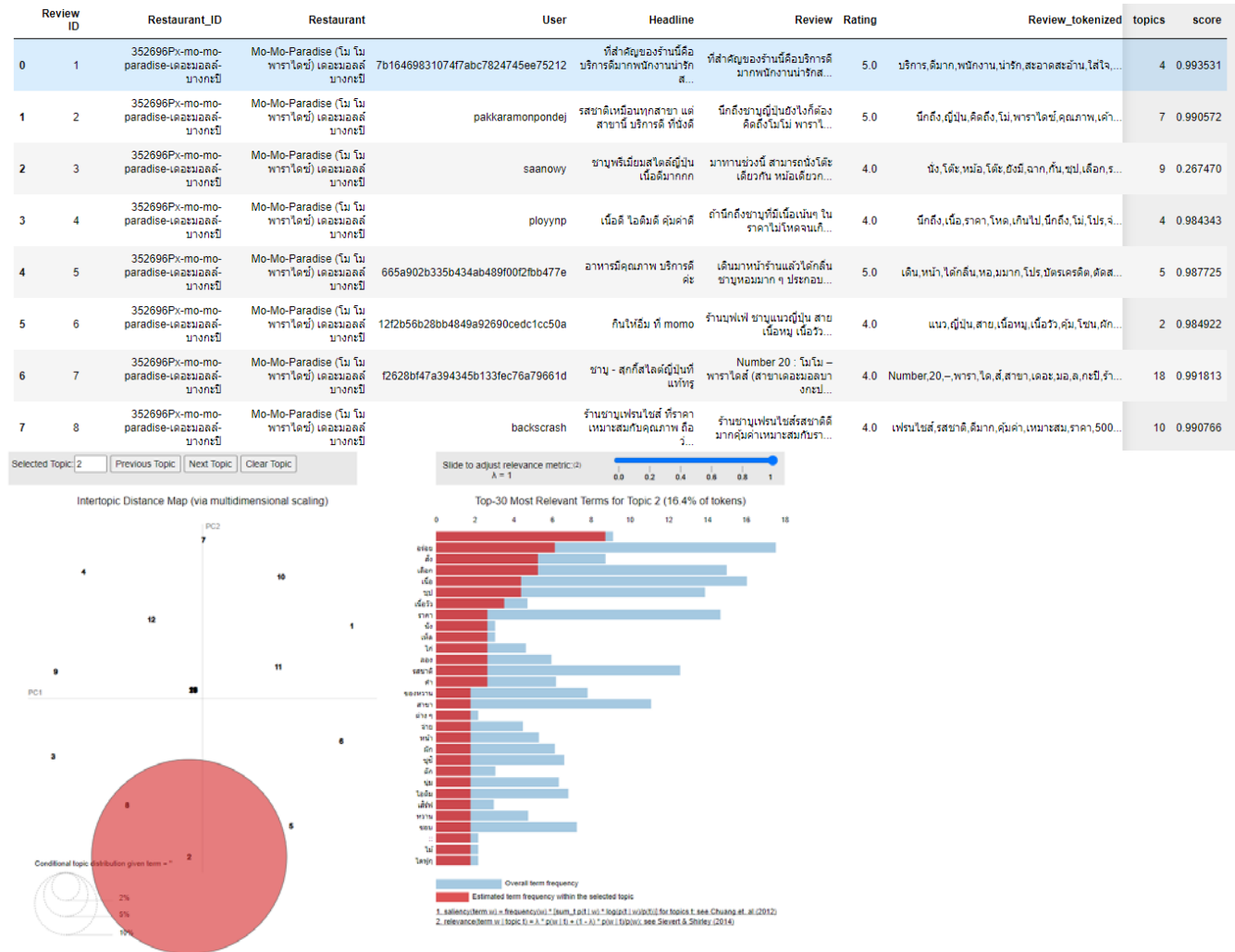
1. นำไฟล์ CustomerReviews.csv มา review แล้วแบ่งประเภทของ keywords ที่ได้ออกเป็น 3 ประเภท คือ
 - i. คำชม positive comment
 - ii. คำติบ่น negative comment
 - iii. คำกลางๆ ไม่ชม ไม่ติ neutral comment
2. นำไฟล์ ดังกล่าวโหลดขึ้น Github พร้อมทั้งนำต้นฉบับไฟล์ Python มาแก้ไข ก่อนที่จะ run บน Jupyter notebook

ความยากของ Assignment นี้คือ

1. คำภาษาไทยหลายคำเป็นคำเว้นวรรคไม่สามารถระบุ sentiment ได้ว่าดี หรือ แย่
2. คำภาษาไทยหลายคำเป็นคำประสม หรือ คำแสลง ทำให้ไม่สามารถระบุ sentiment ได้ว่าดีหรือแย่
3. คำใน social media บ่อยครั้งที่ไม่ได้เกี่ยวข้องกับตัวผลิตภัณฑ์ที่ต้องการนำมาวิเคราะห์

LDA Text Classification(1/3)

Mo-Mo-Paradise



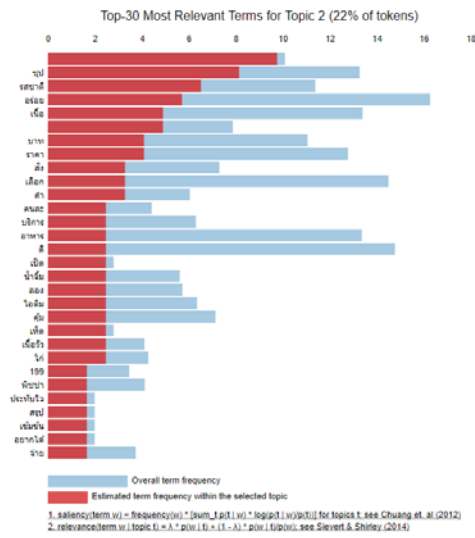
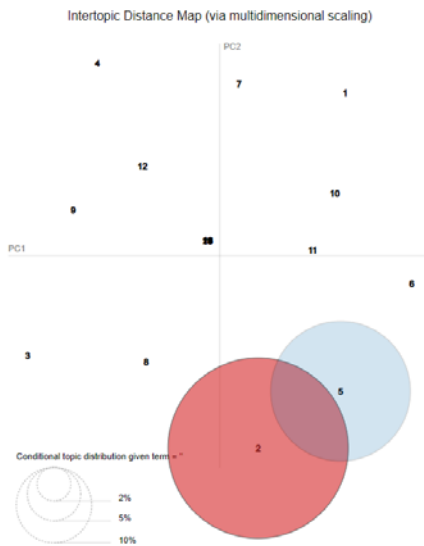
LDA Text Classification(2/3)

Shabushi

	Review ID	Restaurant_ID	Restaurant	User	Headline	Review	Rating	Review_tokenized
9	10	4621tm-shabushi-เดอะมอลล์ บางกะปิ-ชิน-ง	Shabushi (ชาบูชิ) เดอะมอลล์ บางกะปิ ชิน G	fd2ba94ec1bc4948ae109e2cb6c9cbc6	ทานอีก ขอบกึ่งทอดแถมประ	ใครชอบกึ่งทอดแถมประ แด่กึ่ง	3.0	ขอบ,กึ่ง,ทอด,เท,ม,ประ,กึ่ง,เท,ม,ประ,คัม,ดีก...
10	11	4621tm-shabushi-เดอะมอลล์ บางกะปิ-ชิน-ง	Shabushi (ชาบูชิ) เดอะมอลล์ บางกะปิ ชิน G	ploynnp	บุฟเฟ่ต์คุ้มในห้าง	กลับมาฮัฟเตอราคาชาบูชิ ดอนนีอยู่	3.0	กลับมา,ฮัฟเตอ,ราคา,ชาบู,ชิ,ดอนนี,อยู่,นี้,399,บาท,net,1.15,...
11	12	4621tm-shabushi-เดอะมอลล์ บางกะปิ-ชิน-ง	Shabushi (ชาบูชิ) เดอะมอลล์ บางกะปิ ชิน G	c08b2776067a4026b882abfc6fca8330	เปลี่ยนจากกึ่งทอดอาหารเป็นกึ่ง	ห่างหายไปนานสำหรับชาบูชิ ตั้งแต่	3.0	ห่างหาย,สำหรับ,ชาบู,ชิ,ตั้งแต่,ทาง,ร้าน,ได้,รับ,...
12	13	4621tm-shabushi-เดอะมอลล์ บางกะปิ-ชิน-ง	Shabushi (ชาบูชิ) เดอะมอลล์ บางกะปิ ชิน G	a96ec47b6cff4bd1834c90d603ef0a7e	อ้ม อร่อย ฟูตฟูต	เมื่อหลายวันก่อนนัดหาข้าวทานกับ	4.0	วันก่อน,นัด,หา,ข้าว,เดอะ,แกง,ดี,ลง,เออ,อาหาร,บุฟ...
13	14	4621tm-shabushi-เดอะมอลล์ บางกะปิ-ชิน-ง	Shabushi (ชาบูชิ) เดอะมอลล์ บางกะปิ ชิน G	a6991f1e0e9b4916ab3ca91d86f8e481	จัดไว้เต็มคราบกับข้าวโมม	บอกตรงๆว่าหลายครั้งที่เลือกจะกิน	3.0	หลายครั้ง,เลือก,บุฟ,เฟด,อาราม,ณ,รอ,คิว,สุดทำ...
14	15	4621tm-shabushi-เดอะมอลล์ บางกะปิ-ชิน-ง	Shabushi (ชาบูชิ) เดอะมอลล์ บางกะปิ ชิน G	7eb0c7d8683e407f83cc4d18e51d800d	ชาบู ชาบูชิ	สวัสดีครับวันนี้จะขอมารีวิวร้านชาบูชิ	4.0	สวัสดี,รีวิว,ชิ,บุฟ,ขวัญใจ,คน,เมนู,เลือก,ระบบ,...
15	16	4621tm-shabushi-เดอะมอลล์ บางกะปิ-ชิน-ง	Shabushi (ชาบูชิ) เดอะมอลล์ บางกะปิ ชิน G	qvanqvan	บุฟเฟ่ต์น่าเลียด สายพานร้อน	ชาบูชิ สาขาเดอะมอลล์บางกะปิ ตั้ง	4.0	ชิ,สาขา,เดอะ,มอลล์,กะปิ,ตั้ง,อยู่,ประดู,หน้า,ฟ...

Selected Topic:

Slide to adjust relevance metric: $\lambda = 1$



LDA Text Classification(3/3)

ร้านน้อยชามู

Review ID	Restaurant_ID	Restaurant	User	Headline	Review	Rating	Review_tokenized
16	17	436045MJ-ร้านน้อยชามู	plyynp	บุฟเฟ่ต์ข้าวและพืชผักไม่อินในราคา 199 บาท เน้น...	หลังจากที่เคยลองสาขาบูเนียมมอลล์ไป รอบนี้ มาที่...	4.0	หลังจากที่,ลอง,สาขา,บู,เนม,มอลล์,รอบ,สาขา,เดอะ,ม...
17	18	436045MJ-ร้านน้อยชามู	27a91236fe5e4559a4f097c97a480781	ร้านบุฟเฟ่ต์ ราคาดีมีคุณภาพ อยู่ชั้น4 ติดโรงหนึ่ง ...	ร้านบุฟเฟ่ต์ที่มีโปรโมชันหิวละ199บาท ไม่รวม นี...	4.0	บุฟเฟ่ต์,โปรโมชัน,หิว,199,บาท,VAT,ธรรมดา,ชอบ,...
18	19	436045MJ-ร้านน้อยชามู	0b81d251e4db486f9bcdba73b374ed99	ของหลากหลาย ปนๆ งงๆ นิดหน่อย	เคยรู้จักร้านนี้จากที่ union mall ไม่เคยได้ลอง...	3.0	รู้จัก,union,mall,ลอง,จ่ายได้,ขึ้นใจ,ซื้อ,😞,ซื้อ...
19	20	436045MJ-ร้านน้อยชามู	40e0e087f3914fd49a8933b5a29936ca	อร่อยมากค่ะ ค่ำค่าสมราคา บุฟเฟ่ต์มีผักต่างๆ รว...	อร่อยมากค่ะ ค่ำค่าสมราคา บุฟเฟ่ต์มีผักต่างๆ รว...	5.0	อร่อย,ค่ำค่า,สมราคา,หมู,ผัก,จบ,255,ท่าน,เด็ด,...
20	21	436045MJ-ร้านน้อยชามู	41841cb99ea243a8d4b006e946c586	แม้จะแปลกบ้าง แต่ก็ถือว่าอยู่ในเกณฑ์ที่ดี มีอา...	ก็ตามที่เขียนเลยครับ ว่า ถ้าจะจ่ายในราคา 199 บ...	NaN	จ่าย,ราคา,199,บาท,จ่ายเงินสด,ราคา,VAT,7,%,สั...

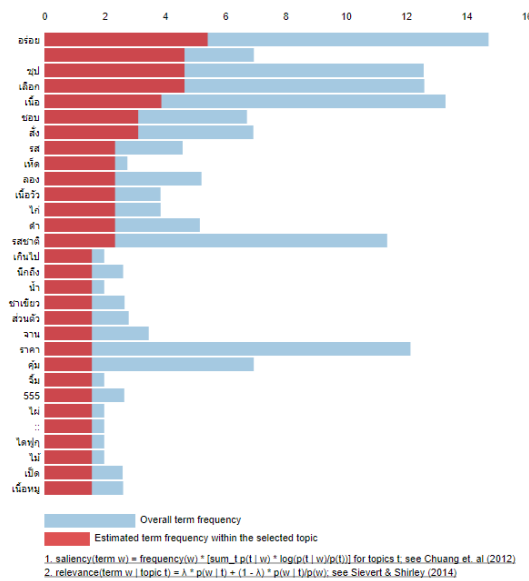
Selected Topic: 2 Previous Topic Next Topic Clear Topic

Slide to adjust relevance metric: (2) $\lambda = 1$ 0.0 0.2 0.4 0.6 0.8 1

Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Relevant Terms for Topic 2 (15.6% of tokens)



บทสรุป Voice of Customer Analytics

จะเห็นว่าการใช้ **LDA Text Classification** โดยโปรแกรม จะมีข้อดีซึ่งได้เปรียบกว่าการทำ **Manual Text Classification** โดยคน ในกรณีที่ข้อมูลมีจำนวนมหาศาล อย่างไรก็ตามการใช้ **LDA Text Classification** จะต้องมีการ **Labelling** ข้อมูลเบื้องต้นก่อนเพื่อตัดคำที่ต้องการสื่อความหมายของ **Sentiment** ออกมา ซึ่งไม่ใช่เรื่องง่ายที่ต้องทำ เพราะ คำบางคำเป็นคำแสดง คำบางคำความหมายดีแต่มื่อนำมารวมกันทั้งประโยคแล้วกลายเป็นคำประชดที่ความหมายแย้งไปซะงั้น หรือคำบางคำมีได้หลายความหมาย

อ้างอิง เพื่อศึกษาเพิ่มเติม

<https://nlpforthai.com/software/>

<https://github.com/PyThaiNLP/pythainlp>