

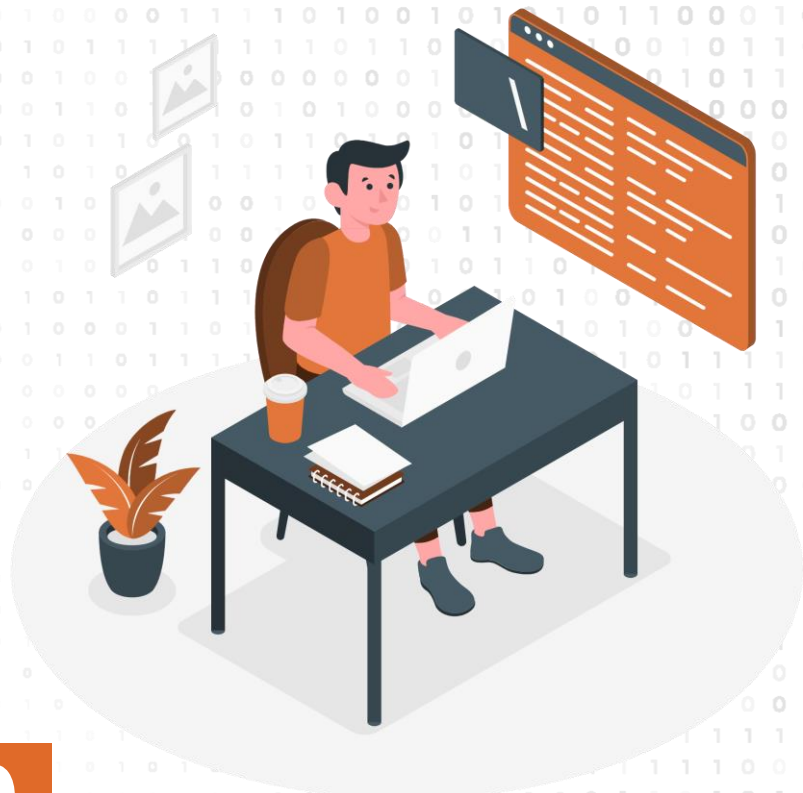
A young woman with long dark hair, wearing a light blue denim jacket over a grey t-shirt, is smiling and looking towards a man in a white shirt. She is holding a black card payment terminal with her right hand, and the man is holding it with his left hand. The background is a blurred indoor setting with warm lighting.

VOICE OF CUSTOMER

AMARA KAWEEKATE 6310412014

Data : Wongnai Reviews

Data : Wongnai Reviews



STEP 0

Install and import dependencies

Import pythainlp (คือ Library แพลตฟอร์มของภาษา Python ใช้สำหรับประมวลผลข้อความ และการวิเคราะห์ทางภาษา คล้ายกับ NLTK แต่ใช้กับภาษาไทยโดยเฉพาะ มีฟังก์ชันการทำงานที่หลากหลาย)



STEP 1

Document embedding and dimension reduction

ใช้ Model ของ Universal Sentence Encoder (USE) ในการสร้าง Vector representation หรือ Embeddings เพื่อนำไปใช้ในหาความเหมือนของประโยคหรือนำไปใช้ต่อในโมเดลอื่น ๆ

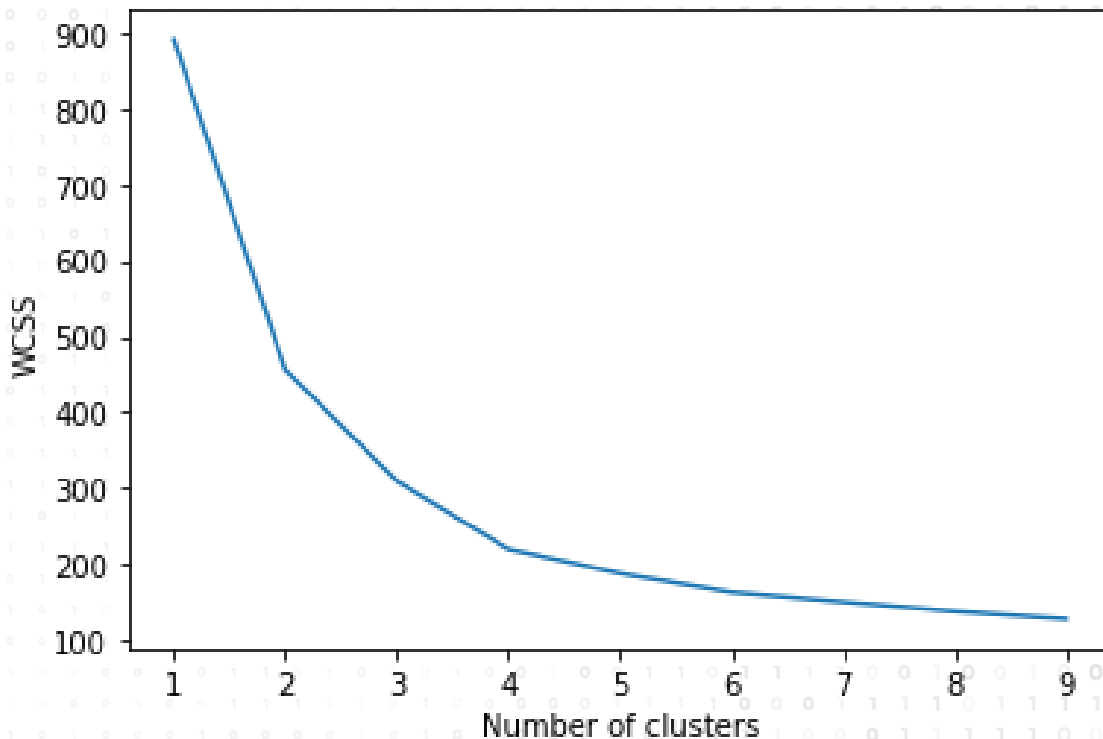


STEP 2

Document clustering using K-means

เพื่อจัดกลุ่มว่า Document ไหนอยู่กลุ่มเดียวกัน โดยใช้ K-means Clustering (เป็นการใช้ค่ายูคลิเดียนในการวัดระยะห่าง) โดยเมื่อดูกราฟ Elbow พบว่าจะได้ **K = 3** ในสไลด์ถัดไป

The Elbow title



K-based on the elbow plot (K=3)

Step 2 - document clustering using K-Means



STEP 3

**Document clustering using Agglomerative Clustering
with cosine similarity**

เพื่อจัดกลุ่มว่า Document ไหนอยู่กลุ่มเดียวกันโดยใช้ Agglomerative Clustering เป็น Hierarchical Method (ค่า cosine)



FINDING

Remove Word, New Word สามารถแบ่งออกเป็นกี่กลุ่ม ?

INTERPRET

เมื่อทำการวิเคราะห์ สามารถแบ่งกลุ่มได้เป็น 3 กลุ่ม ดังนี้



CAFÉ COFFEE

เน้นรีวิวร้านค้าที่เป็น
ร้านกาแฟ มีการพูดถึง
บรรยากาศของร้าน
รสชาติ และราคา



สายแซบ SOCIETY

เน้นรีวิวร้านอาหาร มัก
กินข้าวกับกลุ่มเพื่อน
ส่วนใหญ่เป็นร้านส้มตำ
ชอบทานอาหารรสจัด



รักใครก็ไม่เท่า “ซานมไข่มุก”

เน้นรีวิวเครื่องดื่ม
ประเภทซานมไข่มุก
ซานมไต้หวัน ชาเขียว
เป็นหลัก