

Report twitter sentiment analysis project

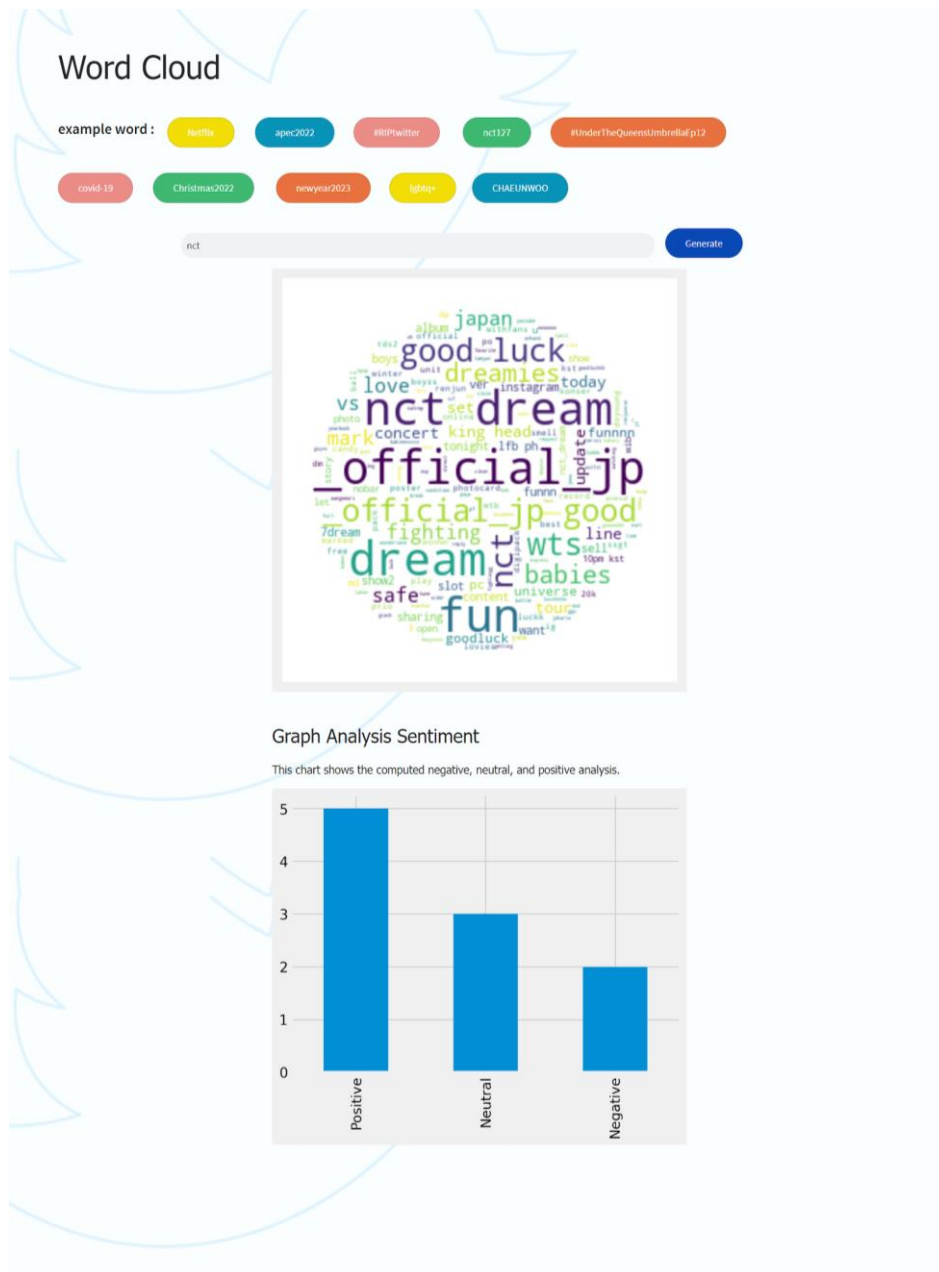
การทำงานของเว็บ

เมื่อเสิร์ชคีย์เวิร์ดลงไปในช่วงค้นหา ในครั้งแรกจะทำการ Authorization กับ Twitter API ก่อน จากนั้นเมื่อได้คีย์เวิร์ดมาแล้ว ก็จะไปเรียกใช้ฟังก์ชัน search_tweets แล้วใส่ arguments เข้าไปซึ่งเป็นคีย์เวิร์ดที่ป้อนเข้าไปในฟังก์ชัน จากนั้นฟังก์ชัน search_tweets ก็จะไปค้นหา tweets ตามคีย์เวิร์ดที่ป้อนเข้ามา และ return ข้อมูลเป็น Data Frame

เมื่อได้ข้อมูลมาแล้ว จะนำ Data Frame ไปเขียนเป็นไฟล์ Json และหลังจากนั้นจะดึงไฟล์ Json ออกมาเพื่อเตรียม Insert ลง Database ซึ่งก่อนที่จะ Insert ข้อมูลเข้าไป จะทำการ Delete ข้อมูลเก่าออกไปก่อน แล้วจะนำข้อมูลใหม่เข้าไปเก็บแทนที่

จากนั้นจะนำ Data Frame ที่ได้มาจากฟังก์ชัน search_tweets ไปทำการ Clean texts ซึ่งก็คือการกำจัดอีโมจิ Link ต่างๆ อย่าง '@' '#' และ 'RT' จากนั้นจะนำข้อมูลที่ Clean แล้ว ไปทำ Sentimental Analysis และ generate Word Cloud ก็จะได้ข้อมูลที่ return ออกมาเป็นรูป แล้วนำรูป Word cloud ที่ generate ได้มา ขึ้นโชว์บนหน้าเว็บ หลังจากนั้นจะนำเอาค่า Sentimental Analysis ไป Count ว่ามี Positive, Neutral, Negative เป็นจำนวนเท่าไร แล้ว return ออกมาในรูปแบบของชาร์ต และนำชาร์ตไปโชว์บนหน้าเว็บ

ตัวอย่าง Output Program



จากคีย์เวิร์ด nct จะได้รูปปกอภมาดังภาพ จากภาพจะสังเกตเห็นคำว่า nct dream_official_jp ตัวใหญ่ที่สุดเป็นเพราะช่วงนี้วงบอยแบนด์ ที่มีชื่อว่า ‘nct dream’ เดินทางไปเดบิวต์และแสดงคอนเสิร์ตที่ประเทศญี่ปุ่นและมีการกล่าวถึงแอคเคาท์ official เป็นจำนวนมาก ซึ่งจะเห็นคำว่า japan, safe, love, good luck, concert, show2, fun, album ในรูปภาพเพราะช่วงนี้แฟนคลับได้โพสต์ให้กำลังใจและสนับสนุนเกี่ยวกับคอนเสิร์ตของ nct dream คำต่อไปที่ขนาดใหญ่รองลงมาก็คือ babies, dreamie, 7dream เป็นคำเรียกของแฟนคลับที่ใช้เรียกศิลปิน และในรูปจะเห็นคำว่า WTS เป็นคีย์เวิร์ดสำหรับการขายสินค้าที่เกี่ยวข้องกับวงศิลปินเกาหลี

สมาชิกในกลุ่ม

63050094 กษภักค์ อภิเดชนาภรณ์

63050115 ชลธิชา ประภาสัย

63050116 ชุตติกาญจน์ แสงแก้ว

63050151 เนตรศิริ รัตนสุวรรณ

63050208 อรรถพงษ์ ขอชัย

63050211 อัญชีษฐา ตินชาตอารักษ์