

บทที่ 5

การประเมินผลการทดลอง

ภายหลังจากที่ได้สร้าง โมเดลเพื่อทดสอบความถูกต้องของการจัดกลุ่มของเอกสารและ การเรียนรู้ของอัลกอริทึม ซัพพอร์ตเวกเตอร์แมชชีน โดยใช้เอกสารข่าวข้อความภาษาอังกฤษ ซึ่งรวบรวมมาจากโดเมน “<http://www.cnn.com>” ในการสร้าง โมเดล 1,000 เอกสาร และเอกสารในการทดสอบ 500 เอกสาร ผลการทดสอบการแบ่งกลุ่มข้อมูล 5 กลุ่มคือ

1. เอกสารข้อความข่าวธุรกิจ
2. เอกสารข้อความข่าวสุขภาพ
3. เอกสารข้อความข่าวกีฬา
4. เอกสารข้อความข่าวท่องเที่ยว
5. เอกสารข้อความข่าวคอมพิวเตอร์

และประสิทธิภาพของการทดสอบ โมเดลการจัดกลุ่มเอกสาร โดยใช้ซัพพอร์ตเวกเตอร์แมชชีน แสดงได้ในตาราง 5.1

ขั้นตอนการวัดประสิทธิภาพเป็นขั้นตอนของการนำเอากลุ่มของเอกสารที่จัดได้มาทำการประเมินประสิทธิภาพ โดยจะตรวจสอบดูว่ากลุ่มของเอกสารที่จัดได้มีค่าเป็นอย่างไร เมื่อเทียบกับกลุ่มของเอกสารที่ถูกตั้งซึ่งวัดจากค่าความระลึก (*Recall*) และค่าความแม่นยำ (*Precision*) ค่าความแม่นยำ จะเป็นค่าที่แสดงว่า การค้นคืนข้อมูลได้ตรงกับความต้องการเพียงใด ส่วนค่าความระลึกจะเป็นค่าที่แสดงถึงความครอบคลุมในการจัดกลุ่มเอกสาร

ตาราง 5.1 ผลการทดสอบโมเดลการจัดกลุ่มเอกสาร โดยใช้ซัพพอร์ตเวกเตอร์แมชชีน

Predict	Business	Health	Sport	Travel	Computer
Business	54	1	0	3	5
Health	3	95	2	13	41
Sport	0	0	92	1	0
Travel	42	4	5	81	7
Computer	1	0	1	2	47

ตาราง 5.2 วัดประสิทธิภาพการจัดกลุ่มเอกสาร โดยใช้ซอฟต์แวร์เวกเตอร์แมชชีน

ประเภทของข่าว	Precision (%)	Recall (%)	F-measure (%)
ธุรกิจ	54	86	66.26
สุขภาพ	98	62	75.72
กีฬา	83	99	90.27
ท่องเที่ยว	81	58	67.78
คอมพิวเตอร์	47	92	62.25

จากผลการทดลองพบว่าสามารถจำแนกได้ 3 กลุ่มคือ

1. กลุ่มวัดประสิทธิภาพการจัดกลุ่มเอกสารได้ดีที่สุด

กลุ่มวัดประสิทธิภาพการจัดกลุ่มเอกสารที่ดีที่สุดคือกลุ่มข่าวกีฬา ซึ่งมีค่าความแม่นยำ

83 เปอร์เซนต์ ค่าความระลึก 99 เปอร์เซนต์ และค่าประสิทธิภาพของงานคิดเป็นร้อยละ 90.72

เปอร์เซนต์ แสดงว่าในตัวเอกสารของกลุ่มข่าวกีฬามีคำเฉพาะของกลุ่ม ทำให้สามารถจัดกลุ่ม

เอกสารอย่างได้ชัดเจน คำเฉพาะที่ปรากฏในเอกสารเช่น golf, sport, ball เป็นต้น

By Anouk Lorie for CNN Golf ball dissolves in water STORY HIGHLIGHTS As sport becomes greener, a Spanish golf ball manufacturer invented an edible golf ball The golf ball is completely biodegradable and dissolves into fish food, said Albus Golf The balls are to be used on piers and cruise liners but cannot be hit as far as normal balls RELATED TOPICS Golf U.S. Open - Golf Women's Golf London, England (CNN) -- For even the most environmentally conscious of players, playing golf and remaining green is a difficult balance to achieve, especially if your aim lands you in water hazards more than on the fairway. Research teams at the Danish Golf Union discovered it takes between 100 to 1,000 years for a golf ball to decompose naturally, a startling fact when it is estimated 300 million balls are lost or discarded each year in the United States alone. Once the misplaced shot settles in the sediment, whether the body of water that consumes your ball is a pond, river or lake; the disintegrating ball will then slowly release a high quantity of the heavy metal zinc as it disintegrates. But it seems there maybe a solution to help offending players keep out of hot water, thanks to a Spanish golf ball manufacturer that has invented a ball which not only dissolves upon contact with water, but is made of fish food. Albert Buscato, CEO of Albus Golf, told CNN he wanted to improve on the range of biodegradable golf equipment on the market. "I thought it would be great to have a ball that not only does not pollute the waters, but also gives something back," he said. ...etc

รูป 5.1 เอกสารกลุ่มข่าวกีฬา

2. กลุ่มวัดประสิทธิภาพการจัดกลุ่มเอกสารได้ปานกลาง

กลุ่มวัดประสิทธิภาพการจัดกลุ่มเอกสารที่ได้ปานกลางคือกลุ่มข่าว สุขภาพ และ กลุ่มข่าวท่องเที่ยว ซึ่งมีค่าความแม่นยำสูงแต่ค่าความระลึกลำ แสดงว่าแบบจำลองสามารถจัดกลุ่มเอกสารตามกลุ่มเอกสารได้แต่เอกสารนั้นมีความคล้ายคลึงกับเอกสารกลุ่มอื่นทำให้ค่าความระลึกลำ จากรูป 5.2 เอกสารกลุ่มข่าวสุขภาพ ซึ่งสังเกตว่าในตัวเอกสารสุขภาพนั้นจะประกอบด้วยคำของสุขภาพและคอมพิวเตอร์ เอกสารนี้กล่าวเกี่ยวกับสุขภาพแต่แบบจำลองทำนายว่ามีความคล้ายคลึงกับกลุ่มข่าวคอมพิวเตอร์ คำในกลุ่มสุขภาพที่ปรากฏในเอกสารนี้เช่น Cancer, Health, HIV ฯลฯ และคำในกลุ่มคอมพิวเตอร์ที่ปรากฏเช่น Hi-Tech, Myspace, Facebook ฯลฯ

Women's Health Men's Health Children's Health Boomer Health Cold & Flu Headache & Migraines Sexual Health Pregnancy & Fertility Beauty & Plastic Surgery Sleep Pain Management Gadgets & Hi-Tech Brain & Nervous System Digestive Health Respiratory Health Infectious Disease Cancer & Chemo Diabetes Heart & Vascular HIV & AIDS Autism Vitamins & Supplements Weight Loss Calories, Carbs & Fats Healthy Eating Recipes Yoga & Meditation Travel Fitness Budget Fitness Strength Training Alternative Exercise Cardio Fitness Exercise Addiction ADHD Alzheimer's Disease Stress Depression Happiness Post Traumatic Stress Disorder Amazing Mind You may not post any unlawful, threatening, libelous, defamatory, obscene, pornographic or other material that would violate the law. By submitting your question, you hereby give CNN the right, but not the obligation, to post, air, edit, exhibit, telecast, cablecast, webcast, re-use, publish, reproduce, use, license, print, distribute or otherwise use your question(s) and accompanying personal identifying and other information you provide via all forms of media now known or hereafter devised, worldwide, in perpetuity. CNN Privacy Statement. It will be reviewed and considered for posting on CNNHealth.com. Questions and comments are moderated by CNN and will not appear until after they have been reviewed and approved. Unfortunately, because of the volume of questions we receive, not all can be posted. Submit another question or Go back to CNNHealth.com Read answers from our experts: Living Well | Diet & Fitness | Mental Health | Conditions Expert Q&A Share this on: Mixx Digg Facebook delicious reddit StumbleUpon MySpace could low iron, protein contribute to soreness after exercise? Asked by Sara, Boston, Massachusetts I seem to get very sore after every workout. My fiance says I'm working out too hard, but I was wondering if there were other reasons. I've been a vegetarian for 13 years and am usually low on protein and iron. Could low levels of protein and/or iron contribute to why I get so sore after working out? Diet and Fitness Expert Dr. Melina Jampolis Physician Nutrition Specialist Expert answer Hi Sara. This a very good question, and to get you the most accurate response, I turned to Steven D. Stovitz, M.D., a fellow of the American College of Sports Medicine and an associate professor in the University of Minnesota Department of Family Medicine and Community Health...etc

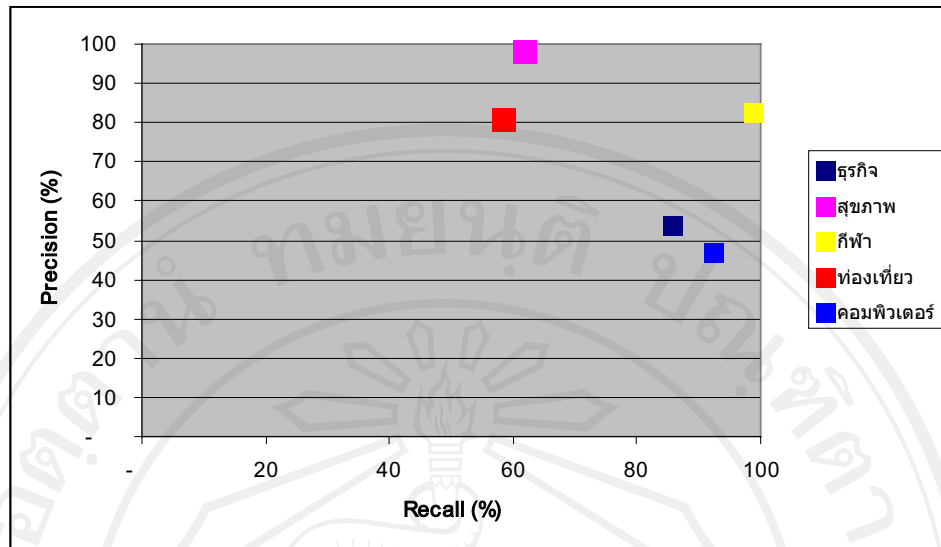
รูป 5.2 เอกสารกลุ่มข่าวสุขภาพ

3.กลุ่มวัดประสิทธิภาพการจัดกลุ่มเอกสารได้ต่ำ

กลุ่มวัดประสิทธิภาพการจัดกลุ่มที่ได้ต่ำคือกลุ่มข่าวธุรกิจและกลุ่มข่าวคอมพิวเตอร์ ซึ่งมีค่าความแม่นยำต่ำ แต่ค่าความระลึกลึกสูง ซึ่งในตัวเอกสารของกลุ่มข่าวธุรกิจและคอมพิวเตอร์อาจจะมีคำที่ใช้ร่วมกับกลุ่มข่าว อื่นและคำที่ใช้เป็นคำที่ใช้ทั่วไปทำให้ตัวแบบจำลองไม่สามารถจัดกลุ่มได้อย่างชัดเจน คำที่ใช้ทั่วไปเช่น model, drivers, connected, services ฯลฯ จากรูป 5.3 เอกสารกลุ่มข่าวคอมพิวเตอร์ ซึ่งสังเกตว่าเอกสารกล่าวถึงเรื่องคอมพิวเตอร์มีเชื่อมโยงกับเรื่องธุรกิจ

STORY HIGHLIGHTS Ford says new technology will first show up in Edge and Focus model cars Ford official says tweeting from behind the wheel expected in a year Ford also includes Pandora music, Stitcher online news partnerships Company says they are just taking what people already do "and make them safer" RELATED TOPICS Consumer Electronics Ford Motor Company Twitter Inc. Las Vegas, Nevada (CNN) -- The debate over distracted drivers ramped up a notch on Thursday as Ford Motor Co. announced technology to let computers read their Twitter feeds to them while behind the wheel. The company has not formally announced technology to let drivers post tweets verbally from behind the wheel, but such an announcement likely will come in the next 12 to 18 months, Doug VanDagens, Ford's global director of connected services, told CNN in an interview after the company's keynote address at the Consumer Electronics Show here. VanDagens said people read Twitter feeds and send text messages while they're driving anyway, and that the automaker is trying to make those activities safer for drivers. "We take what people do -- they talk on the phone, they fumble with mp3 players, they look at maps. We take these activities and make them safer," he said. Safety groups have criticized car companies and technology makers for putting touch-screen devices that surf the Internet in cars. The U.S. is in a distracted driving crisis, and, in general, the more technology electronics makers put into cars, the less drivers are able to pay attention to the road, said Christie Hype, a spokeswoman for AAA, the motor safety group. "The more things that are going on in a vehicle, the more things can distract a driver," she said. "You only have so much attention to give, and we really want everyone to keep their attention on the roadway for safety reasons." More study of the tweeting by voice while driving would need to take place to make an absolute determination as to what effect Ford's announcement will have on distracted driving, she said. Ford's system, which will be installed in some cars later this year, does not let drivers surf the Internet visually while driving. A Wi-Fi hot spot in the new cars generate will let drivers browse the Web when their cars are parked, VanDagens said. The technologies will first show up in Edge and Focus model cars. ...etc

รูป 5.3 เอกสารกลุ่มข่าวคอมพิวเตอร์



รูป 5.4 ประสิทธิภาพของระบบ

ดังนั้นอาจจะกล่าวได้ว่าการจัดกลุ่มเอกสารด้วยอัลกอริทึม ซัพพอร์ตเวกเตอร์แมชชีน นี้
 นำหนักของคำแต่ละคำมีความสำคัญในการจัดกลุ่มเอกสารเป็นอย่างมาก