

การจัดประเภทของเว็บโดยการเชื่อมโยงของเว็บในระดับเพจและบล็อก



นางสาวพรทิพย์ มาลีลัย

สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2549

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

WEB CLASSIFICATION BY CONNECTIVITY AT PAGE AND BLOCK LEVEL



Miss Porntip Maleelai

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Computer Science

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic Year 2006

Copyright of Chulalongkorn University

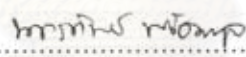
หัวข้อวิทยานิพนธ์ การจัดประเภทของเว็บโดยการเชื่อมโยงของเว็บในระดับเพจและบล็อก
โดย นางสาวพรทิพย์ มาลีลัย
สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษา อาจารย์ นครทิพย์ พร้อมพูล

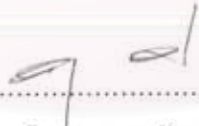
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยานิพนธ์ฉบับนี้
เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต


..... คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.ติเรก ลาวัณย์ศิริ)

คณะกรรมการสอบวิทยานิพนธ์


..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)


..... อาจารย์ที่ปรึกษา
(อาจารย์ นครทิพย์ พร้อมพูล)


..... กรรมการ
(อาจารย์ ดร.พิษณุ คนองชัยยศ)


..... กรรมการ
(อาจารย์ ธงชัย โรจน์กั้งสตาล)

สถาบันวิศวกรรม
จุฬาลงกรณ์มหาวิทยาลัย

พรทิพย์ มาลีสัย : การจัดประเภทของเว็บโดยการเชื่อมโยงของเว็บในระดับเพจและบล็อก. (WEB CLASSIFICATION BY CONNECTIVITY AT PAGE AND BLOCK LEVEL) อ. ที่ปรึกษา : อ. นครทิพย์ พร้อมพูล, 137 หน้า.

การวิจัยนี้มีวัตถุประสงค์เพื่อทำการจัดประเภทของเว็บไซต์โดยการเชื่อมโยงของเว็บในระดับเพจและบล็อกให้เป็นหมวดหมู่เพื่อช่วยด้านการสืบค้นข้อมูล โดยการนำเสนอการประยุกต์ใช้คุณลักษณะระดับเพจและบล็อกมาช่วยในการจัดประเภทของเว็บไซต์

ผู้วิจัยได้ใช้ประเภทของเว็บไซต์จำนวน 11 ประเภท จำนวน 265 เว็บไซต์ รวมทั้งสิ้น 2,650 เพจ โดยใช้อัลกอริทึม VIPS ในการแบ่งเพจเป็นบล็อก แล้วคำนวณค่าคุณลักษณะระดับเพจและบล็อก จากนั้นได้ใช้วิธีการเรียนรู้ของเครื่องแบบ See5 แสดงให้เห็นค่าความถูกต้องในการจัดประเภทของเว็บไซต์

ผลการทดลองที่ได้จากการวิจัยในครั้งนี้ แสดงค่าความผิดพลาดที่เกิดจากการจัดประเภทของเว็บไซต์ที่พิจารณาจากคุณลักษณะระดับเพจ คุณลักษณะระดับบล็อก คุณลักษณะระดับเพจและบล็อก ซึ่งให้ค่าผิดพลาดที่เรียงจากน้อยที่สุดไปหามากที่สุดคือ การจัดประเภทของเว็บไซต์โดยใช้คุณลักษณะระดับเพจเป็น 62.5 เปอร์เซ็นต์ การจัดประเภทของเว็บไซต์โดยใช้คุณลักษณะระดับเพจและระดับบล็อกเป็น 67.8 เปอร์เซ็นต์ และการจัดประเภทของเว็บไซต์โดยใช้คุณลักษณะระดับบล็อกเป็น 75.9 เปอร์เซ็นต์ ทั้งนี้แนวทางของการทดลองในงานวิจัยนี้อาจสามารถนำไปปรับปรุงให้ดีขึ้นโดยการให้ความสำคัญในการเลือกหน่วยตัวอย่าง และเพิ่มจำนวนเพจที่อยู่ในเว็บไซต์เพื่อนำมาซึ่งผลการทดลองที่ดีขึ้นต่อไป

สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา..... วิศวกรรมคอมพิวเตอร์.....ลายมือชื่อนิสิต..... พรทิพย์ มาลีสัย
สาขาวิชา.....วิทยาศาสตร์คอมพิวเตอร์.....ลายมือชื่ออาจารย์ที่ปรึกษา..... นครทิพย์ พร้อมพูล
ปีการศึกษา ...2549.....

46714231 : MAJOR COMPUTER SCIENCE

KEY WORD: WEB CLASSIFICATION / WEB PAGE / WEB BLOCK / DATA MINING

PORNTIP MALEELAI : WEB CLASSIFICATION BY CONNECTIVITY AT PAGE AND BLOCK LEVEL. THESIS ADVISOR : NAKORNTIP PROMPOON, 137 pp.

The purpose of this research is to classify web sites by connectivity at page and block level which helps information searching. This research applies features at page and block level for web classification

In carrying out this research, researcher uses 11 types of web sites, 265 web sites and 2,650 pages. VIPS algorithm is also used to divide page to block . Then page features and block features are computed. After that, See5 is used to classify and evaluate the accuracy of web site classification.

The result of this experiment has shown error of web site classification which considers by features in page level, block level and page&block level. The values of percentage error for using only page features, page and block features and only block features are 62.6%, 67.8% and 75.9% respectively. The experiment from this research can be improved by emphasizing on the selection of the sample units of web site and thoroughly considering to increase more pages in each web site to give more classification accuracy.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department..... Computer Engineering.....Student's.....*นันทน์ ชาติ*
Field of study.....Computer Science.....Advisor's.....*Nakornthip Prompoon*
Academic year ...2006.....

กิตติกรรมประกาศ

ข้าพเจ้าใคร่ขอกราบขอบพระคุณอาจารย์ อ.นครทิพย์ พร้อมพูล อาจารย์ที่ปรึกษาวิทยานิพนธ์ของข้าพเจ้า ที่กรุณาสอนวิชาความรู้ แนะนำให้คำปรึกษา ความช่วยเหลือต่าง ๆ ตลอดจนการทำวิทยานิพนธ์ของข้าพเจ้าจนสำเร็จลุล่วงลงได้ด้วยดี

ขอกราบขอบพระคุณผู้ช่วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล ซึ่งเป็นประธานกรรมการสอบวิทยานิพนธ์ อาจารย์ ดร.พิชณู คนองชัยยศ และอาจารย์ ธงชัย วิจารณ์กั้งสดาด ซึ่งเป็นกรรมการสอบวิทยานิพนธ์ ที่ได้สละเวลาและให้คำแนะนำต่าง ๆ อันเป็นประโยชน์ต่อการจัดทำวิทยานิพนธ์ฉบับนี้

ขอขอบคุณอาจารย์ทุกท่าน ที่ได้ประสิทธิ์ประสาทวิชาให้กับข้าพเจ้า รวมถึงข้อชี้แนะต่าง ๆ ตลอดเวลาที่ข้าพเจ้าได้ศึกษาเล่าเรียนในระดับมหาบัณฑิต ณ สถาบันแห่งนี้

ขอขอบคุณเพื่อน ๆ พี่ ๆ น้อง ๆ ทุกคนในห้องปฏิบัติการชั้น 19 และเพื่อนร่วมรุ่นของข้าพเจ้าทุกคนที่คอยถามไถ่ห่วงใย รวมถึงเป็นแรงกระตุ้นที่ดี โดยเฉพาะนายเมธี จันทรประเสริฐ และนางสาวเสาวลักษณ์ รัตนธรรมสกุล นอกจากนี้ข้าพเจ้าขอขอบคุณสำหรับกำลังใจและความเข้าใจจากพี่ ๆ น้อง ๆ ที่บริษัท ทริมอินเตอร์เนชั่นแนลและครอบครัว

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญภาพ.....	ญ
สารบัญตาราง.....	ฎ
บทที่	
1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์ของการวิจัย	3
1.3 ขอบเขตของการวิจัย.....	3
1.4 ขั้นตอนและวิธีดำเนินการวิจัย.....	4
1.5 ประโยชน์ที่คาดว่าจะได้รับ	4
1.6 ลำดับการจัดเรียงเนื้อหาในวิทยานิพนธ์	4
2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	5
2.1 ทฤษฎีที่เกี่ยวข้อง.....	5
2.1.1 เทคโนโลยีเกี่ยวกับเว็บ	5
2.1.1.1 โครงสร้างของภาษาเซสที่เอ็มแอล	5
2.1.1.2 การเชื่อมโยง.....	7
2.1.2 บล็อก.....	10
2.1.3 การทำเหมืองข้อมูล.....	12
2.1.3.1 การจัดประเภท.....	13
2.1.3.2 การทำนาย	13
2.1.4 ต้นไม้ในการตัดสินใจ (Decision Tree)	13
2.1.5 วิธีการเรียนรู้ของเครื่องแบบ See5.....	14

2.2 งานวิจัยที่เกี่ยวข้อง	16
2.2.1 การเชื่อมต่อโซน่า: การตรวจตราฟังก์ชันของไซต์โดยใช้รูปแบบโครงสร้าง (The Connectivity Sonar: Detecting Site Functionality by Structural Patterns).....	16
2.2.2 การวิเคราะห์การเชื่อมโยงในระดับบล็อก (Block-level Link Analysis).....	18
2.2.3 ขั้นตอนวิธีการแบ่งส่วนเพจด้วยวิชันเบส (VIPS: a Vision-based Page Segmentation Algorithm)	18
2.2.4 ตัวแบบการเรียนรู้ค่าความสำคัญสำหรับบล็อกของเว็บเพจโดยอาศัยการวิเคราะห์การจัดวางหน้าและเนื้อหา (Learning Important Models for Web Page Blocks based on Layout and Content Analysis).....	19
3 แนวทางการวิจัย.....	20
3.1 การประยุกต์ใช้บล็อกเพื่อจัดประเภทของเว็บ	20
3.2 การเลือกเว็บไซต์ในแต่ละประเภท	21
3.3 การเลือกเว็บเพจของแต่ละเว็บไซต์.....	22
3.4 การเลือกบล็อกในแต่ละเว็บเพจ.....	22
3.5 คุณลักษณะระดับเพจ.....	23
3.6 คุณลักษณะระดับบล็อก.....	23
3.7 ขั้นตอนการคำนวณ.....	24
3.7.1 อ่านการเชื่อมโยงของเว็บไซต์เพื่อเลือกเพจ	24
3.7.2 เลือก 10 เพจสำหรับเว็บไซต์นั้น	24
3.7.3 คำนวณคุณลักษณะในระดับเพจในแต่ละเพจ	24
3.7.4 แบ่งเพจเป็นบล็อกโดยอาศัยอัลกอริทึม VIPS	25
3.7.5 คำนวณค่าในแต่ละบล็อกเพื่อเลือก 3 บล็อก	25
3.7.6 คำนวณค่าเฉลี่ยของคุณลักษณะระดับบล็อก 3 บล็อกที่เลือก	25
3.7.7 คำนวณค่าเฉลี่ยของคุณลักษณะระดับบล็อกสำหรับเว็บไซต์	26
3.7.8 คำนวณค่าเฉลี่ยของคุณลักษณะระดับเพจสำหรับเว็บไซต์.....	26
4 การทดลองและผลการทดลอง	27
4.1 ชุดข้อมูลที่ใช้ในการทดลอง	27
4.2 วิธีการทดลอง	28

4.3 ผลการทดลอง.....	29
4.3.1 การใช้คุณลักษณะระดับเพจในการจัดประเภท.....	30
4.3.2 การใช้คุณลักษณะระดับบล็อกในการจัดประเภท	32
4.3.3 การใช้คุณลักษณะระดับเพจและบล็อกในการจัดประเภท.....	34
5 สรุปผลการวิจัยและข้อเสนอแนะ.....	38
5.1 สรุปผลการวิจัย	38
5.2 ปัญหาและข้อจำกัดในการวิจัย	39
5.3 ข้อเสนอแนะ	39
รายการอ้างอิง.....	40
ภาคผนวก	
ภาคผนวก ก เว็บไซต์ที่ใช้ในการเรียนรู้และทดสอบ	42
ภาคผนวก ข เอกซ์เอ็มแอล	52
ภาคผนวก ค ตัวอย่าง See5 และ RULE-BASED CLASSIFIERS	61
ภาคผนวก ง โปรแกรม VIPS.....	68
ภาคผนวก จ ตัวอย่างการคำนวณค่าคุณลักษณะของเว็บไซต์	69
ประวัติผู้เขียนวิทยานิพนธ์	137

สารบัญญภาพ

	หน้า
รูปที่ 1.1 ลักษณะของเว็บประเภทต่าง ๆ	2
รูปที่ 1.2 บล็อกในเว็บเพจ.....	3
รูปที่ 2.1 โครงสร้างของภาษาเอชทีเอ็มแอล.....	5
รูปที่ 2.2 การเชื่อมโยงภายในและการเชื่อมโยงภายนอกของสองไซต์.....	9
รูปที่ 2.3 อัลกอริทึม VIPS	10
รูปที่ 2.4 ตัวอย่างของ Vision-based content structure ในการแบ่งบล็อก	11
รูปที่ 2.5 ลำดับชั้นใน Vision-based content structure	12
รูปที่ 2.6 ต้นไม้ในการตัดสินใจ	13
รูปที่ 2.7 หน้าจอวิธีการเรียนรู้ของเครื่องแบบ See5	15
รูปที่ 2.8 ตัวเลือกต่าง ๆ ในวิธีการเรียนรู้ของเครื่องแบบ See5	15
รูปที่ 3.1 ตัวอย่างของเพจซึ่งสามารถแบ่งออกเป็นบล็อกได้.....	20
รูปที่ 3.2 การเลือก 3 บล็อกในแต่ละเพจ	23
รูปที่ 4.1 กระบวนการเรียนรู้	28
รูปที่ 4.2 กระบวนการแยกประเภท	29

สารบัญตาราง

	หน้า
78ตารางที่ 2.1 ตัวอย่างกฎที่แตกต่างกันของป้ายระบุที่ต่างกัน.....	10
79ตารางที่ 2.2 กฎที่ใช้สำหรับป้ายระบุที่ต่างกันในอัลกอริทึม VIPS.....	11
80ตารางที่ 2.3 ลักษณะการเชื่อมโยง.....	16
81ตารางที่ 3.1 คำสำคัญในการเลือกเว็บไซต์มาใช้ในการเรียนรู้และทดสอบสำหรับเว็บไซต์ในแต่ละประเภท	21
82ตารางที่ 4.1 จำนวนชุดข้อมูลที่ใช้.....	27
83ตารางที่ 4.2 ค่าความผิดพลาดตามกลุ่มข้อมูลโดยใช้คุณลักษณะของเพจ	30
84ตารางที่ 4.3 ผลการจัดประเภทเว็บไซต์ประเภทต่าง ๆ โดยใช้คุณลักษณะระดับเพจ	31
85ตารางที่ 4.4 ค่าความผิดพลาดตามกลุ่มข้อมูลโดยใช้คุณลักษณะระดับบล็อก	32
86ตารางที่ 4.5 ผลการจัดประเภทเว็บไซต์ประเภทต่าง ๆ โดยใช้คุณลักษณะระดับบล็อก	33
87ตารางที่ 4.6 ค่าความผิดพลาดตามกลุ่มข้อมูลโดยใช้คุณลักษณะระดับเพจและบล็อก.....	34
88ตารางที่ 4.7 ผลการจัดประเภทเว็บไซต์ประเภทต่าง ๆ โดยใช้คุณลักษณะระดับเพจและบล็อก	35
89ตารางที่ 4.8 เปอร์เซนต์เปรียบเทียบค่าความผิดพลาดของแต่ละประเภทของเว็บไซต์ของคุณลักษณะระดับเพจ บล็อก เพจและบล็อก	36
90ตารางที่ 5.1 เปอร์เซนต์เปรียบเทียบค่าความผิดพลาดโดยเฉลี่ย (Mean) ของ คุณลักษณะระดับเพจ บล็อก เพจและบล็อก	38

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

อินเทอร์เน็ตถือว่าเป็นเครือข่ายที่เชื่อมโยงข้อมูลและสารสนเทศที่มีขนาดใหญ่ เว็บบนอินเทอร์เน็ตในปัจจุบันมีมากมาย ในการสืบค้นข้อมูลบนอินเทอร์เน็ตทำให้ข้อมูลที่ค้นคืนมาได้มีทั้งข้อมูลที่เกี่ยวข้องและข้อมูลที่ไม่เกี่ยวข้อง เช่น ผู้ใช้ต้องการหาข้อมูลเกี่ยวกับระบบปฏิบัติการ (Operating system) ในโปรแกรมค้นคืน (search engine) จะได้ผลลัพธ์ดังนี้

- หน่วยงานที่ขายระบบปฏิบัติการ
- ไซต์เนื้อหา (Content sites) and developer portals
- มหาวิทยาลัยที่เปิดหลักสูตรสอนเกี่ยวกับระบบปฏิบัติการ

ดังนั้นหากมีการจัดประเภทของเว็บจะช่วยทางด้าน การค้นคืนข้อมูลให้ผู้ใช้ได้รับข้อมูลที่ตรงตามความต้องการของผู้ใช้มากขึ้น

การจัดประเภทของเว็บมีข้อดีหลายประการ คือ ทำให้ผู้ใช้ระบุประเภทของข้อมูลที่ต้องการ ช่วยผู้ใช้ในการเลือกประเภทของข้อมูลที่ต้องการได้ เนื่องจากเว็บแต่ละประเภทจะมีกลุ่มเป้าหมายที่แตกต่างกัน เช่น ผู้ใช้ต้องการหาข้อมูลเกี่ยวกับด้านการศึกษา (Academic) แต่ในปัจจุบันจะได้ผลลัพธ์ทางด้านโดเมน .edu ซึ่งเป็นส่วนของมหาวิทยาลัย ไม่ได้รวมส่วนของสถาบันวิจัย เมื่อทำการค้นหาก็ควรจะได้ข้อมูลของเว็บด้านการศึกษาโดยเฉพาะ

นอกจากนั้นการจัดประเภทของเว็บช่วยในกรณีของเว็บไซต์ที่มีการเปลี่ยนแปลงไม่บ่อยนัก เนื่องจากข้อมูลมีมากมายบนอินเทอร์เน็ต โปรแกรมค้นหาที่มีการปรับ (Update) ข้อมูลอยู่เสมอ ๆ เพื่อให้ทันเหตุการณ์ บางไซต์มีผู้เข้าใช้บ่อย เช่น เว็บของหนังสือพิมพ์ ข่าวประจำวัน ซึ่งเป็นไซต์ที่มีการเปลี่ยนแปลงแก้ไขข้อมูลบ่อย ๆ เพื่อแสดงเหตุการณ์ปัจจุบัน และแนวโน้มต่าง ๆ ดังนั้นจึงมักถูกให้แสดง ณ ตำแหน่งแรก ๆ ของผลลัพธ์การค้นหา ในขณะที่ไซต์ของมหาวิทยาลัยซึ่งมีการเปลี่ยนแปลงแก้ไขข้อมูล ๆ ไม่บ่อยนักเมื่อเทียบกับ .com จะมีประโยชน์ในกรณีที่ผู้ใช้ต้องการให้เว็บไซต์ของมหาวิทยาลัยแสดงที่ตำแหน่งแรก ๆ ของผลลัพธ์การค้นหา

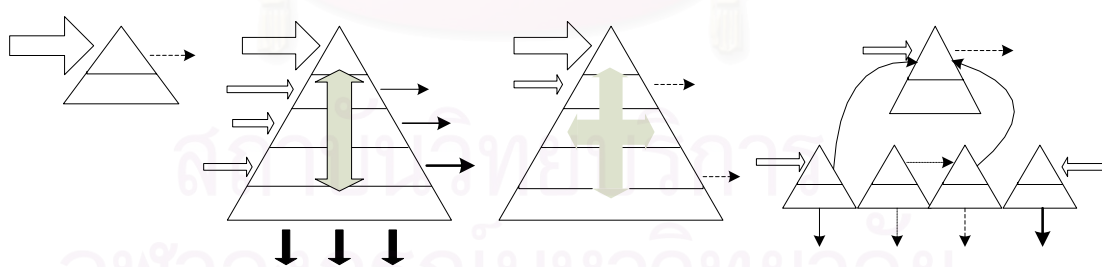
ข้อดีอีกประการหนึ่งของการจัดประเภทของเว็บคือ การสร้างไอคอนข้าง ๆ ผลลัพธ์ของโปรแกรมค้นหาเพื่อบอกประเภทของเว็บ จะช่วยผู้ใช้ในการตัดสินใจในกรณีที่ผู้ใช้มีข้อมูลไม่เพียงพอ

ในเว็บเพจหนึ่ง ๆ ประกอบไปด้วยส่วนต่าง ๆ ที่เรียกว่าบล็อก ซึ่งในแต่ละส่วนมีความสำคัญต่างกัน ถ้าจัดเว็บเพจ ออกเป็นบล็อก ๆ ทำให้เป็นการพิจารณาในการจัดประเภทลงไปยังหน่วยที่เล็กกว่าเพจ ซึ่งน่าจะช่วยให้มีความถูกต้องมากยิ่งขึ้น

การจัดประเภทของเว็บมีหลายวิธี เช่น การจัดประเภทโดยใช้เนื้อหา (Content) โครงสร้างของการเชื่อมโยง (Hyperlinked structure) และรูปแบบการเชื่อมโยงของเว็บ (Connectivity patterns) ซึ่งไม่พิจารณาถึงเนื้อหาของเว็บ

เว็บมีรูปแบบที่หลากหลายและวัตถุประสงค์ของการสร้างที่แตกต่างกัน ถึงแม้รูปแบบการออกแบบจะไม่เหมือนกัน แต่ทำหน้าที่เหมือนกัน ตัวอย่างเช่น เว็บไซต์ของมหาวิทยาลัยต่าง ๆ จำนวนมาก มีความเหมือนกันของโครงสร้างบางอย่างถึงแม้การออกแบบจะไม่เหมือนกัน ลักษณะรูปแบบของการเชื่อมโยง เช่น การเชื่อมโยงภายใน และการเชื่อมโยงภายนอก เป็นต้น

หากพิจารณาโครงสร้างการเชื่อมโยงของเว็บแต่ละประเภทจะพบว่ามีลักษณะที่แตกต่างดังตัวอย่างในรูปที่ 1.1 [1] โปรแกรมค้นหา (Search engine) จะเน้นไปทางด้านการค้นหา เว็บประเภทไดเรกทอรี (Directories) จะมีการเชื่อมโยงภายนอกเป็นจำนวนมาก โดยเฉพาะเพจระดับสุดท้าย (Leaf node) และมีค่าแปรผันโดยตรงตามความลึกของเพจ เว็บไซต์ประเภทบริษัท (Corporate sites) มักจะมีครอสลิงก์ (Cross link) เว็บประเภทแม่ข่ายเสมือน (Virtual hosting) มักจะไม่ค่อยมีการเชื่อมโยงภายใน แต่มีการเชื่อมโยงกลับไปยังโฮสต์ จะเห็นได้ว่าด้วยลักษณะโครงสร้างการเชื่อมโยง และรูปแบบการเชื่อมโยงของเว็บแต่ละประเภทมีความแตกต่างกันจึงสามารถใช้ลักษณะการเชื่อมโยงในการจัดประเภทของเว็บได้



(a) โปรแกรมค้นหา (Search engine) (b) ไดเรกทอรี (Directories) (c) เว็บไซต์บริษัท (Corporate sites) (d) แม่ข่ายเสมือน (Virtual hosting)

รูปที่ 1.1 ลักษณะของเว็บประเภทต่าง ๆ

เพจจะประกอบไปด้วยบล็อกที่แตกต่างกัน ซึ่งก็มีความสำคัญแตกต่างกัน การเชื่อมโยงจากบล็อกที่ต่างกันจะเชื่อมโยงไปยังหัวข้อที่ต่างกัน [2] ดังรูปที่ 1.2 ซึ่งเป็นตัวอย่างแสดงบล็อกบนเพจของเว็บ (news.yahoo.com) จะเห็นว่าเพจสามารถแบ่งออกเป็นบล็อกได้และในแต่ละบล็อกมีเนื้อหาที่ต่างกัน



รูปที่ 1.2 บล็อกในเว็บเพจ

งานวิจัยนี้จะทำการศึกษการจัดประเภทของเว็บ โดยได้แนวคิดมาจากงานวิจัยของ Einat Amitay และคณะ [1] ซึ่งได้ทำการจัดประเภทของเว็บ 8 ประเภท ในงานวิจัยนี้ได้เพิ่มประเภทของเว็บอีก 3 ประเภท คือ เว็บไซต์ดาวน์โหลด (Download) เว็บไซต์หน่วยงานที่ไม่มุ่งหวังผลกำไร (Non-profit organizations) รวมถึงได้นำแนวคิดของเรื่องบล็อกในเพจของ Deng Cai และคณะ [5] มาช่วยในการจัดประเภทของเว็บไซต์ และได้ใช้วิธีการเรียนรู้ของเครื่องแบบ See5 มาช่วยจัดประเภทของเว็บ

1.2 วัตถุประสงค์ของการวิจัย

เพื่อทำการจัดประเภทของเว็บโดยการเชื่อมโยงของเว็บในระดับเพจและบล็อก

1.3 ขอบเขตของการวิจัย

1. จำนวนเว็บที่ใช้ในการเรียนรู้และทดสอบ 265 เว็บไซต์ ตัวอย่างเว็บไซต์ดังแสดงในภาคผนวก ก

2. ใช้ขั้นตอนวิธี VIPS (Vision-Based Page Segmentation Algorithm) ในการแบ่งเพจเป็นบล็อก
3. พิจารณาประเภทของเว็บเพิ่มขึ้น 3 ประเภท คือ เว็บไซต์ดาวน์โหลด (Download) เว็บไซต์หน่วยงานที่ไม่มุ่งหวังผลกำไร (Non-profit organizations) และเว็บไซต์เทคโนโลยี

1.4 ขั้นตอนและวิธีดำเนินการวิจัย

1. ศึกษาทฤษฎีการจัดประเภทของเว็บ บล็อกของเพจ โครงสร้างของเว็บเพจ และประเภทของเว็บที่มีอยู่ในปัจจุบันเพื่อใช้กำหนดประเภทของเว็บ
2. ศึกษางานวิจัยที่เกี่ยวข้อง
3. เลือกชุดข้อมูลที่ใช้ในการทดสอบและทำการเรียนรู้ (Training) การจัดประเภทของเว็บเพจ โดยใช้ผลการสืบค้นจาก กูเกิล (google) ซึ่งระบุค่าสำคัญสำหรับเว็บไซต์แต่ละประเภท
4. ทดสอบวิธีการที่น่าเสนอ
5. วิเคราะห์และเปรียบเทียบผลการทดลอง
6. สรุปผลการวิจัยและจัดทำรายงานวิทยานิพนธ์

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ซอฟต์แวร์ที่ช่วยในการจัดประเภทของเว็บโดยอาศัยการเชื่อมโยงในระดับเพจและบล็อกของเพจ
2. ผลของวิทยานิพนธ์สามารถใช้เป็นแนวทางในการศึกษาบล็อกของเพจในการจัดประเภทของเว็บ

1.6 ลำดับการจัดเรียงเนื้อหาในวิทยานิพนธ์

วิทยานิพนธ์นี้แบ่งเนื้อหาออกเป็น 5 บทดังต่อไปนี้ บทที่ 1 เป็นบทนำซึ่งกล่าวถึงความเป็นมาและความสำคัญของปัญหา รวมถึงวัตถุประสงค์ของการวิจัย บทที่ 2 กล่าวถึงทฤษฎีพื้นฐานและงานวิจัยที่เกี่ยวข้องในงานวิจัยนี้ บทที่ 3 กล่าวถึงรายละเอียดทั้งหมดของการประยุกต์ใช้บล็อกเพื่อจัดประเภทของเว็บ บทที่ 4 แสดงรายละเอียดของการทดลองและผลการทดลอง และบทที่ 5 เป็นข้อสรุปและข้อเสนอแนะจากการวิจัย

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

บทนี้จะแบ่งเนื้อหาออกเป็น 2 ส่วนด้วยกัน โดยเนื้อหาในส่วนแรกจะกล่าวถึงทฤษฎีและแนวคิดพื้นฐาน ซึ่งประกอบด้วย การแบ่งบล็อกและการจัดประเภทของเว็บ เนื้อหาในส่วนที่สองจะกล่าวถึงงานวิจัยที่เกี่ยวข้องกับงานวิจัยนี้

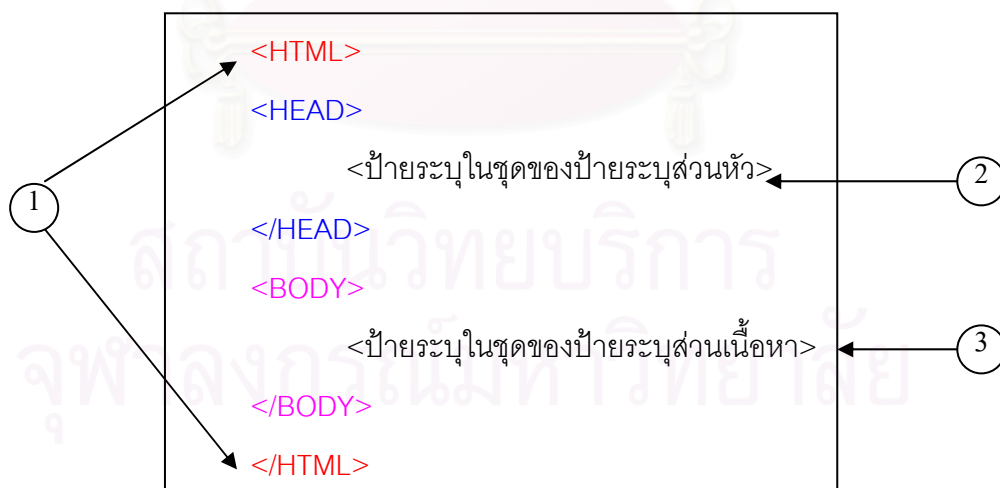
2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 เทคโนโลยีเกี่ยวกับเว็บ

ภาษาเอชทีเอ็มแอล (HyperText Markup Language : HTML) เป็นภาษาคอมพิวเตอร์รูปแบบหนึ่ง [1] ที่มีโครงสร้างการเขียนโดยอาศัยป้ายระบุ (Tag) ควบคุมการแสดงผลข้อความ รูปภาพ หรือวัตถุอื่น ๆ ผ่านโปรแกรมเบราว์เซอร์ แต่ละป้ายระบุอาจจะมีส่วนขยายที่เรียกว่า ลักษณะประจำ (Attribute) สำหรับระบุ หรือควบคุมการแสดงผลของเว็บ ได้ด้วยไฟล์ของโปรแกรมภาษาเอชทีเอ็มแอลเป็นไฟล์ข้อความ (Text file) ธรรมดาที่ใช้นามสกุลว่า .htm หรือ .html

2.1.1.1 โครงสร้างของภาษาเอชทีเอ็มแอล

โครงสร้างของภาษาเอชทีเอ็มแอลมีรูปแบบดังนี้



รูปที่ 2.1 โครงสร้างของภาษาเอชทีเอ็มแอล

1) ป้ายระบุเริ่มต้น

<HTML>.....</HTML>

ป้ายระบุ <HTML> เป็นป้ายระบุเริ่มต้นในการเขียนโปรแกรมและป้ายระบุ </HTML> เป็นการสิ้นสุดโปรแกรม ภาษาเอชทีเอ็มแอลป้ายระบุนี้จะไม่แสดงผลในโปรแกรมเว็บเบราว์เซอร์ แต่ต้องเขียนเพื่อให้เกิดความเป็นระบบของงาน และเพื่อจะให้ออกสารนี้เป็นเอกสารของภาษาเอชทีเอ็มแอล

โครงสร้างของภาษาเอชทีเอ็มแอลจะประกอบไปด้วยส่วนของป้ายระบุ 2 ส่วน คือ ส่วนหัว (Head) และส่วนที่เป็นเนื้อหา (Body) ส่วนหัวเป็นส่วนที่ใช้อธิบายเกี่ยวกับข้อมูลเฉพาะของเพจนั้น ๆ เช่น ชื่อเรื่องของเพจ (Title) ชื่อผู้จัดทำเว็บ (Author) คีย์เวิร์ดสำหรับการค้นหา (Keyword) ส่วนเนื้อหาเป็นส่วนที่เป็นเนื้อหาหลักของเพจ ซึ่งการแสดงผลจะต้องใช้ป้ายระบุจำนวนมาก ขึ้นอยู่กับลักษณะของข้อมูล เช่น ข้อความ รูปภาพ เสียง วิดีโอ หรือไฟล์ต่าง ๆ

2) ส่วนหัวของโปรแกรม

<HEAD>.....</HEAD>

ป้ายระบุ <HEAD> เป็นป้ายระบุที่ใช้กำหนดข้อความในส่วนที่เป็นชื่อเรื่องของไฟล์ภาษาเอชทีเอ็มแอลและภายในป้ายระบุ <HEAD>...</HEAD> จะมีป้ายระบุย่อยอีกป้ายระบุหนึ่งคือ <TITLE>.....</TITLE>

กำหนดข้อความในส่วนหัวของโปรแกรม

<TITLE>.....</TITLE >

ป้ายระบุ <TITLE> เป็นป้ายระบุที่แสดงชื่อของเอกสาร หรือชื่อเรื่องของไฟล์ภาษาเอชทีเอ็มแอลซึ่งข้อความภายในป้ายระบุจะปรากฏหรือแสดงผลในส่วนขอไตเติลบาร์ (Title Bar) ของโปรแกรมเว็บเบราว์เซอร์ แต่จะไม่แสดงในส่วนของการแสดงผลในโปรแกรมเว็บเบราว์เซอร์

3) ส่วนเนื้อหาของโปรแกรม

<BODY>.....</BODY>

ป้ายระบุนี้เป็นส่วนที่สำคัญในการแสดงผลในเว็บเบราว์เซอร์ ซึ่งจะประกอบไปด้วยตัวอักษร รูปภาพกราฟิกต่าง ๆ ส่วนเนื้อหาเอกสารเว็บแบ่งกลุ่มป้ายระบุได้ดังนี้ กลุ่มป้ายระบุเกี่ยวกับการจัดการพารากราฟ กลุ่มป้ายระบุจัดแต่ง/ควบคุมรูปแบบตัวอักษร กลุ่มป้ายระบุการทำเอกสารแบบรายการ (List) กลุ่มป้ายระบุเกี่ยวกับการทำลิงค์ กลุ่มป้ายระบุจัดการรูปภาพ กลุ่มป้ายระบุจัดการตาราง (Table) กลุ่มป้ายระบุควบคุมเฟรม และกลุ่มป้ายระบุอื่น ๆ

ป้ายระบุใช้ในการระบุรูปแบบป้ายระบุ หรือการลงรหัสป้ายระบุภาษาเอชทีเอ็มแอลภายในเครื่องหมายน้อยกว่า (<) และเครื่องหมายมากกว่า (>) โดยแบ่งได้ 2 ลักษณะ คือ ป้าย

ระบุแบบเดี่ยวที่ไม่ต้องมีป้ายระบุปิด เช่น <P>
 เป็นต้น และป้ายระบุแบบเปิด/ปิดที่ประกอบด้วยป้ายระบุเปิดและป้ายระบุปิด โดยป้ายระบุปิดจะมีเครื่องหมาย slash (/) นำหน้าป้ายระบุในป้ายระบุนั้น ๆ เช่น ... <BLINK>...</BLINK> เป็นต้น

ป้ายระบุที่ใช้ในภาษาภาษาเอชทีเอ็มแอลประกอบไปด้วยเครื่องหมายน้อยกว่า <ตามด้วยชื่อป้ายระบุและปิดท้ายด้วยเครื่องหมายมากกว่า>เป็นส่วนที่ทำหน้าที่ตกแต่งข้อความ เพื่อ การแสดงผลข้อมูล โดยทั่วไปป้ายระบุของภาษาเอชทีเอ็มแอลส่วนใหญ่จะอยู่เป็นคู่ มีเพียงบางป้ายระบุเท่านั้น ที่มีรูปแบบป้ายระบุอยู่เพียงตัวเดียว ในแต่ละป้ายระบุ จะมีป้ายระบุเปิดและปิดป้ายระบุปิดของแต่ละ ป้ายระบุจะมี รูปแบบเหมือนป้ายระบุเปิด เพียงแต่จะเพิ่ม / (Slash) นำหน้าป้ายระบุ ปิดให้ดู แตกต่าง เท่านั้น และในป้ายระบุเปิดบางป้ายระบุ อาจมีส่วนขยายอื่นผสมอยู่ด้วย ในการเขียน ด้วยตัวอักษร เล็กหรือใหญ่ ทั้งหมดหรือเขียนปนกันก็ได้

2.1.1.2 การเชื่อมโยง (Link)

เป็นส่วนของเว็บเพจที่เชื่อมโยงไปยังส่วนอื่น ๆ ของเว็บเพจในเว็บไซต์เดิมหรือเว็บไซต์ใหม่ การที่อินเทอร์เน็ตได้รับความนิยมอย่างกว้างขวางอยู่ทั่วโลกนั้น เป็นผลมาจากความสามารถในการเชื่อมโยงข้อมูล (Links) จากข้อมูลหนึ่งไปยังอีกข้อมูลหนึ่งได้อย่างรวดเร็ว โดยเชื่อมโยงข้อความได้ทั้งจากภายในแฟ้มเอกสารข้อมูลของตัวเอง และแฟ้มเอกสารข้อมูลภายนอกที่อยู่ต่างเว็บไซต์กัน

ข้อความที่ถูกกำหนดให้เชื่อมโยงกับฐานข้อมูลอื่น ๆ บนเว็บเบราว์เซอร์จะแสดงผลเป็นตัวอักษรที่มีสีแตกต่างจากตัวอักษรทั่วไป และอาจมีการขีดเส้นใต้ข้อความนั้นด้วย โดยทั่วไปตัวอักษรที่แสดงผลอยู่บนเว็บเบราว์เซอร์จะมีสีดาบนพื้นสีขาว (หรือสีเทา) แต่สำหรับข้อความที่ใช้เป็นตัวเชื่อมโยงข้อมูลนั้น จะมีตัวอักษรเป็นสีน้ำเงิน (หรือสีอื่นตามแต่ที่ผู้สร้างกำหนดขึ้นมา) เมื่อเลื่อนเมาส์ไปที่ข้อความซึ่งมีการเชื่อมโยง รูปแบบของตัวชี้จะเปลี่ยนจาก สัญลักษณ์ลูกศรไปเป็นรูปมือแทน และที่แถบที่แสดงสถานะด้านล่างจะแสดงถึง ตำแหน่งของจุดหมายปลายทางที่ข้อความเชื่อมโยงไปถึง

2.1.2.1 ป้ายระบุของการเชื่อมโยงแบบต่าง ๆ

ภาษาเอชทีเอ็มแอลแบ่งการเชื่อมโยงออกเป็น 2 ประเภท ดังนี้คือ การเชื่อมโยงภายในเว็บไซต์และการเชื่อมโยงนอกเว็บไซต์ การเชื่อมโยงทั้ง 2 ประเภทนี้ แบ่งออกได้เป็นอีก 2 ชนิด คือ การเชื่อมโยงข้อมูลภายในแฟ้มเอกสาร และการเชื่อมโยงข้อมูลต่างแฟ้มเอกสาร มีรูปแบบดังนี้

1) การเชื่อมโยงข้อมูลแฟ้มเอกสารต่างแฟ้มเอกสาร

ข้อความ

ตัวอย่างเช่น การบริการชุมชนและพัฒนาสังคม

2) การเชื่อมโยงข้อมูลภายในแฟ้มเอกสาร

ข้อความ

ข้อความ

ตัวอย่างเช่น ความรู้ทั่วไป

ความรู้ทั่วไป

3) การเชื่อมโยงนอกเว็บไซต์

ข้อความ

ตัวอย่างเช่น วิชาการ

4) การเชื่อมโยงแบบอีเมล

การติดต่อผ่านการส่งอีเมล

ข้อความ

ตัวอย่างเช่น

webmaster@chula.ac.th

5) การเชื่อมโยงแบบ Download

การนำข้อมูลต่าง ๆ เช่น ข้อมูลที่เป็นไฟล์ Word ไฟล์ Power Point ลงสู่เว็บเพจ เพื่อเผยแพร่

ข้อความ

ตัวอย่างเช่น ประโยชน์ของสมุนไพร

6) การเชื่อมโยงไฟล์โดยรูปภาพ

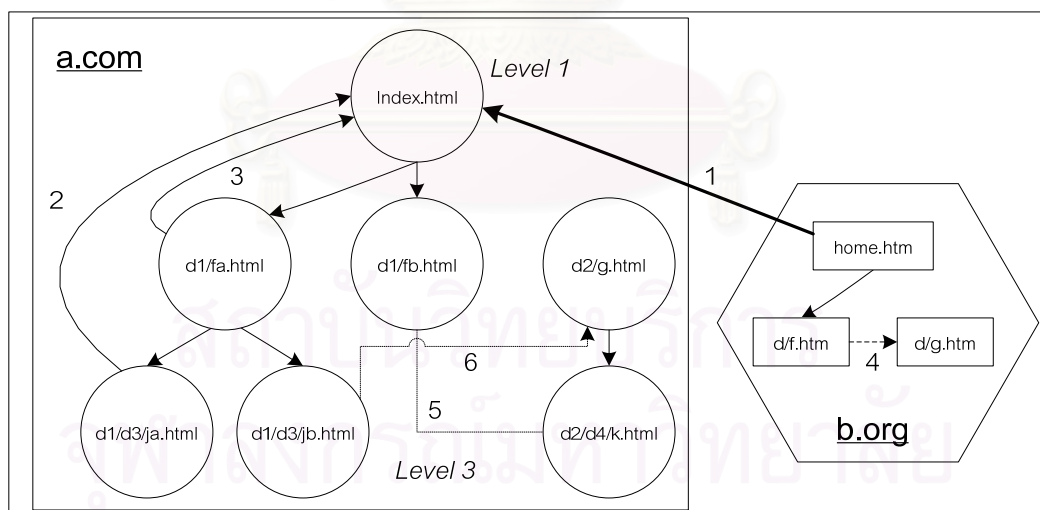
ในการเชื่อมโยงข้อมูลโดยรูปภาพทำให้น่าสนใจกว่าการเชื่อมโยงข้อมูลโดยตัวอักษร

ตัวอย่างเช่น

2.1.2.2 ลักษณะการเชื่อมโยง

ลักษณะการเชื่อมโยงของเว็บ มีทั้งการเชื่อมโยงภายในเว็บไซต์ การเชื่อมโยงออกไปภายนอกเว็บไซต์ ดังรายละเอียดดังนี้

- 1) การเชื่อมโยงภายนอก (External link) คือ การเชื่อมโยงระหว่างสองเพจในต่างแม่ข่าย ลอจิคัล จากรูปที่ 2.1 เป็นการเชื่อมโยง 1
 - 2) การเชื่อมโยงภายใน (Internal link) คือ การเชื่อมโยงระหว่างสองเพจในแม่ข่ายลอจิคัล (logical host) เดียวกัน
 - 3) การเชื่อมโยงขึ้น (Uplink) คือ การเชื่อมโยงภายในที่เชื่อมโยงจากเพจด้านล่างไปยังเพจด้านบน จากรูปคือ การเชื่อมโยง 2 และ 3
 - 4) การเชื่อมโยงลง (Downlink) คือ การเชื่อมโยงภายในที่เชื่อมโยงจากเพจด้านบนไปยังเพจด้านล่าง จากรูปคือ การเชื่อมโยงที่ไม่มีหมายเลขกำกับ
 - 5) ซайдลิงก์ (Sidelink) คือ การเชื่อมโยงภายในที่เชื่อมโยงระหว่างเพจในไต่แรกเทอร์รี่ เดียวกัน จากรูปคือ การเชื่อมโยง 4
 - 6) ครอสลิงก์ (Crosslink) คือ การเชื่อมโยงภายในที่ไม่ได้ทั้งเชื่อมโยงจากเพจด้านล่างไปยังเพจด้านบนและจากเพจด้านบนไปยังเพจทางด้านล่าง จากรูปคือ การเชื่อมโยง 5 และ 6
 - 7) ลีฟเพจ (Leaf page) คือ เพจที่ไม่มีการเชื่อมโยงเชื่อมโยงออกไปที่อื่น
- ลักษณะการเชื่อมโยงของสองไซต์ ทั้งการเชื่อมโยงภายในและการเชื่อมโยงภายนอกของไซต์แสดงดังรูปที่ 2.2 [1]



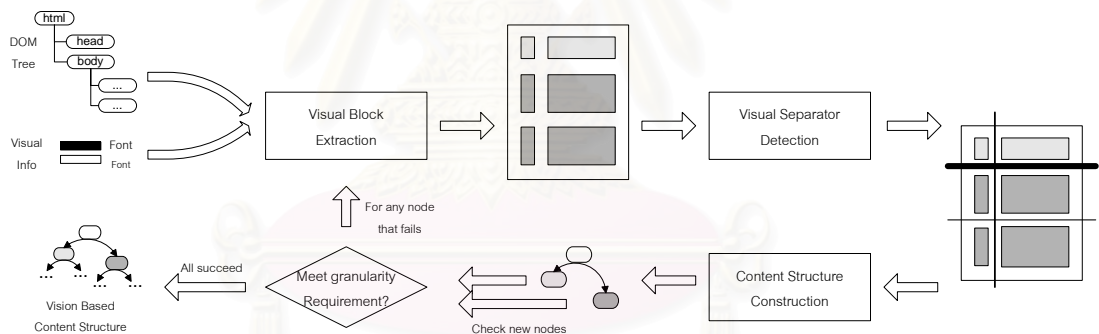
รูปที่ 2.2 การเชื่อมโยงภายในและการเชื่อมโยงภายนอกของสองเว็บไซต์

2.1.2 บล็อก (Block)

ในการแบ่งเพจออกเป็นบล็อกอาศัยหลักการของ VIPS (Vision-based Page Segmentation Algorithm) [3] ซึ่งเป็นการใช้ป้ายระบุในภาษาเอชทีเอ็มแอลในการแบ่งเพจเป็นบล็อก ดังตารางที่ 2.1 และตารางที่ 2.2

ขั้นตอนวิธี VIPS (VIPS Algorithm)

1. การหาบล็อกเสมือน (Visual Block Extraction)
เป็นขั้นตอนในการหาบล็อกที่เหมาะสม โดยอาศัยกฎการแบ่งเพจออกเป็นบล็อก [4]
2. การหาตัวแบ่งเสมือน (Visual Separator Detection)
เป็นขั้นตอนในการหาตัวแบ่ง (Separator) และตั้งค่าน้ำหนักให้ตัวแบ่ง
3. การสร้างโครงสร้างเนื้อหา (Content Structure Construction) เป็นการตั้งค่าน้ำหนัก (Weight) ให้แก่บล็อก



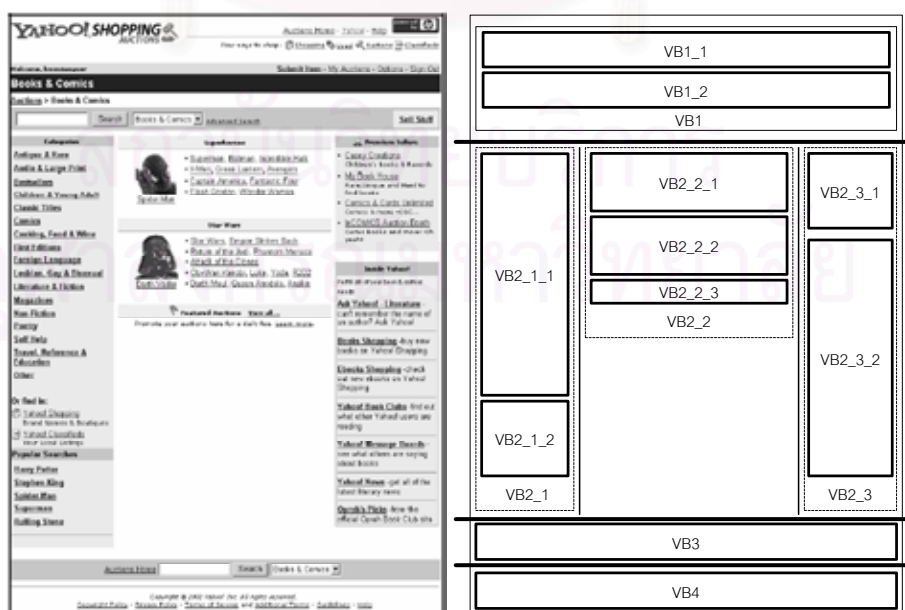
รูปที่ 2.3 อัลกอริทึม VIPS

ตารางที่ 2.1 ตัวอย่างกฎที่แตกต่างกันของป้ายระบุที่ต่างกัน

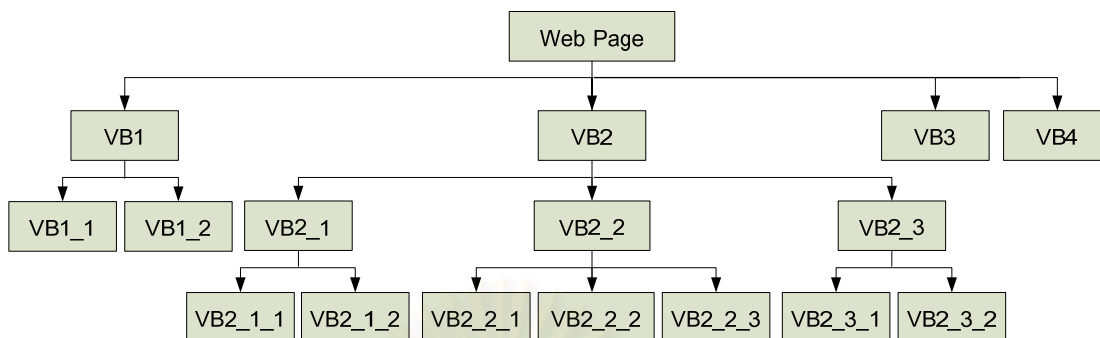
ป้ายระบุ	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13
ป้ายระบุตาราง <TABLE>	/	/	/					/		/			/
ป้ายระบุแถว <TR>	/	/	/				/	/		/			/
ป้ายระบุคอลัมน์ <TD>	/	/	/	/					/	/	/		/
ป้ายระบุย่อหน้า <P>	/	/	/	/	/	/	/		/	/		/	

ตารางที่ 2.2 กฎที่ใช้สำหรับป้ายระบุที่ต่างกันในอัลกอริทึม VIPS

กฎ	คำอธิบาย
R1	ถ้าไม่มีป้ายระบุอยู่ภายในจะไม่พิจารณา
R2	ถ้ามีป้ายระบุอยู่ภายในเพียงหนึ่งและไม่ได้เป็นข้อความ ทำการแบ่งป้ายระบุนี้ ออกเป็นโนนดย่อย
R3	ถ้าเป็นโนนดรากของส่วนกลุ่มย่อย ทำการแบ่งป้ายระบุนี้ออก
R4	ถ้าทุกป้ายระบุที่อยู่ภายในเป็นข้อความ ไม่แบ่ง
R5	ถ้าป้ายระบุที่อยู่ภายในมีป้ายระบุขึ้นบรรทัดใหม่ ทำการแบ่ง
R6	ถ้าป้ายระบุที่อยู่ภายในมีป้ายระบุเส้นแนวนอน <HR> ทำการแบ่ง
R7	ถ้าผลรวมของขนาดทั้งหมดของป้ายระบุที่อยู่ภายในใหญ่กว่า ทำการแบ่ง
R8	ถ้าสีพื้นหลังแตกต่างจากสีพื้นหลังของป้ายระบุที่อยู่ภายใน ทำการแบ่งป้ายระบุนี้ ออก
R9	ถ้ามีอย่างน้อยหนึ่งป้ายระบุที่อยู่ภายในเป็นข้อความและมีขนาดเล็กกว่า ไม่แบ่ง
R10	ถ้ามีป้ายระบุอยู่ภายในที่มีขนาดใหญ่ที่สุดมีขนาดเล็กกว่า ไม่ต้องแบ่ง
R11	ถ้าป้ายระบุก่อนหน้าไม่แบ่ง จะไม่แบ่ง
R12	แบ่งออก
R13	ไม่แบ่ง



รูปที่ 2.4 ตัวอย่างของ Vision-based content structure ในการแบ่งบล็อก



รูปที่ 2.5 ลำดับชั้นใน Vision-based content structure

จากรูปที่ 2.4 เป็นเว็บเพจตัวอย่าง และรูปที่ 2.5 เป็นตัวอย่างของโครงสร้างของ Vision-based content โดยแสดงถึงโครงสร้างเนื้อหา (Content structure) [5]

2.1.3 การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล (Data mining) เป็นดึงหรือขุด (mining) ความรู้ (knowledge) จากข้อมูล (data) ที่มีจำนวนมาก [6] การวิเคราะห์ข้อมูลที่เกิดขึ้นไว้ในแหล่งข้อมูลต่าง ๆ เพื่อค้นหาและตรวจสอบรูปแบบหรือความสัมพันธ์ในชุดข้อมูลดังกล่าว การประยุกต์ใช้ข้อมูลดังกล่าวถึงนี้มีได้หลายแนวทาง เช่น การวิเคราะห์แนวโน้มการเปลี่ยนแปลงของข้อมูล หรือ การค้นหาความสัมพันธ์ที่ซ่อนอยู่ภายในกลุ่มของข้อมูล การทำเหมืองข้อมูลเป็นการนำความรู้ที่เรียนรู้มาจากข้อมูลที่มีอยู่เพื่อประโยชน์ในการทำนายข้อมูลใหม่ที่จะเกิดขึ้นในอนาคต การค้นหารูปแบบที่น่าสนใจจากกลุ่มข้อมูล รูปแบบนี้มักจะเป็นความสัมพันธ์ หรือลักษณะที่เชื่อมโยงกันของข้อมูล

การวิเคราะห์ข้อมูลที่เกิดขึ้นไว้ในแหล่งข้อมูลต่าง ๆ เพื่อค้นหาและตรวจสอบรูปแบบหรือความสัมพันธ์ในชุดข้อมูลดังกล่าว ส่วนการประยุกต์ทางธุรกิจของการทำเหมืองข้อมูลต้องอาศัยการผสมผสานแนวคิดทางสารสนเทศและทางสถิติเพื่อค้นพบความสัมพันธ์ที่ไม่คาดคิดมาก่อน

รูปแบบของการทำเหมืองข้อมูล

2.1.3.1 การจัดประเภท (Classification)

การจัดประเภท (Classification) เป็นกระบวนการสร้างตัวแบบ (Model) จัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้ โดยพิจารณาจากข้อมูลที่มีอยู่ เป็นการหากฎเพื่อระบุประเภทของวัตถุจากคุณสมบัติของวัตถุนั้น ซึ่งกระบวนการจัดประเภทมี 3 ขั้นตอน คือ

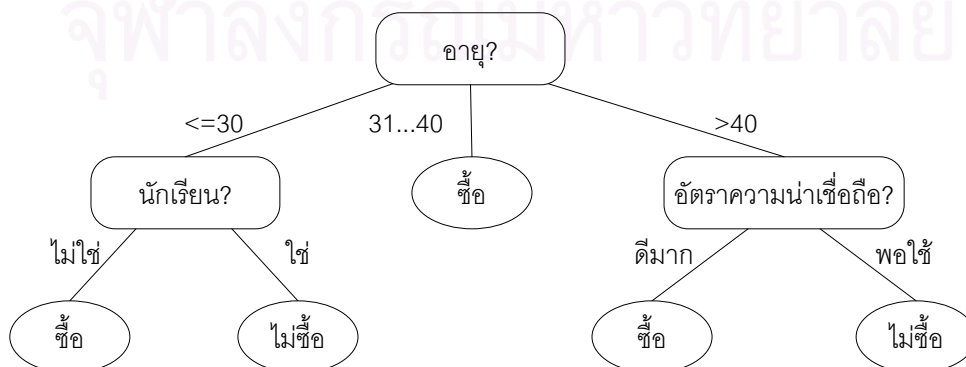
1. ขั้นตอนการสร้างตัวแบบโดยอาศัยการเรียนรู้จากข้อมูลที่ได้กำหนดคลาสไว้เรียบร้อยแล้ว (Training Data)
2. ขั้นตอนการประมาณความถูกต้องโดยอาศัยข้อมูลทดสอบ (Testing Data) ซึ่งคลาสที่แท้จริงของข้อมูลทดสอบนี้จะถูกนำมาเปรียบเทียบกับคลาสที่หามาได้จากตัวแบบเพื่อทดสอบความถูกต้อง
3. ขั้นตอนการกำหนดคลาสให้สำหรับข้อมูลที่ไม่เคยเห็นมาก่อน (Unseen Data) โดยจะกำหนดประเภทให้

2.1.3.2 การทำนาย (Prediction)

เป็นการทำนายค่าที่ต้องการจากข้อมูลที่มีอยู่ ตัวอย่างเช่น หายอดขายของเดือนถัดไปจากข้อมูลที่มีอยู่ หรือทำนายโรคจากอาการของคนไข้ในอดีต เป็นต้น

2.1.4 ต้นไม้ในการตัดสินใจ (Decision Tree)

ต้นไม้ในการตัดสินใจ เป็นผังงานลักษณะโครงสร้างต้นไม้ ซึ่งโหนดภายใน (internal node) คือคุณลักษณะที่ทดสอบ และลีฟโหนด (leaf node) คือประเภทของสิ่งที่เราสนใจ (class) [6] ตัวอย่างต้นไม้ในการตัดสินใจดังในรูปที่ 2.6 เป็นต้นไม้ตัดสินใจเรื่อง ลูกค้าจะซื้อคอมพิวเตอร์ (buy_computer) หรือไม่ โดยโหนดภายในเป็นคุณลักษณะที่ใช้ทดสอบ สำหรับแต่ละลีฟโหนดแทนประเภทของสิ่งที่เราสนใจคือ ซื้อคอมพิวเตอร์หรือไม่ (buy_computer=ใช่ หรือ buy_computer = ไม่ใช่)



รูปที่ 2.6 ต้นไม้ในการตัดสินใจ

ต้นไม้ในการตัดสินใจสามารถแสดงในในรูปแบบของกฎ (rule) โดยแต่ละกฎเกิดจากเส้นทาง (path) จากราก (root) ถึงลีฟโหนด แต่ละค่าของคุณลักษณะจะอยู่ในส่วนของ “ถ้า” (IF) แต่ละค่าของลีฟโหนดจะอยู่ในส่วนของ “แล้ว” (THEN) กฎ “ถ้า-แล้ว” (IF-THEN) อาจจะเข้าใจได้ง่ายสำหรับมนุษย์ โดยเฉพาะอย่างยิ่งถ้าต้นไม้มีขนาดที่ใหญ่มาก จากรูปที่ 2.6 สามารถแทนได้ด้วย

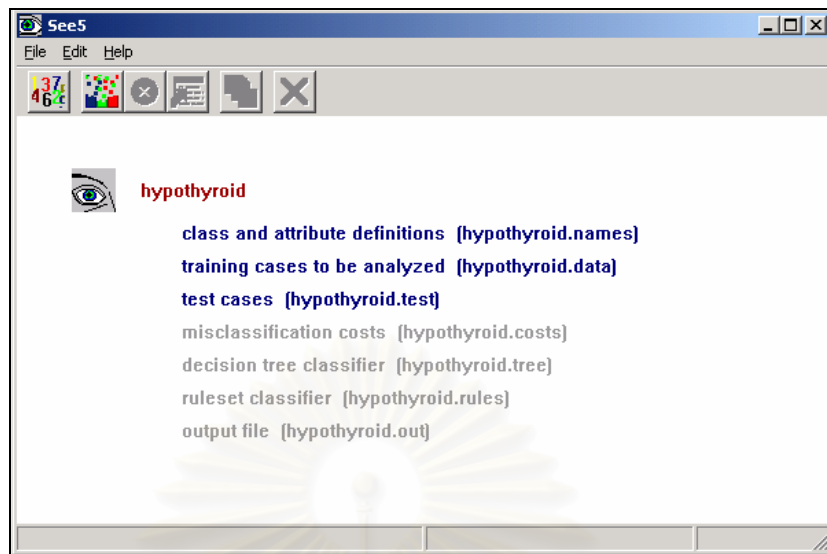
ถ้า อายุ = " ≤ 30 " และ นักเรียน = "ไม่ใช่"	แล้ว ชื่อคอมพิวเตอร์ = "ไม่ใช่"
ถ้า อายุ = " ≤ 30 " และ นักเรียน = "ใช่"	แล้ว ชื่อคอมพิวเตอร์ = "ใช่"
ถ้า อายุ = "31...40"	แล้ว ชื่อคอมพิวเตอร์ = "ใช่"
ถ้า อายุ = " > 40 " และ อัตราความน่าเชื่อถือ = "ดีมาก"	แล้ว ชื่อคอมพิวเตอร์ = "ไม่ใช่"
ถ้า อายุ = " > 40 " และ อัตราความน่าเชื่อถือ = "พอใช้"	แล้ว ชื่อคอมพิวเตอร์ = "ใช่"

2.1.5 วิธีการเรียนรู้ของเครื่องแบบ See5

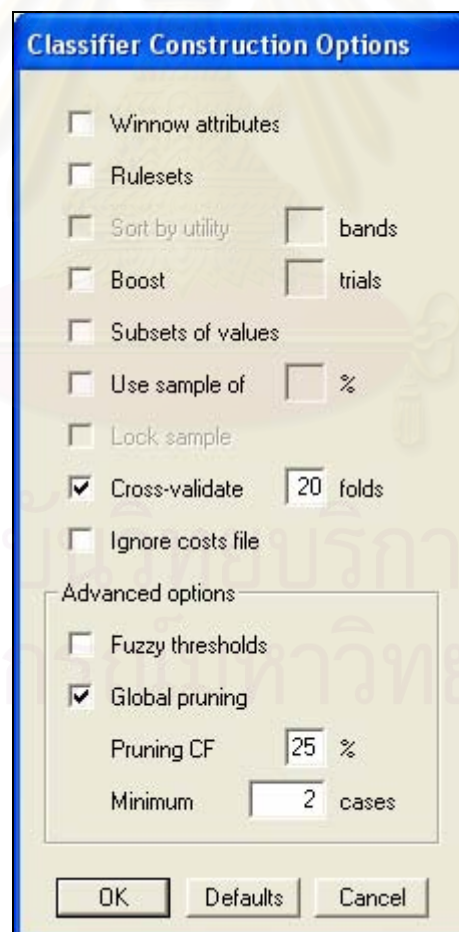
See5 [9] เป็นเครื่องมือหนึ่งที่ใช้ในการศึกษางานด้านการทำเหมืองข้อมูล โดยสร้างตัวจัดประเภท (classifier) ที่ใช้ในการทำนายในรูปแบบของชุดของกฎ (sets of rules) ประกอบไปด้วย

names	ระบุคุณลักษณะ
data	ข้อมูลที่ใช้ในการสร้างตัวจัดประเภท
test	ข้อมูลทดสอบตัวจัดประเภท
case	ข้อมูล
tree	ต้นไม้ในการตัดสินใจ
rules	ชุดของกฎ
out	รายงานแสดงผลลัพธ์ที่เกิดขึ้น
set	การตั้งค่าล่าสุดในการสร้างตัวจัดประเภท

สถาบันวิจัยปัญญาประดิษฐ์
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.7 หน้าจอวิธีการเรียนรู้ของเครื่องแบบ See5



รูปที่ 2.8 ตัวเลือกต่างๆ ในวิธีการเรียนรู้ของเครื่องแบบ See5

2.2 งานวิจัยที่เกี่ยวข้อง

2.2.1 การเชื่อมต่อโซน่า: การตรวจตราฟังก์ชันของไซต์โดยใช้รูปแบบโครงสร้าง (The Connectivity Sonar: Detecting Site Functionality by Structural Patterns)

เป็นงานวิจัยที่ทำการจัดประเภทของเว็บไซต์โดยไม่พิจารณาถึงเนื้อหาของเว็บ[1] ได้จัดประเภทของเว็บไซต์เป็น 8 ประเภทตามฟังก์ชันการทำงาน (Functionality categories) คือ เว็บไซต์ของบริษัท (Corporate sites) เว็บไซต์เนื้อหาและสื่อกลาง (Content & media site) โปรแกรมค้นหา (Search engine) ไดเรกทอรีและลำดับชั้น (Web hierarchies & directories) พอร์ทัล (Portals) อิเล็กทรอนิกส์สโตร์ (E-stores) บริการแม่ข่ายเสมือน (Virtual hosting services) และมหาวิทยาลัย (Universities)

ในการจัดประเภทของเว็บนั้น ใช้ประเภทของการเชื่อมโยง 3 แบบคือ การกระจายของเพจ (Distribution of pages) การเชื่อมโยงภายนอก (External link) การเชื่อมโยงภายใน (Internal link)

ตารางที่ 2.3 ลักษณะของการเชื่อมโยงของเว็บ คือ การกระจายของเพจ การเชื่อมโยงภายนอก การเชื่อมโยงภายใน

ตารางที่ 2.3 ลักษณะการเชื่อมโยง

แบบชนิดของการเชื่อมโยง	คุณลักษณะ (Feature)	คำอธิบาย
การกระจายของเพจ (Distribution of pages)	1. ระดับเฉลี่ยระดับของเพจ (Average level of page)	ระดับของเพจได้จากจำนวนสแลชของยูอาร์แอล
	2. อัตราร้อยละของเพจในระดับที่นิยมมากที่สุด (Percentage of pages in most populated level)	จำนวนเพจในระดับที่มีการเชื่อมโยงเข้ามามากที่สุด
	3. อัตราการขยายของเพจในระดับบนสุด (Top level page expansion ratio)	จำนวนเพจทั้งหมดในระดับที่สองหารด้วยจำนวนเพจทั้งหมดในระดับที่หนึ่ง
การเชื่อมโยงภายนอก (External Link)	4. จำนวนการเชื่อมโยงเข้ามาต่อเพจ (Inlinks per page)	จำนวนการเชื่อมโยงเข้ามาของเพจ
	5. จำนวนการเชื่อมโยงออกไปต่อเพจ (Outlinks per page)	จำนวนการเชื่อมโยงออกไปของเพจ
	6. จำนวนการเชื่อมโยงออกไปต่อจำนวนการเชื่อมโยงเข้ามา (Outlinks	จำนวนการเชื่อมโยงออกไปต่อจำนวนการเชื่อมโยงเข้ามา

	per inlink)	
	7. ส่วนของการเชื่อมโยงเข้ามาในระดับบนสุด (Top level inlink portion)	จำนวนการเชื่อมโยงที่เชื่อมโยงเข้ามาที่เพจในระดับบนสุด
	8. การเชื่อมโยงภายนอกต่อลิฟเพจ (Outlinks per leaf page)	การเชื่อมโยงภายนอกของลิฟเพจ
	9. ค่าเฉลี่ยของการเชื่อมโยงเข้ามา (Average level of inlinks)	ค่าเฉลี่ยของการเชื่อมโยงเข้ามา
	10. ค่าเฉลี่ยของการเชื่อมโยงออกไป (Average level of outlinks)	ค่าเฉลี่ยของการเชื่อมโยงออกไป
	11. ผลต่างของคุณลักษณะที่ 10 และคุณลักษณะที่ 9 (The difference between F10 and F9)	ผลต่างของคุณลักษณะที่ 10 และคุณลักษณะที่ 9
	12. อัตราร้อยละของการเชื่อมโยงเข้ามาในระดับที่นิยมมากที่สุด (Percentage of inlinks to most popular level)	จำนวนการเชื่อมโยงภายในในระดับที่มีการเชื่อมโยงภายในมากที่สุด
	13. อัตราร้อยละของการเชื่อมโยงออกไปจากระดับที่ออกมากที่สุด (Percentage of outlinks from most emitting level)	จำนวนการเชื่อมโยงภายนอกในระดับที่มีการเชื่อมโยงภายนอกมากที่สุด
การเชื่อมโยงภายใน (Internal Link)	14. ครอสลิงก์ต่อเพจ (Crosslinks per page)	จำนวนการเชื่อมโยงภายในเว็บไซต์
	15. สภาพทนทาน (Robustness)	อัตราส่วนระหว่างจำนวนการเชื่อมโยงนำวิเศษ (navigational links) และจำนวนการเชื่อมโยงซีมาติก (semantic links) ในระดับที่ 2-4
	16. จำนวนการเชื่อมโยงเข้ามาในระดับบนสุดต่อเพจของเว็บไซต์ (Top level internal inlinks per page on site)	จำนวนการเชื่อมโยงของการเชื่อมโยงภายในที่มีการเชื่อมโยงไปที่เพจในระดับบนสุด

สิ่งที่จะนำมาใช้คือแนวคิดการจัดประเภทของเว็บโดยไม่คำนึงถึงเนื้อหา แต่ใช้ลักษณะของการเชื่อมโยงในการจัดประเภทของเว็บ โดยที่ผู้วิจัยจะได้นำมาศึกษาหาลักษณะของการเชื่อมโยงอื่น ๆ เพิ่มเติมเพื่อใช้ในการจัดประเภทของเว็บ

2.2.2 การวิเคราะห์การเชื่อมโยงในระดับบล็อก (Block-level Link Analysis)

เป็นงานวิจัยที่ทำการวิเคราะห์การเชื่อมโยงของเว็บเพจ ซึ่งพบว่ามีผลทางด้านการค้นหาข้อมูลของเว็บ ได้เสนอแนวคิดในการแบ่งเว็บเพจออกเป็นบล็อก แล้วหาความสัมพันธ์ระหว่างบล็อกและเว็บเพจ คือ ความสัมพันธ์จากเพจไปยังบล็อก (Page-to-Block) และความสัมพันธ์จากบล็อกไปยังเพจ (Block-to-Page) โดยสร้างกราฟของเพจ (Page graph) และกราฟของบล็อก (Block graph) รวมถึงได้นำเสนอสองขั้นตอนวิธีใหม่คือ Block level PageRank และ Block Level HITS โดยอาศัยแนวคิดจากขั้นตอนวิธีเดิมคือ PageRank และ HITS และแนวคิดการแบ่งเว็บเพจออกเป็นบล็อก [7]

สิ่งที่จะนำมาใช้คือแนวคิดของการหาความสัมพันธ์ระหว่างบล็อกและเว็บเพจ คือ ความสัมพันธ์จากเพจไปยังบล็อกและความสัมพันธ์จากบล็อกไปยังเพจ

2.2.3 ขั้นตอนวิธีการแบ่งส่วนเพจด้วยวิชันเบส (VIPS: a Vision-based Page Segmentation Algorithm)

เป็นงานวิจัยที่ทำการวิเคราะห์ส่วนต่าง ๆ บนเว็บเพจ โดยกำหนดค่าความสำคัญของส่วนต่าง ๆ ของเว็บเพจ ทำการเสนอกฎในการแบ่งเว็บเพจออกเป็นบล็อกโดยอาศัยขั้นตอนวิธี VIPS : vision-based page segmentation algorithm [5] รายละเอียดในหัวข้อที่ 2.1.2

สิ่งที่จะนำมาใช้คือการแบ่งเพจเป็นบล็อกโดยอาศัยขั้นตอนวิธี VIPS : vision-based page segmentation algorithm

2.2.4 ตัวแบบการเรียนรู้ค่าความสำคัญสำหรับบล็อกของเว็บเพจโดยอาศัยการวิเคราะห์การจัดวางหน้าและเนื้อหา (Learning Important Models for Web Page Blocks based on Layout and Content Analysis)

เป็นงานวิจัยที่ทำการวิเคราะห์คุณลักษณะต่าง ๆ ของบล็อกในเว็บเพจ [8] ทำให้เห็นความแตกต่างของแต่ละบล็อกในเว็บเพจ โดยวิเคราะห์ทั้งคุณลักษณะที่เป็นเนื้อหา(content features) และคุณลักษณะที่เป็นระยะห่าง(spatial features) ซึ่งอาศัยขั้นตอนวิธี VIPS : vision-based page segmentation algorithm ในการแบ่งเพจเป็น

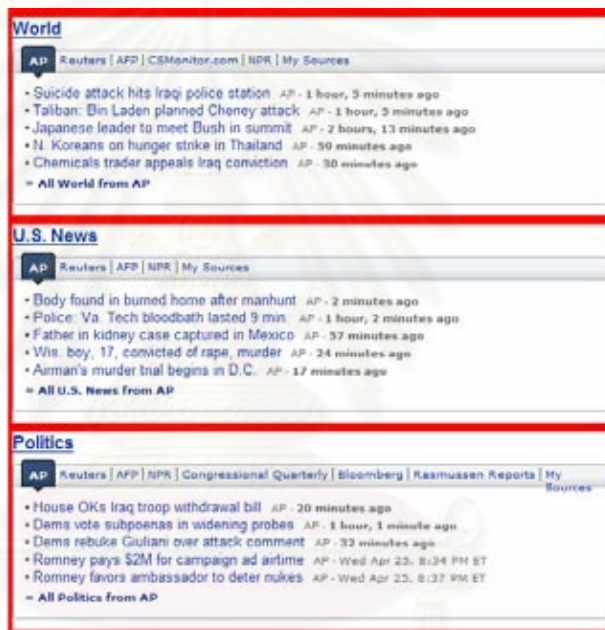
สิ่งที่จะนำมาใช้คือการเลือกคุณลักษณะระดับบล็อกทั้งคุณลักษณะที่เป็นเนื้อหาและคุณลักษณะที่เป็นระยะห่าง มาใช้เป็นคุณลักษณะในระดับบล็อกเพื่อใช้ในการจัดประเภทของเว็บ

บทที่ 3

แนวทางการวิจัย

3.1 การประยุกต์ใช้บล็อกเพื่อจัดประเภทของเว็บ

ในเว็บเพจหนึ่ง ๆ ของเว็บไซต์สามารถแบ่งออกเป็นบล็อกได้หลายบล็อก โดยที่สำหรับเว็บไซต์แต่ละประเภทอาจมีคุณลักษณะระดับบล็อกในแต่ละเพจที่แตกต่างกัน ซึ่งในแต่ละบล็อกของเพจอาจมีความสำคัญที่ต่างกัน จึงเป็นแนวคิดที่จะใช้คุณลักษณะระดับบล็อกในเพจมาวิเคราะห์เพื่อช่วยในการจัดประเภทของเว็บไซต์



รูปที่ 3.1 ตัวอย่างของเพจซึ่งสามารถแบ่งออกเป็นบล็อกได้ (<http://news.yahoo.com>)

ในการวิจัยนี้อาศัยอัลกอริทึม VIPS ในการแบ่งเพจเป็นบล็อกโดยคำนวณค่าความสำคัญของบล็อกจากสูตร

$$f_p(b) = \text{ขนาดของบล็อกในเพจ } p / \text{ระยะระหว่างจุดกึ่งกลางของบล็อกและหน้าจอ} \quad (1)$$

โดย

p คือ เพจ

b คือ บล็อกในเพจ

$f_p(b)$ คือ ฟังก์ชันสำหรับแต่ละบล็อก ซึ่งค่าความสำคัญของบล็อกยิ่งมากหมายถึง บล็อกนั้นมีความสำคัญมาก

การหาค่าน้ำหนักสำหรับแต่ละเพจ

$$W_p = XZ \quad \text{---- (2)}$$

โดย

p คือ เพจ

X คือ ค่าความสำคัญของบล็อก ซึ่งหาค่าได้จากสูตร

$$X_{ij} = \begin{cases} f_{P_i}(b_j) & \text{ถ้า } b_j \in P_i \\ 0 & \text{กรณีอื่นๆ} \end{cases} \quad \text{---- (3)}$$

i คือ เพจใดๆ

j คือ บล็อกใดๆ

Z คือ ค่าความสัมพันธ์จากบล็อกไปยังเพจ ซึ่งหาค่าได้จากสูตร

$$Z_{ij} = \begin{cases} 1/s_i & \text{ถ้ามีการเชื่อมโยงจากบล็อก i ถึงเพจ j} \\ 0 & \text{กรณีอื่นๆ} \end{cases} \quad \text{---- (4)}$$

s_i คือ จำนวนของเพจที่บล็อก i มีการเชื่อมโยงไป

3.2 การเลือกเว็บไซต์ในแต่ละประเภท

ในการเลือกเว็บไซต์มาใช้ในการเรียนรู้และทดสอบสำหรับเว็บไซต์ในแต่ละประเภท จะเลือกจากผลการสืบค้นจากกูเกิล ซึ่งระบุค่าสำคัญสำหรับเว็บไซต์แต่ละประเภทดังรายละเอียดในตารางที่ 3.1

ตารางที่ 3.1 ค่าสำคัญในการเลือกเว็บไซต์มาใช้ในการเรียนรู้และทดสอบสำหรับเว็บไซต์ในแต่ละประเภท

ลำดับ	เว็บไซต์	ค่าสำคัญ
1	เนื้อหาและสื่อกลาง (Content & media site)	Newspaper
2	บริษัท (Corporate sites)	Co., Ltd.
3	ไดเรกทอรีและลำดับชั้น (Web hierarchies & directories)	Web Directory
4	อิเล็กทรอนิกส์สโตร์ (E-stores)	E-store

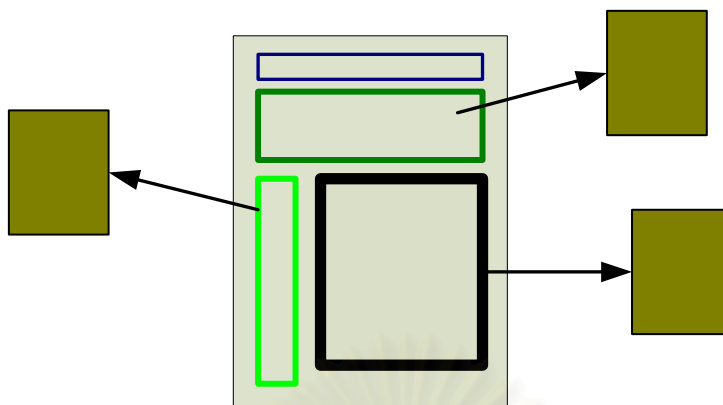
5	พอร์ทัล (Portals)	Web Portal
6	โปรแกรมค้นหา (Search engine)	Search Engines
7	มหาวิทยาลัย (Universities)	University
8	บริการแม่ข่ายเสมือน (Virtual hosting services)	Virtual Hosting
9	ดาวน์โหลด (Download)	Download
10	หน่วยงานที่ไม่มุ่งหวังผลกำไร (Non-profit organizations)	Non-profit
11	เทคโนโลยี (Technology)	Technology

3.3 การเลือกเว็บเพจของแต่ละเว็บไซต์

ในการเลือกเพจของเว็บไซต์ จะเลือก 10 เพจแรกของเว็บไซต์นั้น ๆ โดยที่เพจแรกของเว็บไซต์จะทำการดึงการเชื่อมโยงทั้งหมดมาแล้วเลือกเพจ 10 อันดับแรกของการเชื่อมโยงของเพจนั้น หากเพจแรกมีจำนวนการเชื่อมโยงไม่ถึง 10 เพจ ก็จะทำการดึงการเชื่อมโยงในเพจที่สองเพื่อจะให้ได้จำนวนเพจ 10 เพจ

3.4 การเลือกบล็อกรายในแต่ละเว็บเพจ

ทำการแบ่งเว็บเพจออกเป็นบล็อกโดยอาศัยอัลกอริทึม VIPS แล้วเลือกบล็อกที่มีค่าความสำคัญสูงสุดสามอันดับแรกในแต่ละเว็บเพจของ โดยใช้สูตรในหัวข้อที่ 3.1 ดังสมการที่ 1



รูปที่ 3.2 แสดงการเลือก 3 บล็อกในแต่ละเพจ

3.5 คุณลักษณะระดับเพจ

ในการวิจัยเลือกคุณลักษณะระดับเพจมา 4 คุณลักษณะ

- ระดับของเพจ เป็นการนับจำนวนเครื่องหมายสแลช (slash) ของยูอาร์แอล
- จำนวนการเชื่อมโยงภายในต่อเพจ โดยตรวจสอบว่าการเชื่อมโยงของเพจนั้นและการเชื่อมโยงที่ได้เป็นแม่ข่ายเดียวกัน
- จำนวนการเชื่อมโยงภายนอกต่อเพจ โดยตรวจสอบว่าการเชื่อมโยงของเพจนั้นกับการเชื่อมโยงที่ได้เป็นคนละแม่ข่ายกัน
- อัตราส่วนระหว่างการเชื่อมโยงภายในและการเชื่อมโยงภายนอก

3.6 คุณลักษณะระดับบล็อก

ในการวิจัยเลือกคุณลักษณะระดับบล็อกมา 13 คุณลักษณะ โดยเป็นคุณลักษณะที่ทำให้บล็อกแต่ละบล็อกมีความแตกต่างกัน [8] ค่าต่าง ๆ เป็นค่าที่ได้มาจากเอกซ์เอ็มแอลไฟล์ โดยเป็นค่าเฉลี่ยจาก 3 บล็อกที่มีค่าความสำคัญมากที่สุด

- ค่าเฉลี่ยจำนวนการเชื่อมโยงต่อบล็อก
- ค่าเฉลี่ยจำนวนภาพที่อยู่ในบล็อกต่อบล็อก
- ค่าเฉลี่ยขนาดความยาวของตัวอักษรต่อบล็อก
- ค่าเฉลี่ยขนาดความยาวของตัวอักษรที่ใช้เชื่อมโยงต่อบล็อก
- ค่าเฉลี่ยจำนวนโหนดลูกต่อบล็อก
- ค่าเฉลี่ยขนาดของตัวอักษรต่อบล็อก

- ค่าเฉลี่ยน้ำหนักของตัวอักษรต่อบล็อก
- ค่าเฉลี่ยตำแหน่งด้านซ้ายของบล็อกต่อบล็อก
- ค่าเฉลี่ยตำแหน่งด้านบนของบล็อกต่อบล็อก
- ค่าเฉลี่ยความกว้างของบล็อกต่อบล็อก
- ค่าเฉลี่ยความสูงของบล็อกต่อบล็อก
- ค่าเฉลี่ยจุดศูนย์กลางในแนวนอนของบล็อกต่อบล็อก
- ค่าเฉลี่ยจุดศูนย์กลางในแนวตั้งของบล็อกต่อบล็อก

3.7 ขั้นตอนการคำนวณ

สำหรับแต่ละเว็บไซต์จะมีการคำนวณคุณลักษณะในระดับเพจและบล็อกดังรายละเอียดต่อไปนี้

3.7.1 อ่านการเชื่อมโยงของเว็บไซต์เพื่อเลือกเพจ

อ่านเอชทีเอ็มแอลไฟล์ของเพจแรกของเว็บไซต์แล้วหาการเชื่อมโยงในเพจนี้ทั้งหมดเพื่อให้ได้ 10 หน้าแรกของเว็บไซต์นี้ ถ้าหากเพจแรกมีการเชื่อมโยงไม่ถึง 10 เพจ ก็จะทำการศึกษาการเชื่อมโยงในหน้าถัดไป เพื่อให้ได้เพจทั้งหมดอย่างน้อย 10 เพจ

3.7.2 เลือก 10 เพจสำหรับเว็บไซต์นั้น

เลือกเพจ 10 เพจแรกที่ได้จากขั้นตอนที่ 3.7.1

3.7.3 คำนวณคุณลักษณะในระดับเพจในแต่ละเพจ

ในแต่ละเพจจะทำการอ่านเอชทีเอ็มแอลไฟล์ของเพจนั้นเพื่อหาการเชื่อมโยงในเพจนี้ทั้งหมด โดยมีการเช็คว่าเป็นการเชื่อมโยงภายในหรือการเชื่อมโยงภายนอก โดยเช็คว่าเป็นแม่ข่าย (Host) เดียวกันหรือไม่ โดยที่

- ถ้าการเชื่อมโยงของเพจนั้นกับการเชื่อมโยงที่ได้เป็นแม่ข่ายเดียวกันจะถือว่าเป็นการเชื่อมโยงภายใน เช่น

<http://www.stanford.edu>

<http://www.stanford.edu/home/academics>

- ถ้าการเชื่อมโยงของเพจนั้นกับการเชื่อมโยงที่ได้เป็นคนละแม่ข่ายกันจะถือว่าเป็นการเชื่อมโยงภายนอก เช่น

<http://www.stanford.edu>

<http://gostanford.cstv.com>

3.7.4 แบ่งเพจเป็นบล็อกโดยอาศัยอัลกอริทึม VIPS

ในแต่ละเว็บเพจจะทำการแบ่งเป็นบล็อกโดยอาศัยอัลกอริทึม VIPS โดยจะผ่านโปรแกรมซึ่งกำหนดค่ายูอาร์แอลของเพจนั้น ซึ่งเมื่อผ่านโปรแกรมแล้วจะได้เอกซ์เอ็มแอลไฟล์ ซึ่งแสดงถึงบล็อกทั้งหมดในเพจนั้น หน้าจอโปรแกรม VIPS แสดงที่ภาคผนวก ง และตัวอย่างเอกซ์เอ็มแอลแสดงที่ภาคผนวก ข

3.7.5 คำนวณค่าในแต่ละบล็อกเพื่อเลือก 3 บล็อก

เป็นการเลือกบล็อก 3 บล็อกแรกที่มีค่าความสำคัญมากที่สุด โดยมีขั้นตอนดังนี้คือ

- อ่านไฟล์เอกซ์เอ็มแอลที่ได้จากโปรแกรม VIPS แล้วเลือกเฉพาะลิฟโหนด
- นำลิฟโหนดที่ได้ ซึ่งแทนบล็อกที่ไม่สามารถแบ่งย่อยลงไปได้อีกมาทำการคำนวณหาค่าความสำคัญ จากสูตรในหัวข้อ 3.1
- เลือกบล็อก 3 บล็อกแรกที่มีค่าความสำคัญมากที่สุด

3.7.6 คำนวณค่าเฉลี่ยของคุณลักษณะระดับบล็อก 3 บล็อกที่เลือก

เป็นการหาค่าเฉลี่ยของคุณลักษณะในระดับบล็อกของทั้ง 3 บล็อกใน 1 เพจ

- ค่าเฉลี่ยจำนวนการเชื่อมโยงต่อบล็อก
- ค่าเฉลี่ยจำนวนภาพที่อยู่ในบล็อกต่อบล็อก
- ค่าเฉลี่ยขนาดความยาวของตัวอักษรต่อบล็อก
- ค่าเฉลี่ยขนาดความยาวของตัวอักษรที่ใช้เชื่อมโยงต่อบล็อก
- ค่าเฉลี่ยจำนวนโหนดลูกต่อบล็อก
- ค่าเฉลี่ยขนาดของตัวอักษรต่อบล็อก
- ค่าเฉลี่ยน้ำหนักของตัวอักษรต่อบล็อก
- ค่าเฉลี่ยตำแหน่งด้านซ้ายของบล็อกต่อบล็อก
- ค่าเฉลี่ยตำแหน่งด้านบนของบล็อกต่อบล็อก

- ค่าเฉลี่ยความกว้างของบล็อกต่อบล็อก
- ค่าเฉลี่ยความสูงของบล็อกต่อบล็อก
- ค่าเฉลี่ยจุดศูนย์กลางในแนวนอนของบล็อกต่อบล็อก
- ค่าเฉลี่ยจุดศูนย์กลางในแนวตั้งของบล็อกต่อบล็อก

3.7.7 คำนวณค่าเฉลี่ยของคุณลักษณะบล็อกสำหรับเว็บไซต์

เป็นการหาค่าเฉลี่ยของคุณลักษณะในระดับบล็อกของทั้ง 10 เพจ

- ค่าเฉลี่ยจำนวนการเชื่อมโยงต่อบล็อก
- ค่าเฉลี่ยจำนวนภาพที่อยู่ในบล็อกต่อบล็อก
- ค่าเฉลี่ยขนาดความยาวของตัวอักษรต่อบล็อก
- ค่าเฉลี่ยขนาดความยาวของตัวอักษรที่ใช้เชื่อมโยงต่อบล็อก
- ค่าเฉลี่ยจำนวนโหนดลูกต่อบล็อก
- ค่าเฉลี่ยขนาดของตัวอักษรต่อบล็อก
- ค่าเฉลี่ยน้ำหนักของตัวอักษรต่อบล็อก
- ค่าเฉลี่ยตำแหน่งด้านซ้ายของบล็อกต่อบล็อก
- ค่าเฉลี่ยตำแหน่งด้านบนของบล็อกต่อบล็อก
- ค่าเฉลี่ยความกว้างของบล็อกต่อบล็อก
- ค่าเฉลี่ยความสูงของบล็อกต่อบล็อก
- ค่าเฉลี่ยจุดศูนย์กลางในแนวนอนของบล็อกต่อบล็อก
- ค่าเฉลี่ยจุดศูนย์กลางในแนวตั้งของบล็อกต่อบล็อก

3.7.8 คำนวณค่าเฉลี่ยของคุณลักษณะเพจสำหรับเว็บไซต์

เป็นการหาค่าเฉลี่ยของคุณลักษณะในระดับเพจของทั้ง 10 เพจ

- ระดับของเพจ เป็นการนับจำนวนเครื่องหมายสแลช (slash) ของยูอาร์แอล
- จำนวนการเชื่อมโยงภายในต่อเพจ โดยตรวจสอบว่าการเชื่อมโยงของเพจนั้นและการเชื่อมโยงที่ได้เป็นแม่ข่ายเดียวกัน
- จำนวนการเชื่อมโยงภายนอกต่อเพจ โดยตรวจสอบว่าการเชื่อมโยงของเพจนั้นกับการเชื่อมโยงที่ได้เป็นคนละแม่ข่ายกัน
- อัตราส่วนระหว่างการเชื่อมโยงภายในและการเชื่อมโยงภายนอก

บทที่ 4

การทดลองและผลการทดลอง

4.1 ชุดข้อมูลที่ใช้ในการทดลอง

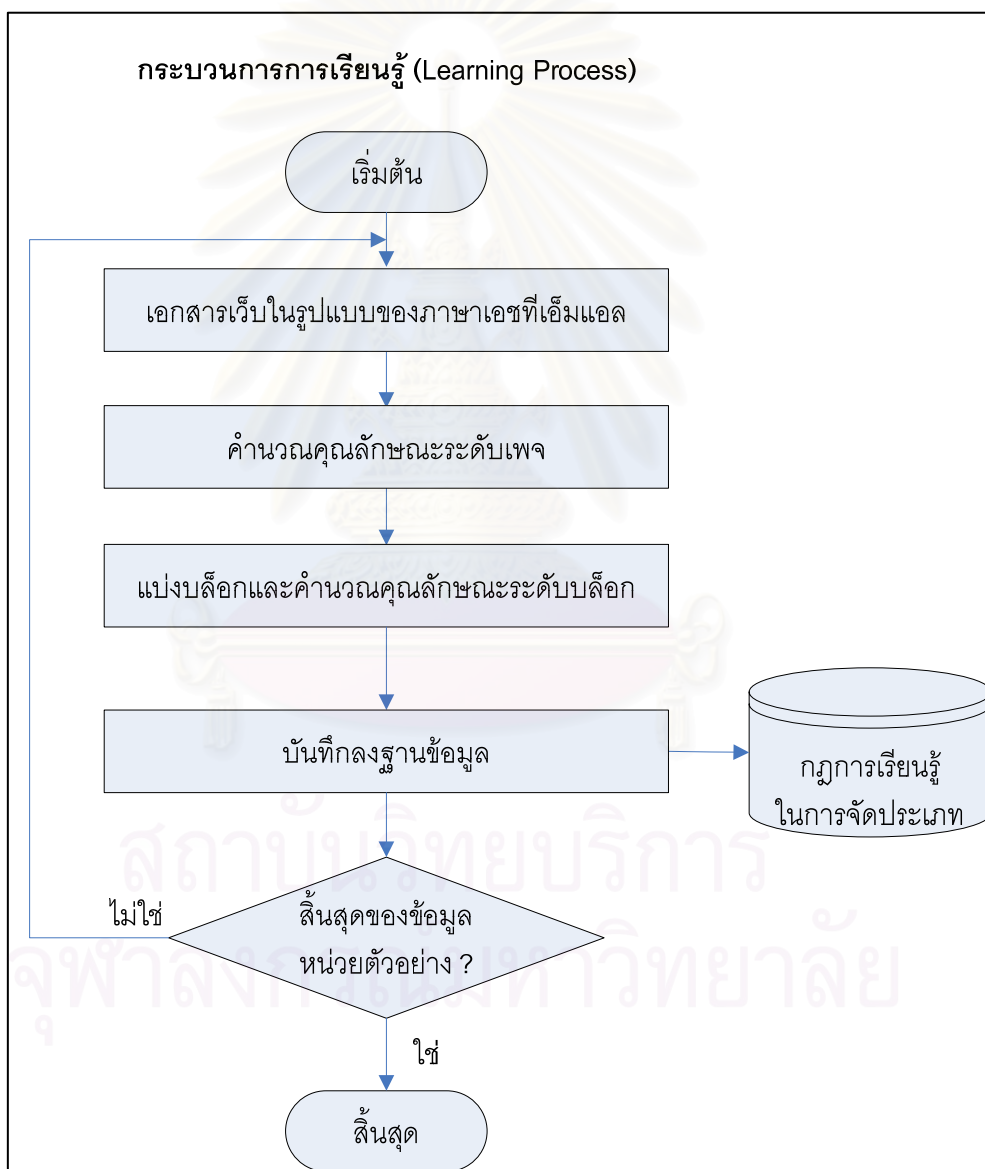
ในการทดลองเพื่อเปรียบเทียบความถูกต้องของการจัดประเภทของเว็บโดยอาศัยคุณลักษณะระดับเพจและบล็อก จำนวนเว็บไซต์ที่ใช้ในการเรียนรู้และทดสอบทั้งหมด 265 เว็บไซต์ โดยในแต่ละเว็บไซต์เลือกเพจมา 10 เพจ ดังนั้นจำนวนเพจทั้งหมดที่ใช้ในการทดลองคือ 2,650 เพจ ในการทดลองมีประเภทของเว็บไซต์มีทั้งหมด 11 ประเภท จำนวนเว็บไซต์ของแต่ละประเภทที่ใช้ในการทดลองดังรายละเอียดในตารางที่ 4.1 และจำนวนเว็บไซต์ของแต่ละประเภทที่ใช้ในการทดลองดังรายละเอียดในตารางที่ 4.2

ตารางที่ 4.1 จำนวนชุดข้อมูลที่ใช้

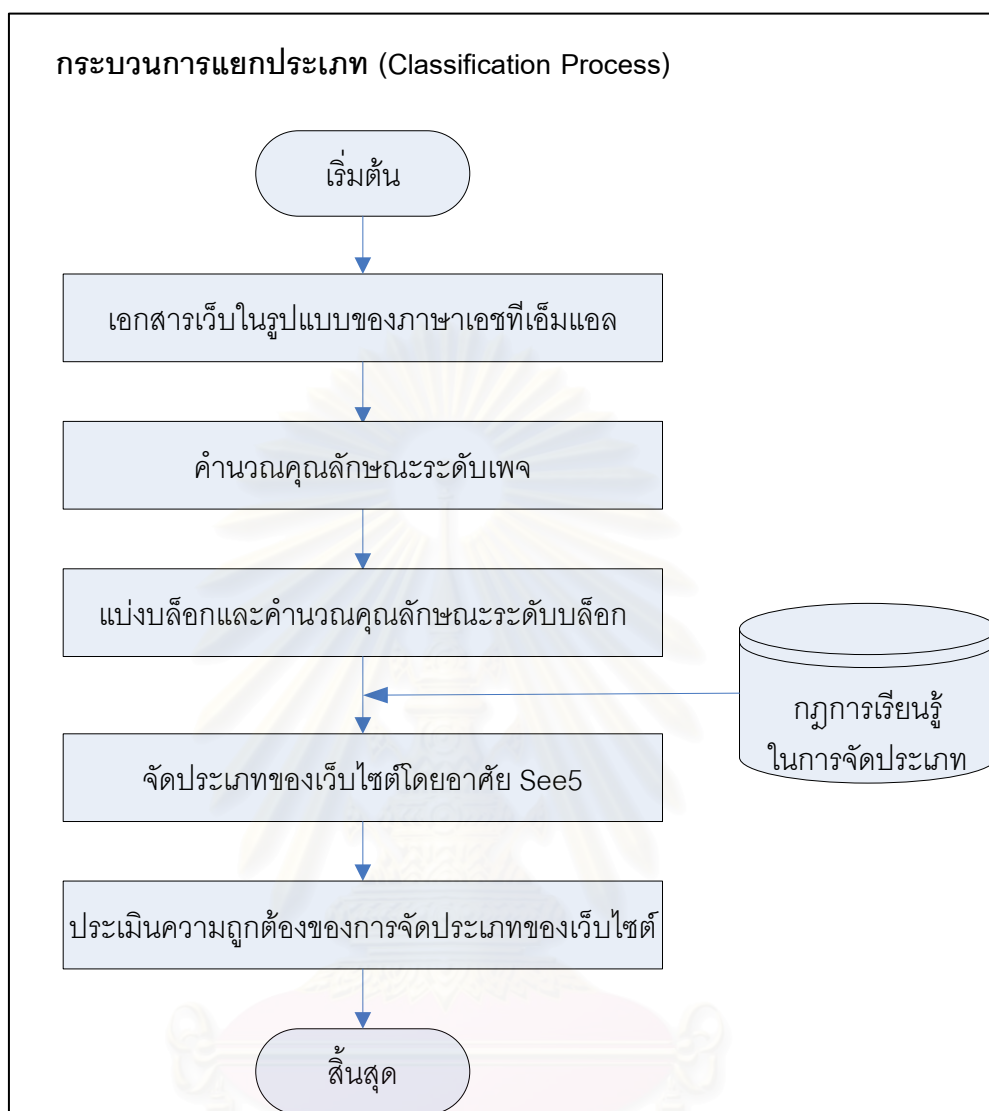
ลำดับ	เว็บไซต์	จำนวนเว็บไซต์	จำนวนเพจ
1	เนื้อหาและสื่อกลาง (Content & media site)	52	520
2	บริษัท (Corporate sites)	38	380
3	ไดเรกทอรีและลำดับชั้น (Web hierarchies & directories)	25	250
4	อิเล็กทรอนิกส์สโตร์ (E-stores)	20	200
5	พอร์ทัล (Portals)	26	260
6	โปรแกรมค้นหา (Search engine)	11	110
7	มหาวิทยาลัย (Universities)	26	260
8	บริการแม่ข่ายเสมือน (Virtual hosting services)	28	280
9	ดาวน์โหลด (Download)	13	130
10	หน่วยงานที่ไม่มุ่งหวังผลกำไร (Non-profit organizations)	13	130
11	เทคโนโลยี (Technology)	13	130
รวม		265	2,650

4.2 วิธีการทดลอง

การทดลองโดยอ่านข้อมูลจากเอชทีเอ็มแอลไฟล์ของเว็บไซต์มาและคำนวณคุณลักษณะในระดับเพจ หลังจากนั้นทำการแบ่งเพจเป็นบล็อกโดยอาศัยอัลกอริทึม VIPS แล้วคำนวณคุณลักษณะในระดับบล็อก ของแต่ละเว็บไซต์เพื่อนำมาเป็นข้อมูลในการเรียนรู้ในการจัดประเภทของเว็บไซต์ ขั้นตอนของกระบวนการการเรียนรู้ดังแสดงในรูปที่ 4.1 และขั้นตอนของกระบวนการแยกประเภทดังแสดงในรูปที่ 4.2



รูปที่ 4.1 กระบวนการเรียนรู้



รูปที่ 4.2 กระบวนการแยกประเภท

4.3 ผลการทดลอง

การประเมินประสิทธิภาพของวิธีการที่นำเสนอในงานวิจัยนี้ จะทำการเปรียบเทียบผลลัพธ์ความถูกต้องของการจัดประเภทเว็บไซต์อาศัยวิธีการเรียนรู้จากเครื่องแบบ See5 ที่ให้ค่าความผิดพลาดของแต่ละกลุ่มข้อมูล(โดยกลุ่มข้อมูลในที่นี้หมายถึงวิธีการเรียนรู้จากเครื่องแบบ See5 โดยใช้เทคนิค 20-fold cross-validate ในการแบ่งกลุ่มการทดลองออกเป็น 20 กลุ่ม คือกลุ่มที่ 0 - 19) ซึ่งในการทดลองนี้ผู้วิจัยได้แบ่งการทดลองออกเป็น 3 กลุ่ม ประกอบไปด้วย

- การใช้คุณลักษณะระดับเพจในการจัดประเภท

- การใช้คุณลักษณะระดับบล็อกในการจัดประเภท
- การใช้คุณลักษณะระดับเพจและบล็อกในการจัดประเภท

โดยมีรายละเอียดของผลการทดลองในแต่ละกลุ่มเป็นไปดังหัวข้อ 4.3.1 – 4.3.3

4.3.1 การใช้คุณลักษณะระดับเพจในการจัดประเภท

จากการทดลอง เมื่อนำคุณลักษณะระดับเพจมาช่วยในการจัดประเภทของเว็บไซต์ ทำให้ค่าความผิดพลาดโดยเฉลี่ย(Mean) อยู่ที่ 62.5% ดังแสดงในตารางที่ 4.2 และจำนวนความถูกต้องของผลการจัดประเภทเว็บไซต์แสดงดังตารางที่ 4.3

โดยคุณลักษณะในระดับเพจที่ใช้คือ

- ระดับของเพจ เป็นการนับจำนวนเครื่องหมายสแลช (/) ของยูอาร์แอล
- จำนวนการเชื่อมโยงภายในต่อเพจ โดยตรวจสอบว่าการเชื่อมโยงของเพจนั้นและการเชื่อมโยงที่ได้เป็นแม่ข่ายเดียวกัน
- จำนวนการเชื่อมโยงภายนอกต่อเพจ โดยตรวจสอบว่าการเชื่อมโยงของเพจนั้นกับการเชื่อมโยงที่ได้เป็นคนละแม่ข่ายกัน
- อัตราส่วนระหว่างการเชื่อมโยงภายในและการเชื่อมโยงภายนอก

ตารางที่ 4.2 ค่าความผิดพลาดตามกลุ่มข้อมูลโดยใช้คุณลักษณะระดับเพจ

กลุ่ม	กฎ	
	จำนวน	ความผิดพลาด
0	30	90.9%
1	27	45.5%
2	24	54.5%
3	32	63.6%
4	31	54.5%
5	31	72.7%
6	26	72.7%
7	25	45.5%
8	25	75.0%
9	30	58.3%
10	27	41.7%

11	28	75.0%
12	26	75.0%
13	30	50.0%
14	25	58.3%
15	30	58.3%
16	25	66.7%
17	26	58.3%
18	27	83.3%
19	28	50.0%
Mean	27.6	62.5%
SE	0.5	3.0%

ตารางที่ 4.3 ผลการจัดประเภทเว็บไซต์ประเภทต่าง ๆ โดยใช้คุณลักษณะระดับเพจ

ประเภทเว็บไซต์	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
C1. เนื้อหาและสื่อกลาง	37	2	1		1		2	2	2		2
C2. บริษัท		21	3	3	2			6			
C3. ไดเรกทอรีและลำดับชั้น	3	3	3		2		4	6	1		
C4. อีเล็กทรอนิกส์ไตร์	1	6	1			1	1	6	1		
C5. พอร์ทัล	3	4	4	1	4		4	3			
C6. โปรแกรมค้นหา	1	1	1		1	2	1	1			
C7. มหาวิทยาลัย	3		4	3	2		7	2	1		1
C8. บริการแม่ข่ายเสมือน	2	5	2	3	2	1	2	7	1		
C9. ดาวโหลด	1	1	2	1					4		1
C10. หน่วยงานไม่หวังผลกำไร	1	2		2	2			2	1		
C11. เทคโนโลยี	5		2	1							2

4.3.2 การใช้คุณลักษณะระดับบล็อกในการจัดประเภท

จากการทดลอง เมื่อนำคุณลักษณะระดับบล็อกมาช่วยในการจัดประเภทของเว็บไซต์ ทำให้ค่าความผิดพลาดโดยเฉลี่ย(Mean) อยู่ที่ 75.9% ดังแสดงในตารางที่ 4.4 และจำนวนความถูกต้องของผลการจัดประเภทเว็บไซต์แสดงดังตารางที่ 4.5

โดยคุณลักษณะในระดับบล็อกที่ใช้คือ

- ค่าเฉลี่ยจำนวนการเชื่อมโยงต่อบล็อก
- ค่าเฉลี่ยจำนวนภาพที่อยู่ในบล็อกต่อบล็อก
- ค่าเฉลี่ยขนาดความยาวของตัวอักษรต่อบล็อก
- ค่าเฉลี่ยขนาดความยาวของตัวอักษรที่ใช้เชื่อมโยงต่อบล็อก
- ค่าเฉลี่ยจำนวนโหนดลูกต่อบล็อก
- ค่าเฉลี่ยขนาดของตัวอักษรต่อบล็อก
- ค่าเฉลี่ยน้ำหนักของตัวอักษรต่อบล็อก
- ค่าเฉลี่ยตำแหน่งด้านซ้ายของบล็อกต่อบล็อก
- ค่าเฉลี่ยตำแหน่งด้านบนของบล็อกต่อบล็อก
- ค่าเฉลี่ยความกว้างของบล็อกต่อบล็อก
- ค่าเฉลี่ยความสูงของบล็อกต่อบล็อก
- ค่าเฉลี่ยจุดศูนย์กลางในแนวนอนของบล็อกต่อบล็อก
- ค่าเฉลี่ยจุดศูนย์กลางในแนวตั้งของบล็อกต่อบล็อก

ตารางที่ 4.4 ค่าความผิดพลาดตามกลุ่มข้อมูลโดยใช้คุณลักษณะระดับบล็อก

กลุ่ม	กฎ	
	จำนวน	ความผิดพลาด
0	21	72.7%
1	26	90.9%
2	28	63.6%
3	34	63.6%
4	31	81.8%
5	26	90.9%
6	26	63.6%
7	28	81.8%
8	38	83.3%

4.3.3 การใช้คุณลักษณะระดับเพจและบล็อกในการจัดประเภท

จากการทดลองในหัวข้อ 4.3.1 และหัวข้อ 4.3.2 เมื่อนำคุณลักษณะระดับทั้งเพจและบล็อกมารวมกันเพื่อช่วยในการจัดประเภทของเว็บไซต์ ทำให้ค่าความผิดพลาดของข้อมูลทั้งหมดอยู่ที่ 67.8% แสดงให้เห็นว่า เมื่อนำคุณลักษณะในระดับบล็อกเข้ามาช่วยในการจัดประเภทของเว็บไซต์ สามารถทำให้ความถูกต้องในการจัดประเภทเว็บไซต์ดีขึ้นกว่าการใช้คุณลักษณะในระดับบล็อกเพียงอย่างเดียว แต่ความถูกต้องในการจัดประเภทเว็บไซต์น้อยกว่าการใช้คุณลักษณะในระดับเพจเพียงอย่างเดียว

ตารางที่ 4.6 ค่าความผิดพลาดตามกลุ่มข้อมูลโดยใช้คุณลักษณะระดับเพจและบล็อก

กลุ่ม	กฎ	
	จำนวน	ความผิดพลาด
0	31	63.6%
1	32	63.6%
2	32	63.6%
3	34	72.8%
4	28	81.8%
5	29	81.8%
6	28	63.6%
7	30	81.8%
8	29	50.0%
9	29	66.7%
10	29	75.0%
11	27	75.0%
12	30	41.7%
13	28	50.0%
14	31	83.3%
15	28	75.0%
16	28	58.3%
17	27	58.3%
18	24	83.3%
19	27	66.7%
Mean	29.1	67.8%
SE	0.5	2.7%

จากตารางที่ 4.6 แสดงค่าความผิดพลาดของทุกกลุ่มข้อมูล ซึ่งมีค่าเฉลี่ยความผิดพลาดของทุกกลุ่ม(Mean) เท่ากับ 67.8% และสำหรับตัวอย่างการคำนวณ รายละเอียดของแต่ละกลุ่ม ซึ่งในที่นี้ได้ยกตัวอย่างการคำนวณสำหรับกลุ่มที่ 0 จะแสดงอยู่ในภาคผนวก ค โดยใช้ข้อมูลดังนี้

- ใช้ข้อมูลในการทดลอง 11 เว็บไซต์
- ใช้กฎในการเรียนรู้จำนวน 31 กฎ
- ทำนายประเภทเว็บไซต์ได้ถูกต้อง 7 เว็บไซต์

ค่าความผิดพลาดสำหรับกลุ่มที่ 0 คือ $(7 \times 100) / 11 = 63.6 \%$

ตารางที่ 4.7 ผลการจัดประเภทเว็บไซต์ประเภทต่าง ๆ โดยใช้คุณลักษณะระดับเพจและบล็อก

ประเภทเว็บไซต์	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
C1. เนื้อหาและสื่อกลาง	29		2	2	2			2	4		8
C2. บริษัท		20		5	2	2	2	3	1		
C3. ไดรากทอรี่และลำดับชั้น	3		5	1	5	1	2	3	1		1
C4. อิเล็กทรอนิกส์สตรี	1	9		3		1	1	2			
C5. พอร์ทัล		5	4	3	2	2	4	1	1	1	
C6. โปรแกรมค้นหา		1	1	1	1		1	1	1		
C7. มหาวิทยาลัย	5	1	1	2	1	1	8	3	1		
C8. บริการแม่ข่ายเสมือน	2	4	2	6			3	6	1	1	
C9. ดาวเน็ต	2	1	1		1			3			2
C10. หน่วยงานไม่หวังผลกำไร	1	3		2			2		2		
C11. เทคโนโลยี	6		1					1			2

ทั้งนี้ถ้าพิจารณาคุณลักษณะระดับเพจและบล็อกจะพบว่า ค่าความถูกต้องสามารถแสดงเป็นเปอร์เซ็นต์เปรียบเทียบค่าความผิดพลาดของแต่ละประเภทเว็บไซต์ดังตารางที่ 4.8

ตารางที่ 4.8 เปอร์เซ็นต์เปรียบเทียบค่าความผิดพลาดของแต่ละประเภทของเว็บไซต์ของคุณลักษณะระดับเพจ บล็อก เพจและบล็อก

ลำดับ	ประเภทเว็บไซต์	คุณลักษณะระดับเพจ			คุณลักษณะระดับบล็อก			คุณลักษณะระดับเพจและบล็อก		
		จำนวนที่จัดประเภทได้ถูกต้อง	จำนวนตัวอย่างทั้งหมด	เปอร์เซ็นต์ความถูกต้อง	จำนวนที่จัดประเภทได้ถูกต้อง	จำนวนตัวอย่างทั้งหมด	เปอร์เซ็นต์ความถูกต้อง	จำนวนที่จัดประเภทได้ถูกต้อง	จำนวนตัวอย่างทั้งหมด	เปอร์เซ็นต์ความถูกต้อง
1	เนื้อหาและสื่อกลาง (Content & media site)	37	52	71	26	52	50	29	52	56
2	บริษัท (Corporate sites)	21	38	55	11	38	29	20	38	53
3	ไดเรกทอรีและลำดับชั้น (Web hierarchies & directories)	3	25	12	5	25	20	5	25	20
4	อิเล็กทรอนิกส์สโตร์ (E-stores)	0	20	0	2	20	10	3	20	15
5	พอร์ทัล (Portals)	4	26	15	0	26	0	2	26	8
6	โปรแกรมค้นหา (Search engine)	2	11	18	2	11	18	0	11	0
7	มหาวิทยาลัย (Universities)	7	26	27	4	26	15	8	26	31
8	บริการแม่ข่ายเสมือน (Virtual hosting services)	7	28	25	5	28	18	6	28	21
9	ดาวน์โหลด (Download)	4	13	31	0	13	0	0	13	0
10	หน่วยงานที่ไม่มุ่งหวังผลกำไร (Non-profit organizations)	0	13	0	0	13	0	0	13	0
11	เทคโนโลยี (Technology)	2	13	15	1	13	8	2	13	15

หมายเหตุ จำนวนเว็บไซต์ใน 8 ประเภทแรกใช้ตามงานวิจัยของ Einat Amitay และคณะ [1] และ 3 ประเภทสุดท้ายที่เพิ่มขึ้นใช้จำนวน 10 เว็บไซต์ โดยทุกเว็บไซต์ได้เพิ่มจำนวนเว็บไซต์อีกประเภทละ 3 เว็บไซต์

เว็บไซต์ที่ให้ค่าความถูกต้องในคุณลักษณะระดับเพจในการจัดประเภทสูงสุด 3 อันดับแรก คือ เว็บไซต์เนื้อหาและสื่อกลางให้ค่าความถูกต้อง 71% เว็บไซต์บริษัทให้ค่าความถูกต้อง 55% และเว็บไซต์ดาวโหลดให้ค่าความถูกต้อง 31%

เว็บไซต์ที่ให้ค่าความถูกต้องในคุณลักษณะระดับบล็อกในการจัดประเภทสูงสุด 3 อันดับแรก คือ เว็บไซต์เนื้อหาและสื่อกลางให้ค่าความถูกต้อง 50% เว็บไซต์บริษัทให้ค่าความถูกต้อง 29% และเว็บไซต์เดสก์ทอปและลำดับชั้นให้ค่าความถูกต้อง 20%

เว็บไซต์ที่ให้ค่าความถูกต้องในคุณลักษณะระดับเพจและบล็อกในการจัดประเภทสูงสุด 3 อันดับแรก คือ เว็บไซต์เนื้อหาและสื่อกลางให้ค่าความถูกต้อง 56% เว็บไซต์บริษัทให้ค่าความถูกต้อง 53% และเว็บไซต์มหาวิทยาลัยให้ค่าความถูกต้อง 31%

สำหรับเว็บไซต์ที่ให้ค่าความถูกต้องน้อย ผู้วิจัยมีความเห็นว่า อาจมีสาเหตุมาจากการเลือกหน่วยตัวอย่างที่ได้มาจากผลการสืบค้นกูเกิล ทำให้ได้บางเว็บไซต์ไม่เป็นประเภทของเว็บไซต์นั้นจริง ๆ และเพจที่ใช้ในการทดลองใช้เพียง 10 เพจแรกของแต่ละเว็บไซต์ รวมถึงโครงสร้างของเว็บไซต์ เช่น เว็บไซต์โปรแกรมค้นหา ซึ่งบล็อกที่มีความสำคัญไม่มีการเชื่อมโยงทำให้บล็อกนั้นไม่ถูกเลือก ในทางกลับกัน บล็อกที่ไม่มีความสำคัญแต่มีการเชื่อมโยงถูกเลือกมาแทน จากการทดลอง มีค่าความผิดพลาดของเว็บไซต์ทั้ง 11 ประเภทอยู่ที่ 67.8%

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

จากการทำการวิจัยและทดลองการจัดประเภทเว็บไซต์โดยใช้คุณลักษณะระดับเพจ บล็อก เพจและบล็อก ผู้วิจัยได้สรุปผลการทดลองและมีข้อเสนอแนะต่าง ๆ ดังนี้

5.1 สรุปผลการวิจัย

แนวคิดในการทำวิจัยเรื่องการจัดประเภทของเว็บไซต์เกิดขึ้นมาจากงานวิจัยของ Einat Amitay และคณะ [1] ซึ่งได้ทำการจัดประเภทของเว็บไซต์ไว้ 8 ประเภท โดยผู้วิจัยได้ประยุกต์ใช้งานวิจัยดังกล่าวโดยได้เพิ่มประเภทของเว็บไซต์อีก 3 ประเภท คือ เว็บไซต์ดาวน์โหลด (Download) เว็บไซต์หน่วยงานที่ไม่มุ่งหวังผลกำไร (Non-profit organizations) และเว็บไซต์เทคโนโลยี จากนั้นผู้วิจัยได้ใช้อัลกอริทึม VIPS มาทำการแบ่งเพจเป็นบล็อก และได้ใช้คุณลักษณะระดับเพจ บล็อก เพจและบล็อกมาช่วยในการจัดประเภทของเว็บไซต์ ซึ่งจากผลการทดลองสามารถสรุปผลการทดลอง ที่แสดงให้เห็นเปอร์เซ็นต์เปรียบเทียบค่าความผิดพลาดโดยเฉลี่ย (Mean) ของ คุณลักษณะระดับเพจ บล็อก เพจและบล็อก ได้ดังตารางที่ 5.1 และแสดงให้เห็นเปอร์เซ็นต์เปรียบเทียบค่าความผิดพลาดของแต่ละประเภทของเว็บไซต์ ได้ดังตารางที่ 5.2 ดังนี้

ตารางที่ 5.1 เปอร์เซ็นต์เปรียบเทียบค่าความผิดพลาดโดยเฉลี่ย (Mean) ของคุณลักษณะระดับเพจ คุณลักษณะระดับบล็อก คุณลักษณะระดับเพจและบล็อก

คุณลักษณะ	เปอร์เซ็นต์ค่าความผิดพลาดโดยเฉลี่ย (Mean)
1. ระดับเพจ	62.5
2. ระดับบล็อก	75.9
3. ระดับเพจและบล็อก	67.8

5.2 ปัญหาและข้อจำกัดในการวิจัย

ผู้วิจัยพบปัญหาและมีข้อจำกัดในการทำวิจัย คือ

1. ในบางขั้นตอนของการทดลองจำเป็นต้องใช้บุคคลในการจัดทำ เช่น การเลือกเว็บไซต์ในแต่ละประเภท และการแบ่งเพจเป็นบล็อกโดยผ่านโปรแกรม VIPS ซึ่งรับค่ายูอาร์แอล แล้วได้ผลลัพธ์เป็นเอกซ์เอ็มแอลไฟล์ในแต่ละเพจของเว็บไซต์ จึงทำให้ใช้เวลามากในการคำนวณผลลัพธ์
2. เนื่องจากข้อมูลเป็นค่าตัวเลขต่อเนื่อง ซึ่งการนำทฤษฎีของเบย์มาใช้ ทำให้มีขั้นตอนที่ซับซ้อน จากการศึกษาเพิ่มเติม ผู้วิจัยจึงได้นำวิธีการเรียนรู้ของเครื่องแบบ See5 มาใช้

5.3 ข้อเสนอแนะ

จากงานวิจัยนี้ สามารถนำไปประยุกต์ใช้เป็นแนวทางในการทำกรวิจัยต่อไปได้ดังนี้คือ

1. ในงานวิจัยนี้ เลือกเว็บไซต์ตามลำดับของผลสืบค้นที่ได้จากกูเกิล โดยใช้คำสำคัญสำหรับเว็บไซต์แต่ละประเภทในการสืบค้น ทำให้บางเว็บไซต์ที่ได้จากผลสืบค้นอาจไม่ใช่เว็บไซต์ประเภทนั้นจริง ๆ
2. ในงานวิจัยนี้ นำเว็บเพจเพียง 10 หน้าในแต่ละเว็บไซต์มาใช้ในการทดลอง อาจทำให้ผลการทดลองที่ได้ ไม่ใช่ค่าที่แท้จริงของเว็บไซต์ ซึ่งค่าที่แท้จริงสำหรับเว็บไซต์ควรได้ค่าจากทุกเพจในเว็บไซด์นั้น ๆ
3. เพิ่มหรือปรับเปลี่ยนคุณลักษณะระดับบล็อกเพื่อให้ได้ค่าความถูกต้องที่ดีขึ้น
4. เป็นแนวทางในการศึกษาประเภทของเว็บไซต์อื่น ๆ เพิ่มเติม เช่น เว็บไซต์ของหน่วยราชการ เว็บไซต์ด้านการศึกษา

รายการอ้างอิง

- [1] Einat Amitay, D.C., Adam Darlow, Ronny Lempel, Aya Soffer, The Connectivity Sonar: Detecting Site Functionality by Structural Patterns. ACM, 2003(HT'03).
- [2] Hung-Yu Kao, S.-H.L., Jan-Ming Ho, Ming-Syan Chen, Mining Web Informative Structures and Contents Based on Entropy Analysis. IEEE, 2004.
- [3] Deng Cai, S.Y., Ji-Rong Wen, Wei-Ying Ma, Block-based Web Search. ACM, 2004.
- [4] Deng Cai, S.Y., Ji-Rong Wen and Wei-Ying Ma, Extracting Content Structure for Web Pages based on Visual Representation. 2003.
- [5] Deng Cai, S.Y., Ji-Rong Wen, Wei-Ying Ma, VIPS: a Vision-based Page Segmentation Algorithm. Microsoft Research, 2003.
- [6] Kamber, J.H.a.M., Data Mining : Concepts and Techniques. 2001: Morgan Kaufmann.
- [7] Deng Cai, X.H., Ji-Rong Wen, Wei-Ying Ma, Block-level Link Analysis. ACM, 2004(SIGIR'04).
- [8] Ruihua Song, Haifeng Liu, Ji-Rong Wen, Wei-Ying Ma, Learning Important Models for Web Page Blocks based on Layout and Content Analysis. SIGKDD, 2004.
- [8] Ruihua Song, Haifeng Liu, Ji-Rong Wen, Wei-Ying Ma, Learning Important Models for Web Page Blocks based on Layout and Content Analysis. SIGKDD, 2004.
- [9] Quinlan, J. R. See5: Programs for Machine Learning.



ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก

เว็บไซต์ที่ใช้ในการเรียนรู้และทดสอบ

เว็บไซต์ที่ใช้ในการเรียนรู้และทดสอบได้จากผลการสืบค้นดังตารางต่อไปนี้

1. เว็บไซต์เนื้อหาและสื่อกลาง (Content & media site)

ลำดับ	ยูอาร์แอล
1	http://edition.cnn.com
2	http://news.bbc.co.uk
3	http://news.yahoo.com
4	http://www.startribune.com
5	http://www.foxnews.com
6	http://www.msnbc.msn.com
7	http://www.bbc.co.uk
8	http://www.wired.com
9	http://www.thestar.com
10	http://abcnews.go.com
11	http://www.chicagotribune.com
12	http://www.cbsnews.com
13	http://www.theage.com.au
14	http://www.reuters.com
15	http://www.usatoday.com
16	http://www.nytimes.com
17	http://www.nzherald.co.nz
18	http://www.guardian.co.uk
19	http://www.fark.com
20	http://www.ananova.com
21	http://news.nationalgeographic.com
22	http://www.theaustralian.news.com.au
23	http://www.cbs.com
24	http://www.abc.net.au/default_800.htm

25	http://news.zdnet.co.uk
26	http://www.dailyherald.com
27	http://www.jpost.com
28	http://www.telegraph.co.uk
29	http://www.signonsandiego.com
30	http://www.chron.com
31	http://www.sportingnews.com
32	http://www.digg.com
33	http://www.thenews.com.pk
34	http://www.iol.co.za
35	http://www.ctv.ca
36	http://www.courant.com
37	http://www.channelnewsasia.com
38	http://www.washingtonpost.com
39	http://news.scotsman.com
40	http://www.theglobeandmail.com
41	http://slashdot.org
42	http://www.cbc.ca
43	http://sciencenow.sciencemag.org
44	http://www.newscientist.com/home.ns
45	http://abc.go.com
46	http://news.aol.com
47	http://allafrica.com
48	http://www.baltimoresun.com
49	http://www.newslink.org
50	http://www.npr.org
51	http://www.canada.com
52	http://uk.news.yahoo.com

2. เว็บไซต์บริษัท (Corporate sites)

ลำดับ	ยูอาร์แอล
1	http://www.ge.com
2	http://www.cexp.com
3	http://www.openpages.com
4	http://www.samsung.com
5	http://www.ediblearrangements.com
6	http://www.ce.com.au
7	http://www.ricoh.com
8	http://www.mhrx.com
9	http://www.murata.com
10	http://www.little.com
11	http://www.dlo.com
12	http://www.sunrx.com
13	http://www.patchlink.com
14	http://www.methodhome.com
15	http://www.cmsfx.com
16	http://globalperformance.net
17	http://www.ticproperties.com
18	http://www.merlints.com
19	http://www.exaktime.com
20	http://www.visicu.com/index_flash.asp
21	http://www.montblancgourmet.com
22	http://www.omniture.com
23	http://www.theplatform.com
24	http://www.masterkeyconsulting.com
25	http://www.issinc.com/home/index.php
26	http://www.odesus.com
27	http://www.mercerstaffing.com
28	http://www.stirrings.com
29	http://www.lancesoft.com
30	http://www.ibiquity.com
31	http://www.rmaland.com

32	http://www.vgt.net
33	http://www.wageworks.com
34	http://www.apshome.com
35	http://www.totalbeveragesolution.com/default.asp
36	http://www.datasynapse.com
37	http://www.insitu.com
38	http://www.fivepointcapital.com

3. ไตรรกทอริและลำดับชั้น (Web hierarchies & directories)

ลำดับ	ยูอาร์แอล
1	http://www.chiff.com
2	http://www.commoncontent.org
3	http://www.uncoverthenet.com
4	http://www.datedex.com/results.asp
5	http://www.gimpsy.com
6	http://www.123world.com
7	http://www.goguides.org
8	http://www.webworldindex.com
9	http://sbd.bcentral.com
10	http://www.andilinks.com
11	http://www.somuch.com
12	http://www.re-request.net
13	http://www.thisisouryear.com
14	http://www.linkopedia.com
15	http://www.webxperience.org
16	http://www.yakeo.com
17	http://aperfectsearch.com
18	http://www.localsites.ca
19	http://www.searchmonster.org
20	http://www.smartlinks.org
21	http://www.massivelinks.com

22	http://www.startingpage.com
23	http://dmoz.org
24	http://www.maven.co.il
25	http://www.nzs.com

4. อิเล็กทรอนิกส์สโตร์ (E-stores)

ลำดับ	ยูอาร์แอล
1	http://www.xynapse.ca/products_estores.htm
2	http://www.estoresnw.com/Home.asp
3	http://www.dollestores.com
4	http://www.flyingjstore.com
5	http://www.storefront.net
6	http://www.xenos.com
7	http://www.actinic.co.uk
8	http://www.butstop.com
9	http://www.intershop.com
10	http://www.mercado.com/1_homepage.shtml
11	http://www.cryptologic.com
12	http://www.zilliant.com
13	http://www.maximizer.com.hk
14	http://www.tumbleweed.com
15	http://www.shift4.com
16	http://www.casewise.com
17	http://www.perwill.com
18	http://www.bullnet.co.uk
19	http://elderestore.net
20	http://www.auspcmarket.com.au

5. พอร์ทัล (Portals)

ลำดับ	ยูอาร์แอล
1	http://www.yahoo.com
2	http://www.lycos.com
3	http://www.nabou.com
4	http://www.newsclipper.com
5	http://www.blackwebportal.com
6	http://www.clickerado.com
7	http://www.beastlair.com
8	http://www.alfy.com
9	http://www.worldhot.com
10	http://www.theallineed.com
11	http://www.zensearch.com
12	http://www.webportal.com
13	http://www.uniraq.org
14	http://www.buzzle.com
15	http://www.skali.net
16	http://www.mreverything.com
17	http://www.yankovic.org
18	http://www.sasmail.net
19	http://www.spinfinder.com
20	http://www.webplaces.com
21	http://www.state.in.us
22	http://www.jiuvalley.com/index1.html
23	http://www.startwebpage.com
24	http://www.webportal.com.my
25	http://www.chinatownconnection.com
26	http://www.dallascityhall.com

6. โปรแกรมค้นหา (Search engine)

ลำดับ	ยูอาร์แอล
1	http://usasearch.gov
2	http://www.objectssearch.com
3	http://www.kartoo.com
4	http://www.searchhippo.com
5	http://www.dogpile.com
6	http://www.scrubtheweb.com
7	http://www.factbites.com
8	http://www.entireweb.com
9	http://searchenginewatch.com
10	http://www.searchengines.com
11	http://www.searchenginecolossus.com

7. มหาวิทยาลัย (Universities)

ลำดับ	ยูอาร์แอล
1	http://www.uga.edu
2	http://www.ox.ac.uk
3	http://www.harvard.edu
4	http://www.stanford.edu
5	http://www.cam.ac.uk
6	http://www.umich.edu
7	http://www.brunel.ac.uk
8	http://www.virginia.edu
9	http://www.uwaterloo.ca
10	http://www.upenn.edu
11	http://www.usyd.edu.au
12	http://www.yale.edu
13	http://www.ufl.edu
14	http://www.cornell.edu

15	http://www.ualberta.ca
16	http://www.staffs.ac.uk
17	http://www.lancs.ac.uk
18	http://www.clemson.edu
19	http://www.uiuc.edu
20	http://www.dur.ac.uk
21	http://www.mcmaster.ca
22	http://www.duke.edu
23	http://www.umd.edu
24	http://www.bham.ac.uk
25	http://www.ucl.ac.uk
26	http://www.coventry.ac.uk

8. บริการแม่ข่ายเสมือน (Virtual hosting services)

ลำดับ	ยูอาร์แอล
1	http://www.hostdime.com
2	http://www.thehostingnews.com
3	http://www.bluehost.com
4	http://www.doteasy.com
5	http://www.directnic.com
6	http://www.fasthosts.co.uk
7	http://www.bizland.com/bizland/index.bml
8	http://www.reliablehosting.com
9	http://www.jaguarpc.com
10	http://www.imhosted.com
11	http://www.cirtexhosting.com
12	http://www.hostrocket.com
13	http://www.micfo.com
14	http://www.canaca.com
15	http://aabox.com/index.htm
16	http://www.enom.com

17	http://www.buydomains.com
18	http://www.omnis.com
19	http://www.dreamhost.com
20	http://www.ipowerweb.com
21	http://www.unitedhosting.co.uk
22	http://www.fluxservices.com
23	http://www.surpasshosting.com
24	http://www.lypha.com
25	http://www.cihost.com
26	http://www.active-venture.com
27	http://www.myvirtualhosting.com
28	http://www.intermedia.net

9. เว็บไซต์ดาวน์โหลด (Download)

ลำดับ	ยูอาร์แอล
1	http://www.tucows.com
2	http://www.sharewareriver.com
3	http://www.download.com
4	http://www.jumbo.com
5	http://www.5star-shareware.com
6	http://www.sharewarejunkies.com
7	http://www.topshareware.com
8	http://www.softpicks.net
9	http://www.simtel.net
10	http://www.download-by.net
11	http://www.downloads.nl
12	http://www.freedownloadscenter.com
13	http://www.speedbit.com

10. เว็บไซต์หน่วยงานที่ไม่มุ่งหวังผลกำไร (Non-profit organizations)

ลำดับ	ยูอาร์แอล
1	http://www.tgci.com
2	http://www.nonprofits.org
3	http://nonprofit.about.com
4	http://www.techsoup.org
5	http://www.boardsource.org
6	http://charitychannel.com
7	http://home.bellanet.org
8	http://www.philanthropicadvisor.com
9	http://www.philanthropyjournal.org
10	http://www.jewishorganizing.org
11	http://www.compasspoint.org
12	http://www.nonprofitfinancefund.org
13	http://www.snpo.org

11. เว็บไซต์เทคโนโลยี (Technology)

ลำดับ	ยูอาร์แอล
1	http://www.newsforge.com
2	http://techrepublic.com.com
3	http://www.microsoft-watch.com
4	http://www.networkworld.com
5	http://www.cioinsight.com
6	http://www.baselinemag.com
7	http://www.fcw.com
8	http://www.technewsworld.com
9	http://www.itworld.com
10	http://www.it-director.com
11	http://www.technologyreview.com
12	http://technology.guardian.co.uk
13	http://java.sun.com

ภาคผนวก ข

เอกซ์เอ็มแอล

XML (Extensible Markup Language) เป็นภาษา Markup ที่มีโครงสร้างเป็นข้อความ (Text-based) ซึ่งจัดเตรียมความสามารถด้านโครงสร้าง และสารสนเทศที่ใช้ นิยามให้กับข้อมูล ภาษา XML เป็นส่วนหนึ่งของภาษา SGML (Standard Generalized Markup Language) จุดมุ่งหมายของภาษา XML คือ ผู้สร้างเอกสารสามารถนำไปใช้งานได้ง่ายโดยใช้เทคนิคง่าย ๆ และไม่ซับซ้อนในขณะที่มีความสามารถในการจัดการนิยาม และโครงสร้างข้อมูลเหมือนกับที่มีใช้ในแอปพลิเคชันที่ใช้เทคนิคขั้นสูง

ตัวอย่างเพิ่มข้อมูลเอกซ์เอ็มแอล

```
- <VIPSPage Url="http://www.stanford.edu/" PageTitle="Stanford University" WindowWidth="433"
WindowHeight="883" PageRectLeft="0" PageRectTop="0" PageRectWidth="997"
PageRectHeight="991" neworder="0" order="0">
- <LayoutNode FrameSourceIndex="0" SourceIndex="13" DoC="2" ContainImg="3" IsImg="false"
ContainTable="false" ContainP="1" TextLen="3072" LinkTextLen="2042" DOMCIdNum="5"
FontSize="13.2" FontWeight="700" BgColor="transparent" ObjectRectLeft="0" ObjectRectTop="0"
ObjectRectWidth="997" ObjectRectHeight="991" ID="1" order="1">
- <LayoutNode FrameSourceIndex="0" SourceIndex="14;16;18;25;290;291;345;346" DoC="4"
ContainImg="4" IsImg="false" ContainTable="false" ContainP="1" TextLen="2905" LinkTextLen="1957"
DOMCIdNum="8" FontSize="13.2" FontWeight="700" BgColor="transparent" ObjectRectLeft="0"
ObjectRectTop="0" ObjectRectWidth="760" ObjectRectHeight="805" ID="1-1" order="2">
- <LayoutNode FrameSourceIndex="0" SourceIndex="14;16;18" DoC="6" ContainImg="2" IsImg="false"
ContainTable="false" ContainP="1" TextLen="15" LinkTextLen="15" DOMCIdNum="3" FontSize="9.6"
FontWeight="400" BgColor="transparent" ObjectRectLeft="0" ObjectRectTop="0"
ObjectRectWidth="760" ObjectRectHeight="180" ID="1-1-1" order="3">
<LayoutNode FrameSourceIndex="0" SourceIndex="14;16" DoC="9" ContainImg="1" IsImg="false"
ContainTable="false" ContainP="1" TextLen="15" LinkTextLen="15" DOMCIdNum="2" FontSize="9.6"
FontWeight="400" BgColor="transparent" ObjectRectLeft="0" ObjectRectTop="0"
ObjectRectWidth="760" ObjectRectHeight="49" Content="Skip to content" SRC="<P class=skip"><A
```

```

href="#content">Skip to content</A></P> <DIV id=sitename><IMG height=46 alt="Stanford
University" src="/home/pics/stanford_1.gif" width=654> </DIV>" ID="1-1-1-1" order="4" />
<LayoutNode FrameSourceIndex="0" SourceIndex="18" DoC="11" ContainImg="1" IsImg="true"
ContainTable="false" ContainP="0" TextLen="0" LinkTextLen="0" DOMCIdNum="1" FontSize="0"
FontWeight="0" BgColor="#900" ObjectRectLeft="0" ObjectRectTop="49" ObjectRectWidth="433"
ObjectRectHeight="131" Content="" SRC="<DIV id=header><DIV> <DIV
id=headerDiv><NOSCRIPT></NOSCRIPT><A onclick="homeslide(); return false;"
href="javascript:homeslide();"><IMG height=111 alt="Stanford Slideshow"
src="/home/pics/banner_760_5.jpg" width=760 border=0></A></DIV></DIV></DIV>" ID="1-1-1-2"
order="5" />
</LayoutNode>
- <LayoutNode FrameSourceIndex="0" SourceIndex="25;290;291;345;346" DoC="5" ContainImg="2"
IsImg="false" ContainTable="false" ContainP="0" TextLen="2890" LinkTextLen="1942"
DOMCIdNum="5" FontSize="13.2" FontWeight="700" BgColor="#fff" ObjectRectLeft="0"
ObjectRectTop="180" ObjectRectWidth="760" ObjectRectHeight="625" ID="1-1-2" order="6">
<LayoutNode FrameSourceIndex="0" SourceIndex="25" DoC="7" ContainImg="0" IsImg="false"
ContainTable="false" ContainP="0" TextLen="1741" LinkTextLen="1715" DOMCIdNum="2"
FontSize="13.2" FontWeight="700" BgColor="#fff" ObjectRectLeft="0" ObjectRectTop="180"
ObjectRectWidth="160" ObjectRectHeight="625" Content="Search: Web People Search for
Academic Programs Overview Schools Departments Interdisciplinary Programs Libraries Course
Catalog Executive, Professional, Continuing, & K-12 Academic Calendar Research Overview Schools
Departments Research Centers Multidisciplinary Research SLAC Libraries Undergraduate Research
Programs Admission Overview About Stanford Undergraduate Admission Graduate Admission
Financial Aid Executive, Professional, Continuing, & K-12 About Stanford Overview Introduction to
Stanford Facts History FAQ Stanford Report The Stanford Challenge Medical Center Overview School
of Medicine Stanford Hospital & Clinics Lucile Packard Children's Hospital Administration Overview
President's Office Provost's Office Trustees & Senior Administrators Faculty Senate Schools
Departments Bondholder Information Computing Arts & Events Overview Events at Stanford Cantor
Center Lively Arts Stanford on iTunes U Athletics Overview Cardinal Sports Recreational Sports
Athletic Department Recruiting Stanford Stadium Students Overview Axxess Undergraduate
Education Graduate Education Student Services Diversity Residential Computing Residential &
Dining Enterprises Student Financial Services Faculty Overview Axxess Center for Teaching &
Learning Research Resources Computing Benefits SU Faculty Statistics Stanford Report Staff &
Employment Overview Stanford Jobs BenefitSU Payroll Computing Parking & Transportation Stanford

```

Report Alumni Overview Alumni Association Volunteer Opportunities Alumni Directory Alumni E-mail Events Stanford Magazine Giving to Stanford Prospective Students Overview About Stanford Undergraduate Admission Graduate Admission Financial Aid Executive, Professional, Continuing & K-12 Visitors & Neighbors Overview Visitor Information Maps Directions Events at Stanford Parents Neighbors Campus Eateries iTunes U » Slideshow » " SRC="<DIV id=sidebar><DIV id=nav> <FORM id=searchform action=/hpcgi/page-route method=post>

<FIELDSET>Search: <INPUT id=web type=radio CHECKED value=web name=search_type> <LABEL for=web>Web</LABEL> <INPUT id=people type=radio value=people name=search_type> <LABEL for=people>People</LABEL>
<LABEL class=skip for=search_string>Search for</LABEL> <INPUT class=searchtext id=search_string name=search_string> <INPUT type=image alt=Go src="/home/pics/go_button.gif"> </FIELDSET>

</FORM> <UL class=bydept> Academic Programs Overview Schools Departments Interdisciplinary Programs Libraries Course Catalog Executive, Professional, Continuing, & K-12 Academic Calendar <LI class="">Research Overview Schools Departments Research Centers Multidisciplinary Research SLAC Libraries Undergraduate Research Programs <LI class="">Admission Overview About Stanford Undergraduate Admission Graduate Admission Financial Aid Executive, Professional, Continuing, & K-12 About Stanford Overview Introduction to Stanford Facts History

FAQ <A href="http://news-
 service.stanford.edu/">Stanford Report The Stanford Challenge Medical Center Overview School of
 Medicine Stanford Hospital & Clinics
 Lucile Packard Children's Hospital Administration Overview President's Office
 Provost's Office Trustees & Senior Administrators Faculty Senate Schools Departments <A href="http://bondholder-
 information.stanford.edu/disclaimer.html">Bondholder Information Computing Arts
 & Events Overview Events at Stanford Cantor Center Lively Arts Stanford on iTunes U Athletics Overview Cardinal Sports Recreational
 Sports Athletic Department Recruiting Stanford Stadium <UL
 class=byaudience> Students Overview Axess
 Undergraduate Education Graduate Education Student
 Services Diversity Residential Computing Residential & Dining Enterprises Student Financial Services <A

[Faculty](/home/faculty/) Overview Axess Center for Teaching & Learning Research Resources Computing Benefits SU Faculty Statistics Stanford Report Staff & Employment Overview Stanford Jobs BenefitSU Payroll Computing Parking & Transportation Stanford Report Alumni Overview Alumni Association Volunteer Opportunities Alumni Directory Alumni E-mail Events Stanford Magazine Giving to Stanford Prospective Students Overview About Stanford Undergraduate Admission Graduate Admission Financial Aid Executive, Professional, Continuing & K-12 Visitors & Neighbors Overview Visitor Information Maps Directions Events at Stanford Parents Neighbors Campus


```

Eateries</A> </LI></UL></LI></UL></DIV> <DIV id=medialinks> <DIV class=medialink><A
href="http://itunes.stanford.edu"><IMG alt="iTunes U" src="home/pics/itunes_icon.gif"> iTunes U
»</A></DIV> <DIV class=medialink><A class=slideshow onclick="homeslide(); return false;"
href="/home/slideshow/slideshow1.html"><IMG alt=Slideshow src="home/pics/slideshow_icon.gif">
Slideshow »</A></DIV> <HR class=clearer> </DIV></DIV> ID="1-1-2-1" order="7" />
- <LayoutNode FrameSourceIndex="0" SourceIndex="290;291;345;346" DoC="5" ContainImg="2"
  IsImg="false" ContainTable="false" ContainP="0" TextLen="1149" LinkTextLen="227"
  DOMCIdNum="4" FontSize="10" FontWeight="400" BgColor="#fff" ObjectRectLeft="183"
  ObjectRectTop="180" ObjectRectWidth="577" ObjectRectHeight="552" ID="1-1-2-2" order="8">
<LayoutNode FrameSourceIndex="0" SourceIndex="290;291" DoC="9" ContainImg="0" IsImg="false"
  ContainTable="false" ContainP="0" TextLen="49" LinkTextLen="49" DOMCIdNum="2" FontSize="12"
  FontWeight="400" BgColor="#fff" ObjectRectLeft="190" ObjectRectTop="191" ObjectRectWidth="321"
  ObjectRectHeight="70" Content="Index of Stanford Sites A B C D E F G H I J K L M N O P Q R S T U
  V W X Y Z" SRC="<A href="/home/atoz/">Index of Stanford Sites</A> <UL><LI><A
  href="/home/atoz/lettera.html">A</A> </LI> <LI><A href="/home/atoz/letterb.html">B</A> </LI>
  <LI><A href="/home/atoz/letterc.html">C</A> </LI> <LI><A href="/home/atoz/letterd.html">D</A>
  </LI> <LI><A href="/home/atoz/lettere.html">E</A> </LI> <LI><A
  href="/home/atoz/letterf.html">F</A> </LI> <LI><A href="/home/atoz/letterg.html">G</A> </LI>
  <LI><A href="/home/atoz/letterh.html">H</A> </LI> <LI><A href="/home/atoz/letteri.html">I</A>
  </LI> <LI><A href="/home/atoz/letterj.html">J</A> </LI> <LI><A
  href="/home/atoz/letterk.html">K</A> </LI> <LI><A href="/home/atoz/letterl.html">L</A> </LI>
  <LI><A href="/home/atoz/letterm.html">M</A> </LI> <LI><A href="/home/atoz/lettern.html">N</A>
  </LI> <LI><A href="/home/atoz/lettero.html">O</A> </LI> <LI><A
  href="/home/atoz/letterp.html">P</A> </LI> <LI><A href="/home/atoz/letterq.html">Q</A> </LI>
  <LI><A href="/home/atoz/letterr.html">R</A> </LI> <LI><A href="/home/atoz/letters.html">S</A>
  </LI> <LI><A href="/home/atoz/lettert.html">T</A> </LI> <LI><A
  href="/home/atoz/letteru.html">U</A> </LI> <LI><A href="/home/atoz/letterv.html">V</A> </LI>
  <LI><A href="/home/atoz/letterw.html">W</A> </LI> <LI><A href="/home/atoz/letterx.html">X</A>
  </LI> <LI><A href="/home/atoz/lettery.html">Y</A> </LI> <LI><A
  href="/home/atoz/letterz.html">Z</A> </LI></UL> ID="1-1-2-2-1" order="9" />
<LayoutNode FrameSourceIndex="0" SourceIndex="345" DoC="11" ContainImg="1" IsImg="true"
  ContainTable="false" ContainP="0" TextLen="0" LinkTextLen="0" DOMCIdNum="0" FontSize="0"
  FontWeight="0" BgColor="#fff" ObjectRectLeft="183" ObjectRectTop="277" ObjectRectWidth="168"

```



```

ObjectRectHeight="265" Content="" SRC="<IMG class=seal height=165 alt="Stanford seal"
src="home/pics/navmap_seal.gif" width=168>" ID="1-1-2-2-2" order="10" />
- <LayoutNode FrameSourceIndex="0" SourceIndex="346" DoC="6" ContainImg="1" IsImg="false"
ContainTable="false" ContainP="0" TextLen="1100" LinkTextLen="178" DOMCIdNum="5"
FontSize="10" FontWeight="400" BgColor="#fff" ObjectRectLeft="375" ObjectRectTop="180"
ObjectRectWidth="385" ObjectRectHeight="552" ID="1-1-2-2-3" order="11">
<LayoutNode FrameSourceIndex="0" SourceIndex="347" DoC="11" ContainImg="0" IsImg="true"
ContainTable="false" ContainP="0" TextLen="0" LinkTextLen="0" DOMCIdNum="1" FontSize="0"
FontWeight="0" BgColor="#fff" ObjectRectLeft="380" ObjectRectTop="190" ObjectRectWidth="375"
ObjectRectHeight="18" Content="" SRC="<H2 title=News><A href="http://news.stanford.edu"><IMG
alt=News src="home/pics/news_title.gif"></A></H2>" ID="1-1-2-2-3-1" order="12" />
- <LayoutNode FrameSourceIndex="0" SourceIndex="350;355;361;363;366;369;371" DoC="6"
ContainImg="1" IsImg="false" ContainTable="false" ContainP="5" TextLen="1100" LinkTextLen="178"
DOMCIdNum="7" FontSize="9.6" FontWeight="400" BgColor="#fff" ObjectRectLeft="380"
ObjectRectTop="208" ObjectRectWidth="375" ObjectRectHeight="518" ID="1-1-2-2-3-2" order="13">
<LayoutNode FrameSourceIndex="0" SourceIndex="350;355;361" DoC="9" ContainImg="0" IsImg="false"
ContainTable="false" ContainP="3" TextLen="601" LinkTextLen="97" DOMCIdNum="3" FontSize="10"
FontWeight="400" BgColor="#fff" ObjectRectLeft="380" ObjectRectTop="208" ObjectRectWidth="375"
ObjectRectHeight="222" Content="4.24.07 New 'Stanford on iTunes U' launched: A new version of
the popular "Stanford on iTunes U" service has been launched, now featuring recorded lectures from
six Stanford courses, in addition to nearly 1,000 other audio and video recordings of campus events,
faculty talks, interviews and music. 4.23.07 New childcare-assistance program for junior faculty: In
an effort to help junior faculty cope with the twin pressures of raising families and winning tenure,
Stanford has developed a new program to provide financial assistance for childcare to families with
children ages 5 and under. more news »" SRC="<P><SPAN
class=dateline>4.24.07</SPAN><BR><STRONG><A href="http://news-
service.stanford.edu/news/2007/april25/itunes-042507.html">New 'Stanford on iTunes U'
launched</A></STRONG>: A new version of the popular "Stanford on iTunes U" service has been
launched, now featuring recorded lectures from six Stanford courses, in addition to nearly 1,000
other audio and video recordings of campus events, faculty talks, interviews and music. </P>
<P><SPAN class=dateline>4.23.07</SPAN><BR><STRONG><A href="http://news-
service.stanford.edu/news/2007/april25/kidcare-042507.html">New childcare-assistance program for
junior faculty</A></STRONG>: In an effort to help junior faculty cope with the twin pressures of
raising families and winning tenure, Stanford has developed a new program to provide financial

```

assistance for childcare to families with children ages 5 and under. </P> <P class=more>more news »</P>" ID="1-1-2-2-3-2-1" order="14" />

- <LayoutNode FrameSourceIndex="0" SourceIndex="363;366;369;371" DoC="6" ContainImg="1" IsImg="false" ContainTable="false" ContainP="2" TextLen="499" LinkTextLen="81" DOMCIdNum="4" FontSize="9.6" FontWeight="400" BgColor="#fff" ObjectRectLeft="380" ObjectRectTop="449" ObjectRectWidth="375" ObjectRectHeight="277" ID="1-1-2-2-3-2-2" order="15">

- <LayoutNode FrameSourceIndex="0" SourceIndex="363;366;369" DoC="6" ContainImg="0" IsImg="false" ContainTable="false" ContainP="2" TextLen="356" LinkTextLen="42" DOMCIdNum="3" FontSize="9.6" FontWeight="400" BgColor="#fff" ObjectRectLeft="380" ObjectRectTop="449" ObjectRectWidth="375" ObjectRectHeight="137" ID="1-1-2-2-3-2-2-1" order="16">

<LayoutNode FrameSourceIndex="0" SourceIndex="363" DoC="11" ContainImg="0" IsImg="true" ContainTable="false" ContainP="0" TextLen="0" LinkTextLen="0" DOMCIdNum="1" FontSize="0" FontWeight="0" BgColor="#fff" ObjectRectLeft="380" ObjectRectTop="449" ObjectRectWidth="375" ObjectRectHeight="18" Content="" SRC="<H2 title=Events></H2>" ID="1-1-2-2-3-2-2-1-1" order="17" />

<LayoutNode FrameSourceIndex="0" SourceIndex="366;369" DoC="9" ContainImg="0" IsImg="false" ContainTable="false" ContainP="2" TextLen="356" LinkTextLen="42" DOMCIdNum="2" FontSize="9.6" FontWeight="400" BgColor="#fff" ObjectRectLeft="380" ObjectRectTop="467" ObjectRectWidth="375" ObjectRectHeight="119" Content="Brainstorms with Pat Hanrahan: Brainstorms is a quarterly series highlighting research by Stanford faculty in the new frontiers of science and technology. Professor Pat Hanrahan will look behind the “scenes” of computer graphics and see how the magic tricks are done. Thursday, April 26 at 7:30 p.m. at the William R. Hewlett Teaching Center. more events »" SRC="<P>Brainstorms with Pat Hanrahan: Brainstorms is a quarterly series highlighting research by Stanford faculty in the new frontiers of science and technology. Professor Pat Hanrahan will look behind the “scenes” of computer graphics and see how the magic tricks are done. Thursday, April 26 at 7:30 p.m. at the William R. Hewlett Teaching Center. </P> <P class=more>more events »</P>" ID="1-1-2-2-3-2-2-1-2" order="18" />

</LayoutNode>

<LayoutNode FrameSourceIndex="0" SourceIndex="371" DoC="8" ContainImg="1" IsImg="false" ContainTable="false" ContainP="0" TextLen="143" LinkTextLen="39" DOMCIdNum="2" FontSize="9.6"

```

FontWeight="400" BgColor="#fff" ObjectRectLeft="389" ObjectRectTop="605" ObjectRectWidth="360"
ObjectRectHeight="121" Content="Diversity Works at Stanford: Learn about the many programs and
initiatives which reflect Stanford's strong commitment to diversity. more sites »" SRC="<DIV
class=promo><IMG height=17 alt="On the Stanford Web" src="home/pics/otsw_title.gif" width=176>
<DIV><A href="http://diversityworks.stanford.edu"><IMG alt=""
src="/dept/ucomm/otsw/images/otsw14.jpg"></A> <P><STRONG><A
href="http://diversityworks.stanford.edu">Diversity Works at Stanford</A></STRONG>: Learn about
the many programs and initiatives which reflect Stanford's strong commitment to diversity.</P> <P
class=more><A href="/home/otsw/">more sites »</A></P></DIV></DIV>" ID="1-1-2-2-3-2-2-2"
order="19" />
</LayoutNode>
</LayoutNode>
</LayoutNode>
</LayoutNode>
</LayoutNode>
</LayoutNode>
</LayoutNode>
<LayoutNode FrameSourceIndex="0" SourceIndex="382" DoC="9" ContainImg="0" IsImg="false"
ContainTable="false" ContainP="1" TextLen="167" LinkTextLen="85" DOMCIdNum="2" FontSize="9.6"
FontWeight="700" BgColor="transparent" ObjectRectLeft="0" ObjectRectTop="821"
ObjectRectWidth="433" ObjectRectHeight="62" Content="Contact Directories Maps & Directions
Giving to Stanford © Stanford University. All Rights Reserved. Stanford, CA 94305. (650) 723-2300.
Terms of Use | Copyright Complaints" SRC="<DIV id=footer><UL> <LI><A
href="/home/atoz/contact.html">Contact</A> <LI><A href="/home/directories/">Directories</A>
<LI><A href="/home/visitors/maps.html">Maps & Directions</A> <LI class=last><A
href="http://givingtostanford.stanford.edu/">Giving to Stanford</A> </LI></UL> <P class=vcard>©
<SPAN class="fn org">Stanford University</SPAN>. All Rights Reserved. <SPAN class=adr><SPAN
class=locality>Stanford</SPAN>, <SPAN class=region>CA</SPAN> <SPAN class=postal-
code>94305</SPAN>. <SPAN class=tel>(650) 723-2300</SPAN></SPAN>. <A
href="home/atoz/terms.html">Terms of Use</A> | <A href="/home/atoz/copyright.html">Copyright
Complaints</A></P></DIV>" ID="1-2" order="20" />
</LayoutNode>
</VIPSPage>

```

ภาคผนวก ค

ตัวอย่าง See5 และ RULE-BASED CLASSIFIERS

ในการจัดประเภทของเว็บไซต์ที่ใช้เครื่องมือในการจัดประเภทคือ See5

[Fold 0]

Rules:

Rule 1: (111/67, lift 1.9)

POutLink > 11

-> class 1 [0.398]

Rule 2: (8, lift 6.0)

POutLink <= 11

FontSize > 6

BlockRectWidth > 688

-> class 2 [0.900]

Rule 3: (4, lift 5.6)

PLevel <= 1

FontSize <= 6

BlockRectToop <= 425

-> class 2 [0.833]

Rule 4: (10/1, lift 5.6)

POutLink <= 11

TextLen > 400

TextLen <= 625

FontSize > 6

FontSize <= 11

BlockRectWidth > 442

BlockCenterX > 312

BlockCenterY <= 1047

-> class 2 [0.833]

Rule 5: (2, lift 5.0)

TextLen <= 704

DOMCIdNum > 6

BlockRectWidth <= 442

-> class 2 [0.750]

Rule 6: (16/4, lift 4.8)

POutLink <= 11

TextLen <= 259

FontSize > 6

BlockRectWidth > 442

BlockCenterX > 312

BlockCenterY <= 1047

-> class 2 [0.722]

Rule 7: (5, lift 9.0)

PInLink <= 55

POutLink > 11

POutLink <= 50

TextLen > 601

FontSize <= 10

BlockRectHeight > 566

-> class 3 [0.857]

Rule 8: (4, lift 8.8)

PLevel > 1

POutLink > 0

POutLink <= 11

BlockRectWidth > 337

BlockCenterX <= 312

-> class 3 [0.833]

Rule 9: (3, lift 8.4)

PInLink <= 55

POutLink > 50

InPerOut > 0

LinkTextLen > 38

BlockRectHeight <= 760

-> class 3 [0.800]

Rule 10: (4/1, lift 7.0)

PInLink <= 33

BlockCenterX > 312

BlockCenterY > 1047

-> class 3 [0.667]

Rule 11: (78/65, lift 2.4)

POutLink <= 11

TextLen > 259

-> class 4 [0.175]

Rule 12: (5, lift 8.6)

PInLink <= 55

POutLink > 11

POutLink <= 50

InPerOut <= 5

LinkTextLen > 59

LinkTextLen <= 130

BlockRectToop <= 878

BlockRectHeight <= 566

-> class 5 [0.857]

Rule 13: (3, lift 8.0)

POutLink <= 11

TextLen > 571

TextLen <= 704

BlockRectWidth <= 442

-> class 5 [0.800]

Rule 14: (2, lift 7.5)

PInLink <= 18

POutLink > 50

InPerOut <= 0

-> class 5 [0.750]

Rule 15: (5/1, lift 7.2)

PLevel > 1

POutLink <= 11

FontSize <= 6

BlockCenterY <= 1047

-> class 5 [0.714]

Rule 16: (2, lift 20.7)

POutLink > 11
 LinkTextLen <= 38
 BlockCenterX <= 332
 -> class 6 [0.750]

Rule 17: (5/1, lift 19.7)

PLevel <= 1
 POutLink > 0
 POutLink <= 11
 DOMCidNum <= 11
 BlockRectToop > 245
 BlockCenterX <= 312
 -> class 6 [0.714]

Rule 18: (3, lift 8.0)

PInLink <= 55
 POutLink > 11
 InPerOut > 5
 TextLen <= 603
 LinkTextLen > 38
 BlockRectHeight <= 566
 -> class 7 [0.800]

Rule 19: (3, lift 8.0)

POutLink <= 11
 BlockRectWidth <= 337
 BlockCenterX <= 312
 -> class 7 [0.800]

Rule 20: (9/2, lift 7.3)

PInLink <= 55
 POutLink <= 50
 TextLen <= 603
 LinkTextLen > 130
 BlockRectHeight <= 566
 -> class 7 [0.727]

Rule 21: (8/3, lift 6.0)

PInLink <= 55
POutLink > 11
FontSize > 10
BlockRectHeight > 566
-> class 7 [0.600]

Rule 22: (3/1, lift 6.0)

PInLink > 55
BlockRectToop <= 218
-> class 7 [0.600]

Rule 23: (4, lift 8.0)

PLevel <= 1
InPerOut > 13
TextLen <= 400
FontSize > 6
BlockRectWidth > 442
BlockRectWidth <= 688
-> class 8 [0.833]

Rule 24: (3, lift 7.7)

POutLink <= 11
TextLen > 400
FontSize > 11
BlockRectWidth > 442
BlockRectWidth <= 688
-> class 8 [0.800]

Rule 25: (7/1, lift 7.5)

POutLink <= 11
TextLen > 259
TextLen <= 400
DOMCIdNum > 2
FontSize > 6
BlockRectWidth > 442
BlockRectWidth <= 688
-> class 8 [0.778]

Rule 26: (4/1, lift 6.4)

PLevel <= 1
 FontSize <= 6
 BlockRectToop > 425
 -> class 8 [0.667]

Rule 27: (2, lift 16.6)

PInLink > 33
 POutLink <= 11
 TextLen <= 704
 BlockCenterY > 1047
 -> class 9 [0.750]

Rule 28: (4/2, lift 11.0)

PInLink > 55
 POutLink > 11
 POutLink <= 15
 -> class 9 [0.500]

Rule 29: (8/5, lift 8.8)

POutLink <= 8
 TextLen > 704
 BlockCenterX > 312
 -> class 9 [0.400]

Rule 30: (2, lift 16.6)

PInLink <= 55
 POutLink > 50
 InPerOut > 0
 LinkTextLen > 38
 BlockRectHeight > 760
 -> class 11 [0.750]

Rule 31: (6/1, lift 16.6)

PInLink > 55
 POutLink > 79
 BlockCenterY > 1396
 -> class 11 [0.750]

Default class: 2

Evaluation on hold-out data (11 cases):



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก จ

ตัวอย่างการคำนวณค่าคุณลักษณะระดับเว็บไซต์

ตัวอย่างเป็นเว็บไซต์ของมหาวิทยาลัยสแตนฟอร์ด (URL : <http://www.stanford.edu>) สำหรับแต่ละเว็บไซต์จะมีการคำนวณดังรายละเอียดดังนี้

3.7.1 อ่านการเชื่อมโยงของเว็บไซต์เพื่อเลือกเพจ

อ่านเอชทีเอ็มแอลไฟล์ของเพจแรกของเว็บไซต์แล้วหาการเชื่อมโยงในเพจนี้ทั้งหมดเพื่อให้ได้ 10 หน้าแรกของเว็บไซต์นี้ ถ้าหากเพจแรกมีการเชื่อมโยงไม่ถึง 10 เพจ ก็จะทำการศึกษาการเชื่อมโยงในหน้าถัดไป เพื่อให้ได้เพจทั้งหมดอย่างน้อย 10 เพจ

- 1 <http://www.stanford.edu/home/academics>
- 2 <http://www.stanford.edu/home/academics>
- 3 <http://www.stanford.edu/home/academics/schools.html>
- 4 <http://www.stanford.edu/home/academics/departments.html>
- 5 <http://www.stanford.edu/home/academics/programs.html>
- 6 <http://library.stanford.edu>
- 7 <http://www.stanford.edu/dept/registrar/bulletin>
- 8 <http://www.stanford.edu/home/academics/continuing.html>
- 9 http://www.stanford.edu/dept/registrar/academic_calendar
- 10 <http://www.stanford.edu/home/research>
- 11 <http://www.stanford.edu/home/research>
- 12 <http://www.stanford.edu/home/academics/schools.html>
- 13 <http://www.stanford.edu/home/academics/departments.html>
- 14 <http://www.stanford.edu/home/research/centers.html>
- 15 <http://multi.stanford.edu>
- 16 <http://www.slac.stanford.edu>
- 17 <http://library.stanford.edu>
- 18 <http://www.stanford.edu/dept/undergrad/urp>
- 19 <http://www.stanford.edu/home/admission>

- 20 <http://www.stanford.edu/home/admission>
- 21 <http://www.stanford.edu/home/stanford>
- 22 <http://www.stanford.edu/dept/uga>
- 23 <http://www.stanford.edu/dept/gradadmissions>
- 24 <http://www.stanford.edu/dept/finaid>
- 25 <http://www.stanford.edu/home/academics/continuing.html>
- 26 <http://www.stanford.edu/home/stanford>
- 27 <http://www.stanford.edu/home/stanford>
- 28 <http://www.stanford.edu/home/welcome>
- 29 <http://www.stanford.edu/home/stanford/facts>
- 30 <http://www.stanford.edu/home/stanford/history>
- 31 <http://www.stanford.edu/home/atoz/faq.html>
- 32 <http://news-service.stanford.edu>
- 33 <http://thestanfordchallenge.stanford.edu>
- 34 <http://www.stanford.edu/home/medcenter>
- 35 <http://www.stanford.edu/home/medcenter>
- 36 <http://med.stanford.edu>
- 37 <http://www.stanfordhospital.com>
- 38 <http://www.lpch.org>
- 39 <http://www.stanford.edu/home/administration>
- 40 <http://www.stanford.edu/home/administration>
- 41 <http://www.stanford.edu/dept/president>
- 42 <http://www.stanford.edu/dept/provost>
- 43 <http://www.stanford.edu/home/stanford/facts/board.html>
- 44 <http://faculty senate.stanford.edu>
- 45 <http://www.stanford.edu/home/academics/schools.html>
- 46 <http://www.stanford.edu/home/academics/departments.html>
- 47 <http://bondholder-information.stanford.edu/disclaimer.html>
- 48 <http://computing.stanford.edu>
- 49 <http://www.stanford.edu/home/arts>

- 50 <http://www.stanford.edu/home/arts>
- 51 <http://events.stanford.edu>
- 52 <http://museum.stanford.edu>
- 53 <http://livelyarts.stanford.edu>
- 54 <http://itunes.stanford.edu>
- 55 <http://www.stanford.edu/home/athletics>
- 56 <http://www.stanford.edu/home/athletics>
- 57 <http://gostanford.cstv.com>
- 58 <http://www.stanford.edu/dept/pe>
- 59 <http://gostanford.cstv.com/school-bio>
- 60 <http://gostanford.cstv.com/genrel/stan-recruiting.html>
- 61 <http://www.stanfordstadium.com>
- 62 <http://www.stanford.edu/home/students>
- 63 <http://www.stanford.edu/home/students>
- 64 <https://axess.stanford.edu>
- 65 <http://www.stanford.edu/dept/undergrad/vpue>
- 66 <http://www.stanford.edu/dept/DoR/GSH/Sec1.html>
- 67 <http://www.stanford.edu/dept/vpsa>
- 68 <http://diversityworks.stanford.edu>
- 69 <http://rescomp.stanford.edu>
- 70 <http://www.stanford.edu/dept/hds>
- 71 <http://co.stanford.edu/students>
- 72 <http://www.stanford.edu/home/faculty>
- 73 <http://www.stanford.edu/home/faculty>
- 74 <https://axess.stanford.edu>
- 75 <http://ctl.stanford.edu>
- 76 <http://www.stanford.edu/home/research>
- 77 <http://computing.stanford.edu>
- 78 <http://benefitsu.stanford.edu>
- 79 <http://www.stanford.edu/home/stanford/facts/faculty.html>

- 80 <http://news-service.stanford.edu>
- 81 <http://www.stanford.edu/home/working>
- 82 <http://www.stanford.edu/home/working>
- 83 <http://jobs.stanford.edu>
- 84 <http://benefitsu.stanford.edu>
- 85 http://co.stanford.edu/payroll_manual
- 86 <http://computing.stanford.edu>
- 87 <http://transportation.stanford.edu>
- 88 <http://news-service.stanford.edu>
- 89 <http://www.stanford.edu/home/alumni>
- 90 <http://www.stanford.edu/home/alumni>
- 91 <http://www.stanfordalumni.org>
- 92 <http://www.stanfordalumni.org/volunteer>
- 93 <https://pgnet.stanford.edu/WebDirectory/search?pgOrg=saa&pgnPageCode=110>
- 94 <https://www.stanfordalumni.org/email>
- 95 <http://www.stanfordalumni.org/erc/regional/listing.html?type=events>
- 96 <http://www.stanfordalumni.org/magazine>
- 97 <http://giving.stanford.edu>
- 98 <http://www.stanford.edu/home/prospective>
- 99 <http://www.stanford.edu/home/prospective>
- 100 <http://www.stanford.edu/home/stanford>
- 101 <http://www.stanford.edu/dept/uga>
- 102 <http://www.stanford.edu/dept/gradadmissions>
- 103 <http://www.stanford.edu/dept/finaid>
- 104 <http://www.stanford.edu/home/academics/continuing.html>
- 105 <http://www.stanford.edu/home/visitors>
- 106 <http://www.stanford.edu/home/visitors>
- 107 <http://www.stanford.edu/dept/visitorinfo>
- 108 <http://www.stanford.edu/search/maps>
- 109 <http://www.stanford.edu/search/maps/directions.html>

- 110 <http://events.stanford.edu>
- 111 <http://parents.stanford.edu>
- 112 <http://www.stanford.edu/dept/ucomm/neighbors>
- 113 <http://www.stanford.edu/home/visitors/eateries>
- 114 <http://itunes.stanford.edu>
- 115 <http://www.stanford.edu/home/slideshow/slideshow1.html>
- 116 <http://www.stanford.edu/home/atoz>
- 117 <http://www.stanford.edu/home/atoz/lettera.html>
- 118 <http://www.stanford.edu/home/atoz/letterb.html>
- 119 <http://www.stanford.edu/home/atoz/letterc.html>
- 120 <http://www.stanford.edu/home/atoz/letterd.html>
- 121 <http://www.stanford.edu/home/atoz/lettere.html>
- 122 <http://www.stanford.edu/home/atoz/letterf.html>
- 123 <http://www.stanford.edu/home/atoz/letterg.html>
- 124 <http://www.stanford.edu/home/atoz/letterh.html>
- 125 <http://www.stanford.edu/home/atoz/letteri.html>
- 126 <http://www.stanford.edu/home/atoz/letterj.html>
- 127 <http://www.stanford.edu/home/atoz/letterk.html>
- 128 <http://www.stanford.edu/home/atoz/letterl.html>
- 129 <http://www.stanford.edu/home/atoz/letterm.html>
- 130 <http://www.stanford.edu/home/atoz/lettern.html>
- 131 <http://www.stanford.edu/home/atoz/lettero.html>
- 132 <http://www.stanford.edu/home/atoz/letterp.html>
- 133 <http://www.stanford.edu/home/atoz/letterq.html>
- 134 <http://www.stanford.edu/home/atoz/letterr.html>
- 135 <http://www.stanford.edu/home/atoz/letters.html>
- 136 <http://www.stanford.edu/home/atoz/lettert.html>
- 137 <http://www.stanford.edu/home/atoz/letteru.html>
- 138 <http://www.stanford.edu/home/atoz/letterv.html>
- 139 <http://www.stanford.edu/home/atoz/letterw.html>

- 140 <http://www.stanford.edu/home/atoz/letterx.html>
- 141 <http://www.stanford.edu/home/atoz/lettery.html>
- 142 <http://www.stanford.edu/home/atoz/letterz.html>
- 143 <http://news.stanford.edu>
- 144 <http://news-service.stanford.edu/news/2007/april25/itunes-042507.html>
- 145 <http://news-service.stanford.edu/news/2007/april25/kidcare-042507.html>
- 146 <http://news.stanford.edu>
- 147 <http://events.stanford.edu>
- 148 <https://events.stanford.edu/events/108/10861>
- 149 <http://events.stanford.edu>
- 150 <http://www.stanford.edu/home/welcome/academics>
- 151 <http://www.stanford.edu/home/welcome/academics>
- 152 <http://www.stanford.edu/home/otsw>
- 153 <http://www.stanford.edu/home/atoz/contact.html>
- 154 <http://www.stanford.edu/home/directories>
- 155 <http://www.stanford.edu/home/visitors/maps.html>
- 156 <http://givingtostanford.stanford.edu>
- 157 <http://www.stanford.edu/home/atoz/terms.html>
- 158 <http://www.stanford.edu/home/atoz/copyright.html>

3.7.2 เลือก 10 เพจสำหรับเว็บไซต์นั้น

เลือกเพจ 10 เพจแรกที่ได้จากการเชื่อมโยงทั้งหมดในเพจแรก

- 1 <http://www.stanford.edu>
- 2 <http://www.stanford.edu/home/academics/schools.html>
- 3 <http://www.stanford.edu/home/academics/departments.html>
- 4 <http://www.stanford.edu/home/academics/programs.html>
- 5 <http://www.stanford.edu/home/academics/continuing.html>
- 6 <http://www.stanford.edu/home/research/centers.html>
- 7 <http://www.stanford.edu/home/atoz/faq.html>
- 8 <http://www.stanford.edu/home/stanford/facts/board.html>

9 <http://www.stanford.edu/dept/DoR/GSH/Sec1.html>

10 <http://www.stanford.edu/home/stanford/facts/faculty.html>

3.7.3 คำวนคุณลักษณะในระดับเพจในแต่ละเพจ

ในแต่ละเพจจะทำการอ่านเอชทีเอ็มแอลพีลล์ของเพจนั้นเพื่อหาการเชื่อมโยงในเพจนั้นทั้งหมด โดยมีการเช็คว่าเป็นการเชื่อมโยงภายในหรือการเชื่อมโยงภายนอก โดยเช็คว่าเป็นแม่ข่าย (Host) เดียวกันหรือไม่ ในการคำนวณคุณลักษณะในระดับเพจจะคำนวณที่ละเพจจากเพจที่ 1 ถึงเพจที่ 10 โดยคุณลักษณะในระดับเพจที่เก็บข้อมูลคือการเชื่อมโยงภายใน การเชื่อมโยงภายนอก และ อัตราส่วนระหว่างการเชื่อมโยงภายในและการเชื่อมโยงภายนอก

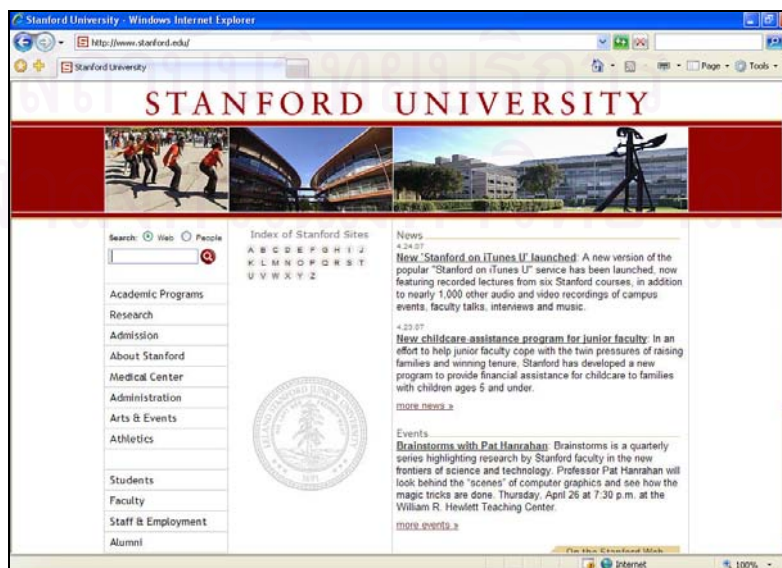
- ระดับของเพจ
- การเชื่อมโยงของเพจนั้นกับการเชื่อมโยงที่ได้เป็นแม่ข่ายเดียวกันจะถือว่าการเชื่อมโยงภายใน
- การเชื่อมโยงของเพจนั้นกับการเชื่อมโยงที่ได้เป็นคนละแม่ข่ายกันจะถือว่าการเชื่อมโยงภายนอก
- อัตราส่วนระหว่างการเชื่อมโยงภายในและการเชื่อมโยงภายนอก

โดย

Link: I คือ การเชื่อมโยงภายใน

Link: X คือ การเชื่อมโยงภายนอก

เว็บเพจที่ 1 : <http://www.stanford.edu>



รูปที่ ๑.1 เว็บไซต์ที่ 1

[Level of page = 0]

Total Links(Each) : 158

- 1 <http://www.stanford.edu/home/academics> Link: I
- 2 <http://www.stanford.edu/home/academics> Link: I
- 3 <http://www.stanford.edu/home/academics/schools.html> Link: I
- 4 <http://www.stanford.edu/home/academics/departments.html> Link: I
- 5 <http://www.stanford.edu/home/academics/programs.html> Link: I
- 6 <http://library.stanford.edu> Link: X
- 7 <http://www.stanford.edu/dept/registrar/bulletin> Link: I
- 8 <http://www.stanford.edu/home/academics/continuing.html> Link: I
- 9 http://www.stanford.edu/dept/registrar/academic_calendar Link: I
- 10 <http://www.stanford.edu/home/research> Link: I
- 11 <http://www.stanford.edu/home/research> Link: I
- 12 <http://www.stanford.edu/home/academics/schools.html> Link: I
- 13 <http://www.stanford.edu/home/academics/departments.html> Link: I
- 14 <http://www.stanford.edu/home/research/centers.html> Link: I
- 15 <http://multi.stanford.edu> Link: X
- 16 <http://www.slac.stanford.edu> Link: X
- 17 <http://library.stanford.edu> Link: X
- 18 <http://www.stanford.edu/dept/undergrad/urp> Link: I
- 19 <http://www.stanford.edu/home/admission> Link: I
- 20 <http://www.stanford.edu/home/admission> Link: I
- 21 <http://www.stanford.edu/home/stanford> Link: I
- 22 <http://www.stanford.edu/dept/uga> Link: I
- 23 <http://www.stanford.edu/dept/gradadmissions> Link: I
- 24 <http://www.stanford.edu/dept/finaid> Link: I
- 25 <http://www.stanford.edu/home/academics/continuing.html> Link: I
- 26 <http://www.stanford.edu/home/stanford> Link: I

- 27 <http://www.stanford.edu/home/stanford> Link: I
- 28 <http://www.stanford.edu/home/welcome> Link: I
- 29 <http://www.stanford.edu/home/stanford/facts> Link: I
- 30 <http://www.stanford.edu/home/stanford/history> Link: I
- 31 <http://www.stanford.edu/home/atoz/faq.html> Link: I
- 32 <http://news-service.stanford.edu> Link: X
- 33 <http://thestanfordchallenge.stanford.edu> Link: X
- 34 <http://www.stanford.edu/home/medcenter> Link: I
- 35 <http://www.stanford.edu/home/medcenter> Link: I
- 36 <http://med.stanford.edu> Link: X
- 37 <http://www.stanfordhospital.com> Link: X
- 38 <http://www.lpch.org> Link: X
- 39 <http://www.stanford.edu/home/administration> Link: I
- 40 <http://www.stanford.edu/home/administration> Link: I
- 41 <http://www.stanford.edu/dept/president> Link: I
- 42 <http://www.stanford.edu/dept/provost> Link: I
- 43 <http://www.stanford.edu/home/stanford/facts/board.html> Link: I
- 44 <http://faculty senate.stanford.edu> Link: X
- 45 <http://www.stanford.edu/home/academics/schools.html> Link: I
- 46 <http://www.stanford.edu/home/academics/departments.html> Link: I
- 47 <http://bondholder-information.stanford.edu/disclaimer.html> Link: X
- 48 <http://computing.stanford.edu> Link: X
- 49 <http://www.stanford.edu/home/arts> Link: I
- 50 <http://www.stanford.edu/home/arts> Link: I
- 51 <http://events.stanford.edu> Link: X
- 52 <http://museum.stanford.edu> Link: X
- 53 <http://livelyarts.stanford.edu> Link: X
- 54 <http://itunes.stanford.edu> Link: X
- 55 <http://www.stanford.edu/home/athletics> Link: I
- 56 <http://www.stanford.edu/home/athletics> Link: I

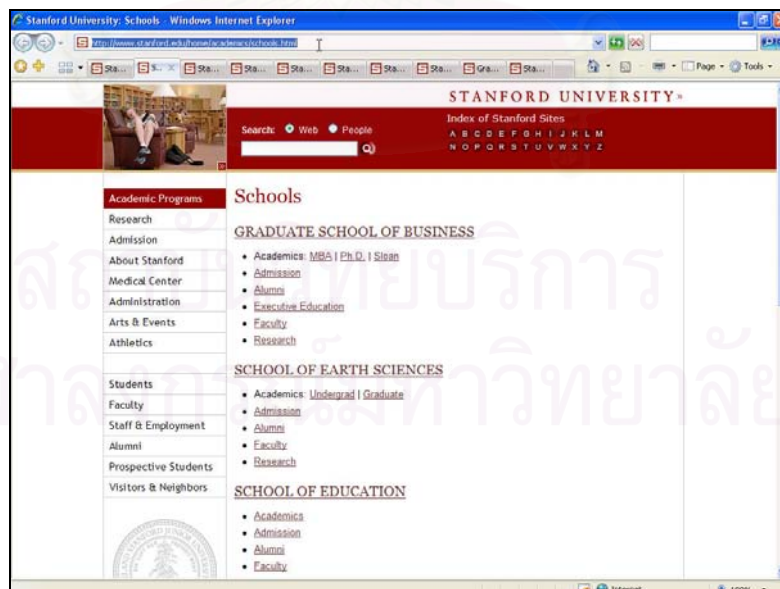
- 57 <http://gostanford.cstv.com> Link: X
- 58 <http://www.stanford.edu/dept/pe> Link: I
- 59 <http://gostanford.cstv.com/school-bio> Link: X
- 60 <http://gostanford.cstv.com/genrel/stan-recruiting.html> Link: X
- 61 <http://www.stanfordstadium.com> Link: X
- 62 <http://www.stanford.edu/home/students> Link: I
- 63 <http://www.stanford.edu/home/students> Link: I
- 64 <https://axess.stanford.edu> Link: X
- 65 <http://www.stanford.edu/dept/undergrad/vpue> Link: I
- 66 <http://www.stanford.edu/dept/DoR/GSH/Sec1.html> Link: I
- 67 <http://www.stanford.edu/dept/vpsa> Link: I
- 68 <http://diversityworks.stanford.edu> Link: X
- 69 <http://rescomp.stanford.edu> Link: X
- 70 <http://www.stanford.edu/dept/hds> Link: I
- 71 <http://co.stanford.edu/students> Link: X
- 72 <http://www.stanford.edu/home/faculty> Link: I
- 73 <http://www.stanford.edu/home/faculty> Link: I
- 74 <https://axess.stanford.edu> Link: X
- 75 <http://ctl.stanford.edu> Link: X
- 76 <http://www.stanford.edu/home/research> Link: I
- 77 <http://computing.stanford.edu> Link: X
- 78 <http://benefitsu.stanford.edu> Link: X
- 79 <http://www.stanford.edu/home/stanford/facts/faculty.html> Link: I
- 80 <http://news-service.stanford.edu> Link: X
- 81 <http://www.stanford.edu/home/working> Link: I
- 82 <http://www.stanford.edu/home/working> Link: I
- 83 <http://jobs.stanford.edu> Link: X
- 84 <http://benefitsu.stanford.edu> Link: X
- 85 http://co.stanford.edu/payroll_manual Link: X
- 86 <http://computing.stanford.edu> Link: X

- 87 <http://transportation.stanford.edu> Link: X
- 88 <http://news-service.stanford.edu> Link: X
- 89 <http://www.stanford.edu/home/alumni> Link: I
- 90 <http://www.stanford.edu/home/alumni> Link: I
- 91 <http://www.stanfordalumni.org> Link: X
- 92 <http://www.stanfordalumni.org/volunteer> Link: X
- 93 <https://pgnet.stanford.edu/WebDirectory/search?pgOrg=saa&pgnPageCode=110>
Link: X
- 94 <https://www.stanfordalumni.org/email> Link: X
- 95 <http://www.stanfordalumni.org/erc/regional/listing.html?type=events> Link: X
- 96 <http://www.stanfordalumni.org/magazine> Link: X
- 97 <http://giving.stanford.edu> Link: X
- 98 <http://www.stanford.edu/home/prospective> Link: I
- 99 <http://www.stanford.edu/home/prospective> Link: I
- 100 <http://www.stanford.edu/home/stanford> Link: I
- 101 <http://www.stanford.edu/dept/uga> Link: I
- 102 <http://www.stanford.edu/dept/gradadmissions> Link: I
- 103 <http://www.stanford.edu/dept/finaid> Link: I
- 104 <http://www.stanford.edu/home/academics/continuing.html> Link: I
- 105 <http://www.stanford.edu/home/visitors> Link: I
- 106 <http://www.stanford.edu/home/visitors> Link: I
- 107 <http://www.stanford.edu/dept/visitorinfo> Link: I
- 108 <http://www.stanford.edu/search/maps> Link: I
- 109 <http://www.stanford.edu/search/maps/directions.html> Link: I
- 110 <http://events.stanford.edu> Link: X
- 111 <http://parents.stanford.edu> Link: X
- 112 <http://www.stanford.edu/dept/ucomm/neighbors> Link: I
- 113 <http://www.stanford.edu/home/visitors/eateries> Link: I
- 114 <http://itunes.stanford.edu> Link: X
- 115 <http://www.stanford.edu/home/slideshow/slideshow1.html> Link: I

- 116 <http://www.stanford.edu/home/atoz> Link: I
- 117 <http://www.stanford.edu/home/atoz/lettera.html> Link: I
- 118 <http://www.stanford.edu/home/atoz/letterb.html> Link: I
- 119 <http://www.stanford.edu/home/atoz/letterc.html> Link: I
- 120 <http://www.stanford.edu/home/atoz/letterd.html> Link: I
- 121 <http://www.stanford.edu/home/atoz/lettere.html> Link: I
- 122 <http://www.stanford.edu/home/atoz/letterf.html> Link: I
- 123 <http://www.stanford.edu/home/atoz/letterg.html> Link: I
- 124 <http://www.stanford.edu/home/atoz/letterh.html> Link: I
- 125 <http://www.stanford.edu/home/atoz/letteri.html> Link: I
- 126 <http://www.stanford.edu/home/atoz/letterj.html> Link: I
- 127 <http://www.stanford.edu/home/atoz/letterk.html> Link: I
- 128 <http://www.stanford.edu/home/atoz/letterl.html> Link: I
- 129 <http://www.stanford.edu/home/atoz/letterm.html> Link: I
- 130 <http://www.stanford.edu/home/atoz/lettern.html> Link: I
- 131 <http://www.stanford.edu/home/atoz/lettero.html> Link: I
- 132 <http://www.stanford.edu/home/atoz/letterp.html> Link: I
- 133 <http://www.stanford.edu/home/atoz/letterq.html> Link: I
- 134 <http://www.stanford.edu/home/atoz/letterr.html> Link: I
- 135 <http://www.stanford.edu/home/atoz/letters.html> Link: I
- 136 <http://www.stanford.edu/home/atoz/lettert.html> Link: I
- 137 <http://www.stanford.edu/home/atoz/letteru.html> Link: I
- 138 <http://www.stanford.edu/home/atoz/letterv.html> Link: I
- 139 <http://www.stanford.edu/home/atoz/letterw.html> Link: I
- 140 <http://www.stanford.edu/home/atoz/letterx.html> Link: I
- 141 <http://www.stanford.edu/home/atoz/lettery.html> Link: I
- 142 <http://www.stanford.edu/home/atoz/letterz.html> Link: I
- 143 <http://news.stanford.edu> Link: X
- 144 <http://news-service.stanford.edu/news/2007/april25/itunes-042507.html> Link: X
- 145 <http://news-service.stanford.edu/news/2007/april25/kidcare-042507.html> Link: X

- 146 <http://news.stanford.edu> Link: X
- 147 <http://events.stanford.edu> Link: X
- 148 <https://events.stanford.edu/events/108/10861> Link: X
- 149 <http://events.stanford.edu> Link: X
- 150 <http://www.stanford.edu/home/welcome/academics> Link: I
- 151 <http://www.stanford.edu/home/welcome/academics> Link: I
- 152 <http://www.stanford.edu/home/otsw> Link: I
- 153 <http://www.stanford.edu/home/atoz/contact.html> Link: I
- 154 <http://www.stanford.edu/home/directories> Link: I
- 155 <http://www.stanford.edu/home/visitors/maps.html> Link: I
- 156 <http://givingtostanford.stanford.edu> Link: X
- 157 <http://www.stanford.edu/home/atoz/terms.html> Link: I
- 158 <http://www.stanford.edu/home/atoz/copyright.html> Link: I
- จำนวนการเชื่อมโยงภายใน : 105, จำนวนการเชื่อมโยงภายนอก : 53

เว็บเพจที่2 : <http://www.stanford.edu/home/academics/schools.html>



รูปที่ ๑.2 เว็บเพจที่ 2

Total Links(Each) : 101

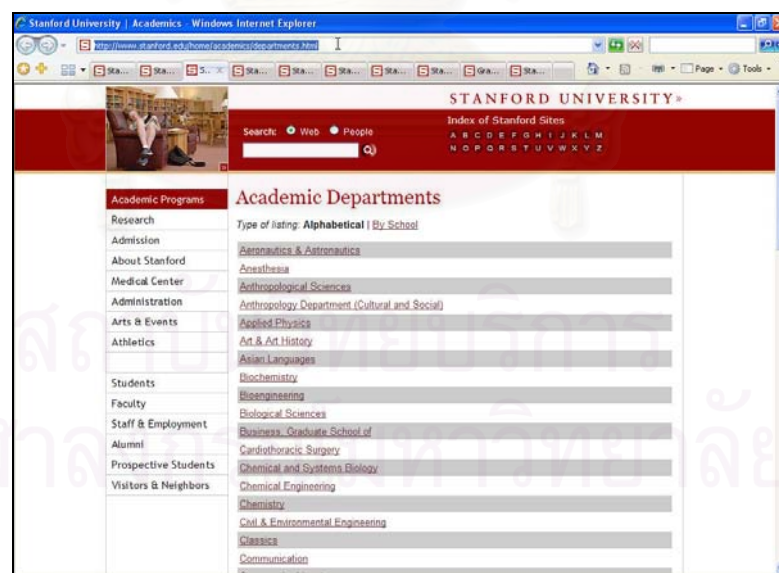
- 1 <http://www.stanford.edu> Link: I
- 2 <http://www.stanford.edu/home/atoz> Link: I
- 3 <http://www.stanford.edu/home/atoz/lettera.html> Link: I
- 4 <http://www.stanford.edu/home/atoz/letterb.html> Link: I
- 5 <http://www.stanford.edu/home/atoz/letterc.html> Link: I
- 6 <http://www.stanford.edu/home/atoz/letterd.html> Link: I
- 7 <http://www.stanford.edu/home/atoz/lettere.html> Link: I
- 8 <http://www.stanford.edu/home/atoz/letterf.html> Link: I
- 9 <http://www.stanford.edu/home/atoz/letterg.html> Link: I
- 10 <http://www.stanford.edu/home/atoz/letterh.html> Link: I
- 11 <http://www.stanford.edu/home/atoz/letteri.html> Link: I
- 12 <http://www.stanford.edu/home/atoz/letterj.html> Link: I
- 13 <http://www.stanford.edu/home/atoz/letterk.html> Link: I
- 14 <http://www.stanford.edu/home/atoz/letterl.html> Link: I
- 15 <http://www.stanford.edu/home/atoz/letterm.html> Link: I
- 16 <http://www.stanford.edu/home/atoz/lettern.html> Link: I
- 17 <http://www.stanford.edu/home/atoz/lettero.html> Link: I
- 18 <http://www.stanford.edu/home/atoz/letterp.html> Link: I
- 19 <http://www.stanford.edu/home/atoz/letterq.html> Link: I
- 20 <http://www.stanford.edu/home/atoz/letterr.html> Link: I
- 21 <http://www.stanford.edu/home/atoz/letters.html> Link: I
- 22 <http://www.stanford.edu/home/atoz/lettert.html> Link: I
- 23 <http://www.stanford.edu/home/atoz/letteru.html> Link: I
- 24 <http://www.stanford.edu/home/atoz/letterv.html> Link: I
- 25 <http://www.stanford.edu/home/atoz/letterw.html> Link: I
- 26 <http://www.stanford.edu/home/atoz/letterx.html> Link: I
- 27 <http://www.stanford.edu/home/atoz/lettery.html> Link: I
- 28 <http://www.stanford.edu/home/atoz/letterz.html> Link: I
- 29 <http://www.stanford.edu/home/slideshow/slideshow23.html> Link: I

- 30 <http://www.stanford.edu/home/academics> Link: I
- 31 <http://www.stanford.edu/home/research> Link: I
- 32 <http://www.stanford.edu/home/admission> Link: I
- 33 <http://www.stanford.edu/home/stanford> Link: I
- 34 <http://www.stanford.edu/home/medcenter> Link: I
- 35 <http://www.stanford.edu/home/administration> Link: I
- 36 <http://www.stanford.edu/home/arts> Link: I
- 37 <http://www.stanford.edu/home/athletics> Link: I
- 38 <http://www.stanford.edu/home/students> Link: I
- 39 <http://www.stanford.edu/home/faculty> Link: I
- 40 <http://www.stanford.edu/home/working> Link: I
- 41 <http://www.stanford.edu/home/alumni> Link: I
- 42 <http://www.stanford.edu/home/prospective> Link: I
- 43 <http://www.stanford.edu/home/visitors> Link: I
- 44 <http://www.stanford.edu> Link: I
- 45 <http://www.gsb.stanford.edu> Link: X
- 46 <http://www.gsb.stanford.edu/mba> Link: X
- 47 <http://www.gsb.stanford.edu/phd> Link: X
- 48 <http://www.gsb.stanford.edu/sloan> Link: X
- 49 <http://www.gsb.stanford.edu/apply> Link: X
- 50 <https://alumni.gsb.stanford.edu> Link: X
- 51 <http://www.gsb.stanford.edu/exed> Link: X
- 52 <https://gsbapps.stanford.edu/facultybios> Link: X
- 53 <http://www.gsb.stanford.edu/research> Link: X
- 54 <http://pangea.stanford.edu> Link: X
- 55 <http://pangea.stanford.edu/academic/undergrad.html> Link: X
- 56 <http://pangea.stanford.edu/academic/grad.html> Link: X
- 57 <http://pangea.stanford.edu/academic/index.html> Link: X
- 58 <http://pangea.stanford.edu/alumni/index.html> Link: X
- 59 <http://pangea.stanford.edu/people/index.html> Link: X

- 60 <http://pangea.stanford.edu/research/index.html> Link: X
- 61 <http://ed.stanford.edu> Link: X
- 62 <http://ed.stanford.edu/suse/programs-degrees/index.html> Link: X
- 63 <http://ed.stanford.edu/suse/admissions/index.html> Link: X
- 64 <http://ed.stanford.edu/suse/alumni/index.html> Link: X
- 65 <http://ed.stanford.edu/suse/faculty/index.html> Link: X
- 66 <http://ed.stanford.edu/suse/faculty/faculty-index-research.html> Link: X
- 67 <http://soe.stanford.edu> Link: X
- 68 http://soe.stanford.edu/prospective_students/under_apply.html Link: X
- 69 http://soe.stanford.edu/prospective_students/grad_apply.html Link: X
- 70 <http://soe.stanford.edu/admissions> Link: X
- 71 <http://soe.stanford.edu/alumni> Link: X
- 72 http://soe.stanford.edu/professional_education Link: X
- 73 <http://soe.stanford.edu/research> Link: X
- 74 <http://www.stanford.edu/dept/humsci> Link: I
- 75 <http://www.stanford.edu/dept/humsci/external/under> Link: I
- 76 <http://www.stanford.edu/dept/humsci/external/grad> Link: I
- 77 <http://www.stanford.edu/dept/humsci/external/grad/admissions.html> Link: I
- 78 <http://www.stanford.edu/dept/humsci/external/departments> Link: I
- 79 <http://www.stanford.edu/dept/humsci/external/faculty> Link: I
- 80 <http://www.stanford.edu/dept/humsci/external/research> Link: I
- 81 <http://www.law.stanford.edu> Link: X
- 82 <http://www.law.stanford.edu/program/courses> Link: X
- 83 <http://www.law.stanford.edu/prospective> Link: X
- 84 <http://www.law.stanford.edu/alumni> Link: X
- 85 <http://www.law.stanford.edu/execed> Link: X
- 86 <http://www.law.stanford.edu/faculty> Link: X
- 87 <http://med.stanford.edu> Link: X
- 88 <http://med.stanford.edu/admissions> Link: X
- 89 <http://med.stanford.edu/alumni> Link: X

- 90 <http://med.stanford.edu/community> Link: X
- 91 <http://med.stanford.edu/departments> Link: X
- 92 <http://med.stanford.edu/education> Link: X
- 93 <http://med.stanford.edu/faculty> Link: X
- 94 http://med.stanford.edu/patient_care Link: X
- 95 <http://med.stanford.edu/research> Link: X
- 96 <http://www.stanford.edu/home/atoz/contact.html> Link: I
- 97 <http://www.stanford.edu/home/directories> Link: I
- 98 <http://www.stanford.edu/home/visitors> Link: I
- 99 <http://givingtostanford.stanford.edu> Link: X
- 100 <http://www.stanford.edu/home/atoz/terms.html> Link: I
- 101 <http://www.stanford.edu/home/atoz/copyright.html> Link: I
- จำนวนการเชื่อมโยงภายใน : 56, จำนวนการเชื่อมโยงภายนอก : 45

เว็บเพจที่3 : <http://www.stanford.edu/home/academics/departments.html>



รูปที่ ๑.3 เว็บเพจที่ 3

[Level of page = 3]

Total Links(Each) : 120

- 1 <http://www.stanford.edu> Link: I
- 2 <http://www.stanford.edu/home/atoz> Link: I
- 3 <http://www.stanford.edu/home/atoz/lettera.html> Link: I
- 4 <http://www.stanford.edu/home/atoz/letterb.html> Link: I
- 5 <http://www.stanford.edu/home/atoz/letterc.html> Link: I
- 6 <http://www.stanford.edu/home/atoz/letterd.html> Link: I
- 7 <http://www.stanford.edu/home/atoz/lettere.html> Link: I
- 8 <http://www.stanford.edu/home/atoz/letterf.html> Link: I
- 9 <http://www.stanford.edu/home/atoz/letterg.html> Link: I
- 10 <http://www.stanford.edu/home/atoz/letterh.html> Link: I
- 11 <http://www.stanford.edu/home/atoz/letteri.html> Link: I
- 12 <http://www.stanford.edu/home/atoz/letterj.html> Link: I
- 13 <http://www.stanford.edu/home/atoz/letterk.html> Link: I
- 14 <http://www.stanford.edu/home/atoz/letterl.html> Link: I
- 15 <http://www.stanford.edu/home/atoz/letterm.html> Link: I
- 16 <http://www.stanford.edu/home/atoz/lettern.html> Link: I
- 17 <http://www.stanford.edu/home/atoz/lettero.html> Link: I
- 18 <http://www.stanford.edu/home/atoz/letterp.html> Link: I
- 19 <http://www.stanford.edu/home/atoz/letterq.html> Link: I
- 20 <http://www.stanford.edu/home/atoz/letterr.html> Link: I
- 21 <http://www.stanford.edu/home/atoz/letters.html> Link: I
- 22 <http://www.stanford.edu/home/atoz/lettert.html> Link: I
- 23 <http://www.stanford.edu/home/atoz/letteru.html> Link: I
- 24 <http://www.stanford.edu/home/atoz/letterv.html> Link: I
- 25 <http://www.stanford.edu/home/atoz/letterw.html> Link: I
- 26 <http://www.stanford.edu/home/atoz/letterx.html> Link: I
- 27 <http://www.stanford.edu/home/atoz/lettery.html> Link: I
- 28 <http://www.stanford.edu/home/atoz/letterz.html> Link: I
- 29 <http://www.stanford.edu/home/slideshow/slideshow23.html> Link: I
- 30 <http://www.stanford.edu/home/academics> Link: I

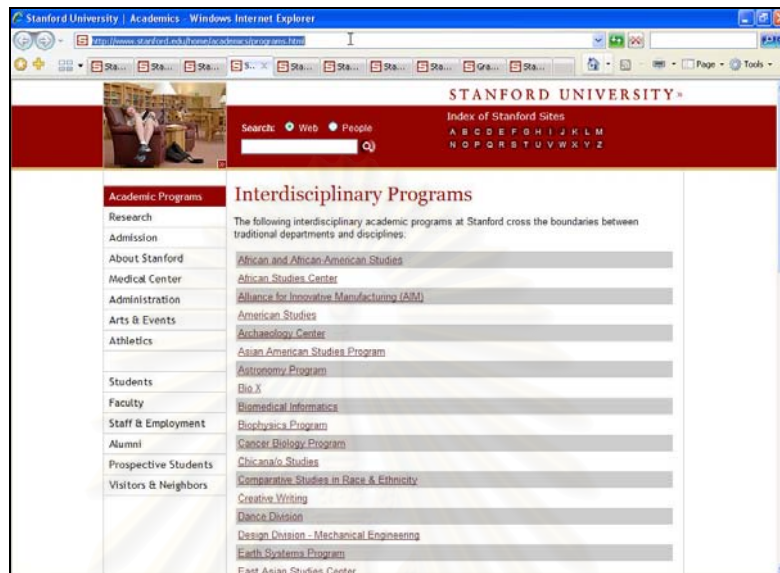
- 31 <http://www.stanford.edu/home/research> Link: I
- 32 <http://www.stanford.edu/home/admission> Link: I
- 33 <http://www.stanford.edu/home/stanford> Link: I
- 34 <http://www.stanford.edu/home/medcenter> Link: I
- 35 <http://www.stanford.edu/home/administration> Link: I
- 36 <http://www.stanford.edu/home/arts> Link: I
- 37 <http://www.stanford.edu/home/athletics> Link: I
- 38 <http://www.stanford.edu/home/students> Link: I
- 39 <http://www.stanford.edu/home/faculty> Link: I
- 40 <http://www.stanford.edu/home/working> Link: I
- 41 <http://www.stanford.edu/home/alumni> Link: I
- 42 <http://www.stanford.edu/home/prospective> Link: I
- 43 <http://www.stanford.edu/home/visitors> Link: I
- 44 <http://www.stanford.edu/home/academics/departments2.html> Link: I
- 45 <http://aa.Stanford.EDU> Link: X
- 46 <http://med.stanford.edu/anesthesia> Link: X
- 47 <http://www.stanford.edu/dept/anthsci> Link: I
- 48 <http://www.stanford.edu/dept/anthroCASA> Link: I
- 49 <http://www.stanford.edu/dept/app-physics> Link: I
- 50 <http://www.stanford.edu/dept/art> Link: I
- 51 <http://www.stanford.edu/dept/asianlang> Link: I
- 52 <http://biochemistry.stanford.edu> Link: X
- 53 <http://bioengineering.stanford.edu> Link: X
- 54 <http://www.stanford.edu/dept/biology> Link: I
- 55 <http://www.gsb.stanford.edu> Link: X
- 56 <http://ctsurgery.stanford.edu> Link: X
- 57 <http://molepharm.stanford.edu> Link: X
- 58 <http://chemeng.stanford.edu> Link: X
- 59 <http://www.stanford.edu/dept/chemistry> Link: I
- 60 <http://www.stanford.edu/dept/cee> Link: I

- 61 <http://www.stanford.edu/dept/classics> Link: I
- 62 <http://communication.stanford.edu> Link: X
- 63 <http://www.stanford.edu/dept/complit> Link: I
- 64 <http://med.stanford.edu/compmed> Link: X
- 65 <http://cs.stanford.edu> Link: X
- 66 <http://dermatology.stanford.edu> Link: X
- 67 <http://devbio.stanford.edu> Link: X
- 68 <http://www.stanford.edu/dept/drama> Link: I
- 69 <http://www-econ.stanford.edu> Link: X
- 70 <http://ed.stanford.edu> Link: X
- 71 <http://ee.stanford.edu> Link: X
- 72 <http://pangea.stanford.edu/ERE> Link: X
- 73 <http://english.stanford.edu> Link: X
- 74 <http://www.stanford.edu/dept/fren-ital> Link: I
- 75 <http://genetics.stanford.edu> Link: X
- 76 <http://pangea.stanford.edu/GES/GES.html> Link: X
- 77 <http://pangea.stanford.edu/GP> Link: X
- 78 <http://www.stanford.edu/dept/german> Link: I
- 79 <http://hrp.stanford.edu> Link: X
- 80 <http://history.stanford.edu> Link: X
- 81 <http://www.law.stanford.edu> Link: X
- 82 <http://www-linguistics.stanford.edu> Link: X
- 83 <http://www.stanford.edu/dept/MSandE> Link: I
- 84 <http://mse.stanford.edu> Link: X
- 85 <http://math.stanford.edu> Link: X
- 86 <http://me.stanford.edu> Link: X
- 87 <http://medicine.stanford.edu> Link: X
- 88 <http://cmgm.stanford.edu/micro> Link: X
- 89 <http://mcp.stanford.edu> Link: X
- 90 <http://music.stanford.edu> Link: X

- 91 <http://www.stanford.edu/dept/nbio> Link: I
- 92 <http://med.stanford.edu/neurology> Link: X
- 93 <http://neurosurgery.stanford.edu> Link: X
- 94 <http://obgyn.stanford.edu> Link: X
- 95 <http://ophthalmology.stanford.edu> Link: X
- 96 <http://ortho.stanford.edu> Link: X
- 97 <http://med.stanford.edu/ohns> Link: X
- 98 <http://pathology.stanford.edu> Link: X
- 99 <http://pediatrics.stanford.edu> Link: X
- 100 <http://www-philosophy.stanford.edu> Link: X
- 101 <http://www.stanford.edu/dept/physics> Link: I
- 102 <http://politicalscience.stanford.edu> Link: X
- 103 <http://psychiatry.stanford.edu> Link: X
- 104 <http://www-psych.stanford.edu> Link: X
- 105 <http://radonc.stanford.edu> Link: X
- 106 <http://radiology.stanford.edu> Link: X
- 107 <http://www.stanford.edu/dept/relstud> Link: I
- 108 <http://www.stanford.edu/dept/slavic> Link: I
- 109 <http://www.stanford.edu/dept/soc> Link: I
- 110 <http://www.stanford.edu/dept/span-port> Link: I
- 111 <http://www-stat.stanford.edu> Link: X
- 112 <http://med.stanford.edu/school/structuralbio> Link: X
- 113 <http://surgery.stanford.edu> Link: X
- 114 <http://urology.stanford.edu> Link: X
- 115 <http://www.stanford.edu/home/atoz/contact.html> Link: I
- 116 <http://www.stanford.edu/home/directories> Link: I
- 117 <http://www.stanford.edu/home/visitors> Link: I
- 118 <http://givingtostanford.stanford.edu> Link: X
- 119 <http://www.stanford.edu/home/atoz/terms.html> Link: I
- 120 <http://www.stanford.edu/home/atoz/copyright.html> Link: I

จำนวนการเชื่อมโยงภายใน : 69, จำนวนการเชื่อมโยงภายนอก : 51

เว็บเพจที่ 4 : <http://www.stanford.edu/home/academics/programs.html>



รูปที่ ๑.4 เว็บเพจที่ 4

[Level of page = 3]

Total Links(Each) : 107

- 1 <http://www.stanford.edu> Link: 1
- 2 <http://www.stanford.edu/home/atoz> Link: 1
- 3 <http://www.stanford.edu/home/atoz/lettera.html> Link: 1
- 4 <http://www.stanford.edu/home/atoz/letterb.html> Link: 1
- 5 <http://www.stanford.edu/home/atoz/letterc.html> Link: 1
- 6 <http://www.stanford.edu/home/atoz/letterd.html> Link: 1
- 7 <http://www.stanford.edu/home/atoz/lettere.html> Link: 1
- 8 <http://www.stanford.edu/home/atoz/letterf.html> Link: 1
- 9 <http://www.stanford.edu/home/atoz/letterg.html> Link: 1
- 10 <http://www.stanford.edu/home/atoz/letterh.html> Link: 1
- 11 <http://www.stanford.edu/home/atoz/letteri.html> Link: 1
- 12 <http://www.stanford.edu/home/atoz/letterj.html> Link: 1
- 13 <http://www.stanford.edu/home/atoz/letterk.html> Link: 1

- 14 <http://www.stanford.edu/home/atoz/letterl.html> Link: I
- 15 <http://www.stanford.edu/home/atoz/letterm.html> Link: I
- 16 <http://www.stanford.edu/home/atoz/lettern.html> Link: I
- 17 <http://www.stanford.edu/home/atoz/lettero.html> Link: I
- 18 <http://www.stanford.edu/home/atoz/letterp.html> Link: I
- 19 <http://www.stanford.edu/home/atoz/letterq.html> Link: I
- 20 <http://www.stanford.edu/home/atoz/letterr.html> Link: I
- 21 <http://www.stanford.edu/home/atoz/letters.html> Link: I
- 22 <http://www.stanford.edu/home/atoz/lettert.html> Link: I
- 23 <http://www.stanford.edu/home/atoz/letteru.html> Link: I
- 24 <http://www.stanford.edu/home/atoz/letterv.html> Link: I
- 25 <http://www.stanford.edu/home/atoz/letterw.html> Link: I
- 26 <http://www.stanford.edu/home/atoz/letterx.html> Link: I
- 27 <http://www.stanford.edu/home/atoz/lettery.html> Link: I
- 28 <http://www.stanford.edu/home/atoz/letterz.html> Link: I
- 29 <http://www.stanford.edu/home/slideshow/slideshow23.html> Link: I
- 30 <http://www.stanford.edu/home/academics> Link: I
- 31 <http://www.stanford.edu/home/research> Link: I
- 32 <http://www.stanford.edu/home/admission> Link: I
- 33 <http://www.stanford.edu/home/stanford> Link: I
- 34 <http://www.stanford.edu/home/medcenter> Link: I
- 35 <http://www.stanford.edu/home/administration> Link: I
- 36 <http://www.stanford.edu/home/arts> Link: I
- 37 <http://www.stanford.edu/home/athletics> Link: I
- 38 <http://www.stanford.edu/home/students> Link: I
- 39 <http://www.stanford.edu/home/faculty> Link: I
- 40 <http://www.stanford.edu/home/working> Link: I
- 41 <http://www.stanford.edu/home/alumni> Link: I
- 42 <http://www.stanford.edu/home/prospective> Link: I
- 43 <http://www.stanford.edu/home/visitors> Link: I

- 44 <http://www.stanford.edu/dept/AAAS> Link: I
- 45 <http://ica.stanford.edu/?q=afr> Link: X
- 46 <http://www.stanford.edu/group/AIM> Link: I
- 47 <http://www.stanford.edu/group/HSP/AmStud> Link: I
- 48 <http://archaeology.stanford.edu> Link: X
- 49 <http://www.stanford.edu/dept/AAS> Link: I
- 50 <http://www.stanford.edu/dept/astro> Link: I
- 51 <http://biox.stanford.edu> Link: X
- 52 <http://bmi.stanford.edu> Link: X
- 53 <http://med.stanford.edu/biophysics> Link: X
- 54 <http://www.stanford.edu/group/cancerbio> Link: I
- 55 http://ccsre.stanford.edu/UE_maj_Chicano.htm Link: X
- 56 <http://ccsre.stanford.edu> Link: X
- 57 <http://www.stanford.edu/dept/english/cw> Link: I
- 58 <http://www.stanford.edu/dept/dance> Link: I
- 59 <http://design.stanford.edu> Link: X
- 60 <http://pangea.stanford.edu/ESYS> Link: X
- 61 <http://ceas.stanford.edu> Link: X
- 62 <http://www.stanford.edu/group/efs> Link: I
- 63 <http://iper.stanford.edu> Link: X
- 64 <http://www.stanford.edu/dept/EIS> Link: I
- 65 <http://www.stanford.edu/dept/femstud> Link: I
- 66 <http://finmath.stanford.edu> Link: X
- 67 <http://HPST.stanford.edu> Link: X
- 68 <http://www.stanford.edu/dept/humbio> Link: I
- 69 <http://hci.stanford.edu> Link: X
- 70 <http://scien.stanford.edu> Link: X
- 71 <http://smi.stanford.edu> Link: X
- 72 <http://icme.stanford.edu> Link: X
- 73 <http://www.stanford.edu/group/HSP/GPH> Link: I

- 74 <http://ica.stanford.edu/?q=ips> Link: X
- 75 <http://ica.stanford.edu/?q=internationalrelations> Link: X
- 76 <http://ihum.stanford.edu> Link: X
- 77 http://www.stanford.edu/dept/islamic_studies Link: I
- 78 <http://www.stanford.edu/group/las> Link: I
- 79 <http://www.stanford.edu/dept/SUSE/ldt> Link: I
- 80 <http://www.stanford.edu/group/mathcompsci> Link: I
- 81 <http://www.stanford.edu/dept/medieval> Link: I
- 82 <http://www.stanford.edu/dept/MTL> Link: I
- 83 <http://cmgm.stanford.edu> Link: X
- 84 http://ccsre.stanford.edu/UE_maj_NativeAmer.htm Link: X
- 85 http://neuroscience.stanford.edu/education/phd_program/program.html Link: X
- 86 <http://osp.stanford.edu> Link: X
- 87 <http://pcap.stanford.edu> Link: X
- 88 <http://design.stanford.edu/PD> Link: X
- 89 <http://www.gsb.stanford.edu/pmp> Link: X
- 90 <http://www.stanford.edu/dept/publicpolicy> Link: I
- 91 <http://pulmonary.stanford.edu> Link: X
- 92 <http://www.stanford.edu/dept/CREES> Link: I
- 93 <http://www.stanford.edu/group/STS> Link: I
- 94 <http://www-sccm.stanford.edu> Link: X
- 95 <http://ed.stanford.edu/suse/programs-degrees/sssep.html> Link: X
- 96 <http://www.stanford.edu/dept/undergrad/areaone/14des.shtml> Link: I
- 97 <http://symsys.stanford.edu> Link: X
- 98 <http://www.stanford.edu/dept/jewishstudies> Link: I
- 99 <http://suse-step.stanford.edu> Link: X
- 100 <http://stvp.stanford.edu> Link: X
- 101 <http://www.stanford.edu/dept/URBS> Link: I
- 102 <http://www.stanford.edu/home/atoz/contact.html> Link: I
- 103 <http://www.stanford.edu/home/directories> Link: I

104 <http://www.stanford.edu/home/visitors> Link: I

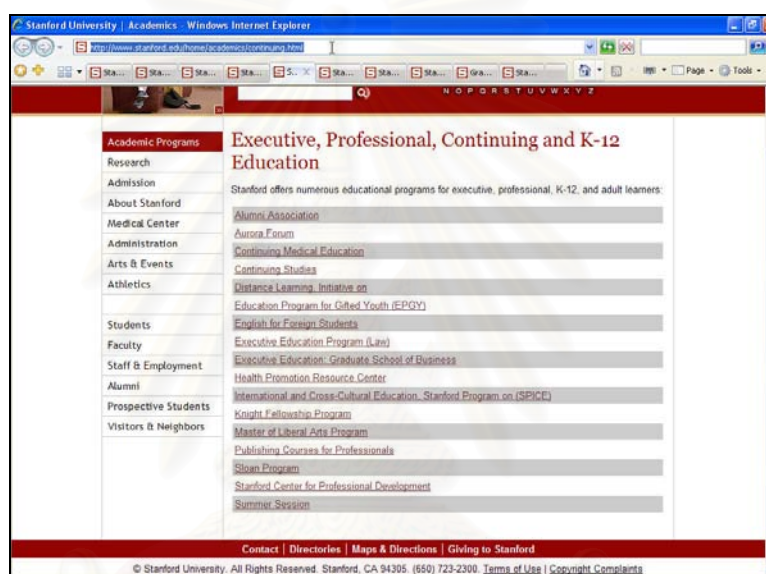
105 <http://givingtostanford.stanford.edu> Link: X

106 <http://www.stanford.edu/home/atoz/terms.html> Link: I

107 <http://www.stanford.edu/home/atoz/copyright.html> Link: I

จำนวนการเชื่อมโยงภายใน : 73, จำนวนการเชื่อมโยงภายนอก : 34

เว็บเพจที่5 : <http://www.stanford.edu/home/academics/continuing.html>



รูปที่ ๑.5 เว็บเพจที่ 5

[Level of page = 3]

Total Links(Each) : 66

1 <http://www.stanford.edu> Link: I

2 <http://www.stanford.edu/home/atoz> Link: I

3 <http://www.stanford.edu/home/atoz/lettera.html> Link: I

4 <http://www.stanford.edu/home/atoz/letterb.html> Link: I

5 <http://www.stanford.edu/home/atoz/letterc.html> Link: I

6 <http://www.stanford.edu/home/atoz/letterd.html> Link: I

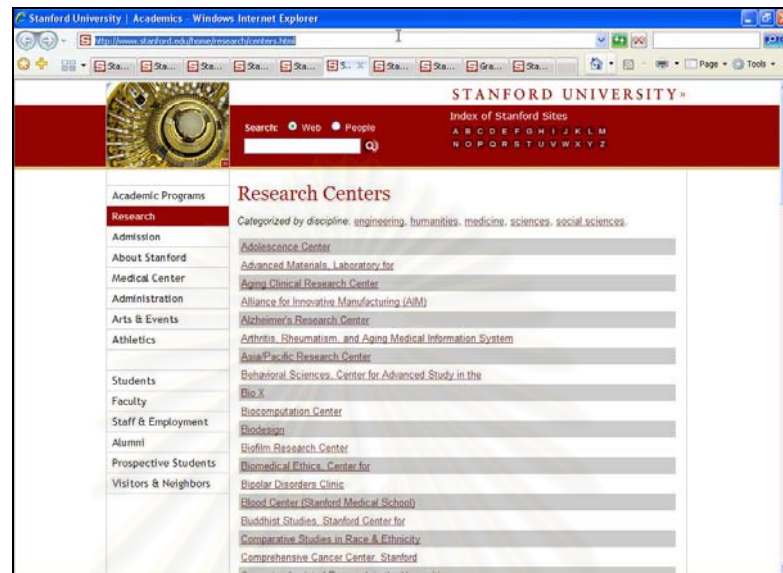
7 <http://www.stanford.edu/home/atoz/lettere.html> Link: I

8 <http://www.stanford.edu/home/atoz/letterf.html> Link: I

- 9 <http://www.stanford.edu/home/atoz/letterg.html> Link: I
- 10 <http://www.stanford.edu/home/atoz/letterh.html> Link: I
- 11 <http://www.stanford.edu/home/atoz/letteri.html> Link: I
- 12 <http://www.stanford.edu/home/atoz/letterj.html> Link: I
- 13 <http://www.stanford.edu/home/atoz/letterk.html> Link: I
- 14 <http://www.stanford.edu/home/atoz/letterl.html> Link: I
- 15 <http://www.stanford.edu/home/atoz/letterm.html> Link: I
- 16 <http://www.stanford.edu/home/atoz/lettern.html> Link: I
- 17 <http://www.stanford.edu/home/atoz/lettero.html> Link: I
- 18 <http://www.stanford.edu/home/atoz/letterp.html> Link: I
- 19 <http://www.stanford.edu/home/atoz/letterq.html> Link: I
- 20 <http://www.stanford.edu/home/atoz/letterr.html> Link: I
- 21 <http://www.stanford.edu/home/atoz/letters.html> Link: I
- 22 <http://www.stanford.edu/home/atoz/lettert.html> Link: I
- 23 <http://www.stanford.edu/home/atoz/letteru.html> Link: I
- 24 <http://www.stanford.edu/home/atoz/letterv.html> Link: I
- 25 <http://www.stanford.edu/home/atoz/letterw.html> Link: I
- 26 <http://www.stanford.edu/home/atoz/letterx.html> Link: I
- 27 <http://www.stanford.edu/home/atoz/lettery.html> Link: I
- 28 <http://www.stanford.edu/home/atoz/letterz.html> Link: I
- 29 <http://www.stanford.edu/home/slideshow/slideshow23.html> Link: I
- 30 <http://www.stanford.edu/home/academics> Link: I
- 31 <http://www.stanford.edu/home/research> Link: I
- 32 <http://www.stanford.edu/home/admission> Link: I
- 33 <http://www.stanford.edu/home/stanford> Link: I
- 34 <http://www.stanford.edu/home/medcenter> Link: I
- 35 <http://www.stanford.edu/home/administration> Link: I
- 36 <http://www.stanford.edu/home/arts> Link: I
- 37 <http://www.stanford.edu/home/athletics> Link: I
- 38 <http://www.stanford.edu/home/students> Link: I

- 39 <http://www.stanford.edu/home/faculty> Link: I
- 40 <http://www.stanford.edu/home/working> Link: I
- 41 <http://www.stanford.edu/home/alumni> Link: I
- 42 <http://www.stanford.edu/home/prospective> Link: I
- 43 <http://www.stanford.edu/home/visitors> Link: I
- 44 <http://www.stanfordalumni.org> Link: X
- 45 <http://auroraforum.stanford.edu> Link: X
- 46 <http://cme.stanford.edu> Link: X
- 47 <http://continuingstudies.stanford.edu> Link: X
- 48 <http://idl.stanford.edu> Link: X
- 49 <http://epgy.stanford.edu> Link: X
- 50 <http://www.stanford.edu/group/efs> Link: I
- 51 <http://lawschool.stanford.edu/execed> Link: X
- 52 <http://www.gsb.stanford.edu/exed> Link: X
- 53 <http://hprc.stanford.edu> Link: X
- 54 <http://spice.stanford.edu> Link: X
- 55 <http://knightfellows.stanford.edu> Link: X
- 56 <http://mla.stanford.edu> Link: X
- 57 <http://publishingcourses.stanford.edu> Link: X
- 58 <http://www.gsb.stanford.edu/sloan> Link: X
- 59 <http://scpd.stanford.edu> Link: X
- 60 <http://summersession.stanford.edu> Link: X
- 61 <http://www.stanford.edu/home/atoz/contact.html> Link: I
- 62 <http://www.stanford.edu/home/directories> Link: I
- 63 <http://www.stanford.edu/home/visitors> Link: I
- 64 <http://givingtostanford.stanford.edu> Link: X
- 65 <http://www.stanford.edu/home/atoz/terms.html> Link: I
- 66 <http://www.stanford.edu/home/atoz/copyright.html> Link: I
- จำนวนการเชื่อมโยงภายใน : 49, จำนวนการเชื่อมโยงภายนอก : 17

เว็บเพจที่ 6 : <http://www.stanford.edu/home/research/centers.html>



รูปที่ ๑.6 เว็บเพจที่ 6

[Level of page = 3]

Total Links(Each) : 155

- 1 <http://www.stanford.edu> Link: 1
- 2 <http://www.stanford.edu/home/atoz> Link: 1
- 3 <http://www.stanford.edu/home/atoz/lettera.html> Link: 1
- 4 <http://www.stanford.edu/home/atoz/letterb.html> Link: 1
- 5 <http://www.stanford.edu/home/atoz/letterc.html> Link: 1
- 6 <http://www.stanford.edu/home/atoz/letterd.html> Link: 1
- 7 <http://www.stanford.edu/home/atoz/lettere.html> Link: 1
- 8 <http://www.stanford.edu/home/atoz/letterf.html> Link: 1
- 9 <http://www.stanford.edu/home/atoz/letterg.html> Link: 1
- 10 <http://www.stanford.edu/home/atoz/letterh.html> Link: 1
- 11 <http://www.stanford.edu/home/atoz/letteri.html> Link: 1
- 12 <http://www.stanford.edu/home/atoz/letterj.html> Link: 1
- 13 <http://www.stanford.edu/home/atoz/letterk.html> Link: 1

- 14 <http://www.stanford.edu/home/atoz/letterl.html> Link: I
- 15 <http://www.stanford.edu/home/atoz/letterm.html> Link: I
- 16 <http://www.stanford.edu/home/atoz/lettern.html> Link: I
- 17 <http://www.stanford.edu/home/atoz/lettero.html> Link: I
- 18 <http://www.stanford.edu/home/atoz/letterp.html> Link: I
- 19 <http://www.stanford.edu/home/atoz/letterq.html> Link: I
- 20 <http://www.stanford.edu/home/atoz/letterr.html> Link: I
- 21 <http://www.stanford.edu/home/atoz/letters.html> Link: I
- 22 <http://www.stanford.edu/home/atoz/lettert.html> Link: I
- 23 <http://www.stanford.edu/home/atoz/letteru.html> Link: I
- 24 <http://www.stanford.edu/home/atoz/letterv.html> Link: I
- 25 <http://www.stanford.edu/home/atoz/letterw.html> Link: I
- 26 <http://www.stanford.edu/home/atoz/letterx.html> Link: I
- 27 <http://www.stanford.edu/home/atoz/lettery.html> Link: I
- 28 <http://www.stanford.edu/home/atoz/letterz.html> Link: I
- 29 <http://www.stanford.edu/home/slideshow/slideshow12.html> Link: I
- 30 <http://www.stanford.edu/home/academics> Link: I
- 31 <http://www.stanford.edu/home/research> Link: I
- 32 <http://www.stanford.edu/home/admission> Link: I
- 33 <http://www.stanford.edu/home/stanford> Link: I
- 34 <http://www.stanford.edu/home/medcenter> Link: I
- 35 <http://www.stanford.edu/home/administration> Link: I
- 36 <http://www.stanford.edu/home/arts> Link: I
- 37 <http://www.stanford.edu/home/athletics> Link: I
- 38 <http://www.stanford.edu/home/students> Link: I
- 39 <http://www.stanford.edu/home/faculty> Link: I
- 40 <http://www.stanford.edu/home/working> Link: I
- 41 <http://www.stanford.edu/home/alumni> Link: I
- 42 <http://www.stanford.edu/home/prospective> Link: I
- 43 <http://www.stanford.edu/home/visitors> Link: I

- 44 <http://www.stanford.edu/home/research/centers-eng.html> Link: I
- 45 <http://www.stanford.edu/home/research/centers-hum.html> Link: I
- 46 <http://www.stanford.edu/home/research/centers-med.html> Link: I
- 47 <http://www.stanford.edu/home/research/centers-sci.html> Link: I
- 48 <http://www.stanford.edu/home/research/centers-socsci.html> Link: I
- 49 <http://www.stanford.edu/group/adolescent.ctr> HTML: false Link: I
- 50 <http://www-lam.stanford.edu> Link: X
- 51 <http://www.stanford.edu/~yesavage/ACRC.html> Link: I
- 52 <http://www.stanford.edu/group/AIM> Link: I
- 53 <http://arcc.stanford.edu> Link: X
- 54 <http://aramis.stanford.edu> Link: X
- 55 <http://aparc.stanford.edu> Link: X
- 56 <http://casbs.stanford.edu> Link: X
- 57 <http://biox.stanford.edu> Link: X
- 58 <http://biocomp.stanford.edu> Link: X
- 59 <http://biodesign.stanford.edu> Link: X
- 60 <http://www.stanford.edu/group/biofilm> Link: I
- 61 <http://bioethics.stanford.edu> Link: X
- 62 <http://bipolar.stanford.edu> Link: X
- 63 <http://bloodcenter.stanford.edu> Link: X
- 64 <http://www.stanford.edu/group/scbs> Link: I
- 65 <http://ccsre.stanford.edu> Link: X
- 66 <http://cancer.stanford.edu> Link: X
- 67 <http://www.ccarh.org> Link: X
- 68 <http://graphics.stanford.EDU> Link: X
- 69 <http://www-ccrma.stanford.edu> Link: X
- 70 <http://www.stanford.edu/group/sccn> Link: I
- 71 <http://www.stanford.edu/group/CCB> Link: I
- 72 <http://cdd.stanford.edu> Link: X
- 73 <http://cdr.stanford.edu> Link: X

- 74 <http://prevention.stanford.edu> Link: X
- 75 <http://blume.stanford.edu> Link: X
- 76 <http://siepr.stanford.edu> Link: X
- 77 <http://www.stanford.edu/group/EMF> Link: I
- 78 <http://www.gsb.stanford.edu/ces> Link: X
- 79 <http://www.stanford.edu/group/efml> Link: I
- 80 <http://cesp.stanford.edu> Link: X
- 81 <http://sepwww.stanford.edu> Link: X
- 82 <http://sfdc.stanford.edu> Link: X
- 83 <http://www-evo.stanford.edu> Link: X
- 84 <http://fsi.stanford.edu> Link: X
- 85 <http://www.stanford.edu/group/gender> Link: I
- 86 <http://www.med.stanford.edu/gcrc> Link: X
- 87 <http://www-sequence.stanford.edu> Link: X
- 88 <http://www.stanford.edu/group/ginzton> Link: I
- 89 <http://www.stanford.edu/group/scforum> Link: I
- 90 <http://einstein.stanford.edu> Link: X
- 91 <http://www.stanford.edu/group/hepl> Link: I
- 92 <http://chppcor.stanford.edu> Link: X
- 93 <http://siher.stanford.edu> Link: X
- 94 <http://www-hoover.stanford.edu> Link: X
- 95 <http://www-marine.stanford.edu> Link: X
- 96 <http://www-shgc.stanford.edu> Link: X
- 97 <http://shc.stanford.edu> Link: X
- 98 <http://shl.stanford.edu> Link: X
- 99 <http://isl.stanford.edu> Link: X
- 100 <http://scil.stanford.edu> Link: X
- 101 <http://www.stanford.edu/group/CIFE> Link: I
- 102 <http://cis.stanford.edu> Link: X
- 103 <http://ctr-sgi1.stanford.edu/CITS> Link: X

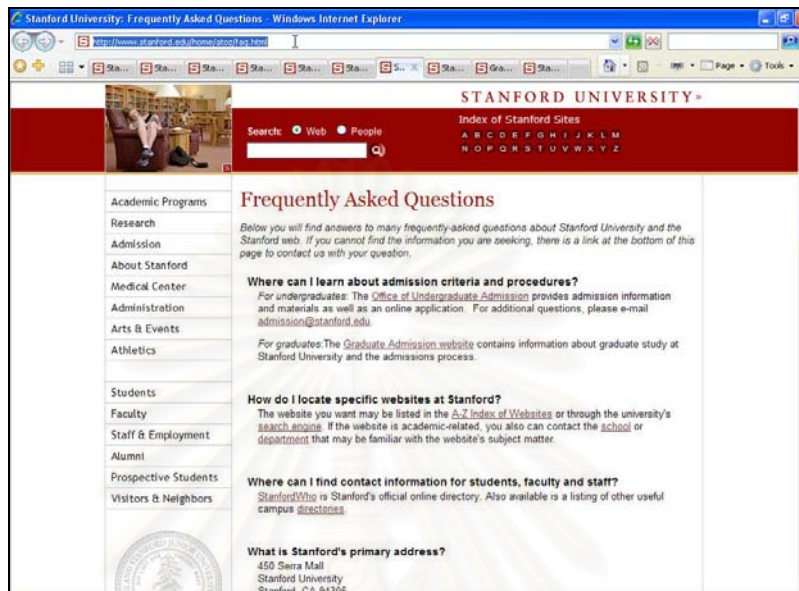
- 104 <http://cisac.stanford.edu> Link: X
- 105 <http://www.stanford-jc.or.jp> Link: X
- 106 <http://jrpb.stanford.edu> Link: X
- 107 <http://www.stanford.edu/group/King> Link: I
- 108 <http://www-csli.stanford.edu> Link: X
- 109 <http://language.stanford.edu> Link: X
- 110 <http://www.stanford.edu/group/las> Link: I
- 111 <http://www.slac.stanford.edu> Link: X
- 112 <http://www.stanford.edu/group/luolab> Link: I
- 113 <http://smrl.stanford.edu> Link: X
- 114 http://rsl.stanford.edu/research/lucas_center.html Link: X
- 115 <http://www.smi.stanford.edu> Link: X
- 116 <http://www.stanford.edu/group/snl> Link: I
- 117 <http://snf.stanford.edu> Link: X
- 118 <http://med.stanford.edu/school/Psychiatry/narcolepsy> Link: X
- 119 <http://nrc.stanford.edu> Link: X
- 120 <http://novel.stanford.edu> Link: X
- 121 <http://patienteducation.stanford.edu> Link: X
- 122 <http://www.stanford.edu/group/SPRC> Link: I
- 123 <http://carnegiedpb.stanford.edu> Link: X
- 124 <http://pcl.stanford.edu> Link: X
- 125 <http://www.stanford.edu/group/CPIMA> Link: I
- 126 <http://www.stanford.edu/group/siqss> Link: I
- 127 <http://www-rpl.stanford.edu> Link: X
- 128 <http://guide.stanford.edu> Link: X
- 129 <http://ekofisk.stanford.edu/SCRF.html> Link: X
- 130 <http://robotics.stanford.edu> Link: X
- 131 <http://www.stanford.edu/dept/CREES> Link: I
- 132 <http://www.scancor.org> Link: X
- 133 <http://med.stanford.edu/school/psychiatry/coe> Link: X

- 134 <http://www.gsb.stanford.edu/csi> Link: X
- 135 <http://www.stanford.edu/group/sshi> Link: I
- 136 <http://iriss.stanford.edu> Link: X
- 137 <http://soi.stanford.edu> Link: X
- 138 <http://www.stanford.edu/group/CSSA> Link: I
- 139 <http://nova.stanford.edu> Link: X
- 140 <http://strokecenter.stanford.edu> Link: X
- 141 <http://summit.stanford.edu> Link: X
- 142 <http://www-ssl.slac.stanford.edu> Link: X
- 143 <http://www.stanford.edu/group/SOL> Link: I
- 144 <http://www.stanford.edu/dept/jewishstudies> Link: I
- 145 <http://stvp.stanford.edu> Link: X
- 146 <http://cmgm.stanford.edu/theriot> Link: X
- 147 <http://ctr.stanford.edu> Link: X
- 148 <http://fuji.stanford.edu> Link: X
- 149 <http://syr.stanford.edu> Link: X
- 150 <http://www.stanford.edu/home/atoz/contact.html> Link: I
- 151 <http://www.stanford.edu/home/directories> Link: I
- 152 <http://www.stanford.edu/home/visitors> Link: I
- 153 <http://givingtostanford.stanford.edu> Link: X
- 154 <http://www.stanford.edu/home/atoz/terms.html> Link: I
- 155 <http://www.stanford.edu/home/atoz/copyright.html> Link: I

จำนวนการเชื่อมโยงภายใน : 79, จำนวนการเชื่อมโยงภายนอก : 76

จุฬาลงกรณ์มหาวิทยาลัย

เว็บเพจที่ 7 : <http://www.stanford.edu/home/atoz/faq.html>



รูปที่ ๑.7 เว็บเพจที่ 7

[Level of page = 3]

Total Links(Each) : 85

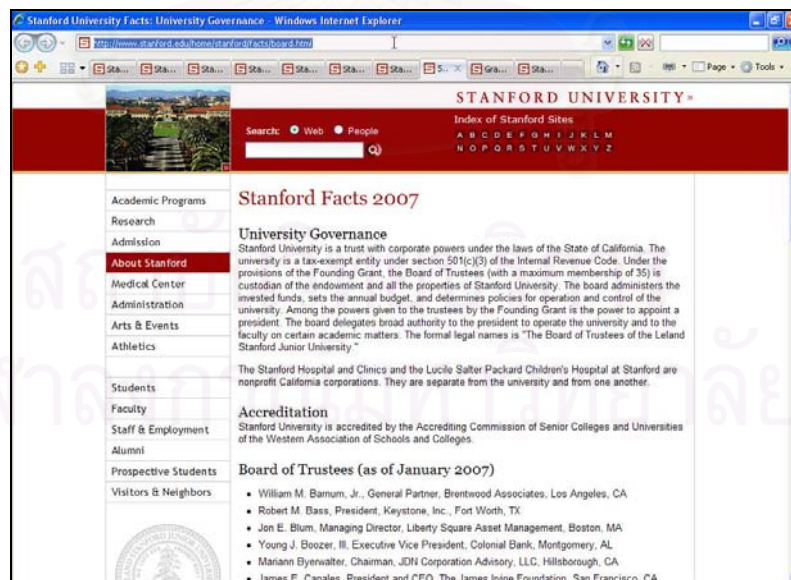
- 1 <http://www.stanford.edu> Link: 1
- 2 <http://www.stanford.edu/home/atoz> Link: 1
- 3 <http://www.stanford.edu/home/atoz/lettera.html> Link: 1
- 4 <http://www.stanford.edu/home/atoz/letterb.html> Link: 1
- 5 <http://www.stanford.edu/home/atoz/letterc.html> Link: 1
- 6 <http://www.stanford.edu/home/atoz/letterd.html> Link: 1
- 7 <http://www.stanford.edu/home/atoz/lettere.html> Link: 1
- 8 <http://www.stanford.edu/home/atoz/letterf.html> Link: 1
- 9 <http://www.stanford.edu/home/atoz/letterg.html> Link: 1
- 10 <http://www.stanford.edu/home/atoz/letterh.html> Link: 1
- 11 <http://www.stanford.edu/home/atoz/letteri.html> Link: 1
- 12 <http://www.stanford.edu/home/atoz/letterj.html> Link: 1
- 13 <http://www.stanford.edu/home/atoz/letterk.html> Link: 1

- 14 <http://www.stanford.edu/home/atoz/letterl.html> Link: I
- 15 <http://www.stanford.edu/home/atoz/letterm.html> Link: I
- 16 <http://www.stanford.edu/home/atoz/lettern.html> Link: I
- 17 <http://www.stanford.edu/home/atoz/lettero.html> Link: I
- 18 <http://www.stanford.edu/home/atoz/letterp.html> Link: I
- 19 <http://www.stanford.edu/home/atoz/letterq.html> Link: I
- 20 <http://www.stanford.edu/home/atoz/letterr.html> Link: I
- 21 <http://www.stanford.edu/home/atoz/letters.html> Link: I
- 22 <http://www.stanford.edu/home/atoz/lettert.html> Link: I
- 23 <http://www.stanford.edu/home/atoz/letteru.html> Link: I
- 24 <http://www.stanford.edu/home/atoz/letterv.html> Link: I
- 25 <http://www.stanford.edu/home/atoz/letterw.html> Link: I
- 26 <http://www.stanford.edu/home/atoz/letterx.html> Link: I
- 27 <http://www.stanford.edu/home/atoz/lettery.html> Link: I
- 28 <http://www.stanford.edu/home/atoz/letterz.html> Link: I
- 29 <http://www.stanford.edu/home/academics> Link: I
- 30 <http://www.stanford.edu/home/research> Link: I
- 31 <http://www.stanford.edu/home/admission> Link: I
- 32 <http://www.stanford.edu/home/stanford> Link: I
- 33 <http://www.stanford.edu/home/medcenter> Link: I
- 34 <http://www.stanford.edu/home/administration> Link: I
- 35 <http://www.stanford.edu/home/arts> Link: I
- 36 <http://www.stanford.edu/home/athletics> Link: I
- 37 <http://www.stanford.edu/home/students> Link: I
- 38 <http://www.stanford.edu/home/faculty> Link: I
- 39 <http://www.stanford.edu/home/working> Link: I
- 40 <http://www.stanford.edu/home/alumni> Link: I
- 41 <http://www.stanford.edu/home/prospective> Link: I
- 42 <http://www.stanford.edu/home/visitors> Link: I
- 43 <http://www.stanford.edu> Link: I

- 44 <http://admission.stanford.edu> Link: X
- 45 <mailto:admission@stanford.edu> HTML: false Link: X
- 46 <http://gradadmissions.stanford.edu> Link: X
- 47 <http://www.stanford.edu/home/atoz/index.html> Link: I
- 48 <http://www.stanford.edu/search> Link: I
- 49 <http://www.stanford.edu/home/academics/schools.html> Link: I
- 50 <http://www.stanford.edu/home/academics/departments.html> Link: I
- 51 <http://stanfordwho.stanford.edu/lookup> Link: X
- 52 <http://www.stanford.edu/home/directories> Link: I
- 53 <http://www.stanford.edu/home/directories> Link: I
- 54 <http://jobs.stanford.edu> Link: X
- 55 <http://www.stanford.edu/home/faculty/positions.html> Link: I
- 56 <http://www.stanfordhospital.com/forPatients/requestAppt/referralCenterConsumerSHC.html> Link: X
- 57 <http://www.med.stanford.edu> Link: X
- 58 <http://gostanford.ocsn.com/school-bio/stan-nickname-mascot.html> Link: X
- 59 <http://www.stanford.edu/home/stanford/history/begin.html#myth> Link: I
- 60 <http://www.stanford.edu/home/stanford/history> Link: I
- 61 <http://summersession.stanford.edu> Link: X
- 62 <http://www-epgy.stanford.edu> Link: X
- 63 <http://summersession.stanford.edu> Link: X
- 64 <http://www-epgy.stanford.edu/summer> Link: X
- 65 <http://www.stanford.edu/group/efs> Link: I
- 66 <http://www.stanford.edu/dept/hds/scs> Link: I
- 67 <http://gostanford.ocsn.com/marketplace/events/events-camps-body.html> Link: X
- 68 <http://www.stanford.edu/dept/hds/scs/youth> Link: I
- 69 <http://scpd.stanford.edu> Link: X
- 70 <http://www-epgy.stanford.edu> Link: X
- 71 <http://continuingstudies.stanford.edu> Link: X
- 72 <http://www.gsb.stanford.edu/exed> Link: X

- 73 <http://www.law.stanford.edu/execed> Link: X
- 74 <http://www.stanford.edu/dept/news/neighbors/learning/children.html> Link: I
- 75 <http://www.stanford.edu/home/admission/continuing.html> Link: I
- 76 <http://www.stanford.edu/group/itss-ccs/security/index.html> Link: I
- 77 <http://www.stanford.edu/home/atoz/copyright.html> Link: I
- 78 <http://helpsu.stanford.edu> Link: X
- 79 <http://www.stanford.edu/home/comment> Link: I
- 80 <http://www.stanford.edu/home/atoz/contact.html> Link: I
- 81 <http://www.stanford.edu/home/directories> Link: I
- 82 <http://www.stanford.edu/home/visitors> Link: I
- 83 <http://givingtostanford.stanford.edu> Link: X
- 84 <http://www.stanford.edu/home/atoz/terms.html> Link: I
- 85 <http://www.stanford.edu/home/atoz/copyright.html> Link: I
- จำนวนการเชื่อมโยงภายใน : 65, จำนวนการเชื่อมโยงภายนอก : 20

เว็บเพจที่ 8 : <http://www.stanford.edu/home/stanford/facts/board.html>



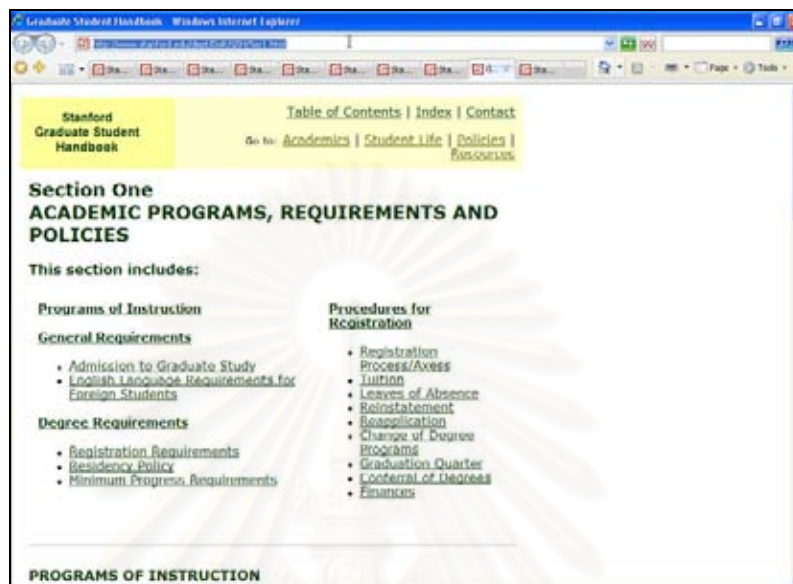
รูปที่ ๑.8 เว็บเพจที่ 8

Total Links(Each) : 52

- 1 <http://www.stanford.edu> Link: I
- 2 <http://www.stanford.edu/home/atoz> Link: I
- 3 <http://www.stanford.edu/home/atoz/lettera.html> Link: I
- 4 <http://www.stanford.edu/home/atoz/letterb.html> Link: I
- 5 <http://www.stanford.edu/home/atoz/letterc.html> Link: I
- 6 <http://www.stanford.edu/home/atoz/letterd.html> Link: I
- 7 <http://www.stanford.edu/home/atoz/lettere.html> Link: I
- 8 <http://www.stanford.edu/home/atoz/letterf.html> Link: I
- 9 <http://www.stanford.edu/home/atoz/letterg.html> Link: I
- 10 <http://www.stanford.edu/home/atoz/letterh.html> Link: I
- 11 <http://www.stanford.edu/home/atoz/letteri.html> Link: I
- 12 <http://www.stanford.edu/home/atoz/letterj.html> Link: I
- 13 <http://www.stanford.edu/home/atoz/letterk.html> Link: I
- 14 <http://www.stanford.edu/home/atoz/letterl.html> Link: I
- 15 <http://www.stanford.edu/home/atoz/letterm.html> Link: I
- 16 <http://www.stanford.edu/home/atoz/lettern.html> Link: I
- 17 <http://www.stanford.edu/home/atoz/lettero.html> Link: I
- 18 <http://www.stanford.edu/home/atoz/letterp.html> Link: I
- 19 <http://www.stanford.edu/home/atoz/letterq.html> Link: I
- 20 <http://www.stanford.edu/home/atoz/letterr.html> Link: I
- 21 <http://www.stanford.edu/home/atoz/letters.html> Link: I
- 22 <http://www.stanford.edu/home/atoz/lettert.html> Link: I
- 23 <http://www.stanford.edu/home/atoz/letteru.html> Link: I
- 24 <http://www.stanford.edu/home/atoz/letterv.html> Link: I
- 25 <http://www.stanford.edu/home/atoz/letterw.html> Link: I
- 26 <http://www.stanford.edu/home/atoz/letterx.html> Link: I
- 27 <http://www.stanford.edu/home/atoz/lettery.html> Link: I
- 28 <http://www.stanford.edu/home/atoz/letterz.html> Link: I
- 29 <http://www.stanford.edu/home/slideshow/slideshow1.html> Link: I

- 30 <http://www.stanford.edu/home/academics> Link: I
 - 31 <http://www.stanford.edu/home/research> Link: I
 - 32 <http://www.stanford.edu/home/admission> Link: I
 - 33 <http://www.stanford.edu/home/stanford> Link: I
 - 34 <http://www.stanford.edu/home/medcenter> Link: I
 - 35 <http://www.stanford.edu/home/administration> Link: I
 - 36 <http://www.stanford.edu/home/arts> Link: I
 - 37 <http://www.stanford.edu/home/athletics> Link: I
 - 38 <http://www.stanford.edu/home/students> Link: I
 - 39 <http://www.stanford.edu/home/faculty> Link: I
 - 40 <http://www.stanford.edu/home/working> Link: I
 - 41 <http://www.stanford.edu/home/alumni> Link: I
 - 42 <http://www.stanford.edu/home/prospective> Link: I
 - 43 <http://www.stanford.edu/home/visitors> Link: I
 - 44 <http://www.stanford.edu> Link: I
 - 45 <http://jobs.stanford.edu> Link: X
 - 46 <http://www.stanford.edu/home/stanford/facts/index.html> Link: I
 - 47 <http://www.stanford.edu/home/atoz/contact.html> Link: I
 - 48 <http://www.stanford.edu/home/directories> Link: I
 - 49 <http://www.stanford.edu/home/visitors> Link: I
 - 50 <http://givingtostanford.stanford.edu> Link: X
 - 51 <http://www.stanford.edu/home/atoz/terms.html> Link: I
 - 52 <http://www.stanford.edu/home/atoz/copyright.html> Link: I
- จำนวนการเชื่อมโยงภายใน : 50, จำนวนการเชื่อมโยงภายนอก : 2

เว็บเพจที่ 9 : <http://www.stanford.edu/dept/DoR/GSH/Sec1.html>



รูปที่ ๑.๙ เว็บเพจที่ ๑

[Level of page = 4]

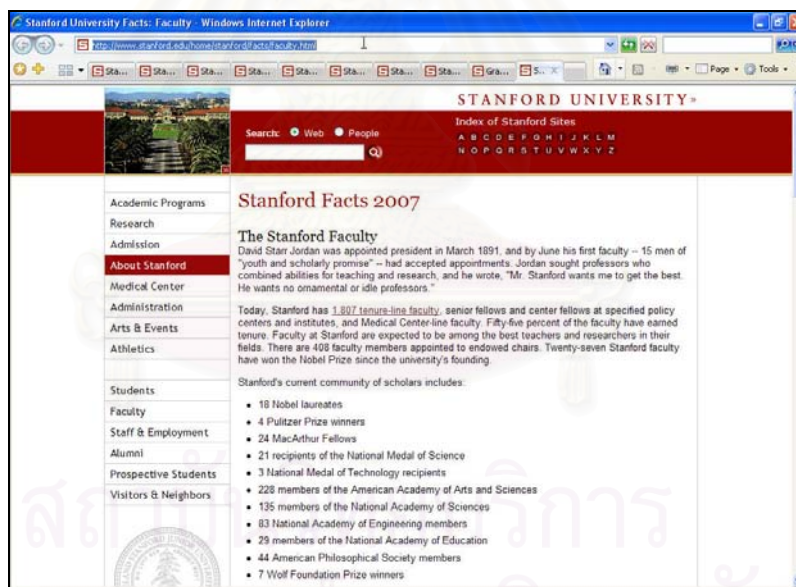
Total Links(Each) : 22

- 1 <http://www.stanford.edu/dept/DoR/GSH/index.html> Link: 1
- 2 <http://www.stanford.edu/dept/DoR/GSH/gshindex.html> Link: 1
- 3 <http://www.stanford.edu/dept/DoR/GSH/contacts.html> Link: 1
- 4 <http://www.stanford.edu/dept/DoR/GSH/Sec1.html> Link: 1
- 5 <http://www.stanford.edu/dept/DoR/GSH/Sec2.html> Link: 1
- 6 <http://www.stanford.edu/dept/DoR/GSH/Sec3.html> Link: 1
- 7 <http://www.stanford.edu/dept/DoR/GSH/Sec4.html> Link: 1
- 8 <http://www.stanford.edu/dept/DoR/GSH/Sec1A.html> Link: 1
- 9 <http://www.stanford.edu/dept/DoR/GSH/Sec1A.html#1e> Link: 1
- 10 <http://www.stanford.edu/dept/DoR/GSH/Sec1A.html#1f> Link: 1
- 11 <http://www.stanford.edu/dept/DoR/GSH/Sec1A.html#1g> Link: 1
- 12 <http://www.stanford.edu/dept/DoR/GSH/Sec1B.html> Link: 1
- 13 <http://www.stanford.edu/dept/DoR/GSH/Sec1B.html#1k> Link: 1

- 14 <http://www.stanford.edu/dept/DoR/GSH/Sec1B.html#1l> Link: I
- 15 <http://www.stanford.edu/dept/DoR/GSH/Sec1B.html#1m> Link: I
- 16 <http://www.stanford.edu/dept/DoR/GSH/Sec1B.html#1n> Link: I
- 17 <http://www.stanford.edu/dept/DoR/GSH/Sec1B.html#1o> Link: I
- 18 <http://www.stanford.edu/dept/DoR/GSH/Sec1B.html#1p> Link: I
- 19 <http://www.stanford.edu/dept/DoR/GSH/Sec1B.html#gq> Link: I
- 20 <http://www.stanford.edu/dept/DoR/GSH/Sec1B.html#1s> Link: I
- 21 <http://www.stanford.edu/dept/DoR/GSH/Sec1C.html> Link: I
- 22 <http://gradadmissions.stanford.edu> Link: X

จำนวนการเชื่อมโยงภายใน : 21, จำนวนการเชื่อมโยงภายนอก : 1

เว็บเพจที่10 : <http://www.stanford.edu/home/stanford/facts/faculty.html>



รูปที่ ๑.10 เว็บเพจที่ 10

[Level of page = 4]

Total Links(Each) : 52

- 1 <http://www.stanford.edu> Link: I
- 2 <http://www.stanford.edu/home/atoz> Link: I
- 3 <http://www.stanford.edu/home/atoz/lettera.html> Link: I

- 4 <http://www.stanford.edu/home/atoz/letterb.html> Link: I
- 5 <http://www.stanford.edu/home/atoz/letterc.html> Link: I
- 6 <http://www.stanford.edu/home/atoz/letterd.html> Link: I
- 7 <http://www.stanford.edu/home/atoz/lettere.html> Link: I
- 8 <http://www.stanford.edu/home/atoz/letterf.html> Link: I
- 9 <http://www.stanford.edu/home/atoz/letterg.html> Link: I
- 10 <http://www.stanford.edu/home/atoz/letterh.html> Link: I
- 11 <http://www.stanford.edu/home/atoz/letteri.html> Link: I
- 12 <http://www.stanford.edu/home/atoz/letterj.html> Link: I
- 13 <http://www.stanford.edu/home/atoz/letterk.html> Link: I
- 14 <http://www.stanford.edu/home/atoz/letterl.html> Link: I
- 15 <http://www.stanford.edu/home/atoz/letterm.html> Link: I
- 16 <http://www.stanford.edu/home/atoz/lettern.html> Link: I
- 17 <http://www.stanford.edu/home/atoz/lettero.html> Link: I
- 18 <http://www.stanford.edu/home/atoz/letterp.html> Link: I
- 19 <http://www.stanford.edu/home/atoz/letterq.html> Link: I
- 20 <http://www.stanford.edu/home/atoz/letterr.html> Link: I
- 21 <http://www.stanford.edu/home/atoz/letters.html> Link: I
- 22 <http://www.stanford.edu/home/atoz/lettert.html> Link: I
- 23 <http://www.stanford.edu/home/atoz/letteru.html> Link: I
- 24 <http://www.stanford.edu/home/atoz/letterv.html> Link: I
- 25 <http://www.stanford.edu/home/atoz/letterw.html> Link: I
- 26 <http://www.stanford.edu/home/atoz/letterx.html> Link: I
- 27 <http://www.stanford.edu/home/atoz/lettery.html> Link: I
- 28 <http://www.stanford.edu/home/atoz/letterz.html> Link: I
- 29 <http://www.stanford.edu/home/slideshow/slideshow1.html> Link: I
- 30 <http://www.stanford.edu/home/academics> Link: I
- 31 <http://www.stanford.edu/home/research> Link: I
- 32 <http://www.stanford.edu/home/admission> Link: I
- 33 <http://www.stanford.edu/home/stanford> Link: I

- 34 <http://www.stanford.edu/home/medcenter> Link: I
- 35 <http://www.stanford.edu/home/administration> Link: I
- 36 <http://www.stanford.edu/home/arts> Link: I
- 37 <http://www.stanford.edu/home/athletics> Link: I
- 38 <http://www.stanford.edu/home/students> Link: I
- 39 <http://www.stanford.edu/home/faculty> Link: I
- 40 <http://www.stanford.edu/home/working> Link: I
- 41 <http://www.stanford.edu/home/alumni> Link: I
- 42 <http://www.stanford.edu/home/prospective> Link: I
- 43 <http://www.stanford.edu/home/visitors> Link: I
- 44 <http://www.stanford.edu> Link: I
- 45 <http://www.stanford.edu/home/stanford/facts/faculty.html#profile> Link: I
- 46 <http://www.stanford.edu/home/stanford/facts/index.html> Link: I
- 47 <http://www.stanford.edu/home/atoz/contact.html> Link: I
- 48 <http://www.stanford.edu/home/directories> Link: I
- 49 <http://www.stanford.edu/home/visitors> Link: I
- 50 <http://givingtostanford.stanford.edu> Link: X
- 51 <http://www.stanford.edu/home/atoz/terms.html> Link: I
- 52 <http://www.stanford.edu/home/atoz/copyright.html> Link: I
- จำนวนการเชื่อมโยงภายใน : 51, จำนวนการเชื่อมโยงภายนอก : 1

3.7.4 แบ่งเพจเป็นบล็อกโดยอาศัยอัลกอริทึม VIPS

ในแต่ละเว็บเพจจะทำการแบ่งเป็นบล็อกโดยอาศัยอัลกอริทึม VIPS โดยจะผ่านโปรแกรมซึ่งกำหนดค่ายูอาร์แอลของเพจนั้น ซึ่งเมื่อผ่านโปรแกรมแล้วจะได้เอกซ์เอ็มแอลไฟล์ ซึ่งแสดงถึงบล็อกทั้งหมดในเพจนั้น หน้าจอโปรแกรม VIPS แสดงที่ภาคผนวก ง และตัวอย่างเอกซ์เอ็มแอลแสดงที่ภาคผนวก ข

3.7.5 คำนวณค่าในแต่ละบล็อกเพื่อเลือก 3 บล็อก

เป็นการเลือกบล็อก 3 บล็อกแรกที่มีค่าความสำคัญมากที่สุด โดยมีขั้นตอนดังนี้คือ

- อ่านไฟล์เอกเอ็มแอลที่ได้จากโปรแกรม VIPS แล้วเลือกเฉพาะลีฟโหนด
- นำลีฟโหนดที่ได้ ซึ่งแทนบล็อกที่ไม่สามารถแบ่งย่อยลงไปได้อีกมาทำการคำนวณหาค่าความสำคัญ จากสูตรในหัวข้อ 3.1
- เลือกบล็อก 3 บล็อกแรกที่มีค่าความสำคัญมากที่สุด

ขนาดของเพจ (หน่วยเป็นพิกเซล)

- ความกว้างของหน้าต่าง (WindowWidth) : 433
- ความสูงของหน้าต่าง (WindowHeight) : 883
- ตำแหน่งมุมด้านซ้ายของหน้าต่าง (PageRectLeft) : 0
- ตำแหน่งมุมด้านบนของหน้าต่าง (PageRectTop) : 0
- ความกว้างของหน้าต่าง (PageRectWidth) : 997
- ความสูงของหน้าต่าง (PageRectHeight) : 991

จำนวนบล็อกทั้งหมดในเพจ (โหนด)

สำหรับตัวอย่างนี้ในเพจสามารถแบ่งออกเป็นมี 20 บล็อก

All of Node : 20

- 1) ID="1"(2.0) order="1"
- 2) ID="1-1"(4.0) order="2"
- 3) ID="1-1-1"(6.0) order="3"
- 4) ID="1-1-1-1"(9.0) order="4"
- 5) ID="1-1-1-2"(11.0) order="5"
- 6) ID="1-1-2"(5.0) order="6"
- 7) ID="1-1-2-1"(7.0) order="7"
- 8) ID="1-1-2-2"(5.0) order="8"
- 9) ID="1-1-2-2-1"(9.0) order="9"
- 10) ID="1-1-2-2-2"(11.0) order="10"
- 11) ID="1-1-2-2-3"(6.0) order="11"
- 12) ID="1-1-2-2-3-1"(11.0) order="12"
- 13) ID="1-1-2-2-3-2"(6.0) order="13"
- 14) ID="1-1-2-2-3-2-1"(9.0) order="14"

- 15) ID="1-1-2-2-3-2-2"(6.0) order="15"
 16) ID="1-1-2-2-3-2-2-1"(6.0) order="16"
 17) ID="1-1-2-2-3-2-2-1-1"(11.0) order="17"
 18) ID="1-1-2-2-3-2-2-1-2"(9.0) order="18"
 19) ID="1-1-2-2-3-2-2-2"(8.0) order="19"
 20) ID="1-2"(9.0) order="20"

จำนวนบล็อกที่ไม่สามารถแบ่งย่อยลงไปได้อีกทั้งหมดในเพจ (ลิฟท์โหมด)

สำหรับตัวอย่างนี้ในเพจมี 11 บล็อกที่เป็นลิฟท์โหมด

คำนวณจุดศูนย์กลางของเพจ

$$\text{จุดศูนย์กลางด้านกว้าง (mPageCentW)} = 0.0 + (997.0/2) = 498.5$$

$$\text{จุดศูนย์กลางด้านสูง (mPageCentH)} = 0.0 + (991.0/2) = 495.5$$

คำนวณทีละบล็อก

บล็อกที่ 1) ID="1-1-1-1"(9.0) order="4" Left = 0, Top = 0, Width = 760, Height = 49

$$\text{ขนาด (Size)} = 760.0 * 49.0 = 37240.0$$

$$\text{จุดศูนย์กลางด้านกว้าง} = 0.0 + (760.0/2) = 380.0$$

$$\text{จุดศูนย์กลางด้านสูง} = 0.0 + (49.0/2) = 24.5$$

$$\text{ระยะห่างระหว่างสองจุดในแนวนอน} = \text{ABS}(380.0 - 498.5) = 118.5,$$

$$\text{ระยะห่างระหว่างสองจุดในแนวตั้ง} = \text{ABS}(24.5 - 495.5) = 471.0$$

$$\text{mPowerWH} = (118.5^2) + (471.0^2) = 235883.25,$$

$$\text{ระยะห่างระหว่างสองจุด} = \text{SQRT}(235883.25) = 485.67$$

$$\text{ค่าความสำคัญ} = (760.0 * 49.0) / 485.67 = 76.68$$

จำนวนการเชื่อมโยงทั้งหมด :

1 #content

All Links : 1

จำนวนเพจทั้งหมด : : 1

จำนวนเพจที่บล็อกมีการเชื่อมโยงไป = 1

Weight = (76.67629522687488 / 1) = 76.67629522687488 ==> 77

Node ID : 1-1-1-1, tcnt : 1

บล็อกที่ 2) ID="1-1-1-2"(11.0) order="5" Left = 0, Top = 49, Width = 433, Height = 131

ขนาด (Size) = $433.0 * 131.0 = 56723.0$

จุดศูนย์กลางด้านกว้าง = $0.0 + (433.0 / 2) = 216.5$

จุดศูนย์กลางด้านสูง = $49.0 + (131.0 / 2) = 114.5$

ระยะห่างระหว่างสองจุดในแนวนอน = $ABS(216.5 - 498.5) = 282.0$,

ระยะห่างระหว่างสองจุดในแนวตั้ง = $ABS(114.5 - 495.5) = 381.0$

mPowerWH = $(282.0^2) + (381.0^2) = 224685.0$

ระยะห่างระหว่างสองจุด = $SQRT(224685.0) = 474.00949357581436$

ค่าความสำคัญ = $(433.0 * 131.0) / 474.00 = 119.67$

จำนวนการเชื่อมโยงทั้งหมด :

1 javascript:homeslide();

All Links : 1

จำนวนเพจทั้งหมด : : 1

1 javascript:homeslide();

จำนวนเพจที่บล็อกมีการเชื่อมโยงไป = 1

Weight = (119.66637961635587 / 1) = 119.66637961635587 ==> 120

Node ID : 1-1-1-2, tcnt : 1

บล็อกที่ 3) ID="1-1-2-1"(7.0) order="7" Left = 0, Top = 180, Width = 160, Height = 625

ขนาด (Size) = $160.0 * 625.0 = 100000.0$

จุดศูนย์กลางด้านกว้าง = $0.0 + (160.0 / 2) = 80.0$

จุดศูนย์กลางด้านสูง = $180.0 + (625.0 / 2) = 492.5$

$$\begin{aligned}
 \text{ระยะห่างระหว่างสองจุดในแนวนอน} &= \text{ABS}(80.0-498.5) = 418.5, \\
 \text{ระยะห่างระหว่างสองจุดในแนวตั้ง} &= \text{ABS}(492.5-495.5) = 3.0 \\
 m\text{PowerWH} &= (418.5^2)+(3.0^2) = 175151.25, \\
 \text{ระยะห่างระหว่างสองจุด} &= \text{SQRT}(175151.25) = 418.510752550039 \\
 \text{ค่าความสำคัญ} &= (160.0*625.0)/418.51 = 238.94
 \end{aligned}$$

จำนวนการเชื่อมโยงทั้งหมด :

- 1 /home/academics/
- 2 /home/academics/
- 3 /home/academics/schools.html
- 4 /home/academics/departments.html
- 5 /home/academics/programs.html
- 6 http://library.stanford.edu/
- 7 /dept/registrar/bulletin/
- 8 /home/academics/continuing.html
- 9 /dept/registrar/academic_calendar/
- 10 /home/research/
- 11 /home/research/
- 12 /home/academics/schools.html
- 13 /home/academics/departments.html
- 14 /home/research/centers.html
- 15 http://multi.stanford.edu
- 16 http://www.slac.stanford.edu/
- 17 http://library.stanford.edu/
- 18 /dept/undergrad/urp/
- 19 /home/admission/
- 20 /home/admission/
- 21 /home/stanford/
- 22 /dept/uga/
- 23 /dept/gradadmissions

- 24 /dept/finaid/
- 25 /home/academics/continuing.html
- 26 /home/stanford/
- 27 /home/stanford/
- 28 /home/welcome/
- 29 /home/stanford/facts/
- 30 /home/stanford/history/
- 31 /home/atoz/faq.html
- 32 <http://news-service.stanford.edu/>
- 33 <http://thestanfordchallenge.stanford.edu>
- 34 /home/medcenter/
- 35 /home/medcenter/
- 36 <http://med.stanford.edu/>
- 37 <http://www.stanfordhospital.com/>
- 38 <http://www.lpch.org/>
- 39 /home/administration/
- 40 /home/administration/
- 41 /dept/president/
- 42 /dept/provost/
- 43 /home/stanford/facts/board.html
- 44 <http://facultysenate.stanford.edu/>
- 45 /home/academics/schools.html
- 46 /home/academics/departments.html
- 47 <http://bondholder-information.stanford.edu/disclaimer.html>
- 48 <http://computing.stanford.edu>
- 49 /home/arts/
- 50 /home/arts/
- 51 <http://events.stanford.edu/>
- 52 <http://museum.stanford.edu/>
- 53 <http://livelyarts.stanford.edu/>

54 <http://itunes.stanford.edu/>

55 </home/athletics/>

56 </home/athletics/>

57 <http://gostanford.cstv.com/>

58 </dept/pe/>

59 <http://gostanford.cstv.com/school-bio/>

60 <http://gostanford.cstv.com/genrel/stan-recruiting.html>

61 <http://www.stanfordstadium.com/>

62 </home/students/>

63 </home/students/>

64 <https://axess.stanford.edu/>

65 </dept/undergrad/vpue/>

66 </dept/DoR/GSH/Sec1.html>

67 </dept/vpsa/>

68 <http://diversityworks.stanford.edu>

69 <http://rescomp.stanford.edu/>

70 </dept/hds/>

71 <http://co.stanford.edu/students/>

72 </home/faculty/>

73 </home/faculty/>

74 <https://axess.stanford.edu/>

75 <http://ctl.stanford.edu/>

76 </home/research/>

77 <http://computing.stanford.edu>

78 <http://benefitsu.stanford.edu/>

79 </home/stanford/facts/faculty.html>

80 <http://news-service.stanford.edu/>

81 </home/working/>

82 </home/working/>

83 <http://jobs.stanford.edu/>

- 84 <http://benefitsu.stanford.edu/>
- 85 http://co.stanford.edu/payroll_manual/
- 86 <http://computing.stanford.edu>
- 87 <http://transportation.stanford.edu/>
- 88 <http://news-service.stanford.edu/>
- 89 </home/alumni/>
- 90 </home/alumni/>
- 91 <http://www.stanfordalumni.org/>
- 92 <http://www.stanfordalumni.org/volunteer/>
- 93 <https://pgnet.stanford.edu/WebDirectory/search?pgOrg=saa&pgnPageCode=110>
- 94 <https://www.stanfordalumni.org/email/>
- 95 <http://www.stanfordalumni.org/erc/regional/listing.html?type=events>
- 96 <http://www.stanfordalumni.org/magazine/>
- 97 <http://giving.stanford.edu/>
- 98 </home/prospective/>
- 99 </home/prospective/>
- 100 </home/stanford/>
- 101 </dept/uga/>
- 102 </dept/gradadmissions>
- 103 </dept/finaid/>
- 104 </home/academics/continuing.html>
- 105 </home/visitors/>
- 106 </home/visitors/>
- 107 </dept/visitorinfo/>
- 108 </search/maps/>
- 109 </search/maps/directions.html>
- 110 <http://events.stanford.edu/>
- 111 <http://parents.stanford.edu/>
- 112 </dept/ucomm/neighbors/>
- 113 </home/visitors/eateries/>

114 <http://itunes.stanford.edu>

115 <http://itunes.stanford.edu>

116 </home/slideshow/slideshow1.html>

117 </home/slideshow/slideshow1.html>

All Links : 117

จำนวนเพจทั้งหมด : : 81

1 </home/academics/>

2 </home/academics/schools.html>

3 </home/academics/departments.html>

4 </home/academics/programs.html>

5 <http://library.stanford.edu/>

6 </dept/registrar/bulletin/>

7 </home/academics/continuing.html>

8 /dept/registrar/academic_calendar/

9 </home/research/>

10 </home/research/centers.html>

11 <http://multi.stanford.edu>

12 <http://www.slac.stanford.edu/>

13 </dept/undergrad/urp/>

14 </home/admission/>

15 </home/stanford/>

16 </dept/uga/>

17 </dept/gradadmissions>

18 </dept/finaid/>

19 </home/welcome/>

20 </home/stanford/facts/>

21 </home/stanford/history/>

22 </home/atoz/faq.html>

23 <http://news-service.stanford.edu/>

24 <http://thestanfordchallenge.stanford.edu>

- 25 /home/medcenter/
- 26 <http://med.stanford.edu/>
- 27 <http://www.stanfordhospital.com/>
- 28 <http://www.lpch.org/>
- 29 /home/administration/
- 30 /dept/president/
- 31 /dept/provost/
- 32 /home/stanford/facts/board.html
- 33 <http://facultysenate.stanford.edu/>
- 34 <http://bondholder-information.stanford.edu/disclaimer.html>
- 35 <http://computing.stanford.edu>
- 36 /home/arts/
- 37 <http://events.stanford.edu/>
- 38 <http://museum.stanford.edu/>
- 39 <http://livelyarts.stanford.edu/>
- 40 <http://itunes.stanford.edu/>
- 41 /home/athletics/
- 42 <http://gostanford.cstv.com/>
- 43 /dept/pe/
- 44 <http://gostanford.cstv.com/school-bio/>
- 45 <http://gostanford.cstv.com/genrel/stan-recruiting.html>
- 46 <http://www.stanfordstadium.com/>
- 47 /home/students/
- 48 <https://axess.stanford.edu/>
- 49 /dept/undergrad/vpue/
- 50 /dept/DoR/GSH/Sec1.html
- 51 /dept/vpsa/
- 52 <http://diversityworks.stanford.edu>
- 53 <http://rescomp.stanford.edu/>
- 54 /dept/hds/

- 55 <http://co.stanford.edu/students/>
- 56 [/home/faculty/](#)
- 57 <http://ctl.stanford.edu/>
- 58 <http://benefitsu.stanford.edu/>
- 59 [/home/stanford/facts/faculty.html](#)
- 60 [/home/working/](#)
- 61 <http://jobs.stanford.edu/>
- 62 http://co.stanford.edu/payroll_manual/
- 63 <http://transportation.stanford.edu/>
- 64 [/home/alumni/](#)
- 65 <http://www.stanfordalumni.org/>
- 66 <http://www.stanfordalumni.org/volunteer/>
- 67 <https://pgnet.stanford.edu/WebDirectory/search?pgOrg=saa&pgnPageCode=110>
- 68 <https://www.stanfordalumni.org/email/>
- 69 <http://www.stanfordalumni.org/erc/regional/listing.html?type=events>
- 70 <http://www.stanfordalumni.org/magazine/>
- 71 <http://giving.stanford.edu/>
- 72 [/home/prospective/](#)
- 73 [/home/visitors/](#)
- 74 [/dept/visitorinfo/](#)
- 75 [/search/maps/](#)
- 76 [/search/maps/directions.html](#)
- 77 <http://parents.stanford.edu/>
- 78 [/dept/ucomm/neighbors/](#)
- 79 [/home/visitors/eateries/](#)
- 80 <http://itunes.stanford.edu>
- 81 [/home/slideshow/slideshow1.html](#)

จำนวนเพจที่บล็อกมีการเชื่อมโยงไป = 81

Weight = (238.94248687921956 / 81) = 2.9499072454224637 ==> 3

Node ID : 1-1-2-1, tcnt : 81

บล็อกที่ 4) ID="1-1-2-2-1"(9.0) order="9" Left = 190, Top = 191, Width = 321, Height =

$$\begin{aligned}
 70 \quad \text{ขนาด (Size)} &= 321.0 * 70.0 = 22470.0 \\
 \text{จุดศูนย์กลางด้านกว้าง} &= 190.0 + (321.0 / 2) = 350.5 \\
 \text{จุดศูนย์กลางด้านสูง} &= 191.0 + (70.0 / 2) = 226.0 \\
 \text{ระยะห่างระหว่างสองจุดในแนวนอน} &= \text{ABS}(350.5 - 498.5) = 148.0, \\
 \text{ระยะห่างระหว่างสองจุดในแนวตั้ง} &= \text{ABS}(226.0 - 495.5) = 269.5 \\
 \text{mPowerWH} &= (148.0^2) + (269.5^2) = 94534.25, \\
 \text{ระยะห่างระหว่างสองจุด} &= \text{SQRT}(94534.25) = 307.4642255612838 \\
 \text{ค่าความสำคัญ} &= (321.0 * 70.0) / 307.46 = 73.08
 \end{aligned}$$

จำนวนการเชื่อมโยงทั้งหมด :

- 1 /home/atoz/
- 2 /home/atoz/lettera.html
- 3 /home/atoz/letterb.html
- 4 /home/atoz/letterc.html
- 5 /home/atoz/letterd.html
- 6 /home/atoz/lettere.html
- 7 /home/atoz/letterf.html
- 8 /home/atoz/letterg.html
- 9 /home/atoz/letterh.html
- 10 /home/atoz/letteri.html
- 11 /home/atoz/letterj.html
- 12 /home/atoz/letterk.html
- 13 /home/atoz/letterl.html
- 14 /home/atoz/letterm.html
- 15 /home/atoz/lettern.html
- 16 /home/atoz/lettero.html
- 17 /home/atoz/letterp.html

18 /home/atoz/letterq.html

19 /home/atoz/letterr.html

20 /home/atoz/letters.html

21 /home/atoz/lettert.html

22 /home/atoz/letteru.html

23 /home/atoz/letterv.html

24 /home/atoz/letterw.html

25 /home/atoz/letterx.html

26 /home/atoz/lettery.html

27 /home/atoz/letterz.html

All Links : 27

จำนวนเพจทั้งหมด : : 27

1 /home/atoz/

2 /home/atoz/lettera.html

3 /home/atoz/letterb.html

4 /home/atoz/letterc.html

5 /home/atoz/letterd.html

6 /home/atoz/lettere.html

7 /home/atoz/letterf.html

8 /home/atoz/letterg.html

9 /home/atoz/letterh.html

10 /home/atoz/letteri.html

11 /home/atoz/letterj.html

12 /home/atoz/letterk.html

13 /home/atoz/letterl.html

14 /home/atoz/letterm.html

15 /home/atoz/lettern.html

16 /home/atoz/lettero.html

17 /home/atoz/letterp.html

18 /home/atoz/letterq.html

19 /home/atoz/letterr.html
 20 /home/atoz/letters.html
 21 /home/atoz/lettert.html
 22 /home/atoz/letteru.html
 23 /home/atoz/letterv.html
 24 /home/atoz/letterw.html
 25 /home/atoz/letterx.html
 26 /home/atoz/lettery.html
 27 /home/atoz/letterz.html

จำนวนเพจที่บล็อกมีการเชื่อมโยงไป = 27

Weight = (73.08167302709914 / 27) = 2.7067286306333016 ==> 3

Node ID : 1-1-2-2-1, tcnt : 27

บล็อกที่ 5) ID="1-1-2-2-2"(11.0) order="10" Left = 183, Top = 277, Width = 168, Height =

265 ขนาด (Size) = 168.0*265.0 = 44520.0

จุดศูนย์กลางด้านกว้าง = 183.0+(168.0/2) = 267.0

จุดศูนย์กลางด้านสูง = 277.0+(265.0/2) = 409.5

ระยะห่างระหว่างสองจุดในแนวนอน = ABS(267.0-498.5) = 231.5,

ระยะห่างระหว่างสองจุดในแนวตั้ง = ABS(409.5-495.5) = 86.0

mPowerWH = (231.5^2)+(86.0^2) = 60988.25,

ระยะห่างระหว่างสองจุด = SQRT(60988.25) = 246.95799237927085

ค่าความสำคัญ = (168.0*265.0)/246.95 = 180.27

จำนวนการเชื่อมโยงทั้งหมด :

All Links : 0

จำนวนเพจทั้งหมด : : 0

จำนวนเพจที่บล็อกมีการเชื่อมโยงไป = 0

Weight = 0

บล็อกที่ 6) ID="1-1-2-2-3-1"(11.0) order="12" Left = 380, Top = 190, Width = 375, Height = 18

ขนาด (Size)	=	$375.0 * 18.0 = 6750.0$
จุดศูนย์กลางด้านกว้าง	=	$380.0 + (375.0/2) = 567.5$
จุดศูนย์กลางด้านสูง	=	$190.0 + (18.0/2) = 199.0$
ระยะห่างระหว่างสองจุดในแนวนอน	=	$ABS(567.5 - 498.5) = 69.0,$
ระยะห่างระหว่างสองจุดในแนวตั้ง	=	$ABS(199.0 - 495.5) = 296.5$
mPowerWH	=	$(69.0^2) + (296.5^2) = 92673.25,$
ระยะห่างระหว่างสองจุด	=	$SQRT(92673.25) = 304.42281451954284$
ค่าความสำคัญ	=	$(375.0 * 18.0) / 304.42 = 22.17$

จำนวนการเชื่อมโยงทั้งหมด :

1 <http://news.stanford.edu>

All Links : 1

จำนวนเพจทั้งหมด : : 1

1 <http://news.stanford.edu>

จำนวนเพจที่บล็อกมีการเชื่อมโยงไป = 1

Weight = $(22.17310818393565 / 1) = 22.17310818393565 ==> 22$

Node ID : 1-1-2-2-3-1, tcnt : 1

บล็อกที่ 7) ID="1-1-2-2-3-2-1"(9.0) order="14" Left = 380, Top = 208, Width = 375,

ขนาด (Size)	=	$375.0 * 222.0 = 83250.0$
จุดศูนย์กลางด้านกว้าง	=	$380.0 + (375.0/2) = 567.5$
จุดศูนย์กลางด้านสูง	=	$208.0 + (222.0/2) = 319.0$
ระยะห่างระหว่างสองจุดในแนวนอน	=	$ABS(567.5 - 498.5) = 69.0,$
ระยะห่างระหว่างสองจุดในแนวตั้ง	=	$ABS(319.0 - 495.5) = 176.5$
mPowerWH	=	$(69.0^2) + (176.5^2) = 35913.25,$
ระยะห่างระหว่างสองจุด	=	$SQRT(35913.25) = 189.50791540196943$
ค่าความสำคัญ	=	$(375.0 * 222.0) / 189.50 = 439.30$

จำนวนการเชื่อมโยงทั้งหมด :

- 1 <http://news-service.stanford.edu/news/2007/april25/itunes-042507.html>
- 2 <http://news-service.stanford.edu/news/2007/april25/kidcare-042507.html>
- 3 <http://news.stanford.edu/>

All Links : 3

จำนวนเพจทั้งหมด : : 3

- 1 <http://news-service.stanford.edu/news/2007/april25/itunes-042507.html>
- 2 <http://news-service.stanford.edu/news/2007/april25/kidcare-042507.html>
- 3 <http://news.stanford.edu/>

จำนวนเพจที่บล็อกมีการเชื่อมโยงไป = 3

Weight = (439.29563482040624 / 3) = 146.43187827346875 ==> 146

Node ID : 1-1-2-2-3-2-1, tcnt : 3

บล็อกที่ 8) ID="1-1-2-2-3-2-2-1-1"(11.0) order="17" Left = 380, Top = 449, Width = 375,

Height = 18 ขนาด (Size) = 375.0*18.0 = 6750.0

จุดศูนย์กลางด้านกว้าง = 380.0+(375.0/2) = 567.5

จุดศูนย์กลางด้านสูง = 449.0+(18.0/2) = 458.0

ระยะห่างระหว่างสองจุดในแนวนอน = ABS(567.5-498.5) = 69.0,

ระยะห่างระหว่างสองจุดในแนวตั้ง = ABS(458.0-495.5) = 37.5

mPowerWH = (69.0^2)+(37.5^2) = 6167.25,

ระยะห่างระหว่างสองจุด = SQRT(6167.25) = 78.53184067625055

ค่าความสำคัญ = (375.0*18.0)/78.53 = 85.95

จำนวนการเชื่อมโยงทั้งหมด :

- 1 <http://events.stanford.edu>

All Links : 1

จำนวนเพจทั้งหมด : : 1

- 1 <http://events.stanford.edu>

จำนวนเพจที่บล็อกมีการเชื่อมโยงไป = 1

Weight = (85.9523976755752 / 1) = 85.9523976755752 ==> 86

Node ID : 1-1-2-2-3-2-2-1-1, tcnt : 1

บล็อกที่ 9) ID="1-1-2-2-3-2-2-1-2"(9.0) order="18" Left = 380, Top = 467, Width = 375,

Height = 119 ขนาด (Size) = 375.0*119.0 = 44625.0

จุดศูนย์กลางด้านกว้าง = 380.0+(375.0/2) = 567.5

จุดศูนย์กลางด้านสูง = 467.0+(119.0/2) = 526.5

ระยะห่างระหว่างสองจุดในแนวนอน = ABS(567.5-498.5) = 69.0,

ระยะห่างระหว่างสองจุดในแนวตั้ง = ABS(526.5-495.5) = 31.0

mPowerWH = (69.0^2)+(31.0^2) = 5722.0,

ระยะห่างระหว่างสองจุด = SQRT(5722.0) = 75.64390259630977

ค่าความสำคัญ = (375.0*119.0)/75.64 = 589.94

จำนวนการเชื่อมโยงทั้งหมด :

1 <https://events.stanford.edu/events/108/10861/>

2 <http://events.stanford.edu/>

All Links : 2

จำนวนเพจทั้งหมด : : 2

1 <https://events.stanford.edu/events/108/10861/>

2 <http://events.stanford.edu/>

จำนวนเพจที่บล็อกมีการเชื่อมโยงไป = 2

Weight = (589.9351893324578 / 2) = 294.9675946662289 ==> 295

Node ID : 1-1-2-2-3-2-2-1-2, tcnt : 2

บล็อกที่ 10) ID="1-1-2-2-3-2-2-2"(8.0) order="19" Left = 389, Top = 605, Width = 360,

Height = 169 ขนาด (Size) = 360.0*169.0 = 60840.0

จุดศูนย์กลางด้านกว้าง = 389.0+(360.0/2) = 569.0

จุดศูนย์กลางด้านสูง = 605.0+(169.0/2) = 689.5

$$\begin{aligned}
 \text{ระยะห่างระหว่างสองจุดในแนวนอน} &= \text{ABS}(569.0-498.5) = 70.5, \\
 \text{ระยะห่างระหว่างสองจุดในแนวตั้ง} &= \text{ABS}(689.5-495.5) = 194.0 \\
 mPowerWH &= (70.5^2)+(194.0^2) = 42606.25, \\
 \text{ระยะห่างระหว่างสองจุด} &= \text{SQRT}(42606.25) = 206.41281452468013 \\
 \text{ค่าความสำคัญ} &= (360.0*169.0)/206.41 = 294.75
 \end{aligned}$$

จำนวนการเชื่อมโยงทั้งหมด :

- 1 <http://folding.stanford.edu/>
- 2 <http://folding.stanford.edu/>
- 3 </home/otsw/>

All Links : 3

จำนวนเพจทั้งหมด : : 2

- 1 <http://folding.stanford.edu/>
- 2 </home/otsw/>

จำนวนเพจที่บล็อกมีการเชื่อมโยงไป = 2

$$\text{Weight} = (294.7491421019578 / 2) = 147.3745710509789 \implies 147$$

Node ID : 1-1-2-2-3-2-2-2, tcnt : 2

บล็อกที่ 11) ID="1-2"(9.0) order="20" Left = 0, Top = 821, Width = 433, Height = 62

$$\begin{aligned}
 \text{ขนาด (Size)} &= 433.0*62.0 = 26846.0 \\
 \text{จุดศูนย์กลางด้านกว้าง} &= 0.0+(433.0/2) = 216.5 \\
 \text{จุดศูนย์กลางด้านสูง} &= 821.0+(62.0/2) = 852.0 \\
 \text{ระยะห่างระหว่างสองจุดในแนวนอน} &= \text{ABS}(216.5-498.5) = 282.0, \\
 \text{ระยะห่างระหว่างสองจุดในแนวตั้ง} &= \text{ABS}(852.0-495.5) = 356.5 \\
 mPowerWH &= (282.0^2)+(356.5^2) = 206616.25, \\
 \text{ระยะห่างระหว่างสองจุด} &= \text{SQRT}(206616.25) = 454.5506022435786 \\
 \text{ค่าความสำคัญ} &= (433.0*62.0)/454.55 = 59.06
 \end{aligned}$$

จำนวนการเชื่อมโยงทั้งหมด :

- 1 /home/atoz/contact.html
- 2 /home/directories/
- 3 /home/visitors/maps.html
- 4 http://givingtostanford.stanford.edu/
- 5 home/atoz/terms.html
- 6 /home/atoz/copyright.html

All Links : 6

จำนวนเพจทั้งหมด : : 6

- 1 /home/atoz/contact.html
- 2 /home/directories/
- 3 /home/visitors/maps.html
- 4 http://givingtostanford.stanford.edu/
- 5 home/atoz/terms.html
- 6 /home/atoz/copyright.html

จำนวนเพจที่บล็อกมีการเชื่อมโยงไป = 6

Weight =(59.06053114327218 / 6) = 9.84342185721203 ==> 10

Node ID : 1-2, tcnt : 6

Total Leave Nodes: 11

- 1 : 76.67629522687488
- 2 : 119.66637961635587
- 3 : 2.9499072454224637
- 4 : 2.7067286306333016
- 5 : 0.0
- 6 : 22.17310818393565
- 7 : 146.43187827346875
- 8 : 85.9523976755752
- 9 : 294.9675946662289
- 10 : 147.3745710509789

11 : 9.84342185721203

Sort : all nodes

[0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 2.7067286306333016,
2.9499072454224637, 9.84342185721203, 22.17310818393565, 76.67629522687488,
85.9523976755752, 119.66637961635587, 146.43187827346875, 147.3745710509789,
294.9675946662289]

1 : 0.0

2 : 0.0

3 : 0.0

4 : 0.0

5 : 0.0

6 : 0.0

7 : 0.0

8 : 0.0

9 : 0.0

10 : 0.0

11 : 2.7067286306333016

12 : 2.9499072454224637

13 : 9.84342185721203

14 : 22.17310818393565

15 : 76.67629522687488

16 : 85.9523976755752

17 : 119.66637961635587

18 : 146.43187827346875

19 : 147.3745710509789

20 : 294.9675946662289

Top Three:

Max : 294.9675946662289 ,Node Id : 1-1-2-2-3-2-2-1-2 ,LkCnt : 2

Max2 : 147.3745710509789 ,Node Id : 1-1-2-2-3-2-2-2 ,LkCnt : 2

Max3 : 146.43187827346875 ,Node Id : 1-1-2-2-3-2-1 ,LkCnt : 3

[MaxWeight : 294.9675946662289], ID:1-1-2-2-3-2-2-1-2(DoC:9.0 ,Level:9)

AllLink	=	2
UniqueLink	=	2
ContainImg	=	0.0
IsImg	=	false
ContainTable	=	false
TextLen	=	356.0
LinkTextLen	=	42.0
DOMCIdNum	=	2.0
FontSize	=	9.6
FontWeight	=	400.0
BlockRectLeft	=	380.0
BlockRectTop	=	467.0
BlockRectWidth	=	375.0
BlockRectHeight	=	119.0
BlockCentX	=	567.5
BlockCentY	=	526.5

[MaxWeight2 : 147.3745710509789], ID:1-1-2-2-3-2-2-2(DoC:8.0 ,Level:8)

AllLink	=	3
UniqueLink	=	2
ContainImg	=	1.0
IsImg	=	false
ContainTable	=	false
TextLen	=	283.0
LinkTextLen	=	19.0

DOMCIdNum = 2.0
 FontSize = 9.6
 FontWeight = 400.0
 BlockRectLeft = 389.0
 BlockRectTop = 605.0
 BlockRectWidth = 360.0
 BlockRectHeight = 169.0
 BlockCentX = 569.0
 BlockCentY = 689.5

[MaxWeight3 : 146.43187827346875], ID:1-1-2-2-3-2-1(DoC:9.0 ,Level:7)

AllLink = 3
 UniqueLink = 3
 ContainImg = 0.0
 IsImg = false
 ContainTable = false
 TextLen = 601.0
 LinkTextLen = 97.0
 DOMCIdNum = 3.0
 FontSize = 10.0
 FontWeight = 400.0
 BlockRectLeft = 380.0
 BlockRectTop = 208.0
 BlockRectWidth = 375.0
 BlockRectHeight = 222.0
 BlockCentX = 567.5
 BlockCentY = 319.0

Total of Top Three Blocks

mGTBIkLevel : 24

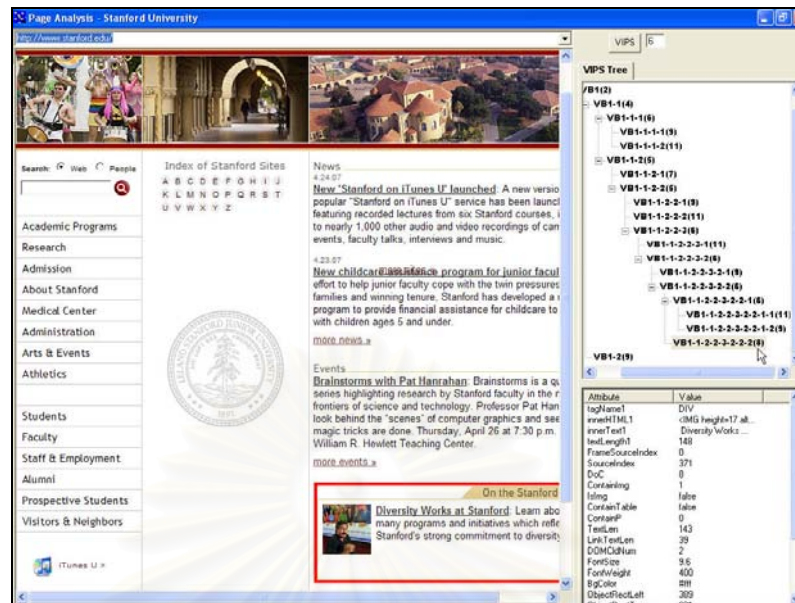
mGTDoC	:	26.0
mGTWeight	:	589
mGTAIlLk	:	8
mGTUniqueLk	:	7
mGTContainImg	:	1.0
mGTTextLen	:	1240.0
mGTLinkTextLen	:	158.0
mGTDOMCIdNum	:	7.0
mGTFontSize	:	29.2
mGTFontWeight	:	1200.0
mGTBlockRectLeft	:	1149.0
mGTBlockRectTop	:	1280.0
mGTBlockRectWidth	:	1110.0
mGTBlockRectHeight	:	510.0
mGTBlockCentX	:	1704.0
mGTBlockCentY	:	1535.0

3.7.6 คำนวณค่าเฉลี่ยของคุณลักษณะระดับบล็อก 3 บล็อกที่เลือก

เป็นการหาค่าเฉลี่ยของคุณลักษณะในระดับบล็อกของทั้ง 3 บล็อกใน 1 เพจ

Average of Top Three Blocks

mGABlkLevel	:	8
mGADoC	:	8.666666666666666
mGAWeight	:	196
mGAAllLk	:	3.0
mGAUniqueLk	:	2.0
mGAContainImg	:	0.0
mGATextLen	:	413
mGALinkTextLen	:	53.0
mGADOMCIdNum	:	2.0
mGAFontSize	:	10



รูปที่ ๑.12 บล็อกที่ 2



รูปที่ ๑.13 บล็อกที่ 3

3.7.7 คำนวณค่าเฉลี่ยของคุณลักษณะบล็อกสำหรับเว็บไซต์

เป็นการหาค่าเฉลี่ยของคุณลักษณะในระดับบล็อกของทั้ง 10 เพจ

3.7.8 คำนวณค่าเฉลี่ยของคุณลักษณะเพจสำหรับเว็บไซต์

เป็นการหาค่าเฉลี่ยของคุณลักษณะในระดับเพจของทั้ง 10 เพจ

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวพรทิพย์ มาลีลัย เกิดเมื่อวันที่ 14 พฤศจิกายน พ.ศ. 2517 ที่จังหวัดอุดรธานี สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น ในปีการศึกษา 2540 และได้เข้าศึกษาในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิทยาศาสตร์คอมพิวเตอร์(ภาคนอกเวลาราชการ) คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2546



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย