การประชุมวิชาการระดับชาติวิทยาลัยเทคโนโลยีสยาม
(Techcon 2023 National Conference)

ระบบแนะนำหลักสูตรสาขาวิชาคอมพิวเตอร์ด้วยเทคนิคเหมืองข้อมูล

ณัฐวรรณ ภูมิพันธ์ ทรงสิทธิ์ อยู่สุข ผศ.ดร.อัจฉรา มหาวีรวัฒน์ และ ผศ.ดร.สาธิษฐ์ นากกระแสร์
คณะวิทยาศาสตร์ มหาวิทยาลัยรามคำแหง

บทนำ

ในปัจจุบันมีหลักสูตรการเรียนการการสอนมากมาย รวมถึงสาขาคอมพิวเตอร์ที่มีความหลากหลาย นักศึกษาส่วนมากเลือกเรียนสาขาที่ตัวเองชอบหรือเรียนตามเพื่อน หรืออาจไม่เข้าใจว่า เรียนสาขานี้จบมาต้องทำอาชีพอะไร บางคนเข้ามาสาขาเลือกเรียนสาขานี้เพราะชอบเล่นเกม บางคนเลือกสาขานี้เพราะต้องการนำวุฒิไปเพิ่มเงินเดือน โดยอาจไม่ได้คำนึงถึงปัญหาระหว่างเรียน อาจสับสนหรือไม่เข้าใจในขณะที่กำลังศึกษาอยู่ซึ่งทำให้นักศึกษากลุ่มนี้หมดกำลังใจ ย่อท้อ เลิกเรียน ย้ายคณะ ส่งผลให้เสียเวลาไปกับสิ่งเหล่านี้ รวมถึงผู้สนใจงานทางด้านคอมพิวเตอร์ที่มีความรู้และทักษะต่างๆ มาบ้างแล้ว แต่ยังไม่ทราบว่าตัวเองเหมาะกับงานสายงานด้านใดและสายงานนั้นต้องใช้ทักษะทางด้านใดบ้าง ดังนั้นเราควรทราบความเหมาะสมของตัวเองก่อน โดยพิจารณาจากความรู้ ทักษะ บุคคลิกภาพ เพื่อให้สอดคล้องกับสาขาที่เรียน และสามารถนำไปประกอบการตัดสินใจในการเลือกสายงานที่เหมาะสมกับตัวเอง โครงงานนี้จึงนำเสนอแนวทางการแก้ปัญหาข้างต้นด้วยการพัฒนาเว็บแอพพลิเคชั่นที่ระบบจะสามารถแนะนำหลักสูตรที่เหมาะสมกับผู้ที่สนใจศึกษา

งานวิจัยนี้ได้ประยุกต์ใช้ K Mean Clustering ซึ่งเป็นหนึ่งในวิธีที่เป็นที่นิยมสำหรับการแบ่งกลุ่มข้อมูล (Cluster Analysis) ในการทำเหมืองข้อมูล (Data Mining) โดยทำการจัดกลุ่มเพื่อหาความเหมาะสมโดยใช้โครงสร้างหลักสูตรสาขาคอมพิวเตอร์ ได้แก่ วิศวกรรมซอฟต์แวร์ (Software Engineering), วิศวกรรมคอมพิวเตอร์ (Computer Engineering), วิทยาศาสตร์คอมพิวเตอร์หรือวิทยาการคอมพิวเตอร์ (Computer Science), เทคโนโลยีสารสนเทศ (Information Technology), คอมพิวเตอร์ธุรกิจ (Business Computer) ของมหาวิทยาลัยต่างๆ มาเปรียบเทียบ ซึ่งจะทำการคัดเลือก วิชาแกนของแต่ละสาขาซึ่งมหาวิทยาลัยแต่ละที่มีเหมือนกัน แล้วนำมาสร้างเป็นมาตรฐานในการกำหนดกฎการแบ่งกลุ่มสาขา และใช้เป็นเครื่องมือในการประมวลผลให้ได้ค่าคำตอบในการแนะนำสาขาที่เหมาะสมต่อไป

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาการแนะนำสาขาวิชาคอมพิวเตอร์ด้วยเทคนิคเหมืองข้อมูล
2. เพื่อสอบถามความคิดเห็นของผู้ใช้ที่มีต่อเว็บแอปพลิเคชั่นแนะนำสาขาวิชาคอมพิวเตอร์ด้วยเทคนิคเหมืองข้อมูล

วิธีดำเนินการวิจัย

การทำงานของ “ระบบแนะนำสาขาวิชาคอมพิวเตอร์ด้วยเทคนิคเหมืองข้อมูล” แบ่งออกเป็น 3 ส่วน ดังนี้ ส่วนความรู้ภูมิหลัง, ส่วนทักษะต่างๆ, ส่วนบุคลิกภาพ

1. ส่วนความรู้ภูมิหลัง

ใช้ความรู้ภูมิหลังที่ได้ทำการเก็บรวบรวมข้อมูลจากการทำสำรวจมาทำการจัดกลุ่มและทำนายข้อมูลที่ผู้ใช้ระบุเข้ามา โดยมีการกำหนดค่าให้กับฟีเจอร์ก่อนนำเข้ากระบวนการ K mean Clustering โดยใช้สูตรปรับค่าดังนี้

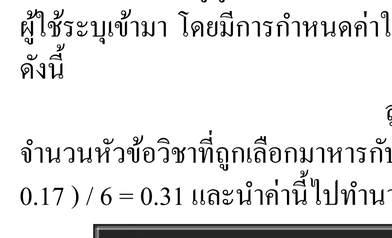
สูตรการปรับค่าวิชา $= (\sum_{k=1}^{n} k)/n$

จำนวนหัวข้อวิชาที่ถูกเลือกมาหารกับจำนวนหัวข้อทั้งหมด ดังภาพที่ 1 จะได้ (0.17 + 0.17 + 1 + 0.17 + 0.17 + 0.17) / 6 = 0.31 และนำค่านี้ไปทำนายใน K-Mean Clustering Model

ข้อมูล	0.166666666666
ตำแหน่งที่ของข้อมูล	0.166666666666
ค่ากลาง ฐานนิยม มัธยฐาน ค่าเฉลี่ยเลขคณิต	1
ค่าการกระจาย พิสัย ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน	0.166666666666
การนำเสนอข้อมูลเชิงคุณภาพและเชิงปริมาณ	0.166666666666
การแปลความหมายของค่าสถิติ	0.166666666666

ภาพที่ 1 แสดงการเลือกหัวข้อเรียนของผู้ใช้

หลังจากการกำหนดค่าแล้ว จะนำข้อมูลของทุกหัวข้อวิชามาทำนายเพื่อจัด Cluster โดยจะต้องกำหนดจำนวน Cluster = 5 ทุกครั้ง ผลลัพธ์การจัด Cluster จะได้ดังภาพที่ 2 ซึ่ง 1 Cluster อาจมีมากกว่า 1 สาขา ซึ่งสนใจเฉพาะข้อมูลที่นำไปทำนายที่อยู่ใน Cluster Y จุดสีแดงดังภาพที่ 2



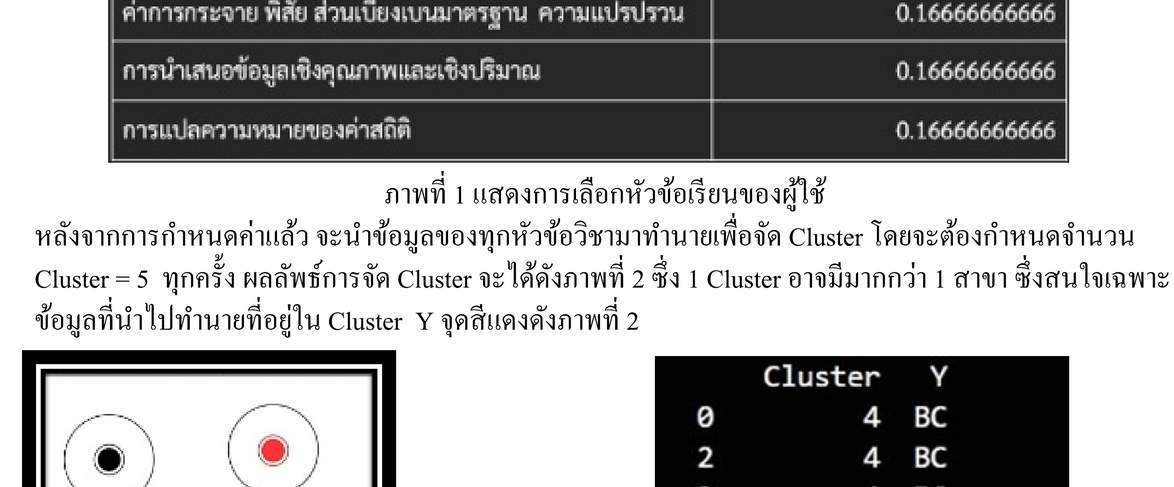
ภาพที่ 2 ตัวอย่างแสดงการจัด Cluster ของ K-Mean Clustering Model

	Cluster	Y
0	4	BC
2	4	BC
3	4	BC
4	4	BC
35	4	CS
62	4	IT
64	4	IT
65	4	IT

ภาพที่ 3 จำนวนสาขาของ Cluster Y (จุดสีแดง)

จากนั้นหาเปอร์เซ็นต์ของแต่ละสาขาใน Cluster Y ซึ่งจะได้ผลลัพธ์ความเหมาะสมดังนี้

CS (Computer Science) 1 / 8 = 0.13, BC (Business Computer) 4 / 8 = 0.50, IT (Information Technology) 3/8 = 0.375 และระบบจะทำการคำนวณหาวิชาที่ต้องเรียนเพิ่มของแต่ละสาขา โดยนำวิชาที่ผู้ใช้กรอกเข้ามาในส่วนของความรู้ภมิหลังไปคำนวน ดังภาพที่ 4



ภาพที่ 4 ผลลัพธ์วิชาที่ต้องเรียนเพิ่มของแต่ละสาขา

2. ส่วนการคำนวนทักษะ

Feature ทักษะประกอบด้วย ทักษะการ Coding, Programming, Artificial Intelligence, Practice, Application, Engineering ภายในจะประกอบด้วยภาษาโปรแกรมหรือความรู้ต่างๆ อ้างอิงจากวิชาเลือกใน มคอ. ยกตัวอย่างโครงสร้างทักษะดังนี้ Coding = 1, Artificial Intelligence = 0, Practice = 1, , Programming = 0, Application/Business = 0, Engineering = 1 โดยทักษะผู้ใช้มีดังนี้ Coding = 0.7, Artificial Intelligence = 0.3, Practice = 0, Programming = 0.7, Application/Business = 0.7, Engineering = 0.2

เงื่อนไขการคำนวณทักษะมีดังนี้

IF B > 0 AND A > 0 return B

ELSE IF B > 0 And A = 0 return 1

ELSE IF B = 0 And A =0 return 1

ELSE return 0

ข้อมูลเทียบกับโครงสร้างวิชาเลือก A[1,0,1,0,0,1] และ ข้อมูลทักษะผู้ใช้ B[0.7,0.3,0,0.7,0,0.2] สำหรับเงื่อนไขคำนวณค่า Result จะได้ผลลัพธ์ดังนี้ Result [0.7,1,0,1,1,0.2] แล้วนำค่าทั้งหมดภายในของ Result มาบวกรวมกันแล้วหารด้วยจำนวนภายใน vector จะได้ 0.6 + 1 + 0 + 1 + 1 + 0.2 = 3.8 / 6 = 0.63 แสดงถึงความเหมาะสมทางด้านทักษะ 63%

3. ส่วนการคำนวณบุคคลิกภาพนักออกแบบ

มีทั้งหมด 60 คำถาม แต่ละคำถามจะมีค่า 1 และ -1 เป็นตัวกำหนดคะแนนว่าควรลดหรือควรเพิ่ม มีตัวเลือกให้เลือก เห็นด้วย ทั้งหมด 7 ระดับ ให้ P เป็น การเลือกระดับความเห็นด้วย

IF P=7 return -4, Else IF P=6 return -3, Else IF P=5 return -1, Else IF P=4 return -1, Else IF P=3 return 1, Else IF P=2 return 3, ELSE P=1, return 4

กำหนดให้ Result = [วิสัยทัศน์, เป็นคนเก็บตัว, ระมัดระวัง, มีเหตุผล, ช่างวางแผน]

46 + 50 + 50 + 50 + 50 = 246 / 5 = 49.2

จากตัวอย่างบุคลิกจะได้ 49.2 % ของการเป็นบุคคลิกภาพนักออกแบบ

ผลการวิจัย

1. ผลการแนะนำข้อมูลหลักสูตรสาขาวิชาคอมพิวเตอร์ด้วยเทคนิคเหมืองข้อมูล จากการดำเนินงานพัฒนาระบบแนะนำสาขาวิชาคอมพิวเตอร์ด้วยเทคนิคเหมืองข้อมูลที่เหมาะสม ทำให้ผู้ใช้สามารถเลือกหัวข้อในวิชาที่มีความเกี่ยวข้องกับสาขาคอมพิวเตอร์ เพื่อให้ระบบนำไปประมวลผล ซึ่งระบบสามารถแนะนำสาขาที่เหมาะสมโดยทำนายจากความรู้ภูมิหลังและแสดงผลเปรียบเทียบเป็นเปอร์เซ็นของแต่ละสาขาดังแสดงเป็นกราฟแท่งตามภาพที่ 6
2. ผลการสอบถามความคิดเห็นของผู้ใช้ที่มีต่อเว็บแอปพลิเคชั่นแนะนำข้อมูลหลักสูตรสาขาวิชาคอมพิวเตอร์ด้วยเทคนิคเหมืองข้อมูลผลการทดสอบจากกลุ่มผู้ใช้ระบบพบว่าเว็บแอปพลิเคชั่นนี้สามารถแนะนำการเลือก สาขาด้านคอมพิวเตอร์สำหรับการเรียนได้ตามที่ต้องการ โดยสามารถแสดงความเหมาะสม ของสาขาต่างๆ ตามที่ผู้ใช้ป้อนข้อมูล ส่งผลให้ผู้ใช้นำไปประกอบการเลือกเรียนได้ ซึ่งเป็นการสะท้อนความรู้ความสามารถของผู้ใช้ให้เลือกเรียนในสาชาที่เหมาะสมมากขึ้น

การอภิปรายผล

ผลการทดสอบจากกลุ่มผู้ใช้ระบบพบว่าเว็บแอปพลิเคชั่นแนะนำสาขาวิชาคอมพิวเตอร์ด้วยเทคนิคเหมืองข้อมูลนี้สามารถแนะนำความเหมาะสมในสาขาวิชาของผู้ใช้งานโดยพิจารณาจากความรู้ ทักษะ บุคคลิกภาพ ที่สอดคล้องกับสาขาที่สนใจ ซึ่งเป็นไปตามทฤษฎีและหลักการทำงานของ K Means Clustering (Google Developers. 2565) (Laura Atema. 2561) โดยการพัฒนาโมเดลดังกล่าว ผู้วิจัยไม่ได้มีการวัดค่าความถูกต้องและความแม่นยำเชิงตัวเลขหรือเปอร์เซนต์ แต่ใช้เป็นแนวทางร่วมกับปัจจัยความเหมาะสมด้านอื่นประกอบกัน เพื่อเป็นแนวทางในการให้คำแนะนำในการเลือกสาขาวิชา และผลการทำนายจากระบบนี้จะเป็นประโยชน์ในการนำไปประกอบการตัดสินใจเพื่อการเลือกสาขาวิชาที่จะศึกษาต่อในระดับมหาวิทยาลัยหรือเลือกสายงานที่เหมาะสมกับตัวเองได้ดียิ่งขึ้น ซึ่งสอดคล้องกับงานวิจัยของ (มาณวิกา กิตติพร. 2562) และ (ปาลิตา แสงศิริ. 2561)

รายการอ้างอิง

กอบเกียรติ สระอุบล. (2563). เรียนรู้ Data Science และ AI Machine Learning ด้วย Python. (พิมพ์ครั้งที่ 1). กรุงเทพฯ : สำนักพิมพ์มีเดีย เนทเวิร์ค.

ชนาธิป ลีนิน. (2563). รู้จักเอกสาร มคอ. สืบค้น 30 มีนาคม 2565. จาก: http://inded.rmutsv.ac.th/qa/?q=th/tqf

ณัฐโชติ พรหมฤทธิ์ และ สัจจาภรณ์ ไวจรรยา. (2563). Fundamental of Deep Learning in Practice. (พิมพ์ครั้งที่ 1). กรุงเทพฯ : สำนักพิมพ์อินโฟเพรส.

ปาลิตา แสงศิริ. (2561). ระบบแนะนำร้านอาหารอัตโนมัติบนสมาร์ทโฟนโดยใช้ข้อมูลเชิงตำแหน่งและรายการอาหาร. สืบค้น 30 มีนาคม 2565. จาก: https://kb.psu.ac.th/psukb/bitstream/2016/12468/1/426822.pdf

มาณวิกา กิตติพร. (2562). ระบบแนะนำทางการศึกษา และเทคนิคการเรียนรู้ของเครื่องจักร (Educational Recommender Systems and Machine Learning). สืบค้น 30 มีนาคม 2565. จาก: https://so03.tci-thaijo.org/index.php/jiskku/article/view/184646/162747

DEMETER ICT (2563). Machine Learning VS Deep Learning. สืบค้น 30 มีนาคม 2565. จาก: https://www.dmit.co.th/th/zendesk-updates-th/machine-learning-vs-deep-learnng/

สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ. (2563). สำนักวิชาการและมาตรฐานการศึกษา (สพฐ.). สืบค้น 30 มีนาคม 2565. จาก: http://academic.obec.go.th/web/home

Google Developers. (2565). k-Means Advantages and Disadvantages. สืบค้น 30 มีนาคม 2565. จาก: https://developers.google.com/machine-learning/clustering/algorithm/advantages-disadvantages

Laura Atema. (2561). Pros and Cons of K-Means Clustering. สืบค้น 30 มีนาคม 2565. จาก: http://prosancons.com/education/pros-and-cons-of-k-means-clustering/

NERIS Analytics Limited. (2565). PERSONALITY TYPES. สืบค้น 30 มีนาคม 2565. จาก: https://www.16personalities.com/