

บทที่ 4

ผลการวิจัย

การเปรียบเทียบประสิทธิภาพของการจำแนกหมวดหมู่ของข้อคิดเห็นในแบบสอบถามปลายเปิด โดยวิธีเนอโฟเบย์และซัพพอร์ตเวกเตอร์แมชชีน เรื่องความคิดเห็นที่มีต่อสถาบันการศึกษาไทยระหว่างภาครัฐและเอกชนของประชาชนในเขตกรุงเทพมหานคร ได้ผลการวิจัยดังต่อไปนี้

- 4.1 ผลการเก็บข้อมูลจากแบบสอบถามและการคัดเลือกข้อมูล
- 4.2 ผลการจำแนกหมวดหมู่ของข้อคิดเห็น
- 4.3 ผลการวัดประสิทธิภาพของการจำแนกหมวดหมู่

4.1 ผลการเก็บข้อมูลจากแบบสอบถามและการคัดเลือกข้อมูล

ประชากรที่ใช้ในการศึกษานี้ คือ ประชากรทั้งเพศชายและเพศหญิงในเขตกรุงเทพมหานคร 5,710,883 คน ทำการสุ่มจำนวนกลุ่มตัวอย่าง 400 คน นั่นคือ ในการเก็บข้อมูลจากแบบสอบถามเป็นสอบถามความคิดเห็นแบบปลายเปิด เรื่อง ความคิดเห็นที่มีต่อสถาบันการศึกษาไทยระหว่างภาครัฐและเอกชนของประชาชนในเขตกรุงเทพมหานคร จำนวน 400 ชุด แบบสอบถามแบ่งออกเป็น 3 ส่วน ดังนี้

ส่วนที่ 1 ข้อมูลพื้นฐานทางประชากรศาสตร์ของกลุ่มตัวอย่าง

ส่วนที่ 2 ความคิดเห็นที่มีต่อสถาบันการศึกษาไทยภาครัฐและเอกชน แบ่งเป็น 4 ด้าน ได้แก่ ด้านราคา ด้านสถานที่ ด้านความรู้และทักษะที่ได้รับ และด้านความยอมรับในสังคม

ส่วนที่ 3 ข้อคิดเห็นเพิ่มเติม

ข้อมูลจากแบบสอบถามที่นำมาใช้ในการวิเคราะห์ คือ ส่วนที่ 2 ความคิดเห็นที่มีต่อสถาบันการศึกษาไทยภาครัฐและเอกชนแบ่งเป็น 4 ด้าน ได้แก่ ด้านราคา ด้านสถานที่ ด้านความรู้ และทักษะที่ได้รับ และด้านความยอมรับในสังคม

ผลการสำรวจเก็บข้อมูล พบว่ามีผู้ตอบแบบสอบถามไม่ครบถ้วนบางด้าน และข้อคิดเห็นบางข้อความไม่สามารถระบุขั้ว (Polar Words) ของข้อคิดเห็นได้ว่าเป็นความคิดเห็นเชิงบวก กลาง หรือลบ จึงจำเป็นต้องมีการคัดเลือกข้อความที่ไม่สามารถระบุขั้วได้ออก ตัวอย่างของลักษณะข้อความที่ไม่สามารถระบุขั้วได้ทางด้านความยอมรับในสังคม (ภาครัฐ) ได้แก่ “ต้องมีความสามารถทางการศึกษา” “ด้านความยอมรับขึ้นอยู่กับคณะที่ศึกษา” “ยอมรับเป็นกรณี” เป็นต้น

หลังการคัดกรองข้อมูลจากแบบสอบถาม 400 ชุด ทำให้ได้ข้อความสำหรับการวิเคราะห์แต่ละด้านไม่เท่ากัน ซึ่งสามารถแบ่งจำนวนข้อมูลสำหรับเรียนรู้ (Train) 80% และข้อมูลทดสอบ (Test) 20% เพื่อใช้ในการวิเคราะห์ได้ ดังตารางที่ 4-1

ตารางที่ 4-1 จำนวนข้อความจากแบบสอบถาม 400 ชุด จำนวนข้อมูลสำหรับเรียนรู้และจำนวนข้อมูลทดสอบ

ข้อคิดเห็น	จำนวนข้อความจากแบบสอบถาม 400 ชุด		จำนวนข้อมูลสำหรับเรียนรู้ (Train)		จำนวนข้อมูลทดสอบ (Test)	
	ภาครัฐ	เอกชน	ภาครัฐ	เอกชน	ภาครัฐ	เอกชน
ด้านราคา	392	389	314	311	78	78
ด้านสถานที่	382	378	306	302	76	76
ด้านความรู้และทักษะที่ได้รับ	387	385	310	308	77	77
ด้านความยอมรับในสังคม	374	388	299	310	75	78

เมื่อเก็บข้อมูลแสดงความคิดเห็นจากแบบทดสอบแบบปลายเปิด เรื่อง ความคิดเห็นที่มีต่อสถาบันการศึกษาไทยระหว่างภาครัฐและเอกชนของประชาชนในเขตกรุงเทพมหานครแล้ว จากนั้นกรอกข้อมูลลงใน Microsoft Office Excel 2007 และทำการระบุขั้ว (Polar Words) เป็นคำที่บ่งบอกถึงระดับความคิดเห็น แบ่งเป็น 3 หมวดหมู่ คือ เชิงบวก (Positive) กลาง (Medium) และลบ (Negative) ซึ่งได้แสดงตัวอย่างของข้อความแสดงความคิดเห็นและการระบุขั้ว ดังตารางที่ 4-2

ตารางที่ 4-2 ตัวอย่างข้อความแสดงความคิดเห็นและการระบุขั้ว

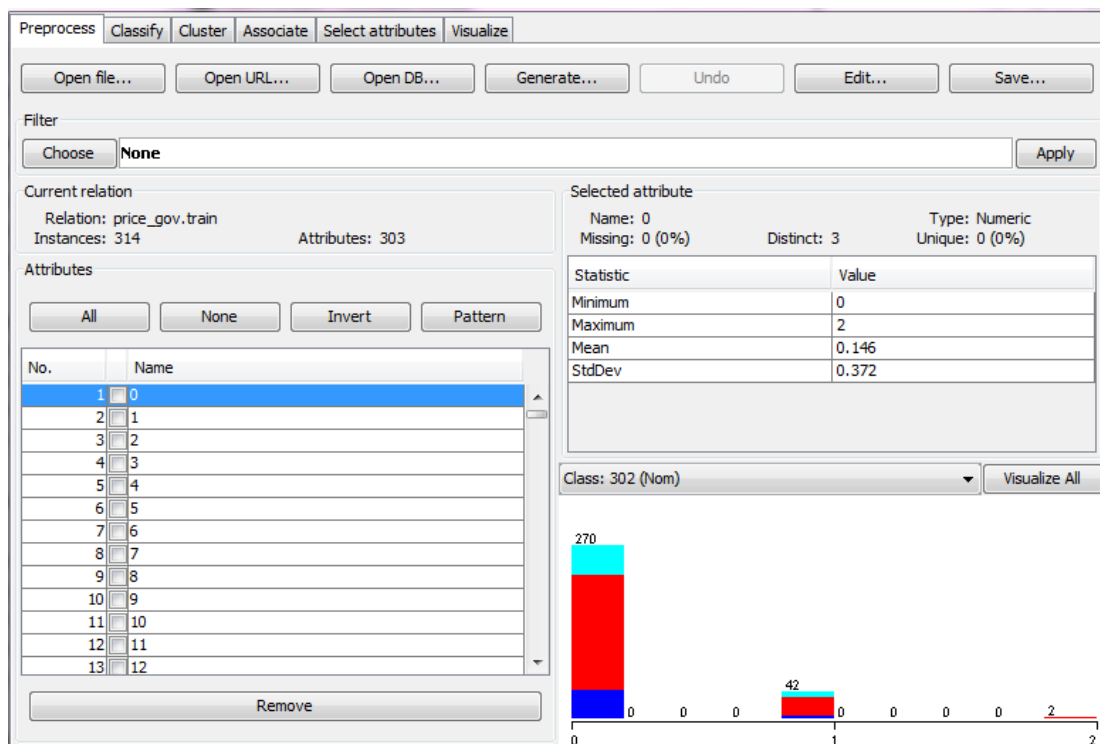
ข้อคิดเห็น	ตัวอย่างข้อความแสดงความคิดเห็นและการระบุขั้ว	
	ภาครัฐ	เอกชน
ด้านราคา	มีราคาเล่าเรียนที่ถูกกว่า<Positive>	ราคาถือว่าไม่แพงเกินไป<Positive>
	ราคาพอใช้ได้<Medium>	พอกับรัฐบาล<Medium>
	บางที่แพงมาก ๆ แพงกว่าเอกชนซะอีกๆ บางที่ราคาเหมาะสม และบางครั้งถูกกว่า ภาครัฐด้วยซ้ำ<Negative>	ค่าเกี่ยวกับการศึกษาก็มีมาก พอสมควร ค่าใช้จ่ายส่วนตัวค่อนข้าง สูง<Negative>

ตารางที่ 4-2 ต่อ

ข้อคิดเห็น	ตัวอย่างข้อความแสดงความคิดเห็นและการระบุข้อ	
	ภาครัฐ	เอกชน
ด้านสถานที่	แน่นอนมหาวิทยาลัยของรัฐจะมีสถานที่ที่ใหญ่เพียงพอต่อจำนวน<Positive>	สถานที่ดูหรูหรา ดูดีกว่าภาครัฐ<Positive>
	มีขนาดกลาง<Medium>	ก็จะคล้ายๆกันกับมหาวิทยาลัยของรัฐ แต่ละที่จะมีขนาดของสถานที่ที่แตกต่างกันไป<Medium>
	สถานที่ที่จะเก่าและไม่ค่อยได้รับการดูแล<Negative>	บางทีก็มีพื้นที่มากอาจจะอยู่ห่างไกลสถานที่สำคัญมาก เดินทางลำบาก<Negative>
ด้านความรู้และทักษะที่ได้รับ	ความรู้มากมาย ทักษะที่ดี<Positive>	บุคลากรเยอะ ได้ความรู้มากมาย<Positive>
	มีความรู้เท่าเทียมกัน ที่อยู่นักศึกษาเองว่ามีความสนใจมากน้อยเพียงใด<Medium>	ให้ความรู้ไม่แตกต่างกัน<Medium>
	ได้รับความรู้เท่าเอกชนแต่เทคโนโลยีจะด้อยกว่า<Negative>	ความรู้และทักษะน้อยกว่า<Negative>
ด้านความยอมรับในสังคม	สังคมจะยอมรับสถาบันของรัฐบาลมากกว่าเอกชน รวมทั้งสถานประกอบการด้วย<Positive>	ยอมรับในสังคมได้ดี ขึ้นกับผู้เรียน<Positive>
	พอๆ กัน<Medium>	เท่าเทียมกัน<Medium>
	ได้รับการยอมรับน้อยถ้าสถาบันไม่มีชื่อ<Negative>	สังคมและสถานประกอบการจะยอมรับสถาบันเอกชนน้อยกว่า<Negative>

4.2 ผลการจำแนกหมวดหมู่ของข้อคิดเห็น

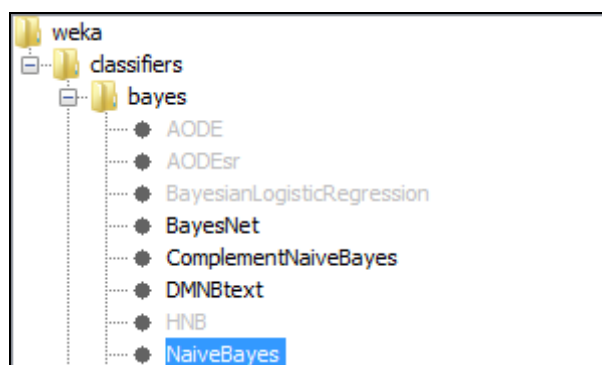
ผลการรันข้อมูลเพื่อการจำแนกหมวดหมู่ด้วยอัลกอริทึม การรันข้อมูลเพื่อการจำแนกหมวดหมู่ด้วยอัลกอริทึมด้วยโปรแกรม WEKA 3.6.2 โดยเปิดไฟล์เพื่อเลือกไฟล์ที่ต้องการวิเคราะห์ ตัวอย่างที่นำมาแสดงเป็นข้อมูลที่ได้จากข้อคิดเห็นที่มีต่อสถาบันการศึกษาไทยด้านราคาของภาครัฐ ดังภาพที่ 4-1



ภาพที่ 4-1 ผลการเปิดไฟล์ที่ต้องการวิเคราะห์ด้วยโปรแกรม WEKA

จากภาพที่ 4-1 Current relation ระบุว่า Relation คือ price_gov.test ซึ่งเป็นการเปิดไฟล์ price_gov.test.arff ซึ่งมี จำนวนเอกสาร (Instance) 314 แถว และมี Attributes จำนวน 303 คอลัมน์

หลังจากนำข้อมูลสำหรับการเรียนรู้เข้าสู่โปรแกรม WEKA แล้วเลือกวิธีการเรียนรู้ข้อมูลและสร้างโมเดลด้วยวิธีเนอฟเบย์ สำหรับการจำแนกหมวดหมู่ข้อคิดเห็นด้วยโปรแกรม WEKA โดยเลือกดังภาพที่ 4-2



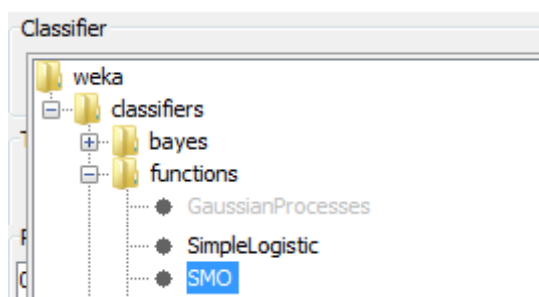
ภาพที่ 4-2 การเลือกใช้อัลกอริทึมเนอฟเบย์ในการเรียนรู้และสร้างโมเดล

ผลการทำนายด้วยข้อมูลทดสอบ โดยตัวอย่างนี้เป็นข้อมูลข้อคิดเห็นที่มีต่อสถาบันการศึกษาไทยด้านราคาของภาครัฐ โดยใช้ข้อมูลทดสอบ จำนวน 78 ตัวอย่าง ได้ผลการทำนายด้วยวิธีเนอ์ฟเบย์ด้วยโปรแกรม WEKA ดังภาพที่ 4-4

=== Predictions on test set ===				
inst#	actual	predicted	error	probability distribution
1	? 2:Positive	+	0.235	*0.694 0.071
2	? 2:Positive	+	0.047	*0.948 0.006
3	? 3:Negative	+	0.04	0.178 *0.783
4	? 2:Positive	+	0.224	*0.536 0.241
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
78	? 2:Positive	+	0.235	*0.694 0.071

ภาพที่ 4-4 ผลการทำนายด้วยวิธีเนอ์ฟเบย์

หลังจากนำข้อมูลสำหรับการเรียนรู้เข้าสู่โปรแกรม WEKA แล้วเลือกวิธีการเรียนรู้ข้อมูลและสร้างโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน สำหรับการจำแนกหมวดหมู่ข้อคิดเห็น ด้วยโปรแกรม WEKA โดยเลือกดังภาพที่ 4-5



ภาพที่ 4-5 การเลือกใช้วิธีซัพพอร์ตเวกเตอร์แมชชีนในการเรียนรู้และสร้างโมเดล

ผลการนำข้อมูลสำหรับเรียนรู้และสร้างโมเดลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน ซึ่งรันในโปรแกรม WEKA ดังภาพที่ 4-6

=== Evaluation on training set ===		
=== Summary ===		
Correctly Classified Instances	301	95.8599 %
Incorrectly Classified Instances	13	4.1401 %
Kappa statistic	0.914	
Mean absolute error	0.2321	
Root mean squared error	0.2898	
Relative absolute error	70.1791 %	
Root relative squared error	71.3763 %	
Total Number of Instances	314	

ภาพที่ 4-6 ผลการรันข้อมูลสำหรับเรียนรู้และสร้างโมเดลด้วยวิธีพอร์ตเวกเตอร์แมชชีน

จากภาพที่ 4-6 เป็นผลการรันข้อมูลสำหรับเรียนรู้ซึ่งเป็นข้อคิดเห็นที่มีต่อสถาบันการศึกษาไทยด้านราคาของภาครัฐ มีข้อมูลจำนวน 314 ตัวอย่าง ด้วยวิธีพอร์ตเวกเตอร์แมชชีนพบว่า มี Attribute จำนวน 3 Attribute ได้แก่ Medium, Positive และ Negative ได้ค่า Correctly Classified Instances เท่ากับ 95.85%

ผลการทำนายด้วยข้อมูลทดสอบ โดยตัวอย่างนี้เป็นข้อมูลข้อคิดเห็นที่มีต่อสถาบันการศึกษาไทยด้านราคาของภาครัฐ โดยใช้ข้อมูลทดสอบ จำนวน 78 ตัวอย่าง ได้ผลการทำนายด้วยวิธีพอร์ตเวกเตอร์แมชชีนด้วยโปรแกรม WEKA ดังภาพที่ 4-7

=== Predictions on test set ===				
inst#,	actual,	predicted,	error,	probability distribution
1	? 2:Positive	+	0.333	*0.667 0
2	? 2:Positive	+	0.333	*0.667 0
3	? 3:Negative	+	0	0.333 *0.66
.	.	.	.	.
78	? 2:Positive	+	0.333	*0.667 0

ภาพที่ 4-7 ผลการทำนายด้วยวิธีพอร์ตเวกเตอร์แมชชีน

4.3 ผลการวัดประสิทธิภาพของการจำแนกหมวดหมู่

การวัดประสิทธิภาพของการจำแนกหมวดหมู่ โดยนำผลการทำนายที่ได้มาคำนวณ ประสิทธิภาพของการจำแนกหมวดหมู่ข้อคิดเห็นในแบบสอบถามปลายเปิด โดยวิธีเอนีฟเบย์ และชัพพอร์ตเวกเตอร์แมชชีน เรื่องความคิดเห็นที่มีต่อสถาบันการศึกษาไทยระหว่างภาครัฐและ เอกชนของประชาชนในเขตกรุงเทพมหานคร คำนวณตามสูตรในบทที่ 3 ข้อที่ 3.6 ได้ผลดังตารางที่ 4-3 และ 4-4

ตารางที่ 4-3 ผลการวัดประสิทธิภาพของการจำแนกหมวดหมู่ด้วยวิธีเอนีฟเบย์

ข้อคิดเห็น	ภาค	วิธีเอนีฟเบย์		
		Precision	Recall	F-Measure
ด้านราคา	รัฐบาล	0.849	0.852	0.850
	เอกชน	0.841	0.844	0.841
ด้านสถานที่	รัฐบาล	0.833	0.836	0.834
	เอกชน	0.834	0.838	0.835
ด้านความรู้และทักษะที่ได้รับ	รัฐบาล	0.800	0.741	0.769
	เอกชน	0.838	0.841	0.838
ด้านความยอมรับในสังคม	รัฐบาล	0.744	0.697	0.719
	เอกชน	0.767	0.717	0.742

จากตารางที่ 4-3 ผลการวัดประสิทธิภาพของการจำแนกหมวดหมู่ด้วยวิธีเอนีฟเบย์ ทั้ง 4 ด้าน พบว่า การจำแนกหมวดหมู่ความคิดเห็นที่มีต่อสถาบันการศึกษาภาครัฐ มีค่า F-Measure อยู่ในช่วง ระหว่าง 0.719-0.850 สำหรับการจำแนกหมวดหมู่ความคิดเห็นที่มีต่อสถาบันการศึกษาภาคเอกชน มีค่า F-Measure อยู่ในช่วงระหว่าง 0.742-0.835

ตารางที่ 4-4 ผลการวัดประสิทธิภาพของการจำแนกหมวดหมู่ด้วยวิธีชัพพอร์ตเวกเตอร์แมชชีน

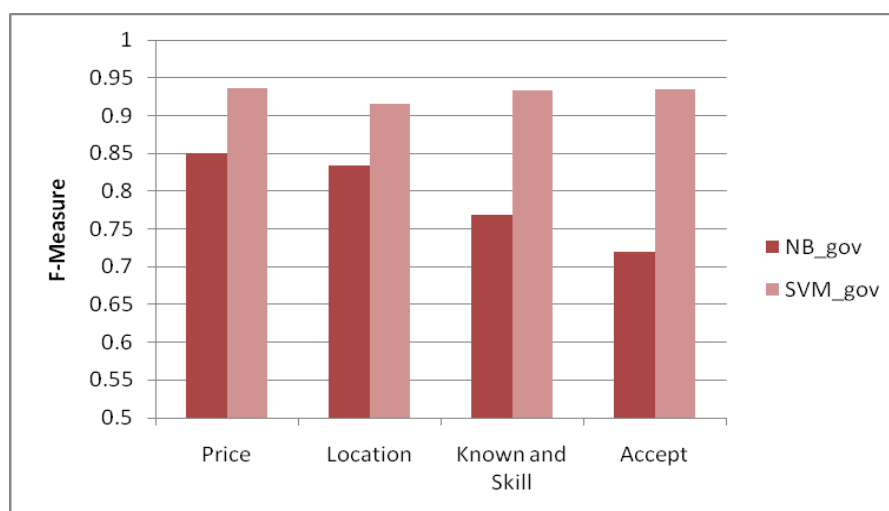
ข้อคิดเห็น	ภาค	วิธีชัพพอร์ตเวกเตอร์แมชชีน		
		Precision	Recall	F-Measure
ด้านราคา	รัฐบาล	0.939	0.938	0.937
	เอกชน	0.959	0.958	0.958
ด้านสถานที่	รัฐบาล	0.927	0.906	0.916
	เอกชน	0.964	0.964	0.963

ตารางที่ 4-4 (ต่อ)

ข้อคิดเห็น	ภาค	วิธีซัพพอร์ตเวกเตอร์แมชชีน		
		Precision	Recall	F-Measure
ด้านความรู้และทักษะที่ได้รับ	รัฐบาล	0.942	0.925	0.933
	เอกชน	0.959	0.958	0.957
ด้านความยอมรับในสังคม	รัฐบาล	0.943	0.926	0.935
	เอกชน	0.913	0.892	0.903

จากตารางที่ 4-4 ผลการวัดประสิทธิภาพของการจำแนกหมวดหมู่ด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนทั้ง 4 ด้าน พบว่า การจำแนกหมวดหมู่ความคิดเห็นที่มีต่อสถาบันการศึกษาภาครัฐ มีค่า F-Measure อยู่ในช่วงระหว่าง 0.916-0.937 สำหรับการจำแนกหมวดหมู่ความคิดเห็นที่มีต่อสถาบันการศึกษาภาคเอกชน มีค่า F-Measure อยู่ในช่วงระหว่าง 0.903-0.963

เมื่อนำค่า F-Measure ของการจำแนกหมวดหมู่ข้อคิดเห็นที่มีต่อสถาบันการศึกษาไทย ระหว่างภาครัฐและเอกชนของประชาชนในเขตกรุงเทพมหานคร ด้วยวิธีเอนีฟเบย์ และวิธีซัพพอร์ตเวกเตอร์แมชชีนทั้ง 4 ด้าน มาสร้างเป็นกราฟเปรียบเทียบประสิทธิภาพระหว่างการจำแนกการจำแนกหมวดหมู่ด้วยวิธีเอนีฟเบย์ และวิธีซัพพอร์ตเวกเตอร์แมชชีนของข้อคิดเห็นที่มีต่อสถาบันการศึกษาไทยภาครัฐ แสดงดังภาพที่ 4-8

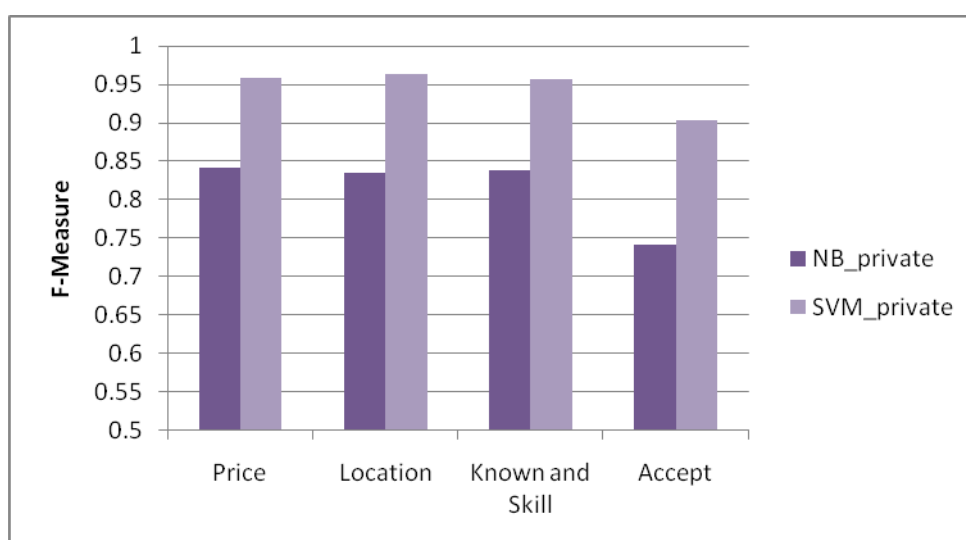


ภาพที่ 4-8 กราฟแสดงผลการเปรียบเทียบประสิทธิภาพการจำแนกหมวดหมู่ของข้อคิดเห็นที่มีต่อสถาบันการศึกษาไทยในภาครัฐทั้ง 4 ด้าน

จากภาพที่ 4-8 เป็นกราฟแสดงผลการเปรียบเทียบประสิทธิภาพการจำแนกหมวดหมู่ของข้อคิดเห็นที่มีต่อสถาบันการศึกษาไทยในภาครัฐทั้ง 4 ด้าน ซึ่งเป็นการวัดประสิทธิภาพด้วยค่า F-Measure โดยพิจารณาเปรียบเทียบระหว่าง 2 วิธีการจำแนกหมวดหมู่ พบว่าประสิทธิภาพของการจำแนกหมวดหมู่ด้วยวิธีเนอโฟเบย์มีค่า F-Measure น้อยกว่าการจำแนกหมวดหมู่ด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน แสดงว่าการจำแนกหมวดหมู่ด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนสามารถระบุข้อความรู้ลึกในเชิงบวก กลางและลบได้ดีกว่าการจำแนกหมวดหมู่ด้วยวิธีเนอโฟเบย์

เมื่อพิจารณาค่า F-Measure โดยเปรียบเทียบข้อคิดเห็นแต่ละด้านทั้ง 4 ด้าน ได้แก่ ด้านราคา (Price) ด้านสถานที่ (Location) ด้านความรู้และทักษะที่ได้รับ (Known and Skill) และด้านความยอมรับในสังคม (Accept) พบว่า การจำแนกหมวดหมู่ด้วยวิธีเนอโฟเบย์ทำให้ข้อคิดเห็นที่มีต่อสถาบันของการศึกษาไทยในภาครัฐ ด้านความยอมรับมีค่า F-Measure ต่ำที่สุดเมื่อเทียบกับด้านอื่น ๆ ส่วนการจำแนกหมวดหมู่ด้วยวิธีการซัพพอร์ตเวกเตอร์แมชชีนทำให้ข้อคิดเห็นที่มีต่อสถาบันของการศึกษาไทยในภาครัฐ ด้านสถานที่ มีค่า F-Measure ต่ำที่สุดเมื่อเทียบกับด้านอื่น ๆ

ผลการสร้างกราฟเปรียบเทียบประสิทธิภาพระหว่างจำแนกหมวดหมู่ข้อคิดเห็นที่มีต่อสถาบันการศึกษาไทยในภาคเอกชนด้วยวิธีเนอโฟเบย์ และวิธีซัพพอร์ตเวกเตอร์แมชชีน แสดงดังภาพที่ 4-9



ภาพที่ 4-9 กราฟแสดงผลการเปรียบเทียบประสิทธิภาพการจำแนกหมวดหมู่ของข้อคิดเห็นที่มีต่อสถาบันการศึกษาไทยในภาคเอกชนทั้ง 4 ด้าน

จากภาพที่ 4-9 เป็นกราฟแสดงผลการเปรียบเทียบประสิทธิภาพการจำแนกหมวดหมู่ของ ข้อคิดเห็นที่มีต่อสถาบันการศึกษาไทยในภาคเอกชนทั้ง 4 ด้านซึ่งเป็นการวัดประสิทธิภาพด้วย ค่า F-Measure เมื่อพิจารณาเปรียบเทียบระหว่าง 2 วิธีการจำแนกหมวดหมู่ พบว่าประสิทธิภาพของการจำแนกหมวดหมู่ด้วยวิธีเอนีฟเบย์มีค่า F-Measure น้อยกว่าการจำแนกหมวดหมู่ด้วยวิธี ชัพพอร์ตเวกเตอร์แมชชีน แสดงว่าการจำแนกหมวดหมู่ด้วยวิธีชัพพอร์ตเวกเตอร์แมชชีนสามารถ ระบุข้อความรู้สึกในเชิงบวก กลางและลบได้ดีกว่าการจำแนกหมวดหมู่ด้วยวิธีเอนีฟเบย์ เมื่อ พิจารณาค่า F-Measure โดยเปรียบเทียบข้อคิดเห็นแต่ละด้านทั้ง 4 ด้าน ได้แก่ ด้านราคา (Price) ด้านสถานที่ (Location) ด้านความรู้และทักษะที่ได้รับ (Known and Skill) และด้านความยอมรับใน สังคม (Accept) พบว่า ส่วนการจำแนกหมวดหมู่ด้วยวิธีทั้งเอนีฟเบย์และชัพพอร์ตเวกเตอร์แมชชีน ทำให้ข้อคิดเห็นที่มีต่อสถาบันของการศึกษาไทยในภาคเอกชน ด้านความยอมรับในสังคมมีค่า F-Measure ต่ำที่สุดเมื่อเทียบกับด้านอื่น ๆ