



การตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอาง
ด้วยวิธีการเปรียบเทียบสายอักขระ



โดย

นางสาวจิราภรณ์ วงษ์เลิศ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรเภสัชศาสตรมหาบัณฑิต
สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ แผน ก แบบ ก 2 ระดับปริญญาโทมหาบัณฑิต

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2565

ลิขสิทธิ์ของมหาวิทยาลัยศิลปากร

การตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้ง
เครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ



โดย
นางสาวจิราภรณ์ วงษ์เลิศ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรเภสัชศาสตรมหาบัณฑิต
สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ แผนก ก แบบ ก 2 ระดับปริญญาโทมหาบัณฑิต

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2565

ลิขสิทธิ์ของมหาวิทยาลัยศิลปากร

DETECTION OF TRADE NAME, COSMETIC NAME, AND INGREDIENTS IN THE COSMETIC NOTIFICATION REQUEST FORM WITH STRING MATCHING METHOD



A Thesis Submitted in Partial Fulfillment of the Requirements
for Master of Pharmacy (HEALTH INFORMATICS)
Graduate School, Silpakorn University
Academic Year 2022
Copyright of Silpakorn University

หัวข้อ การตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบใน
ตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสาย
อักขระ

โดย นางสาวจิราภรณ์ วงษ์เลิศ

สาขาวิชา สารสนเทศศาสตร์ทางสุขภาพ แผนก ก แบบ ก 2 ระดับปริญญา
มหาบัณฑิต

อาจารย์ที่ปรึกษาหลัก รองศาสตราจารย์ ดร. วีรยุทธ์ เลิศนที

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร ได้รับพิจารณาอนุมัติให้เป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรเกสัชศาสตรมหาบัณฑิต

.....คณบดีบัณฑิตวิทยาลัย (ผู้รักษาการแทน)
(ผู้ช่วยศาสตราจารย์ ดร. สานิต นิรัตติศัย)

พิจารณาเห็นชอบโดย

.....ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. สานิต นิรัตติศัย)

.....อาจารย์ที่ปรึกษาหลัก
(รองศาสตราจารย์ ดร. วีรยุทธ์ เลิศนที)

.....ผู้ทรงคุณวุฒิภายนอก
(รองศาสตราจารย์ ดร. อัมรินทร์พันธุ์ รอดทุกข์)

61363309 : สารสนเทศศาสตร์ทางสุขภาพ แผน ก แบบ ก 2 ระดับปริญญาโท

คำสำคัญ : การตัดคำ, การเปรียบเทียบสายอักขระ, ระยะทางเลเวนชเตย์น, เครื่องสำอาง

นางสาว จิราภรณ์ วงษ์เลิศ: การตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก : รองศาสตราจารย์ ดร. วีรยุทธ์ เลิศนที

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) ตรวจสอบคำห้ามใช้ในชื่อการค้าและชื่อเครื่องสำอางในแบบคำขอจดแจ้งเครื่องสำอาง และ 2) ตรวจสอบคำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอางโดยคำนั้นต้องมีความสัมพันธ์สอดคล้องกับส่วนประกอบในตำรับเครื่องสำอาง โปรแกรมที่ถูกพัฒนาขึ้นสำหรับตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับ ในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ ได้รับการประเมินประสิทธิภาพของโปรแกรมโดยเปรียบเทียบผลการพิจารณาชื่อผลิตภัณฑ์เครื่องสำอางกับผลการพิจารณาโดยเจ้าหน้าที่ผู้เชี่ยวชาญ และประเมินความพึงพอใจของผู้ใช้งานที่มีต่อโปรแกรม

กรอบแนวคิดในการพัฒนาโปรแกรมอยู่บนพื้นฐานของทฤษฎีวงจรการพัฒนาระบบ ปัญหาของการตรวจสอบชื่อผลิตภัณฑ์เครื่องสำอางได้ถูกวิเคราะห์พบว่ามี 2 ปัญหา คือ ความล่าช้าและความคลาดเคลื่อนในการตรวจสอบ โปรแกรมประยุกต์บนเว็บสำหรับตรวจสอบชื่อผลิตภัณฑ์เครื่องสำอางได้ถูกพัฒนาขึ้น อัลกอริทึมหลักของโปรแกรมประยุกต์นี้คือการเปรียบเทียบสายอักขระแบบประมาณโดยการหาระยะทางเลเวนชเตย์น การประเมินประสิทธิภาพของโปรแกรมพิจารณาจากค่า F1 การใช้เทคนิค N-gram ในการกำหนดคุณลักษณะของคำเพื่อเพิ่มประสิทธิภาพในการเปรียบเทียบสายอักขระ พบว่าคุณลักษณะของคำที่เหมาะสมกับงานวิจัยนี้คือ รูปแบบ trigram กำหนดค่า threshold ของความคล้ายคลึงกันของสายอักขระที่ 80, 85 และ 90 การกำหนดค่า threshold 80 หมายถึง สายอักขระสองชุดซึ่งมีความคล้ายคลึงกันมากกว่าหรือเท่ากับร้อยละ 80 จะกำหนดว่าเป็นสายอักขระชุดเดียวกัน ในการเปรียบเทียบประสิทธิภาพของโปรแกรมโดยใช้เครื่องมือตัดคำภาษาไทยทั้ง 8 วิธี ร่วมกับ adaptation of LEXITRON dictionary เครื่องมือตัดคำที่มีประสิทธิภาพสูงสุดเมื่อกำหนดค่า threshold เป็น 90 คือ newmm และ newmm-safe ไม่พบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติเมื่อเปรียบเทียบผลการพิจารณาชื่อผลิตภัณฑ์เครื่องสำอางกับผลการพิจารณาโดยเจ้าหน้าที่ผู้เชี่ยวชาญ การประเมินความพึงพอใจของผู้ใช้งานที่มีต่อโปรแกรมทั้ง 3 ด้าน ได้แก่ ด้านตรงตามความต้องการของผู้ใช้งานโปรแกรม, ด้านประสิทธิภาพของโปรแกรม และด้านความสะดวกและง่ายต่อการใช้งาน อยู่ในระดับความพึงพอใจมากที่สุด

โปรแกรมถูกใช้เพื่อการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับ ในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ ผลการพิจารณาชื่อผลิตภัณฑ์เครื่องสำอางมีความถูกต้องและเป็นไปในทิศทางเดียวกันกับการพิจารณาโดยผู้เชี่ยวชาญ จึงสามารถนำไปใช้เพื่อสนับสนุนเจ้าหน้าที่ในการตรวจสอบข้อมูลในแบบคำขอจดแจ้งเครื่องสำอางเบื้องต้น



61363309 : Major (HEALTH INFORMATICS)

Keyword : word segmentation, string matching, Levenstein distance, cosmetic

MISS JIRAPORN WONGLERT :

DETECTION OF TRADE NAME, COSMETIC NAME, AND INGREDIENTS IN THE COSMETIC NOTIFICATION REQUEST FORM WITH STRING MATCHING METHOD THESIS ADVISOR : ASSOCIATE PROFESSOR DR. VERAYUTH LERTNATTEE

The objectives of this study were: 1) to detect the prohibited words for cosmetic trade names and cosmetic names in a cosmetic notification request form, and 2) to detect words that can be used as part of cosmetic trade names and cosmetic names but the cosmetic formulation contains the name-related ingredients. The developed program for checking trade names, cosmetic names, and ingredients in the cosmetic notification request form using the string matching method has been evaluated for the program's performance by comparing its verifying result with that from experts and users' satisfaction.

The program development framework was based on the system development life cycle theory. The problems in verifying cosmetic product names were analyzed with two problems identified including being time consuming and inaccuracy in the verification of product names. A web-based application for verifying cosmetic product names was developed. The main algorithm in the application was an approximate string matching by Levenshtein edit distance. The performance of the application was assessed using F1. The N-gram technique was applied to create a set of appropriate features for increasing the performance in string matching. The appropriate n-gram in this study was trigram with thresholds at 80, 85, and 90. The threshold of 80 meant that two strings with a text similarity greater than or equal to 80 percent were defined as the same string. In comparing the efficiency of the program using the 8 engines of Thai word segmentation with the adaptation of the LEXITRON dictionary, the best engines at thresholds set to 90, were the newmm and newmm-safe. There were no significant differences in the results on the verification of cosmetic product names of the program and those from experts. The satisfaction assessments from users were the

functional requirement satisfaction, the program function satisfaction, and the convenience and ease of use of the program were excellent.

This program was developed for verifying trade names, cosmetic names, and ingredients in the cosmetic notification request form using the string matching method. The verification results on cosmetic product names are accurate and consistent with those of experts. It can be used to support authorities in the preliminary verification of information in the cosmetic notification request form.



กิตติกรรมประกาศ

การวิจัยครั้งนี้สำเร็จลุล่วงได้ด้วยความรู้และความอนุเคราะห์ของรองศาสตราจารย์ ดร. เกษักรวีรยุทธ เลิศนที อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้ความรู้และคำแนะนำอันมีคุณค่า รวมถึงตรวจทานและแก้ไขข้อบกพร่องต่าง ๆ จนทำให้วิทยานิพนธ์เล่มนี้เสร็จสมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง

ผู้วิจัยขอกราบขอบพระคุณผู้ช่วยศาสตราจารย์ ดร. เกษักรสาธิต นิรัติศัย ประธานกรรมการสอบ และรองศาสตราจารย์ ดร.อัมฉินนุพันธ์ รอดทุกข์ กรรมการผู้ทรงคุณวุฒิภายนอก ที่ได้กรุณาให้คำแนะนำและชี้แนะแนวทางการแก้ไขอันเป็นประโยชน์ทำให้งานวิจัยเล่มนี้มีความสมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณอาจารย์ภาควิชาสารสนเทศศาสตร์ทางสุขภาพ คณะเภสัชศาสตร์ มหาวิทยาลัยศิลปากรทุกท่าน ที่ให้ความรู้ คำแนะนำ จากการศึกษาในรายวิชาต่าง ๆ จนทำให้เกิดความคิดต่อยอดนำมาจัดทำวิจัยในครั้งนี้

ขอขอบพระคุณเภสัชกรสุชาติ ถนอมวารภรณ์ หัวหน้ากลุ่มงานคุ้มครองผู้บริโภคและเภสัชสาธารณสุข สำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา ตลอดจนเจ้าหน้าที่กลุ่มงานคุ้มครองผู้บริโภคทุกท่าน ที่ได้ให้ความช่วยเหลือจนงานวิจัยนี้สำเร็จลุล่วงไปได้ด้วยดี

ขอขอบคุณเพื่อนร่วมรุ่น พี่ ๆ น้อง ๆ ภาควิชาสารสนเทศศาสตร์ทางสุขภาพ คณะเภสัชศาสตร์ มหาวิทยาลัยศิลปากร และกัลยาณมิตรทุกท่านที่คอยช่วยเหลือและเป็นกำลังใจให้ผู้วิจัยเสมอมา

ขอขอบคุณการสนับสนุนทุนอุดหนุนการวิจัยประเภทนักศึกษาในระดับบัณฑิตศึกษาจากกองทุนวิจัยและสร้างสรรค์ คณะเภสัชศาสตร์ มหาวิทยาลัยศิลปากร

สุดท้ายนี้ คุณค่าหรือประโยชน์ รวมไปถึงความสำเร็จใดอันเกิดจากวิทยานิพนธ์เล่มนี้ ผู้วิจัยขอมอบให้กับบิดา มารดา ครูอาจารย์ผู้ประสิทธิ์ประสาทวิชา และกัลยาณมิตรทุกท่านที่คอยให้กำลังใจจนวิทยานิพนธ์เล่มนี้สำเร็จลุล่วงไปได้ด้วยดี

นางสาว จิราภรณ์ วงษ์เลิศ

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	ฉ
กิตติกรรมประกาศ.....	ช
สารบัญ.....	ฅ
สารบัญตาราง.....	ฉ
สารบัญภาพ.....	ฐ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของงานวิจัย.....	2
1.3 กรอบแนวคิดงานวิจัย.....	3
1.4 ขอบเขตของงานวิจัย.....	3
1.5 ข้อจำกัดของงานวิจัย.....	3
1.6 นิยามศัพท์เฉพาะ.....	4
1.7 ประโยชน์ที่คาดว่าจะได้รับ.....	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	5
2.1 ทฤษฎีที่เกี่ยวข้อง.....	5
2.1.1 เครื่องสำอาง.....	5
2.1.2 การจดแจ้งเครื่องสำอางระบบอัตโนมัติ.....	6
2.1.3 คำห้ามใช้.....	8
2.1.4 คำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าหรือชื่อเครื่องสำอางได้.....	8
2.1.5 ฐานข้อมูล (database).....	9

2.1.6 การตัดคำภาษาไทย (Thai words segmentation)	12
2.1.7 การเปรียบเทียบสายอักขระ	15
2.2 งานวิจัยที่เกี่ยวข้อง	23
บทที่ 3 วิธีการดำเนินการวิจัย	25
3.1 ประชากร กลุ่มตัวอย่าง วิธีสุ่มตัวอย่าง	25
3.2 แบบแผนการวิจัย	28
3.3 เครื่องมือที่ใช้ในการวิจัย	29
3.4 ขั้นตอนในการดำเนินการวิจัย	29
3.5 การสร้างเครื่องมือที่ใช้ในการศึกษาวิจัย	30
3.6 การเก็บรวบรวมข้อมูล	37
3.7 การวิเคราะห์ข้อมูลและประเมินผล	37
3.8 สถิติที่ใช้ในการวิเคราะห์ข้อมูล	39
บทที่ 4 ผลการวิเคราะห์ข้อมูล	42
4.1 การพัฒนาโปรแกรมสำหรับใช้ในการตรวจสอบซื้อการค้า ชื่อเครื่องสำอาง และส่วนประกอบ ในตำรับเครื่องสำอาง	42
4.2 การประเมินประสิทธิภาพของโปรแกรม	45
4.3 การประเมินความพึงพอใจของผู้ใช้งานโปรแกรม	57
บทที่ 5 สรุป อภิปราย และข้อเสนอแนะ	62
รายการอ้างอิง	71
ประวัติผู้เขียน	94

สารบัญตาราง

หน้า

ตารางที่ 1 ข้อมูลผลการตรวจสอบแบบคำขอดแจ่งเครื่องสำอองโดยใช้เครื่องมือ attacut ในการตัดคำภาษาไทย เมื่อกำหนดคุณลักษณะของคำโดยใช้ trigram และกำหนดค่า threshold 80	27
ตารางที่ 2 ตารางคำห้ามใช้ในชื่อผลิตภัณฑ์เครื่องสำออง (restricted_name).....	31
ตารางที่ 3 ตารางชื่อกลุ่มสารที่สามารถใช้เป็นชื่อผลิตภัณฑ์เครื่องสำอองได้ (allowed_name).....	31
ตารางที่ 4 ตารางคำที่สามารถใช้แทนที่ชื่อกลุ่มสารได้ (allowed_name_synonym).....	32
ตารางที่ 5 ตารางสารที่ใช้เป็นส่วนประกอบในตำรับเครื่องสำออง (ingredient)	32
ตารางที่ 6 ตารางชื่อสารที่สามารถใช้แทนที่สารที่เป็นส่วนประกอบในตำรับเครื่องสำอองได้ (ingredient_synonym).....	32
ตารางที่ 7 ตารางความสัมพันธ์ระหว่างชื่อผลิตภัณฑ์เครื่องสำอองกับสารที่เป็นส่วนประกอบ (name_and_ingredient).....	32
ตารางที่ 8 เกณฑ์ในการแปลความหมายความพึงพอใจตามค่าเฉลี่ย	41
ตารางที่ 9 ค่า F1 ของการตัดคำแต่ละคุณลักษณะที่ค่า threshold 80	45
ตารางที่ 10 ตัวอย่างผลการตัดคำด้วย Library PyThaiNLP โดยใช้เครื่องมือตัดคำทั้ง 8 วิธี	46
ตารางที่ 11 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยทั้ง 8 วิธี และกำหนดค่า threshold ที่ 80	48
ตารางที่ 12 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยทั้ง 8 วิธี และกำหนดค่า threshold ที่ 85	49
ตารางที่ 13 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยทั้ง 8 วิธี และกำหนดค่า threshold ที่ 90	50
ตารางที่ 14 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยร่วมกับการใช้ adaptation of LEXITRON dictionary และกำหนดค่า threshold ที่ 80	53
ตารางที่ 15 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยร่วมกับการใช้ adaptation of LEXITRON dictionary และกำหนดค่า threshold ที่ 85	54

ตารางที่ 16 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยร่วมกับการใช้ adaptation of LEXITRON dictionary และกำหนดค่า threshold ที่ 90 55

ตารางที่ 21 การจัดเรียงข้อมูลผลการตรวจสอบในตาราง two-dimensional contingency 57

ตารางที่ 22 เปรียบเทียบผลการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอาง (n = 400 รายการคำขอ) ระหว่างการตรวจสอบด้วยโปรแกรมและการตรวจสอบโดยเจ้าหน้าที่ผู้เชี่ยวชาญ ด้วยสถิติ McNemar’s Chi-Square test 57

ตารางที่ 17 ภาพรวมผลการประเมินความพึงพอใจของผู้ใช้งานโปรแกรมสำหรับการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอาง 59

ตารางที่ 18 ผลการประเมินความพึงพอใจของผู้ใช้งานด้านตรงตามความต้องการของผู้ใช้งานโปรแกรม..... 59

ตารางที่ 19 ผลการประเมินความพึงพอใจของผู้ใช้งานด้านประสิทธิภาพของโปรแกรม 60

ตารางที่ 20 ผลการประเมินความพึงพอใจของผู้ใช้งานด้านความสะดวกและง่ายต่อการใช้งาน 60



สารบัญภาพ

	หน้า
รูปที่ 1 ตัวอย่างการหาระยะทางเลเวนชเตย์น ชั้นตอนที่ 1: iteration (i) = 0.....	18
รูปที่ 2 ตัวอย่างการหาระยะทางเลเวนชเตย์น ชั้นตอนที่ 2: iteration (i) = 1.....	18
รูปที่ 3 ตัวอย่างการหาระยะทางเลเวนชเตย์น ชั้นตอนที่ 3: iteration (i) = 2.....	19
รูปที่ 4 ตัวอย่างการหาระยะทางเลเวนชเตย์น ชั้นตอนสุดท้าย: iteration (i) = n, j=m.....	19
รูปที่ 5 การจัดเรียงข้อมูลในตาราง two-dimensional contingency สำหรับ McNemar's Test	26
รูปที่ 6 แผนภาพแสดงความสัมพันธ์ระหว่างข้อมูล (entity relationship diagram; ER diagram)	33
รูปที่ 7 ขั้นตอนการทำงานของโปรแกรม	35
รูปที่ 8 ตาราง confusion matrix	38
รูปที่ 9 หน้าจอส่วนของการรับข้อมูล.....	42
รูปที่ 10 ตัวอย่างการกรอกข้อมูลเพื่อทำการตรวจสอบ	43
รูปที่ 11 ตัวอย่างผลลัพธ์จากการตรวจสอบข้อมูล กรณีผลการพิจารณาไม่ผ่าน.....	44
รูปที่ 12 ตัวอย่างผลลัพธ์จากการตรวจสอบข้อมูล กรณีผลการพิจารณาผ่าน.....	44
รูปที่ 13 ผลการเปรียบเทียบค่า F1 ของคำแต่ละคุณลักษณะ ที่ค่า threshold 80, 85 และ 90.....	46
รูปที่ 14 ผลการเปรียบเทียบประสิทธิภาพของโปรแกรมจากค่า F1.....	51
รูปที่ 15 ผลการเปรียบเทียบประสิทธิภาพของโปรแกรมเมื่อใช้เครื่องมือตัดคำวิธีต่าง ๆ ร่วมกับการใช้ adaptation of LEXiTRON dictionary.....	56
รูปที่ 16 ผลการเปรียบเทียบประสิทธิภาพของโปรแกรมระหว่างการใช้เครื่องมือตัดคำเพียงอย่างเดียวกับการใช้เครื่องมือตัดคำร่วมกับการใช้ adaptation of LEXiTRON dictionary ที่ threshold 90	65
รูปที่ 17 ตัวอย่างการเปรียบเทียบสายอักขระและค่าความคล้ายคลึงกันของคำ.....	68
รูปที่ 18 ข้อมูลความยาวอักขระของคำที่พบในงานวิจัย.....	68
รูปที่ 19 ข้อมูลคำที่พบในงานวิจัย	69

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

อุตสาหกรรมเครื่องสำอาง (cosmetics industry) เป็นอุตสาหกรรมหนึ่งที่ยังคงมีแนวโน้มเติบโตอย่างต่อเนื่อง ในปัจจุบันผู้บริโภคทั้งชายและหญิงในทุก ๆ ช่วงวัยต่างให้ความสนใจและใส่ใจในการดูแลตัวเองมากขึ้น ทั้งในด้านความงามและผิวพรรณ ส่งผลให้เกิดการวิจัยและพัฒนาผลิตภัณฑ์ประเภทใหม่ๆ มากขึ้น จากข้อมูลของ Euromonitor [1] อุตสาหกรรมเครื่องสำอางประเทศไทยในปี 2564 เติบโตร้อยละ 5 หรือมีมูลค่ารวม 1,447 แส่นล้านบาท กลุ่มผลิตภัณฑ์ดูแลผิวครองส่วนแบ่งสูงสุดร้อยละ 57.5 รองลงมาคือ กลุ่มผลิตภัณฑ์ดูแลเส้นผม ร้อยละ 21 กลุ่มผลิตภัณฑ์ตกแต่งร้อยละ 15.5 และน้ำหอมร้อยละ 6 ตามลำดับ นอกจากนี้จากข้อมูลของกรมเจรจาการค้าระหว่างประเทศ [2] พบว่า ประเทศไทยเป็นผู้ส่งออกเครื่องสำอางอันดับที่ 2 ของอาเซียน รองจากสิงคโปร์ และเป็นอันดับที่ 10 ของโลก ด้วยเหตุผลข้างต้นจึงทำให้อุตสาหกรรมเครื่องสำอางเป็นอุตสาหกรรมหนึ่งในหลาย ๆ อุตสาหกรรมที่ได้รับความสนใจจากผู้ประกอบการไม่ว่าจะเป็นผู้ประกอบการรายย่อยหรือรายใหญ่ ตามพระราชบัญญัติเครื่องสำอาง พ.ศ. 2558 กำหนดให้ผู้ประกอบการที่ประสงค์จะผลิต (original brand manufacturer: OBM), รับจ้างผลิต (original equipment manufacturer: OEM) หรือนำเข้าเครื่องสำอางจะต้องจดทะเบียนรายละเอียดยของเครื่องสำอางต่อผู้รับจดทะเบียน และเมื่อผู้รับจดทะเบียนออกใบรับจดทะเบียนให้แล้ว จึงจะสามารถผลิตหรือนำเข้าเครื่องสำอางนั้นได้ [3]

สำนักงานคณะกรรมการอาหารและยา กระทรวงสาธารณสุข ได้มีนโยบายให้ดำเนินการยื่นคำขอจดทะเบียนเครื่องสำอางและรับจดทะเบียนเครื่องสำอางทุกชนิดผ่านระบบจดทะเบียนเครื่องสำอางอัตโนมัติ e-Submission ซึ่งเป็นโปรแกรมที่ช่วยอำนวยความสะดวกให้แก่ผู้ประกอบการ ทำให้การยื่นคำขอจดทะเบียนเครื่องสำอางทำได้รวดเร็วมมากขึ้น ส่งผลให้มีคำขอจดทะเบียนเครื่องสำอางในระบบจดทะเบียนเครื่องสำอางอัตโนมัติ e-Submission เพิ่มมากขึ้น จากการศึกษาของเกษรา จันทรวงศ์ไพศาล และชิตชนก เรือนก้อน [4] พบว่าปริมาณคำขอจดทะเบียนเครื่องสำอาง ตั้งแต่เดือนตุลาคม 2558 จนถึงเดือนกันยายน 2562 มีจำนวนทั้งสิ้น 815,627 รายการ เฉลี่ยปีละประมาณ 203,907 รายการ ในการทำงานของพนักงานเจ้าหน้าที่ผู้ทำหน้าที่พิจารณาคำขอจึงต้องมีความรวดเร็วมากยิ่งขึ้นด้วย ดังนั้น การจัดการให้มีระบบที่ดีและดำเนินการอย่างเป็นระบบจะช่วยให้เจ้าหน้าที่สามารถตรวจสอบและพิจารณาคำขอจดทะเบียนเครื่องสำอางได้ถูกต้องและรวดเร็วมยิ่งขึ้น

จากการที่สำนักคณะกรรมการอาหารและยาได้มอบหมายให้พนักงานเจ้าหน้าที่สำนักงานสาธารณสุขจังหวัดเป็นผู้รับจดแจ้งเครื่องสำอางผ่านระบบจดแจ้งเครื่องสำอางอัตโนมัติ e-Submission ตั้งแต่วันที่ 15 สิงหาคม พ.ศ. 2562 เป็นต้นมา ในปี 2564 จังหวัดพระนครศรีอยุธยา มีรายการคำขอจดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ต้องพิจารณาประมาณ 1,200 รายการ หรือเฉลี่ย 100 คำขอต่อเดือน โดยมีระยะเวลาในการพิจารณาสำหรับเครื่องสำอางความเสี่ยงสูงไม่เกิน 3 วันทำการต่อหนึ่งคำขอ สำหรับกระบวนการทำงานของพนักงานเจ้าหน้าที่ในการตรวจสอบและพิจารณาคำขอจดแจ้งเครื่องสำอางจะเกี่ยวข้องกับการตรวจสอบคำห้ามใช้ในชื่อการค้าและชื่อเครื่องสำอาง และตรวจสอบคำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอาง โดยคำนั้นต้องมีความสัมพันธ์สอดคล้องกับส่วนประกอบในตำรับเครื่องสำอาง โดยพนักงานเจ้าหน้าที่จะศึกษาข้อมูลในคู่มือหลักเกณฑ์การจดแจ้งเครื่องสำอาง ที่กลุ่มกำกับดูแลเครื่องสำอางก่อนออกสู่ตลาด สำนักควบคุมเครื่องสำอางและวัตถุอันตราย สำนักคณะกรรมการอาหารและยา จัดทำไว้ให้ในรูปแบบหนังสือหรือเอกสารอิเล็กทรอนิกส์ ในกระบวนการทำงานของพนักงานเจ้าหน้าที่จะอ่านข้อมูลจากในคู่มือและใช้ความจำ ใช้ประสบการณ์จากสิ่งที่ปฏิบัติมาแล้วจนชำนาญ เปิดตำราดูข้อมูล หรือปรึกษาผู้เชี่ยวชาญเพื่อตรวจสอบและพิจารณาคำขอจดแจ้งเครื่องสำอาง ซึ่งการทำงานในลักษณะนี้ทำให้เกิดความผิดพลาดได้ง่าย นอกจากนี้ข้อมูลหรือองค์ความรู้ที่มีอยู่ก็กระจัดกระจาย ไม่มีการจัดเก็บอย่างเป็นระบบ ทำให้ใช้เวลานานในการตรวจสอบ ผู้วิจัยจึงต้องการออกแบบโปรแกรมที่จะช่วยแก้ไขปัญหาเหล่านี้ จึงได้ศึกษาเทคโนโลยีทางคอมพิวเตอร์ในเรื่องการเปรียบเทียบสายอักขระ (string matching method) และนำความรู้มาประยุกต์ใช้ในการช่วยตรวจสอบชื่อการค้าชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางให้มีความถูกต้อง รวดเร็ว และเป็นไปในทิศทางเดียวกัน

1.2 วัตถุประสงค์ของงานวิจัย

1.2.1 เพื่อตรวจสอบคำห้ามใช้ในชื่อการค้าและชื่อเครื่องสำอางในแบบคำขอจดแจ้งเครื่องสำอาง

1.2.2 เพื่อตรวจสอบคำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอาง โดยคำนั้นต้องมีความสัมพันธ์สอดคล้องกับส่วนประกอบในตำรับเครื่องสำอาง

1.3 กรอบแนวคิดงานวิจัย



1.4 ขอบเขตของงานวิจัย

ในงานวิจัยฉบับนี้ผู้วิจัยมุ่งเน้นการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอาง (แบบ จ.ค.1) ที่ถูกส่งเข้ามาในระบบจดแจ้งเครื่องสำอางอัตโนมัติ e-Submission ของสำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยาเท่านั้น

1.5 ข้อยกเว้นของงานวิจัย

การวิจัยครั้งนี้ผู้วิจัยได้ดำเนินการออกแบบและจัดทำฐานข้อมูลของคำที่ห้ามนำมาใช้เป็นชื่อหรือส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอาง และคำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอางได้หากคำนั้นมีความสัมพันธ์สอดคล้องกับรายการสารในตำรับเครื่องสำอาง โดยข้อมูลในส่วน of คำห้ามใช้และคำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอางได้ ทั้งภาษาไทยและภาษาอังกฤษ เป็นคำที่ระบุในคู่มือหลักเกณฑ์การพิจารณาการจดแจ้งเครื่องสำอาง (ฉบับปรับปรุง 2564) ที่กลุ่มกำกับดูแลเครื่องสำอางก่อนออกสู่ตลาด สำนักควบคุมเครื่องสำอางและวัตถุอันตราย สำนักคณะกรรมการอาหารและยาจัดทำไว้ หากมีคำห้ามใช้หรือคำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอางได้ที่นอกเหนือจากในคู่มือดังกล่าว อาจทำให้โปรแกรมไม่สามารถตรวจพบคำห้ามใช้หรือคำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอางได้หากคำนั้นมีความสัมพันธ์สอดคล้องกับรายการสารในตำรับเครื่องสำอาง

1.6 นิยามศัพท์เฉพาะ

1.6.1 ชื่อการค้า (trade name) หมายถึง ชื่อที่ใช้ในการประกอบกิจการ หรือชื่อที่ผู้ประกอบการตั้งขึ้นเพื่อให้ผลิตภัณฑ์เครื่องสำอางเป็นที่รู้จัก มีลักษณะเป็นคำหรือข้อความที่เป็นภาษาไทยและภาษาอังกฤษที่มีความหมายตรงกันหรือสอดคล้องกัน

1.6.2 ชื่อเครื่องสำอาง (cosmetic name) หมายถึง ชื่อที่ผู้ประกอบการตั้งขึ้นเพื่อให้ข้อมูลหรืออธิบายคุณสมบัติของผลิตภัณฑ์เครื่องสำอาง มีลักษณะเป็นคำหรือข้อความที่เป็นภาษาไทยและภาษาอังกฤษที่มีความหมายตรงกันหรือสอดคล้องกัน และการตั้งชื่อเครื่องสำอางจะต้องมีความสอดคล้องกับส่วนประกอบในตำรับเครื่องสำอาง

1.6.3 ส่วนประกอบในตำรับ (ingredients) หมายถึง รายการสารที่ใช้เป็นส่วนผสมในการผลิตหรือนำเข้าเครื่องสำอางตามประกาศกระทรวงสาธารณสุข มีลักษณะเป็นรายการภาษาอังกฤษให้ผู้ประกอบการเลือกในระบบจัดแจ้งเครื่องสำอางอัตโนมัติ e-Submission หากผู้ประกอบการรายใดมีความประสงค์ใช้สารที่ไม่มีให้เลือกระบบจัดแจ้งเครื่องสำอางอัตโนมัติ ต้องทำการยื่นเอกสารขอเพิ่มรายการสารและแนบเอกสารที่เกี่ยวข้อง โดยแจ้งไปที่สำนักควบคุมเครื่องสำอางและวัตถุอันตราย สำนักงานคณะกรรมการอาหารและยา กระทรวงสาธารณสุข

1.6.4 แบบคำขอจัดแจ้งเครื่องสำอาง (แบบ จ.ค.1) หมายถึง แบบฟอร์มสำหรับผู้ประกอบการหรือผู้จัดแจ้งในการแจ้งรายละเอียดเกี่ยวกับเครื่องสำอาง

1.6.5 ใบรับจัดแจ้งเครื่องสำอาง หมายถึง ใบแสดงรายละเอียดของเครื่องสำอางแต่ละผลิตภัณฑ์ที่ผู้รับจัดแจ้งออกให้แก่ผู้จัดแจ้ง

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1.7.1 เพื่อศึกษาความรู้เรื่องการเปรียบเทียบสายอักขระและนำมาพัฒนาโปรแกรมช่วยในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจัดแจ้งเครื่องสำอาง

1.7.2 นำโปรแกรมที่พัฒนาขึ้นมาประยุกต์ใช้ในองค์กรเพื่อช่วยให้เจ้าหน้าที่สามารถตรวจสอบและพิจารณาคำขอจัดแจ้งเครื่องสำอางได้ถูกต้องและรวดเร็วมากยิ่งขึ้น ช่วยลดระยะเวลาในการทำงาน

1.7.3 โปรแกรมสำหรับการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจัดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระที่พัฒนาขึ้นนี้ สามารถนำไปประยุกต์ใช้ในการพิจารณาคำขอจัดแจ้งเครื่องสำอางในสำนักงานสาธารณสุขจังหวัดแห่งอื่น ๆ หรือเพื่อเป็นแนวทางในการนำไปพัฒนาต่อไป

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในงานวิจัยนี้เป็นงานวิจัยที่เกี่ยวข้องกับการตรวจสอบซื้อการค้ำ ชื่อเครื่องสำอาง และ ส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ โดยผู้วิจัยจะทำการเก็บรวบรวมข้อมูลชื่อการค้ำ ชื่อเครื่องสำอาง ส่วนประกอบในตำรับเครื่องสำอาง จากในแบบคำขอจดแจ้งเครื่องสำอาง (แบบ จ.ค.1) ที่ถูกส่งเข้ามาในระบบจดแจ้งเครื่องสำอาง อัตโนมัติ e-Submission ของสำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา และทำการตรวจสอบ หาคำห้ามใช้ในชื่อการค้ำและชื่อเครื่องสำอาง และตรวจสอบหาคำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อ การค้ำและชื่อเครื่องสำอางโดยคำนั้นต้องมีความสัมพันธ์สอดคล้องกับส่วนประกอบในตำรับ เครื่องสำอาง ซึ่งนอกจากการตรวจสอบชื่อการค้ำ ชื่อเครื่องสำอาง และส่วนประกอบในตำรับ เครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระแล้วนั้น ผู้วิจัยยังต้องการที่จะพัฒนาระบบฐานข้อมูล ให้มีประสิทธิภาพ และศึกษาถึงความพึงพอใจของผู้ใช้งานโปรแกรมอีกด้วย ดังนั้นผู้วิจัยจึงจำเป็นต้อง ศึกษาแนวคิด ทฤษฎี บทความทางวิชาการ และงานวิจัยต่าง ๆ ที่เกี่ยวข้อง เพื่อให้ได้องค์ความรู้ที่ เหมาะสมต่อการนำไปใช้สำหรับการดำเนินการวิจัย ซึ่งผู้วิจัยได้ศึกษาแนวคิด ทฤษฎี รวมไปถึง งานวิจัยต่างๆ เพื่อใช้เป็นแนวทางในการศึกษา ดังต่อไปนี้

- 2.1 ทฤษฎีที่เกี่ยวข้อง
 - 2.1.1 เครื่องสำอาง
 - 2.1.2 การจดแจ้งเครื่องสำอางระบบอัตโนมัติ
 - 2.1.3 คำห้ามใช้
 - 2.1.4 คำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อการค้ำหรือชื่อเครื่องสำอางได้
 - 2.1.5 ฐานข้อมูล
 - 2.1.6 การตัดคำ
 - 2.1.7 การเปรียบเทียบสายอักขระ

2.2 งานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 เครื่องสำอาง

พระราชบัญญัติเครื่องสำอาง พ.ศ. 2558 มาตรา 4 [3] ระบุความหมายไว้ว่า เครื่องสำอาง หมายถึง

(1) วัตถุประสงค์มุ่งหมายสำหรับใช้ทา ภู นวด โรอย ฟน หยอด ใส่ อบ หรือกระทำด้วยวิธีอื่นใด กับส่วนภายนอกของร่างกายมนุษย์ และให้หมายความรวมถึงการใช้กับฟันและเยื่อในช่องปาก โดยมีวัตถุประสงค์เพื่อความสะอาด ความสวยงาม หรือเปลี่ยนแปลงลักษณะที่ปรากฏหรือระงับกลิ่นกาย หรือปกป้องดูแลส่วนต่างๆ นั้น ให้อยู่ในสภาพดี และรวมตลอดทั้งเครื่องประทีนต่างๆ สำหรับผิวด้วย แต่ไม่รวมถึงเครื่องประดับและเครื่องแต่งตัวซึ่งเป็นอุปกรณ์ภายนอกร่างกาย

(2) วัตถุประสงค์มุ่งหมายสำหรับใช้เป็นส่วนผสมในการผลิตเครื่องสำอางโดยเฉพาะ หรือ

(3) วัตถุประสงค์ที่กำหนดโดยกฎกระทรวงให้เป็นเครื่องสำอาง

2.1.2 การจดทะเบียนเครื่องสำอางระบบอัตโนมัติ

ในการขออนุญาตเกี่ยวกับเครื่องสำอาง จะมีอยู่ 2 ประเด็นหลัก คือ สถานที่ผลิตนำเข้า หรือเก็บรักษาเครื่องสำอางจะต้องเป็นไปตามมาตรฐานที่สำนักงานคณะกรรมการอาหารและยากำหนด และผลิตภัณฑ์เครื่องสำอางจะต้องมีคุณภาพ ปลอดภัย และมีความเป็นธรรมต่อผู้บริโภค การขออนุญาตเกี่ยวกับเครื่องสำอางในปัจจุบันนั้นใช้ระบบออนไลน์เข้ามาช่วย ทำให้มีความสะดวก รวดเร็วมากยิ่งขึ้น โดยผู้ประกอบการสามารถส่งข้อมูลรายละเอียดเกี่ยวกับผลิตภัณฑ์เครื่องสำอางที่ต้องการขออนุญาตผ่านระบบจดทะเบียนเครื่องสำอางอัตโนมัติ e-Submission ของสำนักงานคณะกรรมการอาหารและยา

สำนักงานคณะกรรมการอาหารและยาได้มีนโยบายให้ระบบจดทะเบียนเครื่องสำอาง ทุกชนิดเป็นระบบอัตโนมัติทั้งหมด โดยผู้ประกอบการจะเป็นผู้ส่งข้อมูลเข้าระบบด้วยตนเอง ผ่านทางเว็บไซต์การจดทะเบียนเครื่องสำอาง <http://privus.fda.moph.go.th> โดยแจ้งรายละเอียดดังนี้ [5]

2.1.2.1 ประเภทคำขอจดทะเบียนเครื่องสำอาง แบ่งเป็น 4 ประเภท คือ

- 1) คำขอจดทะเบียนเครื่องสำอางแบบผลิตเพื่อขาย หรือรับจ้างผลิตเพื่อขาย
- 2) คำขอจดทะเบียนเครื่องสำอางแบบนำเข้าเพื่อขาย
- 3) คำขอจดทะเบียนเครื่องสำอางแบบผลิตเพื่อการส่งออกเท่านั้น
- 4) คำขอจดทะเบียนเครื่องสำอางแบบผลิตเพื่อขาย กรณีแบ่งบรรจุ

2.1.2.2 สถานที่ผลิต/นำเข้า/เก็บรักษาเครื่องสำอาง

2.1.2.3 ผู้ว่าจ้างผลิตเครื่องสำอาง (ถ้ามี)

2.1.2.4 รูปแบบผลิตภัณฑ์ แบ่งเป็น 4 รูปแบบ คือ

- 1) ผลิตภัณฑ์เดี่ยว

- 2) ผลิตภัณฑ์ชนิดเดียวกันที่ประกอบด้วยหลายชนิดสี บรรจุรวมในบรรจุภัณฑ์เดียวกัน ไม่สามารถแยกจำหน่ายได้
- 3) ผลิตภัณฑ์ต่างชนิดกันบรรจุในภาชนะบรรจุเดียวกัน ไม่สามารถแยกจำหน่ายได้
- 4) ชุดผลิตภัณฑ์ที่อยู่ใน 1 ชุด ประกอบด้วยผลิตภัณฑ์เดี่ยวประเภทเดียวกัน หรือผลิตภัณฑ์เดี่ยวหลายประเภท บรรจุรวมในบรรจุภัณฑ์เดียวกัน (ชุดเซ็ต)

2.1.2.5 ชื่อการค้า ภาษาไทยและภาษาอังกฤษ

2.1.2.6 ชื่อเครื่องสำอาง ภาษาไทยและภาษาอังกฤษ

2.1.2.7 รูปแบบการใช้ผลิตภัณฑ์ แบ่งเป็น 2 รูปแบบ คือ

- 1) ผลิตภัณฑ์ใช้แล้วล้างออก (rinse-off products)
- 2) ผลิตภัณฑ์ใช้แล้วไม่ต้องล้างออก (leave-on products)

2.1.2.8 บริเวณที่ใช้ผลิตภัณฑ์

2.1.2.9 วัตถุประสงค์ในการใช้งานของผลิตภัณฑ์

2.1.2.10 วิธีใช้

2.1.2.11 ลักษณะทางกายภาพของผลิตภัณฑ์

2.1.2.12 ลักษณะทางกายภาพของภาชนะบรรจุ

2.1.2.13 เงื่อนไขการใช้ผลิตภัณฑ์ เช่น

- 1) ห้ามใช้ในเด็กอายุต่ำกว่า 3 ปี
- 2) ห้ามใช้ในผลิตภัณฑ์ประเภทฉีดพ่นหรือสเปรย์ (spray)
- 3) ต้องมีการผสมผลิตภัณฑ์อื่นก่อนใช้ โดยอัตราส่วนผสมของผลิตภัณฑ์นี้ คือ อัตราส่วน 1 :

2.1.2.14 รายการสารที่ใช้เป็นส่วนผสมในการผลิตเครื่องสำอาง โดยการแจ้งชื่อสารจะต้องแจ้งเป็นชื่อตาม International Nomenclature of Chemical Ingredient (INCI) หรือชื่อที่ได้รับการยอมรับตามเอกสารอ้างอิงมาตรฐาน หรือระบุเป็นชื่อ Color Index Number (CI No.) โดยสามารถสืบค้นได้จากฐานข้อมูลเครื่องสำอาง ได้แก่ ฐานข้อมูล CosIng (<http://ec.europa.eu/growth/tools-databases/cosing/>), CTFA เป็นต้น

ผลิตภัณฑ์เครื่องสำอางที่ต้องการผลิตหรือนำเข้าทุกรายการก่อนออกจำหน่ายสู่ท้องตลาด จะต้องทำการจดแจ้งผ่านระบบจดแจ้งเครื่องสำอางอัตโนมัติ e-Submission ซึ่งจะต้องเป็นไปตามหลักเกณฑ์ในการจดแจ้งเครื่องสำอางที่สำนักงานคณะกรรมการอาหารและยา กำหนด พนักงานเจ้าหน้าที่จะพิจารณารายละเอียดที่แจ้งในแบบแจ้งทุกหัวข้อว่ามีความสอดคล้องกัน และไม่ขัดต่อข้อกำหนดกฎระเบียบด้านเครื่องสำอาง เช่น การแสดงชื่อการค้าและชื่อเครื่องสำอางจะต้องไม่ก่อให้เกิดความเข้าใจผิดเกี่ยวกับเครื่องสำอาง รวมทั้งไม่มีคำที่ห้ามนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอาง รายการสารที่เป็นส่วนประกอบในตำรับเครื่องสำอางจะต้องเป็นไปตามประกาศกระทรวงสาธารณสุข และไม่พบสารที่ห้ามใช้เป็นส่วนผสมในเครื่องสำอาง นอกจากนี้ยังต้องตรวจสอบในเรื่องของการนำคำที่อาจนำมาใช้เป็นชื่อหรือส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอางโดยคำนั้นจะต้องมีความสัมพันธ์สอดคล้องกับรายการสารที่เป็นส่วนประกอบในตำรับเครื่องสำอางด้วย

2.1.3 คำห้ามใช้

คำที่ห้ามนำมาใช้เป็นชื่อหรือส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอาง รวมถึงคำพ้องรูป/พ้องเสียง ที่อาจทำให้เข้าใจผิดในสาระสำคัญของเครื่องสำอาง เช่น acne control, anti-inflammation, medication, คลายกล้ามเนื้อ, ลดการอักเสบ, รักษาสิว, รักษาฝ้า เป็นต้น [6]

2.1.4 คำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าหรือชื่อเครื่องสำอางได้

คำที่สามารถนำมาใช้เป็นชื่อหรือส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอาง โดยคำนั้นจะต้องมีความสัมพันธ์สอดคล้องกับสารที่เป็นส่วนประกอบในผลิตภัณฑ์เครื่องสำอาง [6] เช่น

- สามารถใช้คำว่า “VITAMIN C” เป็นส่วนหนึ่งของชื่อได้ เมื่อในสูตรมีสารที่สื่อถึงวิตามินซีหรืออนุพันธ์ของวิตามินซี เช่น ascorbic acid, ascorbyl palmitate
- สามารถใช้คำว่า “SUNSCREEN” หรือ “UV PROTECTION” เป็นส่วนหนึ่งของชื่อได้ เมื่อในสูตรมีส่วนผสมของสารป้องกันแสงแดด (28 รายการตามประกาศฯ)

เนื่องจากข้อมูลคำห้ามใช้และคำที่สามารถนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอางมีลักษณะเป็นคู่มือหรือไฟล์อิเล็กทรอนิกส์ ถูกจัดเก็บในแหล่งข้อมูลหลายแห่ง ทำให้ใช้เวลานานในการสืบค้นข้อมูล รวมถึงการเพิ่ม ลบ และแก้ไขข้อมูลทำได้ค่อนข้างยาก ดังนั้นจึงจำเป็นต้องรวบรวมข้อมูลคำห้ามใช้และคำที่สามารถนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอางดังกล่าวเพื่อนำเข้าสู่ฐานข้อมูล ซึ่งการจัดเก็บข้อมูลไว้ในฐานข้อมูลเดียวกันจะทำให้ผู้ใช้งานสามารถสืบค้นข้อมูลได้โดยง่าย การเพิ่ม ลบ และแก้ไขข้อมูลทำได้สะดวก และสามารถนำไปใช้ประโยชน์ในการเปรียบเทียบสายอักขระต่อไป

2.1.5 ฐานข้อมูล (database)

2.1.5.1 ความหมายของฐานข้อมูล

ฐานข้อมูล หมายถึง กลุ่มข้อมูลขนาดใหญ่ที่รวบรวมไว้ที่ใดที่หนึ่งอย่างเป็นระบบ โดยเป็นข้อมูลที่จัดเก็บอาจมีความสัมพันธ์กันหรือไม่สัมพันธ์กันก็ได้ มีซอฟต์แวร์เข้ามาช่วยควบคุมการจัดเก็บ ค้นหา และประมวลผลข้อมูล ซึ่งทำให้ผู้ใช้สามารถเรียกใช้งานได้อย่างมีประสิทธิภาพ โดยซอฟต์แวร์ที่ใช้ควบคุมฐานข้อมูล เรียกว่าระบบจัดการฐานข้อมูล (database management system: DBMS) [7]

กิตติ ภัคดีวัฒนกุล และทวีศักดิ์ กาญจนสุวรรณ [8] ระบุว่า ฐานข้อมูล หมายถึง การจัดเก็บข้อมูลที่มีความสัมพันธ์กันไว้ด้วยกัน เพื่อลดปัญหาความซ้ำซ้อนของข้อมูลและการที่ไม่สามารถใช้ข้อมูลร่วมกันได้ เพื่อให้เป็นข้อมูลที่ใช้สนับสนุนการดำเนินงานอย่างใดอย่างหนึ่งขององค์กร

อาภากร ธาตุโลหะ [9] ได้ให้ความหมายของฐานข้อมูลไว้ว่า เป็นแหล่งรวมของสารสนเทศที่มีความสัมพันธ์กัน และนำมาบันทึกไว้ในรูปแบบที่อ่านได้ด้วยคอมพิวเตอร์ ด้วยโปรแกรมการจัดการ และสามารถค้นคืนข้อมูลต่าง ๆ เช่น สื่อสิ่งพิมพ์ สื่อโสตทัศน์ ที่ได้บันทึกไว้มาแสดงผลทางหน้าจอ ฐานข้อมูลหนึ่งอาจประกอบด้วยแฟ้มข้อมูลหลาย ๆ แฟ้ม หรือเพียงแฟ้มข้อมูลเดียวก็ได้ โดยมีระบบจัดการข้อมูลเป็นตัวกำหนดโครงสร้างที่ใช้ในการจัดเก็บและค้นคืน

โอภาส เอี่ยมสิริวงศ์ ระบุว่าฐานข้อมูลเป็นศูนย์รวมของข้อมูลต่าง ๆ ที่มีความสัมพันธ์กัน (relationship) โดยจะมีกระบวนการจัดหมวดหมู่ข้อมูลอย่างมีระเบียบแบบแผน ก่อให้เกิดฐานข้อมูลที่เป็นแหล่งรวมของข้อมูลจากแผนกต่าง ๆ ซึ่งถูกจัดเก็บไว้อย่างมีระบบภายในฐานข้อมูลชุดเดียว โดยผู้ใช้งานแต่ละแผนก สามารถเข้าถึงข้อมูลส่วนกลางนี้เพื่อนำไปประมวลผลร่วมกันได้ [10]

จากที่นักวิชาการหลายท่านได้ให้ความหมายของฐานข้อมูลไว้ สามารถสรุปได้ว่า ฐานข้อมูล หมายถึง กลุ่มข้อมูลที่มีความสัมพันธ์กันและถูกนำมารวบรวมไว้ในที่เดียวกันอย่างเป็นระบบเพื่อนำไปใช้ในวัตถุประสงค์อย่างใดอย่างหนึ่ง มีระบบการจัดการฐานข้อมูลซึ่งทำหน้าที่เป็นตัวกลางในการติดต่อระหว่างผู้ใช้กับฐานข้อมูล ทำให้ผู้ใช้สามารถใช้ข้อมูลได้อย่างมีประสิทธิภาพ

2.1.5.2 แบบจำลองฐานข้อมูล

แบบจำลองฐานข้อมูลโดยทั่วไปที่รู้จักกันในปัจจุบันมี 4 แบบ [11] คือ

1) แบบจำลองฐานข้อมูลลำดับชั้น (hierarchical database model)

มีโครงสร้างแบบต้นไม้ (tree structure) สนับสนุนความสัมพันธ์ของข้อมูลเป็นแบบหนึ่งต่อหนึ่ง (1:1) หรือหนึ่งต่อกลุ่ม (1:M) แถวบนเรียกว่าระเบียบพ่อแม่ (parent record) แถว

ถัดมาเรียกว่าระเบียบลูก (child record) แฉวที่อยู่สูงสุดจะเรียกว่า root และในระดับล่างสุดจะเรียกว่า leaves มีโครงสร้างซับซ้อนน้อยที่สุด เหมาะกับข้อมูลที่มีการเรียงลำดับอย่างต่อเนื่อง

2) แบบจำลองฐานข้อมูลเครือข่าย (network database model)

เป็นแบบจำลองที่สนับสนุนความสัมพันธ์ของข้อมูลทุกแบบ ได้แก่ แบบหนึ่งต่อหนึ่ง (1:1), แบบหนึ่งต่อกลุ่ม (1:M) และแบบกลุ่มต่อกลุ่ม (M:N) ลักษณะแบบจำลองฐานข้อมูลเครือข่ายจะมีโครงสร้างของข้อมูลแต่ละแฟ้มข้อมูลที่มีความสัมพันธ์คล้ายร่างแห สามารถเชื่อมโยงข้อมูลแบบไป-กลับได้ มีความยืดหยุ่นในด้านของการค้นหาข้อมูลดีกว่าแบบจำลองลำดับชั้น ใช้พอยน์เตอร์ในการเข้าถึงข้อมูลได้ทันที สามารถเข้าถึงเรคคอร์ดได้โดยตรง ทำให้การป้องกันความปลอดภัยข้อมูลมีน้อย

3) แบบจำลองฐานข้อมูลเชิงสัมพันธ์ (relational database model)

เป็นแบบจำลองที่ได้รับความนิยมกันแพร่หลายในปัจจุบัน โปรแกรมระบบจัดการฐานข้อมูลส่วนมากออกแบบมาสนับสนุนกับแบบจำลองชนิดนี้และนำมาใช้เป็นแนวคิดสำหรับฐานข้อมูลเชิงสัมพันธ์ (relational database) ที่เป็นฐานข้อมูลชนิดหนึ่งที่ยอมรับใช้กันมากที่สุด แบบจำลองฐานข้อมูลเชิงสัมพันธ์จะนำเสนอข้อมูลและความสัมพันธ์ระหว่างข้อมูลในรูปปริเลชัน (relation) หรือเรียกอีกอย่างหนึ่งว่าตาราง (table) ประกอบด้วย แฉว และคอลัมน์ มีความสัมพันธ์กับตารางอื่นได้ทั้งแบบหนึ่งต่อหนึ่ง (1:1), แบบหนึ่งต่อกลุ่ม (1:M) และแบบกลุ่มต่อกลุ่ม (M:N) ใช้คีย์ (key) ในการเชื่อมโยงเพื่ออ้างอิงถึงตารางอื่นๆที่เกี่ยวข้องกัน

4) แบบจำลองฐานข้อมูลเชิงวัตถุ (object-oriented database model)

เกิดจากแนวคิดการเขียนโปรแกรมเชิงวัตถุ (object-oriented program: OOP) โดยมองทุกสิ่งเป็นวัตถุ แต่ละวัตถุจะเป็นแหล่งรวมของข้อมูลและการปฏิบัติ มีคลาสเป็นตัวกำหนดคุณสมบัติหรือรายละเอียดของวัตถุ รวมทั้งคุณสมบัติการปกปิดความลับของวัตถุ การเข้าถึงข้อมูล จะต้องมีการตอบรับจากเมธอดในวัตถุนั้นว่าจะอนุญาตหรือไม่ ที่จะให้วัตถุที่ส่งข้อความร้องขอเพื่อเข้าถึงข้อมูลตน แบบจำลองฐานข้อมูลเชิงวัตถุจัดการกับข้อมูลชนิดต่าง ๆ ที่มีความซับซ้อนได้ดีไม่ว่าจะเป็นภาพกราฟฟิก วิดีโอ และเสียง

2.1.5.3 รูปแบบความสัมพันธ์ข้อมูล

รูปแบบความสัมพันธ์ของข้อมูล แบ่งเป็น 3 รูปแบบ [12] ได้แก่

1) ความสัมพันธ์แบบหนึ่งต่อหนึ่ง (one to one relationships)

เป็นการแสดงความสัมพันธ์ของข้อมูลของเอนทิตี (entity) หนึ่งว่ามีความสัมพันธ์กับข้อมูลในอีกเอนทิตีหนึ่ง ในลักษณะหนึ่งต่อหนึ่ง หรือเขียนได้เป็น 1:1 แสดงว่าข้อมูลเพียงหนึ่ง

ข้อมูลของเอนทิตีแรกมีความสัมพันธ์กับข้อมูลในเอนทิตีที่สองเพียงข้อมูลเดียว เช่น นักศึกษาหนึ่งคน จะมีสูติบัตรเพียงใบเดียวเท่านั้น สูติบัตรหนึ่งใบก็เป็นของนักศึกษาได้เพียงคนเดียวเท่านั้นเช่นกัน

2) ความสัมพันธ์แบบหนึ่งต่อกลุ่ม (one to many relationships)

เป็นการแสดงความสัมพันธ์ของข้อมูลของเอนทิตีหนึ่งว่ามีความสัมพันธ์กับข้อมูลในอีกเอนทิตีหนึ่ง ในลักษณะหนึ่งต่อกลุ่ม หรือเขียนได้เป็น 1:M แสดงว่าข้อมูลเพียงหนึ่งข้อมูลของเอนทิตีแรกมีความสัมพันธ์กับข้อมูลในเอนทิตีที่สองหลายข้อมูล เช่น ลูกค้าหนึ่งคนมีใบเสร็จได้หลายใบ เนื่องจากลูกค้าหนึ่งคนอาจมาซื้อสินค้าหลายครั้ง แต่ใบเสร็จหนึ่งใบต้องเป็นของลูกค้าเพียงคนเดียวเท่านั้น

3) ความสัมพันธ์แบบกลุ่มต่อกลุ่ม (many to many relationships)

เป็นการแสดงความสัมพันธ์ของข้อมูลของเอนทิตีหนึ่งว่ามีความสัมพันธ์กับข้อมูลในอีกเอนทิตีหนึ่ง ในลักษณะกลุ่มต่อกลุ่ม หรือเขียนได้เป็น M:N แสดงว่าข้อมูลเพียงหนึ่งข้อมูลของเอนทิตีแรก มีความสัมพันธ์กับข้อมูลในเอนทิตีที่สองหลายข้อมูล และทำนองเดียวกัน ข้อมูลเพียงหนึ่งข้อมูลของเอนทิตีที่สองมีความสัมพันธ์ในทางกลับกันกับข้อมูลในเอนทิตีที่แรกหลายข้อมูล เช่น นักศึกษาสามารถลงทะเบียนเรียนได้หลาย ๆ วิชาพร้อมกันในแต่ละครั้ง และในทางตรงกันข้ามในแต่ละวิชา 1 วิชา จะประกอบด้วยนักศึกษาหลายคนมาลงทะเบียนในวิชานั้น

2.1.5.4 ประโยชน์ของระบบฐานข้อมูล

การจัดเก็บข้อมูลรวมกันเป็นฐานข้อมูลจะก่อให้เกิดประโยชน์ดังนี้ [8]

1) สามารถลดความซ้ำซ้อนของข้อมูล (data redundancy) โดยจัดเก็บข้อมูลให้รวมอยู่ในที่ ๆ เดียวกัน โดยมีความสัมพันธ์กันของแต่ละตารางที่เกี่ยวข้องกันโดยข้อมูลของระเบียบหนึ่ง ๆ จะไม่ซ้ำซ้อนในอีกตารางหนึ่ง ทำให้การปรับปรุง แก้ไข หรือการควบคุมข้อมูลกระทำได้ง่าย และสามารถแบ่งข้อมูลกันใช้ได้ระหว่างผู้ใช้หลาย ๆ คน รวมทั้งการใช้ข้อมูลเดียวกันในเวลาพร้อม ๆ กันได้อีกด้วย

2) สามารถหลีกเลี่ยงความขัดแย้งของข้อมูล (data inconsistency) เนื่องจากไม่ต้องจัดเก็บข้อมูลที่ซ้ำซ้อนกันในหลายแฟ้มข้อมูล ดังนั้นการแก้ไขข้อมูลในแต่ละชุดจะไม่ก่อให้เกิดค่าที่แตกต่างกัน

3) แต่ละหน่วยงานในองค์กรสามารถใช้ข้อมูลร่วมกันได้

4) สามารถกำหนดให้ข้อมูลมีรูปแบบที่เป็นมาตรฐานเดียวกันได้ เพื่อให้ผู้ใช้ข้อมูลในฐานข้อมูลชุดเดียวกัน สามารถเข้าใจและสื่อสารถึงความหมายเดียวกัน

5) สามารถกำหนดระบบความปลอดภัยให้กับข้อมูลได้ โดยกำหนดระดับความสามารถในการเรียกใช้ข้อมูลของผู้ใช้แต่ละคนได้แตกต่างกันตามความรับผิดชอบ

6) สามารถรักษาความถูกต้องของข้อมูลได้ โดยระบุกฎเกณฑ์ในการควบคุมความผิดพลาดที่อาจเกิดขึ้นจากการป้อนข้อมูลผิด

7) สามารถตอบสนองต่อความต้องการใช้ข้อมูลในหลายรูปแบบ

8) ทำให้ข้อมูลเป็นอิสระจากโปรแกรมที่ใช้งานข้อมูลนั้น (data independence) ซึ่งส่งผลให้ผู้พัฒนาโปรแกรมสามารถแก้ไขโครงสร้างข้อมูล โดยไม่กระทบต่อโปรแกรมที่เรียกใช้งานข้อมูลนั้น เช่น ในกรณีที่ต้องเปลี่ยนขนาดของฟิลด์ (filed) สำหรับระบบแฟ้มข้อมูลจะกระทำได้อย่าง เนื่องจากต้องเปลี่ยนแปลงตัวโปรแกรมที่อ้างถึงฟิลด์นั้นทั้งหมด ซึ่งต่างจากการใช้ระบบฐานข้อมูลที่ การอ้างอิงข้อมูลจะไม่ขึ้นอยู่กับโครงสร้างทางกายภาพของข้อมูลจึงไม่ส่งผลให้ต้องแก้ไขโปรแกรมที่ เรียกใช้ข้อมูลนั้นมากนัก

จากข้อมูลในแบบคำขอจดแจ้งเครื่องสำอาง ชื่อการค้าและชื่อเครื่องสำอางประกอบด้วยข้อความที่เป็นภาษาไทยและภาษาอังกฤษ แต่เนื่องจากภาษาไทยเป็นภาษาที่มีการเขียนคำในลักษณะที่ยาวต่อเนื่องกันโดยไม่มีการใช้ตัวอักษรหรือสัญลักษณ์ที่นำมาใช้คั่นระหว่างคำ ไม่มีการใช้เครื่องหมายวรรคตอนแสดงการแบ่งคำดังเช่นภาษาอังกฤษ นอกจากนี้ยังมีโครงสร้างของคำที่ไม่แน่นอน ซึ่งเป็นอุปสรรคอย่างหนึ่งในการนำคำภาษาไทยไปใช้ในการเปรียบเทียบสายอักขระ ในการประมวลผลทางภาษานั้นจำเป็นต้องทราบขอบเขตของคำก่อนถึงจะสามารถนำไปประมวลผลต่อไปได้ การตัดคำจึงมีความสำคัญเป็นอย่างยิ่งในการแบ่งคำแต่ละคำในประโยคออกจากกันและนำคำเหล่านั้นไปตรวจสอบกับฐานข้อมูลที่ได้จัดทำไว้ เพื่อตรวจสอบหาคำห้ามใช้ในชื่อการค้าและชื่อเครื่องสำอาง และตรวจสอบหาคำที่อาจนำมาใช้เป็นชื่อหรือส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอางได้โดยคำนั้นจะต้องมีความสัมพันธ์สอดคล้องกับรายการสารที่เป็นส่วนประกอบในตำรับเครื่องสำอาง

2.1.6 การตัดคำภาษาไทย (Thai words segmentation)

ภาษาไทยเป็นภาษาที่มีการเขียนติดกันไปทั้งประโยค จึงต้องมีการประมวลผลทางภาษาที่เรียกว่าการตัดคำ (word segmentation หรือ word separation) ซึ่งเป็นการแบ่งคำแต่ละคำในประโยคออกจากกันเพื่อวัตถุประสงค์ในการใช้ประโยชน์ในด้านต่างๆ เช่น การจัดรูปแบบเอกสารในการประมวลผลคำ (word processing) การตรวจสอบตัวสะกดภาษาไทย (spelling check) การวิเคราะห์ไวยากรณ์ (syntax analysis) การแปลภาษาด้วยเครื่องจักร (machine translation) การทำดัชนีสำหรับเอกสาร (document indexing) การเชื่อมโยงความหมายของคำ (thesaurus) การประมวลผลภาษาธรรมชาติ (natural language processing: NLP) การสังเคราะห์เสียงพูด (speech synthesis) การวิเคราะห์กฎเกณฑ์ในการสร้างประโยค (syntactic rules analysis) เป็นต้น [13]

วิธีที่ใช้ตัดคำ

วิธีที่ใช้ตัดคำแบ่งออกเป็น 3 ประเภทใหญ่ๆ ได้แก่ การใช้กฎ การใช้พจนานุกรม และการใช้คลังข้อความ

1) การใช้กฎ (rule-based approach)

การตัดคำโดยการตรวจสอบกฎเกณฑ์ทางอักขระวิธีโดยอาศัยหลักไวยากรณ์ภาษาไทยที่มีการกำหนดลักษณะการประสมอักษร ลักษณะการเว้นวรรค และการขึ้นย่อหน้า โดยวิธีการนี้เริ่มจากการตัดพยางค์เนื่องจากพยางค์มีรูปแบบที่แน่นอนมากกว่าคำ จากนั้นนำพยางค์ที่ได้มาใช้เป็นเกณฑ์ในการกำหนดขอบเขตของคำแต่ละคำ ซึ่งวิธีการนี้มีข้อดีคือ มีความรวดเร็วในการทำงานและใช้ทรัพยากรในการประมวลผลน้อย แต่มีข้อจำกัดในการทำงานคือ จะมีความถูกต้องของการตัดคำในระดับพยางค์สูง แต่ความถูกต้องของการตัดคำค่อนข้างต่ำ และผลของการตัดคำอาจได้เป็นกลุ่มคำที่สามารถตัดคำออกไปได้อีก

2) การใช้พจนานุกรม (dictionary approach)

การตัดคำโดยใช้พจนานุกรมเป็นการตัดคำโดยการนำสายอักขระ (string) มาเปรียบเทียบกับคำศัพท์ที่มีอยู่ในพจนานุกรม เพื่อหาว่าสายอักขระดังกล่าวควรตัดคำที่บริเวณใด และประกอบด้วยคำใดบ้าง ซึ่งวิธีนี้จะต้องทำการจัดเก็บคำไว้ในพจนานุกรมก่อน วิธีการตัดคำโดยใช้พจนานุกรมนี้ทำให้ได้ความถูกต้องในการตัดคำสูงกว่าวิธีการใช้กฎแต่จะใช้เวลามากกว่า และมีข้อจำกัดบางประการ เช่น อาจมีคำบางคำในสายอักขระ แต่ไม่มีในพจนานุกรม ทำให้ได้ผลของการตัดคำไม่ถูกต้อง

3) การใช้คลังข้อความ (corpus-based approach)

การตัดคำโดยใช้คลังข้อมูลเป็นการตัดคำโดยนำวิธีการทางสถิติเข้ามาใช้ในการประมวลภาษา โดยใช้คลังข้อมูลทางภาษาเป็นฐานความรู้เกี่ยวกับค่าความถี่ที่ใช้ในการตัดคำ ซึ่งการตัดคำโดยใช้คลังข้อมูลแบ่งออกเป็น 2 วิธี คือการตัดคำโดยอาศัยความน่าจะเป็น (probabilistic word segmentation) และวิธีการตัดคำโดยอาศัยคุณลักษณะของคำ (feature-based word segmentation)

วิธีการตัดคำโดยอาศัยค่าความน่าจะเป็นจะเป็นการตัดคำโดยใช้แบบจำลองเอ็นแกรม (N-gram model) ในการหารูปแบบของการตัดคำและลำดับคำที่เป็นไปได้มากที่สุด โดยวิธีการนี้จะต้องมีการใช้คลังข้อมูลที่มีการตัดคำและกำกับหมวดคำที่เตรียมเอาไว้แล้ว ซึ่งวิธีการนี้ผลลัพธ์ที่ได้จะเป็นการเลือกรูปแบบการตัดคำที่มีความน่าจะเป็นมากที่สุด

วิธีการตัดคำโดยอาศัยคุณลักษณะของคำ จะเป็นการแก้ข้อผิดพลาดของการตัดคำโดยอาศัยค่าความน่าจะเป็นของการจำกัดหมวดคำที่จะเป็นแบบจำลอง

ในการตัดคำ ซึ่งวิธีการตัดคำโดยอาศัยคุณลักษณะของคำจะเป็นวิธีการแบบผสม (hybrid approach)

เทคนิคที่ช่วยในการตัดคำ

เทคนิคที่ใช้ในการตัดคำที่นิยมใช้กันทั่วไปคือ วิธีการเทียบคำที่ยาวที่สุด วิธีการเทียบคำที่สั้นที่สุด วิธีการตัดคำที่ใช้ความถี่ของคำ และวิธีการย้อนรอยกลับ

1) วิธีการเทียบคำที่ยาวที่สุด (longest word pattern matching)

วิธีนี้จะทำการตรวจสอบสายอักขระที่นำเข้ามาจากซ้ายไปขวา จากนั้นนำไปเปรียบเทียบกับคำที่มีอยู่ในพจนานุกรม หากตรวจสอบพบว่าพบพยางค์มากกว่า 1 พยางค์ในพจนานุกรม จะทำการเลือกพยางค์ที่ยาวที่สุดแล้วทำต่อไปเรื่อย ๆ จนจบสายอักขระ วิธีการนี้ให้ความถูกต้องหลังการตัดคำสูงกว่าวิธีการอื่นโดยเฉพาะเมื่อใช้ร่วมกับวิธีย้อนรอยกลับ

2) วิธีการเทียบคำที่สั้นที่สุด (shortest word pattern matching)

วิธีการนี้คล้ายกับวิธีการเทียบคำที่ยาวที่สุด เพียงแต่จะเลือกคำที่สั้นที่สุดที่พบก่อน แต่วิธีนี้พบว่าได้จำนวนคำมากที่สุดแต่ความถูกต้องของคำหลังทำการตัดค่าน้อยกว่าการใช้วิธีเทียบคำที่ยาวที่สุด วิธีนี้ใช้เวลาน้อยกว่าการเทียบคำยาวที่สุด แต่ความถูกต้องที่ได้จากการตัดคำแบบเทียบคำยาวที่สุดจะมากกว่า

3) วิธีการตัดคำที่ใช้ความถี่ของคำ (word usage frequency)

วิธีการนี้เป็นแนวทางหนึ่งในการแก้ปัญหาคำกำกวมของประโยคภาษาไทย โดยการวิเคราะห์ความถี่ของการใช้คำในชีวิตประจำวันโดยจัดเรียงคำในพจนานุกรมตามความถี่ที่พบ และใช้วิธีการตัดคำแบบเดียวกับข้อ 1) และ 2)

4) วิธีการย้อนรอยกลับ (back tracking)

เมื่อทำการเปรียบเทียบคำที่นำมาตัดคำกับคำที่มีอยู่ในพจนานุกรม อาจพบกรณีที่คำที่พบมีมากกว่า 1 คำแล้วทำการเลือกคำที่ยาวที่สุดทำให้สายอักขระที่ตามมาจากคำนั้นไม่สามารถตัดคำได้ เนื่องจากไม่พบตามพจนานุกรม กรณีนี้จะทำการย้อนไปอีกคำที่ไม่ถูกเลือกแล้วทำการตัดคำต่อไป

ในกระบวนการทำงานของพนักงานเจ้าหน้าที่ผู้ทำหน้าที่ตรวจสอบคำขอจดทะเบียนเครื่องหมายการค้าของกรมทรัพย์สินทางปัญญา กระทรวงพาณิชย์ การจดทะเบียนเครื่องหมายการค้าของกรมทรัพย์สินทางปัญญาในปัจจุบันมีโปรแกรมจดทะเบียนเครื่องหมายการค้าอัตโนมัติ e-Submission ช่วยทำงานโดยการตรวจสอบคำห้ามใช้เบื้องต้นแล้วก็ตาม แต่เนื่องจากโปรแกรมยังทำงานได้ไม่สมบูรณ์ ตรวจสอบได้เฉพาะคำที่เหมือนกับคำห้ามใช้ทุกประการ แต่ไม่สามารถตรวจสอบคำที่มีความใกล้เคียงหรือมีความคล้ายคลึงกับคำห้ามใช้ได้ ดังนั้นหากมีผู้ประกอบการที่

ตั้งใจใช้คำห้ามใช้เป็นชื่อการค้าหรือชื่อเครื่องสำอาง จึงหลีกเลี่ยงโดยใช้วิธีการปรับเปลี่ยนตัวอักษร การสะกดคำใหม่ หรือแทรกคำอื่นเข้าไป ทำให้ระบบไม่สามารถตรวจสอบคำห้ามใช้ดังกล่าวได้ ดังนั้น เพื่อป้องกันไม่ให้เกิดเหตุการณ์ดังกล่าว จึงต้องใช้วิธีการที่เรียกว่าการเปรียบเทียบสายอักขระแบบ ประมาณ (approximately string matching) เข้ามาใช้ในการช่วยตรวจสอบชื่อการค้าและชื่อ เครื่องสำอางว่ามีความคล้ายคลึงกับคำห้ามใช้หรือไม่

2.1.7 การเปรียบเทียบสายอักขระ

ชื่อการค้าและชื่อเครื่องสำอางมีลักษณะเป็นข้อความ ประกอบด้วยคำสำคัญ (keywords) ที่ใช้เป็นคำหลักในการตรวจสอบว่าชื่อการค้าและชื่อเครื่องสำอางมีถูกต้องและ สอดคล้องกับสารที่เป็นส่วนประกอบหรือไม่ จากนั้นจะนำคำสำคัญเหล่านั้นไปทำการเปรียบเทียบกับ คำสำคัญในฐานข้อมูล โดยจากการทบทวนวรรณกรรมมีวิธีการที่ใช้ในการเปรียบเทียบความเหมือน ระหว่างคำที่เรียกว่า การเปรียบเทียบสายอักขระ (string matching) โดยมีอัลกอริทึม (algorithm) ที่ใช้เปรียบเทียบประกอบด้วยการทำงานสองส่วนที่สำคัญ ส่วนแรกคือการเปรียบเทียบระหว่าง อักขระที่ต้องการค้นหากับสายอักขระย่อยในแฟ้มข้อมูล ส่วนที่สองจะทำการกำหนดตำแหน่งที่จะ เปรียบเทียบ [14]

การเปรียบเทียบจะทำการหาความคล้ายคลึงกันระหว่างคำ (texts similarity) ซึ่งคำมีความคล้ายคลึงกันได้ 2 รูปแบบ คือ คล้ายคลึงกันในรูปแบบคำศัพท์และคล้ายคลึงกันใน รูปแบบความหมาย สำหรับอัลกอริทึมการหาความคล้ายคลึงกันระหว่างคำที่เกี่ยวข้องในงานวิจัยฉบับ นี้ ได้แก่ string-based algorithms ใช้สำหรับวัดความคล้ายคลึงกันของคำศัพท์

String-based algorithms

string-based algorithms [15] จะเป็นการวัดลำดับของอักขระและองค์ประกอบ ของตัวอักษร ซึ่งจะใช้ string metric ในการเปรียบเทียบหรือการจับคู่สายอักขระ และได้ เป็นค่าประมาณ การวัดความคล้ายคลึงกันของคำศัพท์ประกอบด้วย 2 ลักษณะ ได้แก่

1) character-based similarity measure เป็นการวัดความคล้ายคลึงกันของ อักขระ ซึ่งขึ้นกับความยาวของตัวอักษรทั้งสองสาย

1.1.1 longest common substring (LCS) algorithm คืออัลกอริทึมที่ใช้ ค้นหาลำดับความคล้ายระหว่างสองสายอักขระ โดยพิจารณาจากความ ยาวของอักขระที่ต่อเนื่องซึ่งมีอยู่ในสายอักขระทั้งสองและมีการ ประยุกต์ใช้ในงานด้านชีวสารสนเทศศาสตร์

1.1.2 Levenshtein distance (LD) เป็นการคำนวณหาค่าระยะทางการ เปรียบเทียบสายอักขระซึ่งเรียกว่า ‘ระยะทางเลเวนชเตย์น’ ประกอบด้วย

การปฏิบัติการ (operations) 3 รูปแบบ คือ การเพิ่ม (insertion) การลบ (deletion) และการแทนที่ (substitution)

1.1.3 Jaro distance เป็นการวัดความต่างกันของสายอักขระจากจำนวนตัวอักษรร่วม (common character) และจำนวนการเพิ่ม การลบ หรือการย้ายตัวอักษร โดยตัวอักษรร่วมในที่นี้คือตัวอักษรที่เหมือนกันและอยู่ในตำแหน่งใกล้เคียงกัน [16]

1.1.4 Needleman-Wunsch algorithm คืออัลกอริทึมที่ใช้เปรียบเทียบจัดลำดับทางชีวภาพ เช่น โปรตีน นิวคลีโอไทป์ ซึ่งมีความเหมาะสมในการใช้เปรียบเทียบสายอักขระที่มีความยาวใกล้เคียงกันและใช้อย่างแพร่หลายในชีวสารสนเทศศาสตร์

1.1.5 Smith-Waterman คืออัลกอริทึมที่ใช้เปรียบเทียบสายอักขระและค้นหาลำดับ ใช้ประโยชน์สำหรับลำดับที่แตกต่างกันและมีความสงสัยว่าจะมีความคล้ายกัน

1.1.6 N-gram คืออัลกอริทึมการเปรียบเทียบสายอักขระโดยคำนวณความน่าจะเป็นที่จะเกิดรูปแบบอักขระจากจำนวนคำ N คำก่อนหน้า

2) **term-based similarity measures** เป็นการใช้อัลกอริทึมและค่าสัมประสิทธิ์ความเหมือนในรูปแบบต่างๆ ซึ่งมีการให้น้ำหนักความสำคัญของสายอักขระที่แตกต่างกัน

2.1.1 Manhattan distance เป็นการวัดระยะทางที่วัดตามระบบพิกัดฉาก XY ระหว่างข้อมูลสองข้อมูล หรือข้อมูลศูนย์กลางของพื้นที่

2.1.2 Cosine similarity ค่าความเหมือนโคไซน์เป็นการหาความเหมือนจากค่ามุมของข้อมูลบนพื้นที่เวกเตอร์ (vector space) มีค่าอยู่ระหว่าง 0-1

2.1.3 Dice's coefficient คือค่าสัมประสิทธิ์ที่ใช้หาความเหมือนของสายอักขระ คำนวณจากจำนวนสองเท่าของคำที่ใช้ร่วมกันต่อจำนวนคำในทั้งสองสายอักขระ

2.1.4 Euclidean distance เป็นมาตรวัดระยะพื้นฐานสำหรับวัดค่าความห่างระหว่างข้อมูล 2 ข้อมูล ซึ่งเป็นที่นิยมใช้ในงานประเภทต่าง ๆ เป็นอย่างมาก เพราะง่ายต่อความเข้าใจ และมีลักษณะเป็นการคำนวณที่ใช้ระบบพิกัดคาร์ทีเซียนที่มีที่มาจากทฤษฎีพีทาโกรัส ถ้าข้อมูล 2 ข้อมูลมีความคล้ายกันมาก ข้อมูลแต่ละข้อมูลจะอยู่ใกล้กันมาก จะทำให้ค่ายูคลิดีเนียนมีค่าน้อย เข้าใกล้ศูนย์

- 2.1.5 Jaccard similarity คือค่าความเหมือน Jaccard เป็นอัลกอริทึมที่ใช้หาความเหมือนของสายอักขระ คำนวณจากจำนวนคำที่ใช้ร่วมกันต่อจำนวนคำทั้งหมดในทั้งสองสายอักขระ
- 2.1.6 simple matching coefficient (SMC) คือค่าสัมประสิทธิ์การจับคู่ซึ่งเป็นวิธีการแบบเวกเตอร์ที่นับจำนวนสายอักขระที่คล้ายกันในทั้งสองเวกเตอร์
- 2.1.7 overlap coefficient คือค่าสัมประสิทธิ์การทับซ้อนมีค่าใกล้เคียงกับค่าสัมประสิทธิ์ของ Dice แต่พิจารณาสองสายอักขระที่มีการจับคู่แบบเต็มรูปแบบ

2.1.7.1 ระยะทางเลเวนชเตย์น (Levenshtein distance; LD หรือ edit distance)

ระยะทางเลเวนชเตย์นถูกพัฒนาโดย Vladimir Levenshtein ในปี 1996 เป็นขั้นตอนวิธีการวัดหาค่าความต่างกันของสายอักขระสองชุด ระหว่างชุดแรกที่เป็นต้นแบบและชุดที่สองที่เป็นชุดเปรียบเทียบ โดยค่าความต่างกันจะวัดจากจำนวนของการที่จะต้องทำการตัดออก การแทรก และการแทนที่อักขระในชุดที่นำมาเปรียบเทียบจนกระทั่งมีลักษณะเหมือนชุดอักขระที่เป็นต้นแบบทุกประการ [17] และ [18]

2.1.7.1.1 การดำเนินการในการหาระยะทางเลเวนชเตย์น

1) การแทรก (insertions) เป็นการนำเอาอักขระตัวใด ๆ มา เพื่อให้ชุดอักขระชุดนั้นเหมือนกับอีกชุดอักขระหนึ่งในภายหลัง เช่น run \rightarrow ruin จะแทรกตัว i ให้กับ run เพื่อให้ run กลายเป็น ruin เป็นต้น

2) การตัดออก (deletions) เป็นการตัดอักขระออกครั้งละ 1 ตัว จากชุดอักขระตัวหนึ่ง เพื่อให้ชุดอักขระชุดนั้นเหมือนกับอีกชุดอักขระหนึ่งในภายหลัง เช่น dog \rightarrow do จะตัดตัว g ออก เพื่อให้ dog กลายเป็น do เป็นต้น

3) การแทนที่ (substitutions) เป็นการนำอักขระของชุดอักขระหนึ่งไปแทนอักขระของอีกชุดอักขระหนึ่ง เพื่อให้ชุดอักขระชุดนั้นเหมือนกับอีกชุดอักขระหนึ่งในภายหลัง เช่น cat \rightarrow rat จะแทนที่ r ด้วย c เพื่อให้ cat กลายเป็น rat หรืออาจมองในทางกลับกันก็ได้ เป็นต้น

2.1.7.1.2 ตัวอย่างการหาระยะทางเลเวนชเตย์น

อักขระชุดที่ 1: LIGHTENING

อักขระชุดที่ 2: WHITENING

		W	H	I	T	E	N	I	N	G
	0	1	2	3	4	5	6	7	8	9
L	1									
I	2									
G	3									
H	4									
T	5									
E	6									
N	7									
I	8									
N	9									
G	10									

รูปที่ 1 ตัวอย่างการหาระยะทางเลขคณิต ขั้นตอนที่ 1: iteration (i) = 0

		W	H	I	T	E	N	I	N	G
	0	1	2	3	4	5	6	7	8	9
L	1	1	2	3	4	5	6	7	8	9
I	2	2								
G	3	3								
H	4	4								
T	5	5								
E	6	6								
N	7	7								
I	8	8								
N	9	9								
G	10	10								

รูปที่ 2 ตัวอย่างการหาระยะทางเลขคณิต ขั้นตอนที่ 2: iteration (i) = 1

		W	H	I	T	E	N	I	N	G
	0	1	2	3	4	5	6	7	8	9
L	1	1	2	3	4	5	6	7	8	9
I	2	2	2	2	3	4	5	6	7	8
G	3	3	3							
H	4	4	3							
T	5	5	4							
E	6	6	5							
N	7	7	6							
I	8	8	7							
N	9	9	8							
G	10	10	9							

รูปที่ 3 ตัวอย่างการหาระยะทางเลเวนชเตย์น ขั้นตอนที่ 3: iteration (i) = 2

		W	H	I	T	E	N	I	N	G
	0	1	2	3	4	5	6	7	8	9
L	1	1	2	3	4	5	6	7	8	9
I	2	2	2	2	3	4	5	6	7	8
G	3	3	3	3	3	4	5	6	7	7
H	4	4	3	4	4	4	5	6	7	8
T	5	5	4	4	4	5	5	6	7	8
E	6	6	5	5	5	4	5	6	7	8
N	7	7	6	6	6	5	4	5	6	7
I	8	8	7	6	7	6	5	4	5	6
N	9	9	8	7	7	7	6	5	4	5
G	10	10	9	8	8	8	7	6	5	4

รูปที่ 4 ตัวอย่างการหาระยะทางเลเวนชเตย์น ขั้นตอนสุดท้าย: iteration (i) = n, j=m
ค่าความต่างกันของสายอักขระสองชุดที่น้อยที่สุด (minimum edit distance) เท่ากับ 4

2.1.7.1.3 รหัสจำลองการคำนวณหาระยะทางเลเวนชเต๋น

```

function LevenshteinDistance(char s[1..m], char t[1..n]):
    // for all i and j, d[i,j] will hold the Levenshtein distance between
    // the first i characters of s and the first j characters of t
    declare int d[0..m, 0..n

    set each element in d to zero
    // source prefixes can be transformed into empty string by
    // dropping all characters
    for i from 1 to m:
        d[i, 0] := i
    // target prefixes can be reached from empty source prefix
    // by inserting every character
    for j from 1 to n:
        d[0, j] := j
    for j from 1 to n:
        for i from 1 to m:
            if s[i] = t[j]:
                substitutionCost := 0
            else:
                substitutionCost := 1
            d[i, j] := minimum(d[i-1, j] + 1,           // การลบ (deletions)
                               d[i, j-1] + 1,         // การแทรก (insertions)
                               d[i-1, j-1] + substitutionCost) // การแทนที่ (substitutions)

    return d[m, n]
  
```

2.1.7.1.4 การนำระยะทางเลเวนชเตย์นไปประยุกต์ใช้

- ใช้ในการตรวจคำสะกด
- พิสูจน์ประโยคข้อความ
- ตรวจสอบการคัดลอกข้อความ
- วิเคราะห์รหัสพันธุกรรม
- การใช้เพื่อแก้คำผิดสำหรับการค้นหาของเครื่องมือค้นหา
- ใช้พิสูจน์เนื้อหาที่เลียนแบบมาและการลอกเลียนแบบ [18]

2.1.7.2 Damerau–Levenshtein distance

Damerau–Levenshtein distance ตั้งตามชื่อผู้คิดค้น Frederick J. Damerau และ Vladimir I. Levenshtein ระยะทางระหว่างสองสายอักขระ ซึ่งสามารถหาได้จากจำนวนการกระทำที่น้อยที่สุดในการแปลงสายอักขระหนึ่งมาเป็นอีกสายอักขระหนึ่ง เป็นการแก้ไขระยะทางเลเวนชเตย์นให้มีการกระทำสลับอักขระเพิ่มขึ้น [19]

2.1.7.2.1 การดำเนินการในการหา Damerau–Levenshtein distance

- 1) การแทรก (insertions)
- 2) การตัดออก (deletions)
- 3) การแทนที่ (substitutions)
- 4) การสลับอักขระ (transpositions) อาจกำหนดให้สามารถสลับได้เฉพาะอักขระที่อยู่ติดกันหรือไม่จำเป็นต้องอยู่ติดกันก็ได้ ขึ้นอยู่กับการใช้งาน

2.1.7.2.2 ตัวอย่างการดำเนินการของ Damerau–Levenshtein distance

- การแทรก: rain > train
- การตัดออก: party > part
- การแทนที่: river > liver
- การสลับอักขระ: house > houes

2.1.7.2.3 รหัสจำลองการคำนวณหา Damerau–Levenshtein distance

```

algorithm OSA-distance is
  input: strings a[1..length(a)], b[1..length(b)]
  output: distance, integer
  let d[0..length(a), 0..length(b)] be a 2-d array of integers, dimensions length(a)+1,
length(b)+1
  // note that d is zero-indexed, while a and b are one-indexed.
  for i := 0 to length(a) inclusive do
    d[i, 0] := i
  for j := 0 to length(b) inclusive do
    d[0, j] := j
  for i := 1 to length(a) inclusive do
    for j := 1 to length(b) inclusive do
      if a[i] = b[j] then
        cost := 0
      else
        cost := 1
      d[i, j] := minimum(d[i-1, j] + 1, // การลบ (deletions)
                        d[i, j-1] + 1, // การแทรก (insertions)
                        d[i-1, j-1] + cost) // การแทนที่ (substitutions)

```

ส่วนที่เพิ่มการสลับอักขระ (transpositions) ของ Damerau–Levenshtein distance

```

if i > 1 and j > 1 and a[i] = b[j-1] and a[i-1] = b[j] then
  d[i, j] := minimum(d[i, j],
                    d[i-2, j-2] + 1)
return d[length(a), length(b)]

```

2.1.7.2.4 การนำ Damerau–Levenshtein distance ไปประยุกต์ใช้

- เพื่อแก้ไขการสะกดคำให้มีความถูกต้อง
- ใช้ในชีววิทยา ในการหาความแตกต่างระหว่างสายอักขระดีเอ็นเอ (DNA)

2.1.7.3 Normalised Levenshtein distance for Origin-Destination matrices (NLOD)

เป็นการปรับเปลี่ยนระยะทางเลเวนชเตย์นเพื่อใช้ในการหาจำนวนโครงสร้างเปรียบเทียบโครงสร้างของเมทริกซ์ origin-destination (OD) มีการนำวิธีการดังกล่าวไปประยุกต์ใช้งานแอปพลิเคชันกรณีศึกษาเกี่ยวกับเมทริกซ์ OD ที่ใช้บลูทอร์ที่เมืองบรีสเบน ในประเทศออสเตรเลีย [18]

2.2 งานวิจัยที่เกี่ยวข้อง

ปัจจุบันมีนักวิจัยหลายกลุ่มใช้ประโยชน์จากการเปรียบเทียบสายอักขระเพื่อแก้ไขปัญหาในงานต่าง ๆ อย่างแพร่หลาย

Yulianto Muhamad M., Arifudin Riza และ Alamsyah Alamsyah [20] ทำการศึกษาการประยุกต์ใช้อัลกอริทึมของเลเวนชเตย์นในการเติมข้อความอัตโนมัติและการตรวจสอบตัวสะกดเพื่อแก้ไขข้อความที่ผิดพลาดในการค้นหาข้อมูลในท้องสมุดพบว่า มีความถูกต้องของอัลกอริทึมร้อยละ 86 อ้างอิงจากแหล่งข้อมูล 1,055 เคส และกลุ่มเป้าหมาย 100 เคส

Haldar Rishin และ Mukhopadhyay Debajyoti [21] ศึกษาการนำอัลกอริทึมของเลเวนชเตย์นมาประยุกต์ใช้ร่วมกับการแบ่งกลุ่มอักขระตามความคล้ายคลึงกันเพื่อนำไปใช้ในการค้นหาคำในดิกชันนารี พบว่าทำให้การค้นหาคำในดิกชันนารีมีประสิทธิภาพมากยิ่งขึ้น และในการศึกษาของ Nagowah Leckraj และคณะ [22] ยังพบว่าอัลกอริทึมของเลเวนชเตย์นช่วยแก้ไขข้อผิดพลาดทางไวยากรณ์แบบอัตโนมัติในไฟล์ต้นฉบับที่ผิดพลาดได้

Christanti Viny M., Rudy และ Naga Dali S. [19] ศึกษาการประยุกต์ใช้อัลกอริทึมเอ็น-แกรม (N-gram) และ Damerau-Levenshtein distance ในการแก้ไขการสะกดคำให้มีความถูกต้องและรวดเร็ว จากผลการศึกษาพบคำที่มีความถูกต้องร้อยละ 84.62 และมีความเร็วในการประมวลผลเฉลี่ย 18.89 มิลลิวินาทีต่อหนึ่งข้อความ

สำหรับการศึกษาในประเทศไทยที่ได้ศึกษาการเปรียบเทียบสายอักขระโดยนำอัลกอริทึมของเลเวนชเตย์นมาใช้ในงานต่าง ๆ ยกตัวอย่างเช่น ในการออกแบบและพัฒนาระบบการจำแนกข้อร้องเรียนรถโดยสารสาธารณะ เพื่อติดแท็กจำแนกปัญหาในการให้บริการของรถโดยสารสาธารณะโดยจักรินทร์ สันติรัตนภักดี และศุภกฤษฎี นิวัฒนากุล [23] ใช้วิธีการจับคู่กับคำศัพท์ในคลังคำศัพท์กับข้อร้องเรียนจากผู้ใช้ และเพิ่มความถูกต้องของผลลัพธ์ด้วยการเปรียบเทียบความคล้ายคลึงของคำด้วยเทคนิคการวัดระยะทางเลเวนชเตย์น จากการศึกษาพบว่ากระบวนการเปรียบเทียบความคล้ายคลึงกันของอักขระสองชุดด้วยเทคนิคการวัดระยะทางเลเวนชเตย์นสามารถให้ผลลัพธ์เปรียบเทียบกับคำศัพท์ที่ถูกต้องได้อย่างมีประสิทธิภาพ

นอกจากนี้การศึกษาของฐิติมา พลิกเพชร และวีรยุทธ์ เลิศนที ได้นำอัลกอริทึมเลเวนชเตย์
นมาประยุกต์ใช้เพื่อหาชื่อพ้องรูปและชื่อคล้ายของพืชสมุนไพร โดยผลจากการใช้อัลกอริทึมดังกล่าว
สามารถจำแนกชื่อพืชสมุนไพรได้ 5 กลุ่ม ตามค่าร้อยละความเหมือนระหว่างสายอักขระของชื่อพืช
สมุนไพร [24]



บทที่ 3

วิธีการดำเนินการวิจัย

การวิจัยนี้ผู้วิจัยมุ่งที่จะศึกษาการตรวจสอบซื้อการค่า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ ซึ่งในบทนี้จะได้กล่าวถึงสาระสำคัญเกี่ยวกับ วิธีการดำเนินการศึกษาค้นคว้าคือ ประชากร กลุ่มตัวอย่าง วิธีสุ่มตัวอย่าง แบบแผนการวิจัย เครื่องมือที่ใช้ในการวิจัย การเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูลตามลำดับ ดังนี้

3.1 ประชากร กลุ่มตัวอย่าง วิธีสุ่มตัวอย่าง

3.1.1 ประชากรและกลุ่มตัวอย่างสำหรับการศึกษาประสิทธิภาพของโปรแกรม

ประชากรที่ใช้ในการศึกษา ได้แก่ แบบคำขอจดแจ้งเครื่องสำอาง (แบบ จ.ค.1) ที่ถูกส่งเข้ามาในระบบจดแจ้งเครื่องสำอางอัตโนมัติ e-Submission ของสำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา ตั้งแต่วันที่ 1 ตุลาคม พ.ศ. 2559 จนถึงวันที่ 30 พฤศจิกายน พ.ศ. 2564 จำนวน 3,100 รายการคำขอ แบ่งประเภทแบบคำขอจดแจ้งเครื่องสำอางตามผลการตรวจสอบได้เป็น 2 ประเภท คือ แบบคำขอจดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าผ่าน จำนวน 1,590 รายการคำขอ และแบบคำขอจดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าไม่ผ่าน 1,510 รายการคำขอ

การศึกษานี้เป็นการวิจัยเชิงทดลอง (experimental research) ศึกษาเปรียบเทียบผลการตรวจสอบซื้อการค่า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับ ในแบบคำขอจดแจ้งเครื่องสำอาง ระหว่างการใช้โปรแกรมกับการตรวจสอบโดยเจ้าหน้าที่ผู้เชี่ยวชาญ ซึ่งจะทำการศึกษาในกลุ่มตัวอย่าง 2 กลุ่มที่สัมพันธ์กันหรือไม่เป็นอิสระต่อกัน (dependent samples) คำนวณหาขนาดกลุ่มตัวอย่าง โดยใช้สูตรการคำนวณที่มีแนวคิดพื้นฐานจากการทดสอบของแมคเนียร์ (McNemar's Test) [25] ดังนี้

$$n = \frac{(Z_{\alpha}\sqrt{p_{disc}} + Z_{\beta}\sqrt{p_{disc} - p_{diff}^2})^2}{(p_{diff})^2} \quad (1)$$

เมื่อ n แทน ขนาดกลุ่มตัวอย่าง

Z_{α} แทน ระดับความเชื่อมั่นที่กำหนด

Z_{β} แทน อำนาจการทดสอบ

กำหนดให้

$$p_{\text{disc}} = p_{12} + p_{21} = \frac{b}{n} + \frac{c}{n}$$

$$p_{\text{diff}} = p_{12} - p_{21} = \frac{b}{n} - \frac{c}{n}$$

Group A	Group B		Row total
	Outcome 1	Outcome 2	
Outcome 1	a	b	a+b
Outcome 2	c	d	c+d
Column total	a+c	b+d	n = a+b+c+d

รูปที่ 5 การจัดเรียงข้อมูลในตาราง two-dimensional contingency สำหรับ McNemar's Test

ที่มา: www.e-ijd.org/viewimage.asp?img=IndianJ Dermatol_2016_61_4_385_185700_t5.jpg

จากการวิจัยเบื้องต้น (preliminary research) ในการตรวจสอบ ชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ ผู้วิจัยศึกษาผลการตรวจสอบ ชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอาง ระหว่างการใช้โปรแกรมกับการตรวจสอบโดยเจ้าหน้าที่ผู้เชี่ยวชาญ

สำหรับการตรวจสอบโดยใช้โปรแกรม ผู้วิจัยใช้ trigram ในการกำหนดคุณลักษณะของคำ และกำหนดค่า threshold 80 ในการเปรียบเทียบความคล้ายคลึงกันของข้อความ ในขั้นตอนการตัดคำภาษาไทย ด้วย Library PyThaiNLP ที่สามารถปรับเปลี่ยนเครื่องมือตัดคำ (engine) ได้ ผู้วิจัยได้ศึกษาผลการตรวจสอบ ชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอาง โดยใช้เครื่องมือตัดคำด้วยวิธีต่าง ๆ 8 วิธี ได้แก่ newmm, longest, multi_cut, deepcut, attacut, newmm-safe, tltk และ sefr_cut ทำการศึกษาโดยใช้กลุ่มตัวอย่าง คือ แบบคำขอจดแจ้งเครื่องสำอาง ที่ถูกส่งเข้ามาในระบบจดแจ้งเครื่องสำอางอัตโนมัติ e-Submission ของสำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา จำนวน 200 รายการคำขอ ผลการศึกษาพบว่า การใช้เครื่องมือ attacut ในการตัดคำภาษาไทย ช่วยให้การตรวจสอบ ชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางมีประสิทธิภาพสูงสุด ผู้วิจัยจึงใช้ข้อมูลผลการตรวจสอบแบบคำขอจดแจ้งเครื่องสำอางโดยใช้เครื่องมือ attacut ในการตัดคำภาษาไทย มาใช้ในการคำนวณหาขนาดตัวอย่างที่เหมาะสมเพื่อทำการวิจัยต่อไป รายละเอียดข้อมูลดังแสดงในตารางที่ 1

ตารางที่ 1 ข้อมูลผลการตรวจสอบแบบคำขอจัดแจ้งเครื่องสำอางโดยใช้เครื่องมือ attacut ในการตัดคำภาษาไทย เมื่อกำหนดคุณลักษณะของคำโดยใช้ trigram และกำหนดค่า threshold 80

ผลการตรวจสอบ โดยเจ้าหน้าที่ผู้เชี่ยวชาญ	ผลการตรวจสอบ โดยใช้โปรแกรม		row total
	ผ่าน	ไม่ผ่าน	
ผ่าน	88	12	100
ไม่ผ่าน	4	96	100
column total	92	108	n = 200

การทดสอบใช้สมมติฐานสองทาง (two-sided test) โดยกำหนด type I error หรือระดับนัยสำคัญ (α) มีค่า .05 และ Power of test หรืออำนาจการทดสอบ ($1-\beta$) มีค่า .80

$$\text{กำหนด } Z_\alpha = Z_{.05/2} = 1.96$$

$$Z_\beta = .84$$

เมื่อแทนค่า จะได้ขนาดของกลุ่มตัวอย่าง ดังนี้

$$n = \frac{(Z_\alpha \sqrt{p_{disc}} + Z_\beta \sqrt{p_{disc} - p_{diff}^2})^2}{(p_{diff})^2}$$

$$n = \frac{(1.96 \sqrt{\frac{12}{200} + \frac{4}{200}} + .84 \sqrt{(\frac{12}{200} + \frac{4}{200}) - (\frac{12}{200} - \frac{4}{200})^2})^2}{(\frac{12}{200} - \frac{4}{200})^2}$$

$$n = 389.6397$$

ขนาดของกลุ่มตัวอย่างที่ได้จากการคำนวณ เท่ากับ 389.6397 เพื่อลดความคลาดเคลื่อน และเพิ่มความเชื่อมั่นในการเก็บข้อมูล ดังนั้นในงานวิจัยครั้งนี้จะใช้กลุ่มตัวอย่าง คือ แบบคำขอจัดแจ้งเครื่องสำอาง (แบบ จ.ค.1) ที่ถูกส่งเข้ามาในระบบจัดแจ้งเครื่องสำอางอัตโนมัติ e-Submission ของสำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา จำนวน 400 รายการคำขอ โดยแบ่งเป็นแบบคำขอจัดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าผ่าน และแบบคำขอจัดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าไม่ผ่าน จำนวนกลุ่มละ 200 รายการคำขอ ในการเลือกกลุ่มตัวอย่างผู้วิจัยใช้วิธีการสุ่มตัวอย่างแบบง่าย (simple random sampling; SRS) ด้วยการใช้โปรแกรมคอมพิวเตอร์

3.1.2 ประชากรและกลุ่มตัวอย่างสำหรับการศึกษาความพึงพอใจของผู้ใช้งานที่มีต่อโปรแกรม

ประชากร ได้แก่ บุคลากรตำแหน่งเภสัชกรในกลุ่มงานคุ้มครองผู้บริโภคและเภสัชสาธารณสุข สำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา จำนวน 16 ท่าน

กลุ่มตัวอย่าง ได้แก่ บุคลากรตำแหน่งเภสัชกรในกลุ่มงานคุ้มครองผู้บริโภคและเภสัชสาธารณสุข สำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา จำนวน 10 ท่าน ที่ปฏิบัติงานหรือเคยปฏิบัติงานเกี่ยวข้องกับการตรวจสอบผลิตภัณฑ์เครื่องสำอางก่อนออกสู่ท้องตลาด (pre-marketing control) หรือการตรวจสอบผลิตภัณฑ์เครื่องสำอางหลังออกสู่ท้องตลาด (post-marketing control) โดยกำหนดคุณสมบัติ ดังนี้

1. คุณสมบัติของกลุ่มตัวอย่างที่เลือกเข้าศึกษา (inclusion criteria)

1.1 เป็นบุคลากรตำแหน่งเภสัชกร ปฏิบัติงานในกลุ่มงานคุ้มครองผู้บริโภคและเภสัชสาธารณสุข สำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา

1.2 ปฏิบัติงานหรือเคยปฏิบัติงานเกี่ยวข้องกับการตรวจสอบผลิตภัณฑ์เครื่องสำอางก่อนออกสู่ท้องตลาด หรือการตรวจสอบผลิตภัณฑ์เครื่องสำอางหลังออกสู่ท้องตลาด

1.3 สามารถอ่าน ฟัง และเขียนหนังสือได้

1.4 ไม่มีการเจ็บป่วยรุนแรงที่เป็นอุปสรรคต่อการเข้าร่วมการวิจัย

1.5 ยินยอมเข้าร่วมการวิจัย

2. เกณฑ์การคัดออกของกลุ่มตัวอย่าง (exclusion criteria)

เกณฑ์การคัดออกของกลุ่มตัวอย่างสำหรับการวิจัยครั้งนี้ คือ เภสัชกรในกลุ่มงานคุ้มครอง ผู้บริโภคและเภสัชสาธารณสุข สำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยาที่ไม่สะดวกในการเข้าใช้งาน โปรแกรม และ/หรือตอบแบบสอบถามประเมินความพึงพอใจไม่ครบตามที่กำหนดไว้

ผู้วิจัยทำการเลือกกลุ่มตัวอย่างใช้วิธีการเลือกตัวอย่างแบบเจาะจง (purposive sampling)

3.2 แบบแผนการวิจัย

การวิจัยนี้เป็นการศึกษาวิจัยเชิงทดลอง โดยศึกษาการตรวจสอบซื้อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดทะเบียนเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระเมื่อเทียบกับการพิจารณาโดยเจ้าหน้าที่ผู้เชี่ยวชาญ ทำการวัดผลโดยวัดประสิทธิภาพของโปรแกรม ได้แก่ ค่าความถูกต้อง ค่าความระลึก ค่าความเที่ยง และค่า F1 และทำการประเมินความพึงพอใจต่อโปรแกรมโดยผู้ปฏิบัติงาน โดยบุคลากรตำแหน่งเภสัชกรในกลุ่มงานคุ้มครองผู้บริโภคและเภสัชสาธารณสุข สำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา จำนวน 16 ท่าน ที่ปฏิบัติงานหรือเคย

ปฏิบัติงานเกี่ยวข้องกับการตรวจสอบผลิตภัณฑ์เครื่องสำอางก่อนออกสู่ท้องตลาด หรือการตรวจสอบผลิตภัณฑ์เครื่องสำอางหลังออกสู่ท้องตลาด

3.3 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ ประกอบด้วย โปรแกรมสำหรับตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ และแบบประเมินความพึงพอใจของผู้ใช้งานโปรแกรม

3.3.1 ภาษาและเครื่องมือที่ใช้ในการพัฒนาโปรแกรม

3.3.1.1 ภาษาคอมพิวเตอร์

- 1) ภาษาพีเอชพี (PHP)
- 2) ภาษาไพธอน (Python)
- 3) ภาษาเอชทีเอ็มแอล (HyperText Markup Language: HTML)
- 4) ภาษาเอสคิวแอล (Structured Query Language: SQL)

3.3.1.2 โปรแกรมคอมพิวเตอร์

- 1) โปรแกรม text-editor: Visual Studio code
- 2) โปรแกรมจัดการฐานข้อมูล MySQL: phpMyAdmin

3.3.1.3 เครื่องมืออื่น ๆ

- 1) เครื่องมือที่ใช้ในการตัดคำภาษาไทย PyThaiNLP v3.0.8
- 2) หลักเกณฑ์การพิจารณาการจดแจ้งเครื่องสำอาง (ฉบับปรับปรุง 2564)
- 3) ระบบจดแจ้งเครื่องสำอางอัตโนมัติ e-Submission

3.3.2 แบบประเมินความพึงพอใจของผู้ใช้งานโปรแกรม

3.4 ขั้นตอนในการดำเนินการวิจัย

3.4.1 ศึกษาและวิเคราะห์กระบวนการตรวจสอบและพิจารณาคำขอจดแจ้งเครื่องสำอาง

3.4.2 ศึกษาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องกับการพัฒนาโปรแกรมสำหรับตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอาง

3.4.3 รวบรวมข้อมูลองค์ความรู้ที่เกี่ยวข้อง เพื่อนำมาประยุกต์ใช้ในการพัฒนาโปรแกรมสำหรับตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอาง

3.4.4 กำหนดประชากรและคำนวณหากกลุ่มตัวอย่างเพื่อใช้ในการศึกษาวิจัย

3.4.5 สร้างเครื่องมือที่ใช้ในการศึกษาวิจัย

1) โปรแกรมสำหรับตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับ ในแบบคำขอจดแจ้งเครื่องสำอาง ด้วยวิธีการเปรียบเทียบสายอักขระ

2) แบบประเมินความพึงพอใจของผู้ใช้งานโปรแกรม

3.4.6 ทดสอบโปรแกรมกับกลุ่มตัวอย่างตามวิธีการดำเนินการวิจัย และเก็บรวบรวมข้อมูลที่ได้จากการทดสอบ

3.4.7 ทำการประเมินความพึงพอใจของผู้ใช้งานที่มีต่อโปรแกรมสำหรับตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอาง ด้วยวิธีการเปรียบเทียบสายอักขระ และเก็บรวบรวมข้อมูลที่ได้จากการประเมิน

3.4.8 นำข้อมูลที่ได้ไปทำการวิเคราะห์ค่าทางสถิติ อภิปรายและสรุปผลการวิจัย

3.5 การสร้างเครื่องมือที่ใช้ในการศึกษาวิจัย

3.5.1 การออกแบบและพัฒนาโปรแกรมสำหรับใช้ในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ

การพัฒนาโปรแกรมสำหรับใช้ในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ ผู้วิจัยมุ่งเน้นที่กระบวนการตรวจสอบชื่อผลิตภัณฑ์เครื่องสำอาง และความสอดคล้องระหว่างชื่อกับสารประกอบในตำรับเครื่องสำอาง ซึ่งได้กำหนดกรอบแนวคิดในการพัฒนาโปรแกรมโดยใช้ทฤษฎีวงจรการพัฒนา ระบบ (system development life cycle; SDLC) ซึ่งประกอบด้วย 5 ขั้นตอน ได้แก่ การวิเคราะห์ (analysis), การออกแบบ (design), การพัฒนา (development), การทดสอบ (testing) และการบำรุงรักษา (maintenance) [26]

3.5.1.1 การวิเคราะห์ (analysis)

ผู้วิจัยศึกษาปัญหาและความต้องการของผู้ใช้งาน วิเคราะห์กระบวนการทำงานในด้านต่างๆที่เกี่ยวข้อง ซึ่งจากการศึกษาวิเคราะห์ระบบงานในปัจจุบัน พบปัญหาที่อาจทำให้การตรวจสอบชื่อผลิตภัณฑ์เครื่องสำอางและการพิจารณาค่าขอจดแจ้งเครื่องสำอางล่าช้าหรือเกิดความคลาดเคลื่อน โดยสามารถวิเคราะห์และจำแนกปัญหาได้ ดังนี้

1. ปัญหาความล่าช้าในการตรวจสอบและพิจารณาค่าขอ เนื่องจากข้อมูลคำห้ามใช้เป็นชื่อผลิตภัณฑ์เครื่องสำอางและคำที่สามารถใช้เป็นชื่อผลิตภัณฑ์เครื่องสำอางได้โดยต้องมีสารที่สอดคล้องในสูตรตำรับเดิมมีลักษณะเป็นคู่มือ ทำให้การค้นหาและตรวจสอบข้อมูลที่ต้องการทำได้ยาก และใช้เวลานาน

2. ขาดความคล่องตัวในการใช้และเรียกดูข้อมูล เนื่องจากข้อมูลมีลักษณะเป็นคู่มือ อีกทั้งยังมีการจัดเก็บที่กระจัดกระจายและมีแหล่งข้อมูลหลายแห่ง เช่น ในแอปพลิเคชันไลน์ (line), เว็บไซต์กลุ่มควบคุมเครื่องสำอางของสำนักงานคณะกรรมการอาหารและยา เป็นต้น ทำให้ไม่สามารถเรียกดูข้อมูลทั้งหมดได้ทันที

3. เนื่องจากข้อมูลเกี่ยวกับชื่อผลิตภัณฑ์เครื่องสำอางและสารที่เป็นส่วนประกอบในตำรับมีการปรับปรุงอยู่ตลอดเวลา ทำให้ข้อมูลที่มีอยู่ไม่เป็นปัจจุบัน ส่งผลให้การพิจารณาคำขอจดแจ้งเครื่องสำอางมีความคลาดเคลื่อนได้

4. อาจเกิดความคลาดเคลื่อนในการพิจารณาจากการตรวจสอบข้อมูลโดยพนักงานเจ้าหน้าที่ (human error) เนื่องจากกระบวนการทำงานเดิมที่อาศัยความรู้ ประสบการณ์ และความจำ ในการพิจารณาคำขอจดแจ้งเครื่องสำอาง

5. ถึงแม้ว่าจะมีระบบจดแจ้งเครื่องสำอางอัตโนมัติ e-Submission ช่วยตรวจสอบคำห้ามใช้ในชื่อผลิตภัณฑ์เครื่องสำอางได้ แต่การตรวจสอบดังกล่าวสามารถตรวจได้เฉพาะคำห้ามใช้ที่มีการสะกดเหมือนกับคำห้ามใช้ในฐานข้อมูลทุกประการ หากมีการสะกดคำผิด ทั้งที่ไม่ได้ตั้งใจ หรือเป็นความตั้งใจของผู้ประกอบการในการหลีกเลี่ยงคำห้ามใช้ หรือต้องการใช้ชื่อที่มีความเป็นเอกลักษณ์ ระบบจดแจ้งเครื่องสำอางอัตโนมัติ e-Submission อาจไม่สามารถตรวจสอบได้ ส่งผลให้การพิจารณาคำขอจดแจ้งเครื่องสำอางมีความคลาดเคลื่อนได้

3.5.1.2 การออกแบบ (design)

1. ผู้วิจัยทำการออกแบบฐานข้อมูลที่ใช้สำหรับเก็บค่าที่เกี่ยวข้อง ประกอบด้วยโครงสร้างตารางข้อมูล ดังต่อไปนี้

ตารางที่ 2 ตารางคำห้ามใช้ในชื่อผลิตภัณฑ์เครื่องสำอาง (restricted_name)

ลำดับ	ชื่อฟิลด์	คำอธิบาย	ชนิดข้อมูล	คีย์
1	restricted_name_id	รหัสคำห้ามใช้	INT(10)	PK
2	restricted_name_list	คำห้ามใช้	VARCHAR(100)	
3	lang	ภาษา	VARCHAR(20)	

ตารางที่ 3 ตารางชื่อกลุ่มสารที่สามารถใช้เป็นชื่อผลิตภัณฑ์เครื่องสำอางได้ (allowed_name)

ลำดับ	ชื่อฟิลด์	คำอธิบาย	ชนิดข้อมูล	คีย์
1	allowed_name_id	รหัสชื่อกลุ่มสาร	INT(10)	PK
2	allowed_name_list	ชื่อกลุ่มสาร	VARCHAR(100)	

ตารางที่ 4 ตารางคำที่สามารถใช้แทนที่ชื่อกลุ่มสารได้ (allowed_name_synonym)

ลำดับ	ชื่อฟิลด์	คำอธิบาย	ชนิดข้อมูล	คีย์
1	allowed_name_id	รหัสชื่อกลุ่มสาร	INT(10)	FK
2	allowed_name_syn_id	รหัสคำที่สามารถใช้แทนที่ชื่อกลุ่มสารได้	INT(10)	PK
3	allowed_name_syn	คำที่สามารถใช้แทนที่ชื่อกลุ่มสารได้	VARCHAR(100)	

ตารางที่ 5 ตารางสารที่ใช้เป็นส่วนประกอบในตำรับเครื่องสำอาง (ingredient)

ลำดับ	ชื่อฟิลด์	คำอธิบาย	ชนิดข้อมูล	คีย์
1	ingredient_id	รหัสสาร	INT(10)	PK
2	ingredient_list	รายชื่อสาร	VARCHAR(100)	

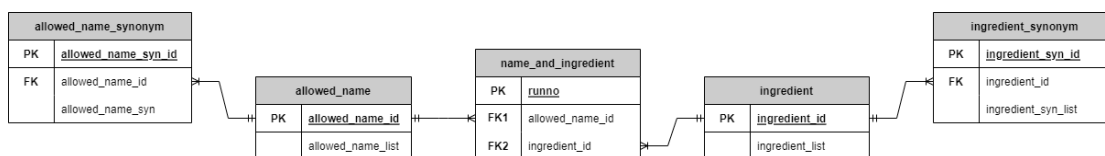
ตารางที่ 6 ตารางชื่อสารที่สามารถใช้แทนที่สารที่เป็นส่วนประกอบในตำรับเครื่องสำอางได้ (ingredient_synonym)

ลำดับ	ชื่อฟิลด์	คำอธิบาย	ชนิดข้อมูล	คีย์
1	ingredient_id	รหัสสาร	INT(10)	FK
2	ingredient_syn_id	รหัสสารที่สามารถใช้แทนที่สารที่เป็นส่วนประกอบในตำรับเครื่องสำอาง	INT(10)	PK
3	ingredient_syn_list	รายชื่อสารที่สามารถใช้แทนที่สารที่เป็นส่วนประกอบในตำรับเครื่องสำอาง	VARCHAR(200)	

ตารางที่ 7 ตารางความสัมพันธ์ระหว่างชื่อผลิตภัณฑ์เครื่องสำอางกับสารที่เป็นส่วนประกอบ (name_and_ingredient)

ลำดับ	ชื่อฟิลด์	คำอธิบาย	ชนิดข้อมูล	คีย์
1	runno	ลำดับข้อมูล	INT(10)	PK
2	allowed_name_id	รหัสชื่อกลุ่มสาร	INT(10)	FK1
3	ingredient_id	รหัสสาร	INT(10)	FK2

จากตารางข้างต้นผู้วิจัยเขียนโครงสร้างฐานข้อมูลที่เกี่ยวข้องด้วยภาษา SQL ใช้ระบบจัดการฐานข้อมูล MySQL และใช้โปรแกรม phpMyAdmin เป็นเครื่องมือช่วยในการจัดการฐานข้อมูล ซึ่งสามารถแสดงความสัมพันธ์ของข้อมูลดังรูปที่ 5



รูปที่ 6 แผนภาพแสดงความสัมพันธ์ระหว่างข้อมูล (entity relationship diagram; ER diagram)

2. ออกแบบหน้าจอร์ับข้อมูลและแสดงผลข้อมูล

3.5.1.3 การพัฒนา (development)

1) จัดทำฐานข้อมูล

จัดทำฐานข้อมูลคำที่ห้ามนำมาใช้เป็นชื่อหรือส่วนหนึ่งของชื่อการค้าหรือชื่อเครื่องสำอาง และคำที่สามารถนำมาใช้เป็นส่วนหนึ่งของชื่อผลิตภัณฑ์เครื่องสำอางได้ โดยคำนั้นจะต้องมีความสอดคล้องกับส่วนประกอบในตำรับของเครื่องสำอาง อ้างอิงข้อมูลตามหลักเกณฑ์การพิจารณาการจดแจ้งเครื่องสำอาง สำนักควบคุมเครื่องสำอางและวัตถุอันตราย สำนักงานคณะกรรมการอาหารและยา กระทรวงสาธารณสุข และรวบรวมข้อมูลเพิ่มเติมจากการสอบถามพนักงานเจ้าหน้าที่กลุ่มกำกับดูแลเครื่องสำอางก่อนออกสู่ตลาดในแอปพลิเคชันไลน์ รายละเอียดข้อมูลดังแสดงในภาคผนวก ก

สำหรับคำที่สามารถนำมาใช้เป็นส่วนหนึ่งของชื่อผลิตภัณฑ์เครื่องสำอาง และสารที่เป็นส่วนประกอบในตำรับ บางส่วนมีลักษณะเป็นคำพ้องความหมาย (synonym) ซึ่งในงานวิจัยนี้หมายความว่ารวมถึงคำแปลจากภาษาอังกฤษเป็นภาษาไทย และคำทับศัพท์ภาษาอังกฤษ ผู้วิจัยจึงได้จัดทำฐานข้อมูลสำหรับเก็บค่าดังกล่าว ข้อมูลแสดงในภาคผนวก ข

2) พัฒนาโปรแกรมสำหรับใช้ในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอาง ด้วยวิธีการเปรียบเทียบสายอักขระ

ผู้วิจัยพัฒนาโปรแกรมโดยอาศัยความรู้ด้านการออกแบบและพัฒนาเว็บไซต์ด้วยภาษา PHP ซึ่งเป็นภาษาสคริปต์ (scripting language) ที่ทำงานแบบ server-side script ภาษา PHP มีลักษณะเป็น open source ที่มีความสามารถเป็นแบบไดนามิก (dynamic language) สามารถใช้งานได้ทั้งบนระบบปฏิบัติการยูนิกซ์ (Unix) และวินโดวส์ (Windows) และยังสามารถติดต่อกับโปรแกรมฐานข้อมูล เช่น MySQL ได้ [27] โดยพัฒนาร่วมกับภาษา HTML ซึ่งเป็นภาษาที่ใช้ในการพัฒนาเว็บเพจ ช่วยให้โปรแกรมมีการตอบสนองกับผู้ใช้งานได้ดียิ่งขึ้น และสามารถเรียกใช้งานได้จากเว็บเบราว์เซอร์ [28]

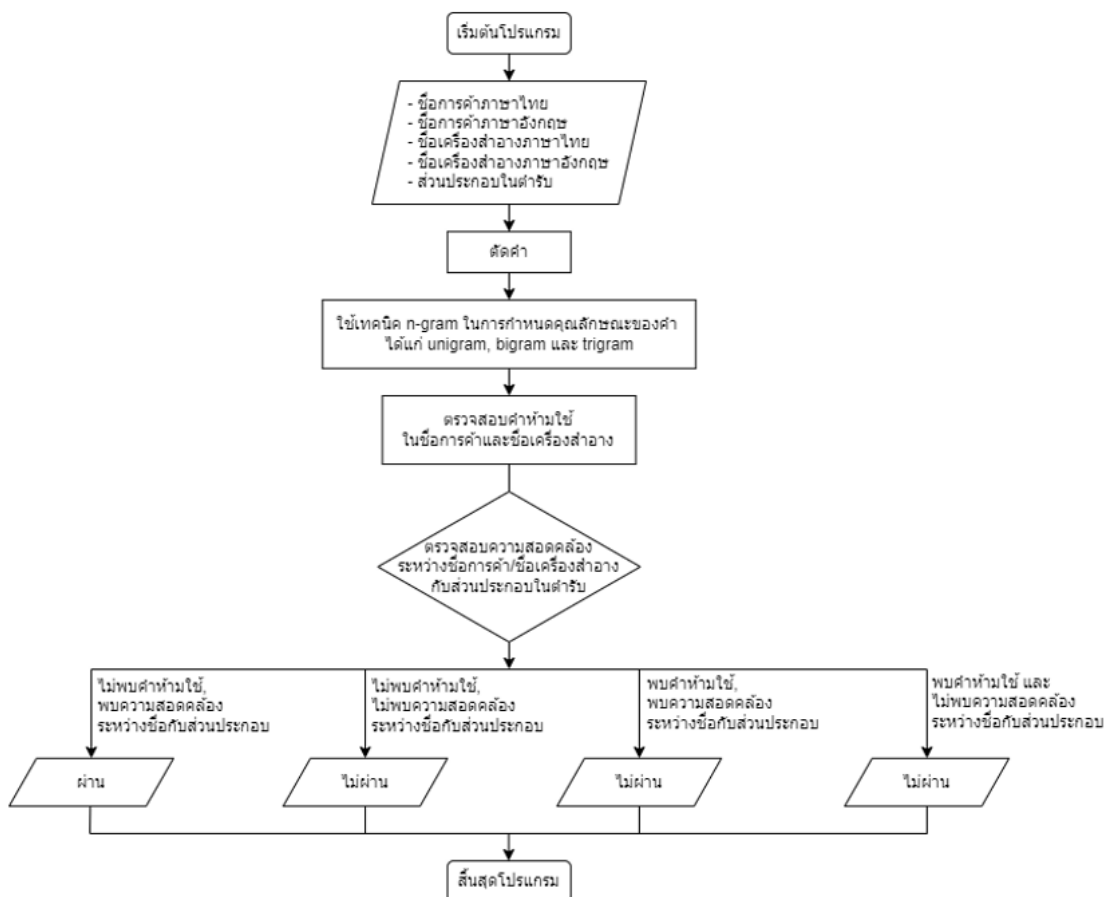
เนื่องจากข้อมูลชื่อผลิตภัณฑ์เครื่องสำอาง ประกอบด้วยชื่อการค้าและชื่อเครื่องสำอาง ซึ่งมีทั้งภาษาอังกฤษและภาษาไทย ในส่วนที่เป็นภาษาไทยมีลักษณะเป็นข้อความที่ยาวต่อเนื่องกันโดยไม่มีการใช้ตัวอักษรหรือสัญลักษณ์ที่นำมาใช้คั่นระหว่างคำ ไม่มีการใช้เครื่องหมายวรรคตอนแสดงการแบ่งคำดังเช่นภาษาอังกฤษ นอกจากนี้ยังมีโครงสร้างของคำที่ไม่แน่นอน ซึ่งเป็นอุปสรรคในการนำคำภาษาไทยไปใช้ในการเปรียบเทียบสายอักขระ ดังนั้นในการพัฒนาโปรแกรมจึงต้องมีขั้นตอนการตัดคำก่อนที่จะนำคำภาษาไทยไปใช้ในกระบวนการต่อไป เริ่มต้นผู้วิจัยใช้ภาษา PHP ในการตัดคำภาษาไทย ซึ่งเป็นการตัดคำโดยใช้พจนานุกรมเป็นหลัก เมื่อพัฒนาโปรแกรมพบปัญหาในเรื่องของคำศัพท์ในพจนานุกรมที่มีอยู่ค่อนข้างจำกัด ทำให้การตัดคำไม่ได้ผลตามที่ควรจะเป็น และส่งผลให้การเปรียบเทียบสายอักขระมีประสิทธิภาพลดลง ผู้วิจัยจึงได้ศึกษาข้อมูลเพิ่มเติมในส่วนของ การตัดคำภาษาไทยด้วยวิธีการต่าง ๆ และพบว่าสำหรับงานด้านการประมวลผลข้อมูลภาษาไทย Library PythaiNLP [29] เป็นที่นิยมใช้ในการตัดคำภาษาไทยเป็นอย่างมาก เนื่องจากเป็นไลบรารีประมวลผลภาษาไทยแบบ open source ในภาษา Python มีฟังก์ชันการตัดคำที่ถูกออกแบบมาให้รองรับการตัดคำที่หลากหลายอัลกอริทึม (algorithm) และมีเครื่องมือการตัดคำ (engine) หลายรูปแบบ โดย engine ที่เป็นค่าเริ่มต้น (default) ของ PyThaiNLP tokenize ในขณะนี้ คือ newmm ซึ่งเป็นอัลกอริทึมการตัดคำอ้างอิงจากพจนานุกรมโดยใช้เทคนิค maximal matching algorithm และ Thai character cluster (TCC) [30] ที่ทำงานได้อย่างรวดเร็วและถูกต้อง นอกจากนี้จะสามารถปรับเปลี่ยนเครื่องมือการตัดคำในรูปแบบต่าง ๆ ได้แล้ว ยังสามารถตัดคำจากการเลือกคลังคำศัพท์หรือพจนานุกรม (custom dictionary) และเพิ่มคำศัพท์ที่ต้องการได้อีกด้วย ซึ่งในงานวิจัยนี้ใช้คลังคำศัพท์จากพจนานุกรมเล็กชิตรอน (LEXITRON dictionary) พัฒนาโดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ [31] เนื่องจากมีคำศัพท์ที่เป็นคำทับศัพท์ภาษาอังกฤษเป็นจำนวนมาก และได้ทำการเพิ่มข้อมูลคำทับศัพท์ภาษาอังกฤษที่เกี่ยวข้องกับผลิตภัณฑ์เครื่องสำอางเข้าไป (adaptation of LEXITRON dictionary)

สำหรับแนวทางการทำงานของโปรแกรมฯ มี 2 ส่วนหลัก คือ ส่วนของการตรวจสอบคำห้ามใช้ในชื่อผลิตภัณฑ์เครื่องสำอางทั้งภาษาไทยและภาษาอังกฤษ และส่วนของการตรวจสอบความสอดคล้องกันระหว่างชื่อผลิตภัณฑ์เครื่องสำอางกับส่วนประกอบในตำรับ ซึ่งจะใช้วิธีการเปรียบเทียบสายอักขระโดยการหาระยะทางเลเวนชเตย์น (Levenshtein distance; LD) ที่พัฒนาโดยใช้ภาษา PHP ทำการเปรียบเทียบกับคำในฐานข้อมูลที่ได้จัดทำไว้ เมื่อได้ค่า LD โปรแกรมจะนำค่าที่ได้ไปคำนวณหาความคล้ายคลึงกันของข้อความ (text similarity) ค่าที่ได้จากการคำนวณแสดงถึงลักษณะความคล้ายซึ่งมีค่าระหว่าง 0 ถึง 1 หากมีค่าเท่ากับ 0 หมายถึงสายอักขระทั้งสองสายไม่มีความคล้ายกันเลย แต่หากมีค่าเท่ากับ 1 หมายถึงสายอักขระทั้งสองสายเหมือนกันทุกประการ [32] สูตรการคำนวณนำเสนอในสมการที่ (2)

$$\text{text similarity (t1, t2)} = 1 - \frac{\text{Levenshtein distance (t1, t2)}}{\max(n(t1), n(t2))} \quad (2)$$

กำหนดให้

- t1, t2 แทนสายอักขระชุดต้นแบบและสายอักขระชุดเปรียบเทียบ
- Levenshtein distance แทนค่าความต่างกันระหว่างสายอักขระสองชุด โดยวัดจากจำนวนของการที่ต้องทำการตัดออก แทรก แทนที่อักขระในชุดที่นำมาเปรียบเทียบจนกระทั่งมีลักษณะเหมือนชุดอักขระที่เป็นต้นแบบทุกประการ ซึ่งการตัดออก แทรก หรือแทนที่อักขระ 1 ครั้ง จะมีค่าน้ำหนักเท่ากับ 1
- n แทนความยาวของสายอักขระ
- $\max(n(t1), n(t2))$ แทนค่าความยาวสูงสุดของสายอักขระใดสายอักขระหนึ่งจากสายอักขระสองชุด



รูปที่ 7 ขั้นตอนการทำงานของโปรแกรม

3.5.1.4 การทดสอบ (testing)

ทำการทดสอบโปรแกรมโดยการเปรียบเทียบผลการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้ง กับการตรวจสอบโดยเจ้าหน้าที่ผู้เชี่ยวชาญ ทำการทดสอบกับกลุ่มตัวอย่างที่ได้จากการคำนวณ ได้แก่ แบบคำขอจดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าผ่าน และแบบคำขอจดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าไม่ผ่าน จำนวนกลุ่มละ 200 รายการคำขอ วัดผลโดยการประเมินประสิทธิภาพของโปรแกรม

3.5.1.5 การบำรุงรักษา (maintenance)

เมื่อนำโปรแกรมไปใช้งาน ต้องมีการติดตามและบำรุงรักษาเพื่อให้โปรแกรมสามารถใช้งานได้มีประสิทธิภาพ รวมถึงการแก้ไขปรับปรุงโปรแกรมกรณีที่เกิดข้อผิดพลาดที่เพิ่งค้นพบ และการเขียนโมดูลการทำงานเพิ่ม หากผู้ใช้งานพบว่ามีปัญหาเกิดขึ้น อาจจะมีการทำการบันทึกปัญหาเหล่านั้นไว้และส่งต่อไปให้ผู้วิจัยเพื่อทำการแก้ไขโปรแกรมต่อไป

3.5.2 ออกแบบและพัฒนาแบบประเมินความพึงพอใจของผู้ใช้งานโปรแกรม

3.5.2.1 ศึกษาทฤษฎีวิธีการสร้างแบบประเมินความพึงพอใจจากตำราและเอกสารที่เกี่ยวข้อง เพื่อเป็นแนวทางในการสร้างแบบประเมินความพึงพอใจ

3.5.2.2 กำหนดหัวข้อและสร้างแบบประเมินความพึงพอใจของผู้ใช้งานโปรแกรม โดยแบ่งออกเป็น 3 ด้าน คือ

- 1) ตรงตามความต้องการของผู้ใช้งานโปรแกรม
- 2) การทำงานได้ตามฟังก์ชันงานของโปรแกรม
- 3) มีความสะดวกและง่ายต่อการใช้งาน

3.5.2.3 สร้างแบบสอบถามประเมินความพึงพอใจ ผู้วิจัยได้สร้างแบบสอบถามประเมินความพึงพอใจที่มีต่อโปรแกรมการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ ลักษณะของแบบสอบถามเป็นแบบมาตราส่วนประมาณค่า (rating scale) ของลิเคิร์ต (Likert scale) ซึ่งแบ่งเกณฑ์การให้คะแนนเป็น 5 ระดับ คือ มากที่สุด มาก ปานกลาง น้อย และน้อยที่สุด ในการประเมินแต่ละข้อจะใช้เกณฑ์การให้คะแนน ดังนี้

ระดับการประเมินความพึงพอใจ ระดับ 5 หมายถึง มากที่สุด

ระดับ 4 หมายถึง มาก

ระดับ 3 หมายถึง ปานกลาง

ระดับ 2 หมายถึง น้อย

ระดับ 1 หมายถึง น้อยที่สุด

3.5.2.4 การตรวจสอบ นำแบบประเมินความพึงพอใจที่ผู้วิจัยสร้างขึ้นไปให้อาจารย์ที่ปรึกษาตรวจสอบความถูกต้องเหมาะสม รวมถึงความสมบูรณ์ของแบบประเมินความพึงพอใจ

3.5.2.5 แก้ไขปรับปรุง นำแบบประเมินความพึงพอใจที่ได้รับการตรวจสอบมาปรับปรุงแก้ไขเพื่อนำไปใช้กับผู้ใช้งานโปรแกรมต่อไป

3.5.2.6 การนำไปใช้ นำแบบประเมินความพึงพอใจไปใช้กับผู้ใช้งานโปรแกรมเพื่อประเมินความพึงพอใจหลังจากที่ทดลองใช้โปรแกรม และทำการประเมินผลต่อไป

3.6 การเก็บรวบรวมข้อมูล

1) แบบคำขอจดแจ้งเครื่องสำอาง (แบบ จ.ค.1)

รวบรวมแบบคำขอจดแจ้งเครื่องสำอางที่ถูกส่งเข้ามาในระบบจดแจ้งเครื่องสำอางอัตโนมัติ e-Submission ของสำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา และคัดเลือกข้อมูลจากแบบคำขอจดแจ้งเครื่องสำอางเพื่อนำไปใช้ในการศึกษา ได้แก่ 1) ชื่อการค้าภาษาไทย 2) ชื่อการค้าภาษาอังกฤษ 3) ชื่อเครื่องสำอางภาษาไทย 4) ชื่อเครื่องสำอางภาษาอังกฤษ และ 5) ส่วนประกอบในตำรับเครื่องสำอาง

2) ข้อมูลความพึงพอใจของผู้ใช้งานที่มีต่อโปรแกรม

ผู้วิจัยได้ดำเนินการติดตั้งโปรแกรมสำหรับตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระไว้ที่เครื่องแม่ข่าย (Server) ผู้ใช้งานสามารถเข้าใช้โปรแกรมได้ที่ <http://45.91.135.31/~jiraporn/cosmetic5/basic/index.html> เมื่อผู้ใช้งานเข้าใช้งานโปรแกรมแล้ว ผู้วิจัยจะดำเนินการเก็บรวบรวมข้อมูลความพึงพอใจ จากแบบประเมินความพึงพอใจของผู้ใช้งานที่มีต่อโปรแกรมสำหรับตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ เพื่อนำไปวิเคราะห์ข้อมูลต่อไป

3.7 การวิเคราะห์ข้อมูลและประเมินผล

แบบคำขอจดแจ้งเครื่องสำอางที่ใช้ในการวิจัยนี้ จะถูกทำการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอาง และประเมินผลการตรวจสอบใหม่อีกครั้งโดยพนักงานเจ้าหน้าที่ผู้เชี่ยวชาญ จำนวน 3 ท่าน ประเมินอย่างเป็นอิสระต่อกัน หากผลการประเมินผ่านทั้ง 3 ท่าน จะถือว่าเป็นแบบคำขอจดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าผ่าน

ผู้วิจัยทำการประเมินประสิทธิภาพของโปรแกรมในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ

เปรียบเทียบกับ การตรวจสอบโดยเจ้าหน้าที่ผู้เชี่ยวชาญ เพื่อให้ได้โปรแกรมที่เหมาะสมต่อการนำไปใช้ ประโยชน์ตามที่ต้องการ ดังนั้นเพื่อให้ได้โปรแกรมที่มีประสิทธิภาพ ทำงานได้อย่างถูกต้องและแม่นยำ จึงได้ทำการนำค่าต่าง ๆ จากในตาราง confusion matrix ไปคำนวณหาค่าความถูกต้อง (accuracy), ค่าความระลึก (recall), ค่าความเที่ยง (precision) และค่า F1 [33] สูตรการคำนวณนำเสนอใน สมการที่ (3) - (6)

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

รูปที่ 8 ตาราง confusion matrix

true positive (TP)	คือ โปรแกรมแสดงผลการพิจารณา “ผ่าน” และเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณา “ผ่าน”
true negative (TN)	คือ โปรแกรมแสดงผลการพิจารณา “ไม่ผ่าน” และเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณา “ไม่ผ่าน”
false positive (FP)	คือ โปรแกรมแสดงผลการพิจารณา “ผ่าน” แต่เจ้าหน้าที่ผู้เชี่ยวชาญพิจารณา “ไม่ผ่าน”
false negative (FN)	คือ โปรแกรมแสดงผลการพิจารณา “ไม่ผ่าน” แต่เจ้าหน้าที่ผู้เชี่ยวชาญพิจารณา “ผ่าน”

การประเมินประสิทธิภาพของโปรแกรม

1. ค่าความถูกต้อง (accuracy)

ค่าความถูกต้อง หมายถึง ค่าที่วัดได้เข้าใกล้ค่าจริงมากเพียงใด ถือว่า มีความถูกต้องมากเท่านั้น

$$\text{accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (3)$$

2. ค่าความระลึก (recall)

ค่าความระลึก หมายถึง อัตราส่วนของการค้นพบข้อมูลที่ต้องการจากจำนวนข้อมูลที่ถูกต้องทั้งหมด

$$\text{recall} = \frac{TP}{(TP + FN)} \quad (4)$$

3. ค่าความเที่ยง (precision)

ค่าความเที่ยง หมายถึงค่าที่วัดเข้าใกล้กันเพียงใด ซึ่งจะหมายถึงความสามารถซ้ำค่าเดิม (repeatability) ของผลการวัดเป็นอัตราส่วนของการค้นพบข้อมูลที่ต้องการจากจำนวนข้อมูลที่ทำ การค้นคืนมา

$$\text{precision} = \frac{TP}{(TP + FP)} \quad (5)$$

4. ค่า F1

ค่า F1 เป็นค่าเฉลี่ยฮาร์โมนิก (harmonic mean) ระหว่างค่าความเที่ยง (precision) และค่าความระลึก (recall)

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (6)$$

3.8 สถิติที่ใช้ในการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้ผู้วิจัยทำการวิเคราะห์ข้อมูลในส่วนของ การประเมินประสิทธิภาพโปรแกรมโดย ใช้สถิติเชิงอนุมาน (inferential statistics) สถิติที่ใช้คือ McNemar's Chi-Square และการวิเคราะห์ ข้อมูลความพึงพอใจของผู้ใช้งานที่มีต่อโปรแกรมใช้สถิติเชิงพรรณนา (descriptive statistics) ค่าสถิติที่ใช้ ได้แก่ ค่าความถี่ (frequency), ค่าร้อยละ (percentage), ค่าเฉลี่ย (mean) และส่วน เบี่ยงเบนมาตรฐาน (standard deviation) [34]

ค่าสถิติพื้นฐาน ได้แก่

- 1) การหาค่าร้อยละ มีสูตรดังนี้

$$P = \frac{F \times 100}{N} \quad (7)$$

เมื่อ $P =$ ค่าร้อยละ

$F =$ จำนวนหรือความถี่ที่ต้องการหาค่าร้อยละ

$N =$ จำนวนข้อมูลทั้งหมด

2) การหาค่าเฉลี่ย (\bar{X}) มีสูตรดังนี้

$$\bar{X} = \frac{\sum x}{n} \quad (8)$$

เมื่อ $\bar{X} =$ ค่าเฉลี่ย

$\sum x =$ ผลรวมความถี่ของคะแนนทั้งหมด

$n =$ จำนวนผู้ทำแบบสอบถามทั้งหมด

3) การหาค่าส่วนเบี่ยงเบนมาตรฐาน (S.D.) มีสูตรดังนี้

$$S.D. = \frac{\sqrt{n\sum x^2 - (\sum x)^2}}{n(n-1)} \quad (9)$$

เมื่อ S.D. = ค่าส่วนเบี่ยงเบนมาตรฐาน

$\sum x =$ ผลรวมของคะแนนในกลุ่ม

$\sum x^2 =$ ผลรวมของคะแนนแต่ละตัวยกกำลังสอง

$n =$ จำนวนผู้ทำแบบสอบถาม

เกณฑ์ในการแปลผล

ข้อมูลความพึงพอใจของผู้ใช้งานที่มีต่อโปรแกรมสำหรับตรวจสอบซื้อการคำ ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ โดยการหาค่าเฉลี่ย (arithmetic mean) จากมาตรวัดของลิเคิร์ต หรือ Linkert rating scales ซึ่งเป็นเครื่องมือในการวัดเชิงคุณภาพ โดยวัดเป็นระดับต่าง ๆ จากสูงถึงต่ำ ซึ่งในการวิจัยนี้จะใช้ 5 ระดับ โดยกำหนดคะแนนเป็น 1 – 5 คะแนน ดังนั้นช่วงความกว้างของแต่ละชั้นเป็น 0.8 ซึ่งได้จากการคำนวณโดยใช้สูตร

$$\text{ความกว้างของอันตรภาคชั้น} = (\text{คะแนนสูงสุด} - \text{คะแนนต่ำสุด}) / \text{จำนวนชั้น}$$

$$\text{ความกว้างของอันตรภาคชั้น} = (5 - 1) / 5 = 0.8$$

ทำให้ได้เกณฑ์ในการแปลความหมายความพึงพอใจตามค่าเฉลี่ย (\bar{X}) ดังนี้

ตารางที่ 8 เกณฑ์ในการแปลความหมายความพึงพอใจตามค่าเฉลี่ย

ค่าเฉลี่ย (\bar{x})	ความหมาย
4.21 – 5.00	มีความพึงพอใจ มากที่สุด
3.41 – 4.20	มีความพึงพอใจ มาก
2.61 – 3.40	มีความพึงพอใจ ปานกลาง
1.81 – 2.60	มีความพึงพอใจ น้อย
1.00 – 1.80	มีความพึงพอใจ น้อยที่สุด



บทที่ 4

ผลการวิเคราะห์ข้อมูล

ผลการวิเคราะห์ข้อมูลในการศึกษาวิจัยเรื่อง การตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และ ส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ ผู้วิจัยแบ่ง การรายงานผลออกเป็น 3 ส่วน ประกอบด้วย

1. การพัฒนาโปรแกรมสำหรับใช้ในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และ ส่วนประกอบในตำรับเครื่องสำอาง
2. การประเมินประสิทธิภาพโปรแกรม
3. การประเมินความพึงพอใจของผู้ใช้งานโปรแกรม

4.1 การพัฒนาโปรแกรมสำหรับใช้ในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และ ส่วนประกอบในตำรับเครื่องสำอาง

ผู้ใช้งานสามารถเข้าใช้งานโปรแกรมได้ที่ <http://45.91.135.31/~jiraporn/cosmetic5/basic/index.html> โดยผู้วิจัยได้ออกแบบหน้าจอซึ่งเป็นส่วนของการรับข้อมูล ได้แก่

- 1) ชื่อการค้าภาษาไทย
- 2) ชื่อการค้าภาษาอังกฤษ
- 3) ชื่อเครื่องสำอางภาษาไทย
- 4) ชื่อเครื่องสำอางภาษาอังกฤษ
- 5) สารที่เป็นส่วนประกอบในตำรับเครื่องสำอาง

โปรแกรมตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอาง คิวอีทีเปรียบเทียบสายอักขระ

ชื่อการค้าภาษาไทย

ชื่อการค้าภาษาอังกฤษ

ชื่อเครื่องสำอางภาษาไทย

ชื่อเครื่องสำอางภาษาอังกฤษ

รายการสาร

ตัวอักษร
1 NIACINAMIDE
2 1197 15 6 TRANEXAMIC ACID
3 9064 41 9 HYALURONIC ACID 0.500000000000

ตรวจสอบ เพิ่มข้อมูล

Jiraporn Wongkiet
Department of Biomedicine and Health Informatics, Faculty of Pharmacy, Sripatum University
Health Consumer Protection and Pharmacy Department, Pira Nareson to Apurthaya Provincial Public Health Office

รูปที่ 9 หน้าจอส่วนของการรับข้อมูล

การทำงานของโปรแกรมประกอบด้วย 2 ส่วนหลัก ได้แก่ การตรวจสอบคำห้ามใช้ในชื่อการค้าและชื่อเครื่องสำอาง และการตรวจสอบความสอดคล้องกันระหว่างชื่อผลิตภัณฑ์เครื่องสำอางกับส่วนประกอบในตำรับ ในการใช้งานโปรแกรมผู้ใช้งานสามารถกรอกข้อมูลที่ต้องการตรวจสอบได้แก่ ชื่อการค้าภาษาไทย ชื่อการค้าภาษาอังกฤษ ชื่อเครื่องสำอางภาษาไทย ชื่อเครื่องสำอางภาษาอังกฤษ และสารที่เป็นส่วนประกอบในตำรับเครื่องสำอาง ผลลัพธ์จากการตรวจสอบข้อมูลจะแสดงผลการพิจารณาค่าขจัดแจ้งเครื่องสำอางเป็น 4 กรณี ดังนี้

- 1) ผ่าน: ไม่พบคำห้ามใช้ และมีความสอดคล้องกันระหว่างชื่อผลิตภัณฑ์เครื่องสำอางกับส่วนประกอบในตำรับ
- 2) ไม่ผ่าน: ไม่พบคำห้ามใช้ แต่ชื่อผลิตภัณฑ์เครื่องสำอางกับส่วนประกอบในตำรับไม่สอดคล้องกัน
- 3) ไม่ผ่าน: พบคำห้ามใช้ แต่มีความสอดคล้องกันระหว่างชื่อผลิตภัณฑ์เครื่องสำอางกับส่วนประกอบในตำรับ
- 4) ไม่ผ่าน: พบคำห้ามใช้ และชื่อผลิตภัณฑ์เครื่องสำอางกับส่วนประกอบในตำรับไม่สอดคล้องกัน

โปรแกรมตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอาง ด้วยวิธีเปรียบเทียบสายอักขระ

ชื่อบริษัท
AROKA REALTY

ชนิดสินค้า โปรดเลือกให้เรียบร้อย
VITAMIN C WHITENING LOTION

1 7782-18-5 AQUA
2 56-81-5 GLYCERIN 3.0000000000000000
3 98-92-0 NIACINAMIDE 3.0000000000000000
4 9005-00-9 STEARETH-21
5 50-81-7 ASCORBIC ACID 0.0500000000000000
6 FRAGRANCE 0.0300000000000000
7 107-88-0 BUTYLENE GLYCOL 1.5000000000000000
8 97-59-6 ALLANTOIN 0.1500000000000000

ตรวจสอบ ลบข้อมูล

Jiraporn Wongtan
Department of Biomedicine and Health Informatics, Faculty of Pharmacy, Silpakorn University
Health Consumer Protection and Pharmacy Department, Phra Nakhon Si Ayutthaya Provincial Public Health Office

รูปที่ 10 ตัวอย่างการกรอกข้อมูลเพื่อทำการตรวจสอบ


Not secure | 45.91.135.31/~jgspom/cosmetics/basic/demo_javan_con.php

ไปเก็บตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในสินค้าเครื่องสำอาง ด้วยวิธีป้อนชื่อการค้า

- ชื่อการค้า (ภาษาไทย) : ลาล่า บีบี
- ชื่อการค้า (ภาษาอังกฤษ) : AROKA BEAUTY
- ชื่อเครื่องสำอาง (ภาษาไทย) : วิตามิน ซี โฟมหน้าใส โกลด์
- ชื่อเครื่องสำอาง (ภาษาอังกฤษ) : VITAMIN C WHITENING LOTION

ชื่อการค้า/ชื่อเครื่องสำอาง-สารที่เป็นส่วนประกอบ

ชื่อค้น	สารที่เป็นส่วนประกอบ	ประเภทของชื่อ
WHITENING	NIACINAMIDE	ชื่อเครื่องสำอาง
โฟมหน้าใส	NIACINAMIDE	ชื่อเครื่องสำอาง
VITAMIN C	ASCORBIC ACID	ชื่อเครื่องสำอาง
WHITENING	ASCORBIC ACID	ชื่อเครื่องสำอาง
C	ASCORBIC ACID	ชื่อเครื่องสำอาง
โฟมหน้าใส	ASCORBIC ACID	ชื่อเครื่องสำอาง



ไม่ผ่านการตรวจสอบ เนื่องจาก
ชื่อการค้าภาษาอังกฤษพบคำห้ามใช้, ชื่อการค้าภาษาไทยพบคำ
ห้ามใช้
พบคำที่อาจพิมพ์ผิดคือ WHITENING

รูปที่ 11 ตัวอย่างผลลัพธ์จากการตรวจสอบข้อมูล กรณีผลการพิจารณาไม่ผ่าน


Not secure | 45.91.135.31/~jgspom/cosmetics/basic/demo_javan_con.php

ไปเก็บตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในสินค้าเครื่องสำอาง ด้วยวิธีป้อนชื่อการค้า

- ชื่อการค้า (ภาษาไทย) : เอบีบี
- ชื่อการค้า (ภาษาอังกฤษ) : ABC
- ชื่อเครื่องสำอาง (ภาษาไทย) : โยเกิร์ต สดส์ สบู่ โยเกิร์ต สบู่ล้าง
- ชื่อเครื่องสำอาง (ภาษาอังกฤษ) : YOGURT MILK CREAM SOAP PLUS Q10

ชื่อการค้า/ชื่อเครื่องสำอาง-สารที่เป็นส่วนประกอบ

ชื่อค้น	สารที่เป็นส่วนประกอบ	ประเภทของชื่อ
โยเกิร์ต	YOGURT POWDER	ชื่อเครื่องสำอาง
YOGURT	YOGURT POWDER	ชื่อเครื่องสำอาง
MILK	HYDROLYZED MILK PROTEIN	ชื่อเครื่องสำอาง



ผ่านการตรวจสอบ

รูปที่ 12 ตัวอย่างผลลัพธ์จากการตรวจสอบข้อมูล กรณีผลการพิจารณาผ่าน

4.2 การประเมินประสิทธิภาพของโปรแกรม

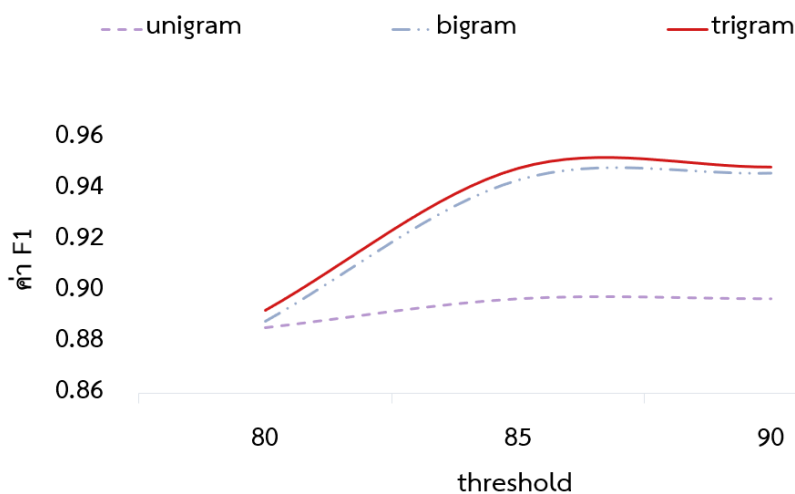
เนื่องจากคำห้ามใช้และคำที่สามารถนำมาใช้เป็นชื่อผลิตภัณฑ์เครื่องสำอางได้โดยต้องมีสารที่สอดคล้องในสูตรตำรับที่แสดงในฐานข้อมูล มีลักษณะเป็นทั้งคำเดี่ยว และกลุ่มคำที่ประกอบด้วยคำตั้งแต่สองคำขึ้นไป แต่คำที่ได้จากขั้นตอนการตัดคำจะมีลักษณะเป็นเดี่ยวหลาย ๆ คำ ซึ่งอาจไม่สื่อความหมายหรือตรงกับคำที่ต้องการเปรียบเทียบ เพื่อเป็นการเพิ่มความถูกต้องของคำ และเพิ่มประสิทธิภาพในการเปรียบเทียบสายอักขระ ผู้วิจัยจึงได้เลือกใช้เทคนิคเอ็นแกรม (N-gram) [35] เข้ามาใช้ในการกำหนดคุณลักษณะของคำ โดยกำหนดเป็น 3 รูปแบบ ได้แก่ 1-แกรม (unigram), 2-แกรม (bigram) และ 3-แกรม (trigram) ซึ่งคุณลักษณะของคำรูปแบบ bigram จะเป็นการรวมกันของคำในรูปแบบ unigram และ bigram และคุณลักษณะของคำรูปแบบ trigram จะเป็นการรวมกันของคำในรูปแบบ unigram, bigram และ trigram มีการใช้เครื่องมือตัดคำ newmm ซึ่งเป็นอัลกอริทึมที่เป็นค่าเริ่มต้นของ PyThaiNLP ในการตัดคำภาษาไทย

การกำหนดค่าเทรชโฮลด์ (threshold) ซึ่งเป็นค่าที่ใช้ในการเปรียบเทียบความคล้ายคลึงกันของคำหรือข้อความโดยการหารระยะทางเลเวนชเตย์น การกำหนดค่า threshold 80 หมายถึง สายอักขระสองชุดที่มีความคล้ายคลึงกันมากกว่าหรือเท่ากับร้อยละ 80 จะกำหนดว่าเป็นสายอักขระชุดเดียวกัน สำหรับงานวิจัยนี้กำหนดค่า threshold เริ่มต้นที่ 80 ผลการเปรียบเทียบประสิทธิภาพของโปรแกรมในการพิจารณาค่าขอจดแจ้งเครื่องสำอาง พบว่าคุณลักษณะของคำรูปแบบ trigram มีค่า F1 สูงสุด เท่ากับ 0.8926 รองลงมาคือคุณลักษณะของคำรูปแบบ bigram และ unigram โดยมีค่า F1 เท่ากับ 0.8883 และ 0.8855 ตามลำดับ

ตารางที่ 9 ค่า F1 ของการตัดคำแต่ละคุณลักษณะที่ค่า threshold 80

คุณลักษณะของคำ	ค่า F1
unigram	0.8855
bigram	0.8883
trigram	0.8926

เพื่อหาค่า threshold ที่เหมาะสมที่สุดที่ทำให้การเปรียบเทียบสายอักขระให้มีประสิทธิภาพสูงสุดในกระบวนการพิจารณาค่าขอจดแจ้งเครื่องสำอาง ผู้วิจัยได้ทำการทดลองปรับค่า threshold เป็น 85 และ 90 ผลการทดลองพบว่า คุณลักษณะของคำรูปแบบ trigram ยังคงมีค่า F1 สูงสุด และมีค่าเพิ่มขึ้น เป็น 0.9482 และ 0.9487 ตามลำดับ



รูปที่ 13 ผลการเปรียบเทียบค่า F1 ของคำแต่ละคุณลักษณะ ที่ค่า threshold 80, 85 และ 90

ดังนั้นเพื่อให้การเปรียบเทียบสายอักขระครอบคลุมคุณลักษณะของคำ ทั้งคำเดี่ยวและกลุ่มคำที่ประกอบด้วยคำตั้งแต่สองคำขึ้นไป เป็นการเพิ่มความถูกต้องของคำและเพิ่มประสิทธิภาพในการเปรียบเทียบสายอักขระ ผู้วิจัยจึงได้เลือกใช้คุณลักษณะคำรูปแบบ trigram ไปใช้ในการเปรียบเทียบสายอักขระต่อไป

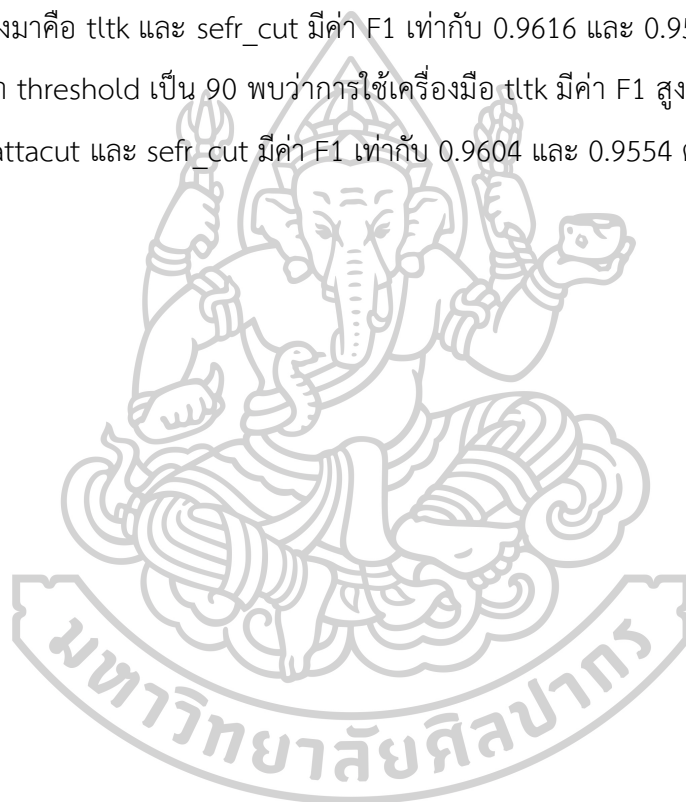
สำหรับฟังก์ชันการตัดคำภาษาไทย (word tokenization) ด้วย Library PyThaiNLP ในการศึกษาวิจัยนี้ผู้วิจัยทดลองใช้เครื่องมือตัดคำ (engine) ทั้งหมด 8 วิธี ได้แก่ newmm, longest, multi_cut, deepcut, attacut, newmm-safe, tltk และ sefr_cut เพื่อศึกษาเปรียบเทียบว่าเครื่องมือตัดคำใดเหมาะสมที่สุดที่ทำให้การเปรียบเทียบสายอักขระมีประสิทธิภาพสูงสุดในกระบวนการพิจารณาคำขอจดแจ้งเครื่องหมาย

ตารางที่ 10 ตัวอย่างผลการตัดคำด้วย Library PyThaiNLP โดยใช้เครื่องมือตัดคำทั้ง 8 วิธี

ข้อความตัวอย่าง: สบู่ก้อนมะขามผสมน้ำผึ้ง

ลำดับ	engine	ผลลัพธ์ที่ได้
1	newmm	สบู่, ก้อน, มะขาม, ผสม, น้ำผึ้ง
2	longest	สบู่, ก้อน, มะขาม, ผสม, น้ำผึ้ง
3	multi_cut	สบู่, ก้อนมะขาม, ผสม, น้ำผึ้ง
4	deepcut	สบู่ก้อน, มะขาม, ผสม, น้ำผึ้ง
5	attacut	สบู่, ก้อน, มะขามผสม, น้ำ, ผึ้ง
6	newmm-safe	สบู่, ก้อน, มะขาม, ผสม, น้ำผึ้ง
7	tltk	สบู่, ก้อน, มะขาม, ผสม, น้ำผึ้ง
8	sefr_cut	สบู่ก้อน, มะขาม, ผสมน้ำผึ้ง

ผลการประเมินประสิทธิภาพของโปรแกรมในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และ ส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ โดยใช้ เครื่องมือตัดคำทั้ง 8 วิธี เมื่อกำหนดค่า threshold ที่ระดับ 80 ทำการคำนวณหาค่าความถูกต้อง, ค่า ความระลึก, ค่าความเที่ยง และค่า F1 พบว่าการตัดคำภาษาไทยด้วยเครื่องมือตัดคำ attacut มีค่า F1 สูงสุด โดยมีค่าเท่ากับ 0.9195 รองลงมาคือ sefr_cut และ tltk มีค่า F1 เท่ากับ 0.9077 และ 0.9076 ตามลำดับ เมื่อทำการปรับค่า threshold เป็น 85 การตัดคำภาษาไทยด้วยเครื่องมือตัดคำทั้ง 8 วิธี พบว่ามีค่า F1 เพิ่มมากขึ้น โดยพบว่าการใช้เครื่องมือ attacut มีค่า F1 สูงสุด มีค่าเท่ากับ 0.9628 รองลงมาคือ tltk และ sefr_cut มีค่า F1 เท่ากับ 0.9616 และ 0.9552 ตามลำดับ และเมื่อ ทำการปรับค่า threshold เป็น 90 พบว่าการใช้เครื่องมือ tltk มีค่า F1 สูงสุด มีค่าเท่ากับ 0.9645 รองลงมาคือ attacut และ sefr_cut มีค่า F1 เท่ากับ 0.9604 และ 0.9554 ตามลำดับ



ตารางที่ 11 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยทั้ง 8 วิธี และกำหนดค่า threshold ที่ 80

engine		ค่าประสิทธิภาพของโปรแกรม			
		accuracy	precision	recall	F1
threshold 80					
newmm					
TP = 162	FN = 38	0.9025	0.9939	0.8100	0.8926
FP = 1	TN = 199				
longest					
TP = 163	FN = 37	0.9050	0.9939	0.8150	0.8956
FP = 1	TN = 199				
multi_cut					
TP = 163	FN = 37	0.9000	0.9819	0.8150	0.8907
FP = 3	TN = 197				
deepcut					
TP = 158	FN = 42	0.8800	0.9634	0.7900	0.8681
FP = 6	TN = 194				
newmm-safe					
TP = 162	FN = 38	0.9025	0.9939	0.8100	0.8926
FP = 1	TN = 199				
attacut					
TP = 177	FN = 23	0.9225	0.9568	0.8850	0.9195
FP = 8	TN = 192				
sefr_cut					
TP = 172	FN = 28	0.9125	0.9609	0.8600	0.9077
FP = 7	TN = 193				
tltk					
TP = 167	FN = 33	0.9150	0.9940	0.8350	0.9076
FP = 1	TN = 199				

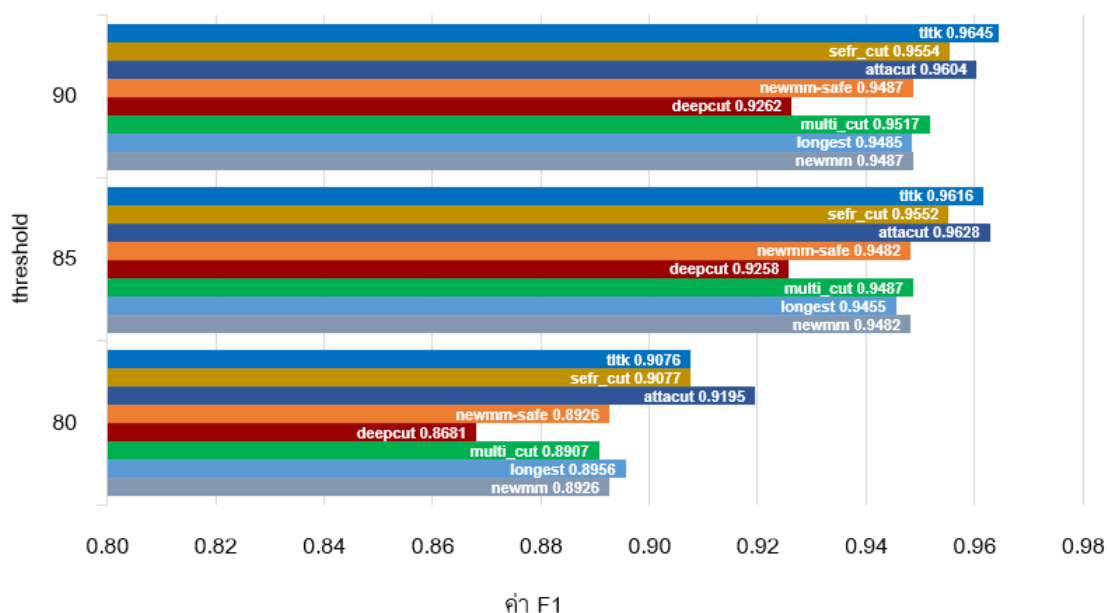
ตารางที่ 12 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยทั้ง 8 วิธี และกำหนดค่า threshold ที่ 85

engine		ค่าประสิทธิภาพของโปรแกรม			
		accuracy	precision	accuracy	F1
threshold 85					
newmm					
TP = 183	FN = 17	0.9500	0.9839	0.9150	0.9482
FP = 3	TN = 197				
longest					
TP = 182	FN = 18	0.9475	0.9838	0.9100	0.9455
FP = 3	TN = 197				
multi_cut					
TP = 185	FN = 15	0.9500	0.9737	0.9250	0.9487
FP = 5	TN = 195				
deepcut					
TP = 181	FN = 19	0.9275	0.9476	0.9050	0.9258
FP = 10	TN = 190				
newmm-safe					
TP = 183	FN = 17	0.9500	0.9839	0.9150	0.9482
FP = 3	TN = 197				
attacut					
TP = 194	FN = 6	0.9625	0.9557	0.9700	0.9628
FP = 9	TN = 191				
sefr_cut					
TP = 192	FN = 8	0.9550	0.9505	0.9600	0.9552
FP = 10	TN = 190				
tltk					
TP = 188	FN = 12	0.9625	0.9843	0.9400	0.9616
FP = 3	TN = 197				

ตารางที่ 13 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยทั้ง 8 วิธี และกำหนดค่า threshold ที่ 90

engine		ค่าประสิทธิภาพของโปรแกรม			
		accuracy	precision	accuracy	F1
threshold 90					
newmm					
TP = 185	FN = 15	0.9500	0.9737	0.9250	0.9487
FP = 5	TN = 195				
longest					
TP = 184	FN = 16	0.9500	0.9787	0.9200	0.9485
FP = 4	TN = 196				
multi_cut					
TP = 187	FN = 13	0.9525	0.9689	0.9350	0.9517
FP = 6	TN = 194				
deepcut					
TP = 182	FN = 18	0.9275	0.9430	0.9100	0.9262
FP = 11	TN = 189				
newmm-safe					
TP = 185	FN = 15	0.9500	0.9737	0.9250	0.9487
FP = 5	TN = 195				
attacut					
TP = 194	FN = 6	0.9600	0.9510	0.9700	0.9604
FP = 10	TN = 190				
sefr_cut					
TP = 193	FN = 7	0.9550	0.9461	0.9650	0.9554
FP = 11	TN = 189				
tltk					
TP = 190	FN = 10	0.9650	0.9794	0.9500	0.9645
FP = 4	TN = 196				

จากการตัดคำภาษาไทยโดยใช้เครื่องมือตัดคำทั้ง 8 วิธี ที่ค่า threshold 80, 85 และ 90 เปรียบเทียบกับค่า F1 ได้ผลการทดลองดังรูปที่ 13



รูปที่ 14 ผลการเปรียบเทียบประสิทธิภาพของโปรแกรมจากค่า F1

เนื่องจาก Library PyThaiNLP มีฟังก์ชันการตัดคำจากการเลือกคลังคำศัพท์หรือพจนานุกรม (custom dictionary) และสามารถเพิ่มคำศัพท์ที่ต้องการได้ ซึ่งฟังก์ชันดังกล่าวสามารถใช้งานร่วมกับเครื่องมือตัดคำต่าง ๆ ได้ ในการวิจัยนี้ผู้วิจัยจึงใช้คลังคำศัพท์จาก adaptation of LEXITRON dictionary

ผลการประเมินประสิทธิภาพของโปรแกรมในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และ ส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ โดยใช้เครื่องมือตัดคำทั้ง 8 วิธี ร่วมกับการใช้ adaptation of LEXITRON dictionary เมื่อกำหนดค่า threshold ที่ระดับ 80 ทำการคำนวณหาค่าความถูกต้อง, ค่าความระลึก, ค่าความเที่ยง และค่า F1 พบว่าการตัดคำภาษาไทยด้วยเครื่องมือตัดคำ newmm และ newmm-safe ร่วมกับการใช้พจนานุกรม มีค่า F1 สูงสุด โดยมีค่าเท่ากับ 0.9223 รองลงมาคือ attacut และ longest มีค่า F1 เท่ากับ 0.9195 และ 0.9164 ตามลำดับ เมื่อทำการปรับค่า threshold เป็น 85 การตัดคำภาษาไทยด้วยเครื่องมือตัดคำทั้ง 8 วิธี ร่วมกับการใช้ adaptation of LEXITRON dictionary มีค่า F1 เพิ่มขึ้น โดยพบว่าการใช้เครื่องมือ newmm และ newmm-safe ร่วมกับการใช้พจนานุกรม มีค่า F1 สูงสุด โดยมีค่าเท่ากับ 0.9772 รองลงมาคือ longest และ multi_cut มีค่า F1 เท่ากับ 0.9746 และ

0.9723 ตามลำดับ และเมื่อทำการปรับค่า threshold เป็น 90 พบว่าการใช้เครื่องมือ newmm และ newmm-safe ร่วมกับการใช้พจนานุกรม มีค่า F1 สูงสุด โดยมีค่าเท่ากับ 0.9773 รองลงมาคือ longest และ multi_cut มีค่า F1 เท่ากับ 0.9747 และ 0.9724 ตามลำดับ



ตารางที่ 14 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยร่วมกับการใช้ adaptation of LEXiTRON dictionary และกำหนดค่า threshold ที่ 80

engine		ค่าประสิทธิภาพของโปรแกรม			
		accuracy	precision	accuracy	F1
threshold 80					
newmm ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 172	FN = 28	0.9275	0.9942	0.8600	0.9223
FP = 1	TN = 199				
longest ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 170	FN = 30	0.9225	0.9942	0.8500	0.9164
FP = 1	TN = 199				
multi_cut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 171	FN = 29	0.9200	0.9828	0.8550	0.9144
FP = 3	TN = 197				
deepcut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 166	FN = 34	0.8925	0.9486	0.8300	0.8853
FP = 9	TN = 191				
newmm-safe ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 172	FN = 28	0.9275	0.9942	0.8600	0.9223
FP = 1	TN = 199				
attacut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 177	FN = 23	0.9225	0.9568	0.8850	0.9195
FP = 8	TN = 192				
sefr_cut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 172	FN = 28	0.9125	0.9609	0.8600	0.9077
FP = 7	TN = 193				
tltk ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 167	FN = 33	0.9150	0.9940	0.8350	0.9076
FP = 1	TN = 199				

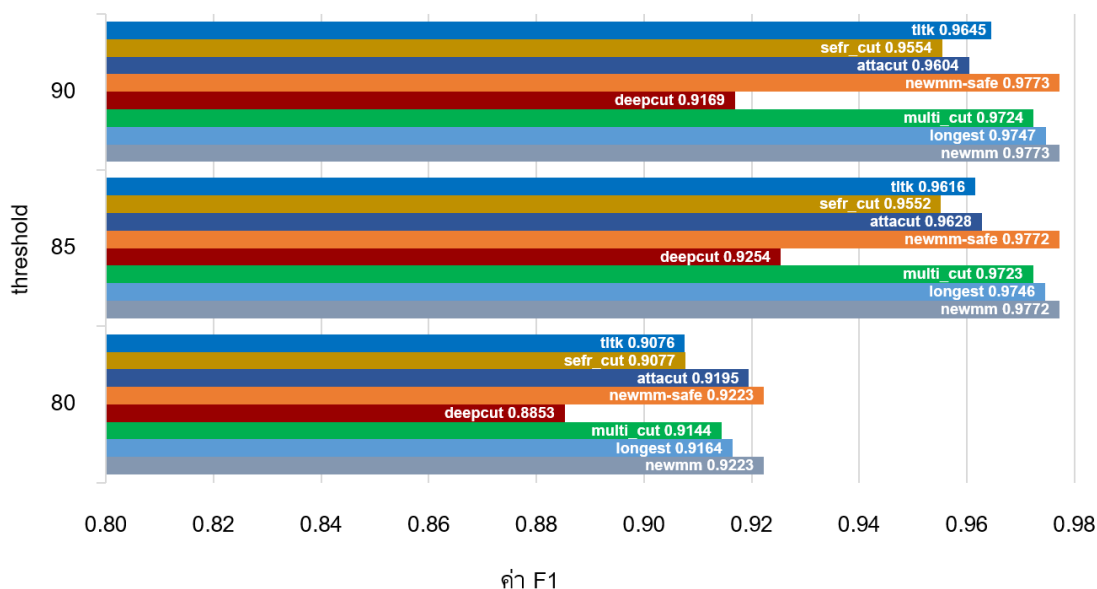
ตารางที่ 15 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยร่วมกับการใช้ adaptation of LEXiTRON dictionary และกำหนดค่า threshold ที่ 85

engine	ค่าประสิทธิภาพของโปรแกรม				
	accuracy	precision	accuracy	F1	
threshold 85					
newmm ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 193	FN = 7	0.9775	0.9897	0.9650	0.9772
FP = 2	TN = 198				
longest ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 192	FN = 8	0.9750	0.9897	0.9600	0.9746
FP = 2	TN = 198				
multi_cut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 193	FN = 7	0.9725	0.9797	0.9650	0.9723
FP = 4	TN = 196				
deepcut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 180	FN = 20	0.9275	0.9524	0.9000	0.9254
FP = 9	TN = 191				
newmm-safe ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 193	FN = 7	0.9775	0.9897	0.9650	0.9772
FP = 2	TN = 198				
attacut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 194	FN = 6	0.9625	0.9557	0.9700	0.9628
FP = 9	TN = 191				
sefr_cut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 192	FN = 8	0.9550	0.9505	0.9600	0.9552
FP = 10	TN = 190				
tltk ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 188	FN = 12	0.9625	0.9843	0.9400	0.9616
FP = 3	TN = 197				

ตารางที่ 16 ผลการประเมินประสิทธิภาพของโปรแกรม เมื่อใช้เครื่องมือตัดคำภาษาไทยร่วมกับการใช้ adaptation of LEXiTRON dictionary และกำหนดค่า threshold ที่ 90

engine		ค่าประสิทธิภาพของโปรแกรม			
		accuracy	precision	accuracy	F1
threshold 90					
newmm ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 194	FN = 6	0.9775	0.9848	0.9700	0.9773
FP = 3	TN = 197				
longest ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 193	FN = 7	0.9750	0.9847	0.9650	0.9747
FP = 3	TN = 197				
multi_cut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 194	FN = 6	0.9725	0.9749	0.9700	0.9724
FP = 5	TN = 195				
deepcut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 182	FN = 18	0.9175	0.9239	0.9100	0.9169
FP = 15	TN = 185				
newmm-safe ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 194	FN = 6	0.9775	0.9848	0.9700	0.9773
FP = 3	TN = 197				
attacut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 194	FN = 6	0.9600	0.9510	0.9700	0.9604
FP = 10	TN = 190				
sefr_cut ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 193	FN = 7	0.9500	0.9461	0.9650	0.9554
FP = 11	TN = 189				
tltk ร่วมกับ adaptation of LEXiTRON dictionary					
TP = 190	FN = 10	0.9650	0.9794	0.9500	0.9645
FP = 4	TN = 196				

ผลการเปรียบเทียบประสิทธิภาพของโปรแกรมโดยใช้เครื่องมือตัดคำวิธีต่าง ๆ ร่วมกับการใช้ adaptation of LEXITRON dictionary แสดงดังรูปที่ 14



รูปที่ 15 ผลการเปรียบเทียบประสิทธิภาพของโปรแกรมเมื่อใช้เครื่องมือตัดคำวิธีต่าง ๆ ร่วมกับการใช้ adaptation of LEXITRON dictionary

จากการเปรียบเทียบประสิทธิภาพของโปรแกรมในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ พบว่าที่ค่า threshold 90 การใช้เครื่องมือตัดคำภาษาไทย newmm ร่วมกับ Adaptation of LEXITRON dictionary ทำให้การทำงานของโปรแกรมมีประสิทธิภาพมากที่สุด อีกทั้งเครื่องมือตัดคำ newmm ยังเป็น engine ที่เป็นค่าเริ่มต้นของ PyThaiNLP tokenize อีกด้วย ผู้วิจัยจึงนำผลการศึกษาดังกล่าวมาทำการวิเคราะห์ข้อมูลการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำของจดแจ้งเครื่องสำอางระหว่างการตรวจสอบด้วยโปรแกรมและตรวจสอบโดยเจ้าหน้าที่ผู้เชี่ยวชาญ โดยใช้ค่าสถิติ McNemar's Chi-Square test

ผู้วิจัยได้กำหนดสมมติฐานในการทดสอบ ดังนี้

H_0 : การตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำของจดแจ้งเครื่องสำอางระหว่างการตรวจสอบด้วยโปรแกรมและการตรวจสอบโดยเจ้าหน้าที่ผู้เชี่ยวชาญไม่แตกต่างกัน

H_1 : การตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำของจดแจ้งเครื่องสำอางระหว่างการตรวจสอบด้วยโปรแกรมและตรวจสอบโดยเจ้าหน้าที่ผู้เชี่ยวชาญแตกต่างกัน

ตารางที่ 17 การจัดเรียงข้อมูลผลการตรวจสอบในตาราง two-dimensional contingency

ผลการตรวจสอบ โดยเจ้าหน้าที่ผู้เชี่ยวชาญ	ผลการตรวจสอบโดยโปรแกรม		รวม
	ผ่าน	ไม่ผ่าน	
ผ่าน	194 (a)	6 (b)	200
ไม่ผ่าน	3 (c)	197 (d)	200
รวม	197	203	

เนื่องจากค่า $b+c < 20$ และค่า $b > (b+c)/2$ ใช้สูตรคำนวณหาค่า p-value ดังนี้

$$\text{สูตรคำนวณ } p = 2 \sum_{i=b}^{b+c} \binom{b+c}{i} (0.5)^{b+c}$$

$$p = 2 \left[\binom{9}{6} (0.5)^9 + \binom{9}{7} (0.5)^9 + \binom{9}{8} (0.5)^9 + \binom{9}{9} (0.5)^9 \right]$$

$$p = .508$$

ตารางที่ 18 เปรียบเทียบผลการตรวจสอบซื้อการคำ ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอาง ($n = 400$ รายการคำขอ) ระหว่างการตรวจสอบด้วยโปรแกรมและการตรวจสอบโดยเจ้าหน้าที่ผู้เชี่ยวชาญ ด้วยสถิติ McNemar's Chi-Square test

ผลการตรวจสอบ โดยเจ้าหน้าที่ผู้เชี่ยวชาญ	ผลการตรวจสอบโดยโปรแกรม		p-value
	ผ่าน	ไม่ผ่าน	
ผ่าน	194	6	.508*
ไม่ผ่าน	3	197	

*Binomial distribution

จากผลลัพธ์ได้ค่า p-value เท่ากับ .508 ซึ่งมากกว่าระดับนัยสำคัญที่กำหนด ($\alpha = .05$) จึงไม่สามารถปฏิเสธ H_0 ได้ สรุปได้ว่า การตรวจสอบซื้อการคำ ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางระหว่างการตรวจสอบด้วยโปรแกรมและการตรวจสอบโดยเจ้าหน้าที่ผู้เชี่ยวชาญไม่แตกต่างกัน

4.3 การประเมินความพึงพอใจของผู้ใช้งานโปรแกรม

ผลการประเมินความพึงพอใจของผู้ใช้งานโปรแกรม ผู้วิจัยได้นำโปรแกรมสำหรับการตรวจสอบซื้อการคำ ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอาง ให้ผู้ใช้งานซึ่งเป็นเภสัชกรในกลุ่มงานคุ้มครองผู้บริโภคและเภสัชสาธารณสุข สำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา จำนวน 10 ท่าน ที่ปฏิบัติงานหรือเคยปฏิบัติงานเกี่ยวข้องกับการตรวจสอบผลิตภัณฑ์เครื่องสำอางก่อนออกสู่ท้องตลาด หรือการตรวจสอบผลิตภัณฑ์เครื่องสำอางหลังออกสู่ท้องตลาด ประเมินความพึง

พอใจที่มีต่อโปรแกรม 3 ด้าน ได้แก่ 1. ด้านตรงตามความต้องการของผู้ใช้งานโปรแกรม 2. ด้านประสิทธิภาพของโปรแกรม และ 3. ด้านความสะดวกและง่ายต่อการใช้งาน

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม

ข้อมูลทั่วไป	จำนวน (N = 10)	ร้อยละ
เพศ		
1. ชาย	4	40
2. หญิง	6	60
รวม	10	100
อายุ		
1. 25 ปี หรือต่ำกว่า	0	0
2. ระหว่าง 26 - 30 ปี	5	50
3. ระหว่าง 31 - 35 ปี	2	20
4. ระหว่าง 36 - 40 ปี	0	0
5. 41 ปี ขึ้นไป	3	30
รวม	10	100
ระดับการศึกษา		
1. ต่ำกว่าปริญญาตรี	0	0
2. ปริญญาตรี	7	70
3. ปริญญาโท	3	30
4. ปริญญาเอก	0	0
รวม	10	100

ตอนที่ 2 ผลการประเมินความพึงพอใจของผู้ใช้งานที่มีต่อโปรแกรม

การวิเคราะห์ผลการประเมินความพึงพอใจของผู้ใช้โปรแกรมสำหรับการตรวจสอบซื้อการค้าซื้อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอาง พบว่า ด้านตรงตามความต้องการของผู้ใช้งานโปรแกรม มีความพึงพอใจอยู่ในระดับมากที่สุด ($\bar{x} = 4.80$, S.D. = 0.40) ด้านประสิทธิภาพของโปรแกรม มีความพึงพอใจอยู่ในระดับมากที่สุด ($\bar{x} = 4.67$, S.D. = 0.49) และด้านความสะดวกและง่ายต่อการใช้งาน มีความพึงพอใจอยู่ในระดับมากที่สุด ($\bar{x} = 4.55$, S.D. = 0.50) สรุปได้ว่า ความ

พึงพอใจของของผู้ใช้โปรแกรมสำหรับการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอางซึ่งทั้ง 3 ด้านที่กล่าวมา มีความพึงพอใจอยู่ในระดับมากที่สุด

ตารางที่ 19 ภาพรวมผลการประเมินความพึงพอใจของผู้ใช้งานโปรแกรมสำหรับการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอาง

รายการประเมิน	\bar{x}	S.D.	ระดับคุณภาพ
1. ด้านตรงตามความต้องการของผู้ใช้งานโปรแกรม	4.80	0.40	มากที่สุด
2. ด้านประสิทธิภาพของโปรแกรม	4.67	0.49	มากที่สุด
3. ด้านความสะดวกและง่ายต่อการใช้งาน	4.55	0.50	มากที่สุด

ผลการประเมินความพึงพอใจของผู้ใช้โปรแกรมสำหรับการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอาง ด้านตรงตามความต้องการของผู้ใช้งานโปรแกรม รายละเอียดการประเมินแสดงในตารางที่ 18

ตารางที่ 20 ผลการประเมินความพึงพอใจของผู้ใช้งานด้านตรงตามความต้องการของผู้ใช้งานโปรแกรม

รายการประเมิน	\bar{x}	S.D.	ระดับคุณภาพ
1. ความสามารถในการตรวจสอบคำห้ามใช้ในชื่อการค้าและชื่อเครื่องสำอาง	4.90	0.32	มากที่สุด
2. ความสามารถในการตรวจสอบความสอดคล้องระหว่างชื่อการค้า/ชื่อเครื่องสำอางและส่วนประกอบในตำรับเครื่องสำอาง	4.70	0.48	มากที่สุด
3. โปรแกรมมีความรวดเร็วในการตอบสนองต่อผู้ใช้งาน	4.70	0.48	มากที่สุด
4. โปรแกรมช่วยลดขั้นตอนการทำงานที่เคยปฏิบัติ	4.90	0.32	มากที่สุด
5. โปรแกรมช่วยลดระยะเวลาในการทำงาน	4.80	0.42	มากที่สุด
ภาพรวม	4.80	0.40	มากที่สุด

ผลการประเมินความพึงพอใจของผู้ใช้โปรแกรมสำหรับการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอาง ด้านประสิทธิภาพของโปรแกรม รายละเอียดการประเมินแสดงในตารางที่ 19

ตารางที่ 21 ผลการประเมินความพึงพอใจของผู้ใช้งานด้านประสิทธิภาพของโปรแกรม

รายการประเมิน	\bar{x}	S.D.	ระดับคุณภาพ
1. ความถูกต้องของผลลัพธ์ที่ได้จากการประมวลผลในโปรแกรม	4.60	0.52	มากที่สุด
2. ความเร็วในการประมวลผลของโปรแกรม	4.80	0.42	มากที่สุด
3. โปรแกรมสามารถแจกแจงรายละเอียดของผลลัพธ์ได้อย่างครบถ้วน	4.60	0.52	มากที่สุด
ภาพรวม	4.67	0.49	มากที่สุด

ผลการประเมินความพึงพอใจของผู้ใช้โปรแกรมสำหรับการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอาง ด้านความสะดวกและง่ายต่อการใช้งาน รายละเอียดการประเมินแสดงในตารางที่ 20

ตารางที่ 22 ผลการประเมินความพึงพอใจของผู้ใช้งานด้านความสะดวกและง่ายต่อการใช้งาน

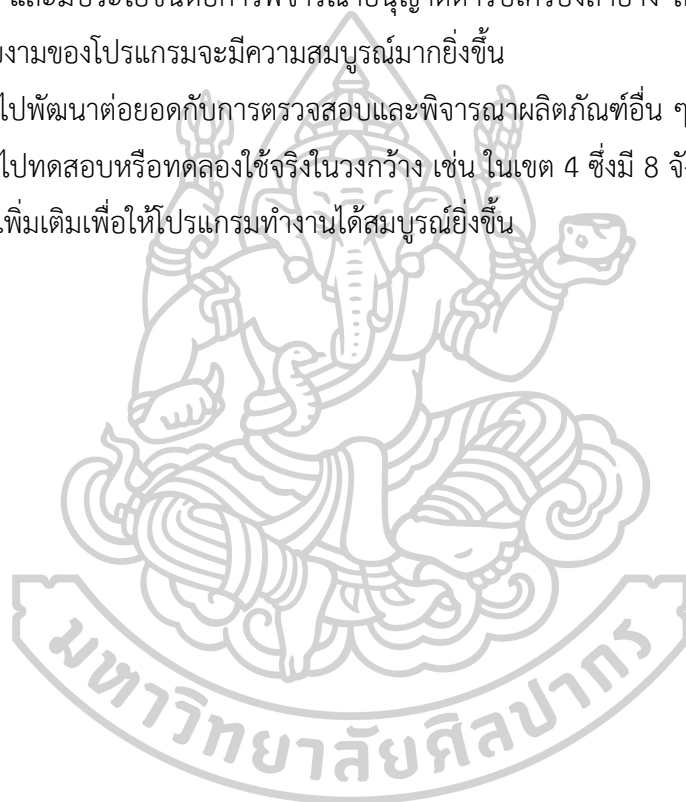
รายการประเมิน	\bar{x}	S.D.	ระดับคุณภาพ
1. การเข้าใช้โปรแกรมมีความสะดวก ง่ายต่อการใช้งาน	4.70	0.48	มากที่สุด
2. มีการออกแบบให้ใช้งานโปรแกรมได้ง่าย เมนูไม่ซับซ้อน	4.60	0.52	มากที่สุด
3. ความสวยงาม ความทันสมัย และน่าสนใจของหน้าโปรแกรม	4.30	0.48	มากที่สุด
4. ขนาดตัวอักษรและรูปแบบตัวอักษรอ่านได้ง่าย และสวยงาม	4.60	0.52	มากที่สุด
ภาพรวม	4.55	0.50	มากที่สุด

ตอนที่ 3 ข้อเสนอแนะอื่นๆ

ข้อเสนอแนะเกี่ยวกับโปรแกรมสำหรับการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอาง ด้วยวิธีการเปรียบเทียบสายอักขระ สรุปได้ดังนี้

1. อยากให้โปรแกรมสามารถดึงข้อมูลจากไฟล์ PDF ที่ผู้ประกอบการยื่นเข้าระบบ e-Submission มาที่โปรแกรมได้เลย โดยไม่ต้องคัดลอกข้อมูลมาวางอีกครั้ง เพื่อความสะดวกต่อการใช้งานมากขึ้นในอนาคต

2. อาจขยายการตรวจสอบข้อมูลเกี่ยวกับเครื่องสำอางไปยังส่วนที่ต้องพิจารณาอื่นๆ เช่น วัตถุประสงค์การใช้ บริเวณที่ใช้ ลักษณะทางกายภาพ ให้สอดคล้องกับชื่อการค้า ชื่อเครื่องสำอาง
3. หากมีการประสานการทำงานของโปรแกรมนี้เข้ากับระบบยื่นคำขออัตโนมัติ e-Submission ของสำนักงานคณะกรรมการอาหารและยา จะช่วยงานเจ้าหน้าที่ในการตรวจสอบ ดำรับได้ดียิ่งขึ้น
4. อยากให้มีการอัปเดตข้อมูลในฐานข้อมูลเพิ่มมากขึ้น
5. โปรแกรมสะดวกต่อการใช้งาน, ช่วยอำนวยความสะดวกต่อการทำงานของเจ้าหน้าที่ได้ เป็นอย่างมาก และมีประโยชน์ต่อการพิจารณาอนุญาตดำรับเครื่องสำอาง ถ้ามีการปรับปรุงเพิ่มเติม เรื่องความสวยงามของโปรแกรมจะมีความสมบูรณ์มากยิ่งขึ้น
6. สามารถนำไปพัฒนาต่อยอดกับการตรวจสอบและพิจารณาผลิตภัณฑ์อื่น ๆ ได้เช่น อาหาร และ น่าจะมีการนำไปทดสอบหรือทดลองใช้จริงในวงกว้าง เช่น ในเขต 4 ซึ่งมี 8 จังหวัด เพื่อความีประเด็น ใดที่ต้องแก้ไขเพิ่มเติมเพื่อให้โปรแกรมทำงานได้สมบูรณ์ยิ่งขึ้น



บทที่ 5

สรุป อภิปราย และข้อเสนอแนะ

จากการศึกษาวิจัยการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ มีรายละเอียดในการศึกษาวิจัยดังนี้

วัตถุประสงค์การวิจัย

1. เพื่อตรวจสอบคำห้ามใช้ในชื่อการค้าและชื่อเครื่องสำอางในแบบคำขอจดแจ้งเครื่องสำอาง
2. เพื่อตรวจสอบคำที่อาจนำมาใช้เป็นส่วนหนึ่งของชื่อการค้าและชื่อเครื่องสำอาง โดยคำนั้นต้องมีความสัมพันธ์สอดคล้องกับส่วนประกอบในตำรับเครื่องสำอาง

ประชากรและกลุ่มตัวอย่าง

1. ประชากรและกลุ่มตัวอย่างสำหรับการศึกษาประสิทธิภาพของโปรแกรม

ประชากร ได้แก่ แบบคำขอจดแจ้งเครื่องสำอาง (แบบ จ.ค.1) ที่ถูกส่งเข้ามาในระบบจดแจ้งเครื่องสำอางอัตโนมัติ e-Submission ของสำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา ตั้งแต่วันที่ 1 ตุลาคม พ.ศ. 2559 จนถึงวันที่ 30 พฤศจิกายน พ.ศ. 2564 เป็นแบบคำขอจดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าผ่าน จำนวน 1,590 รายการคำขอ และแบบคำขอจดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าไม่ผ่าน จำนวน 1,510 รายการคำขอ รวมทั้งสิ้น 3,100 รายการคำขอ

กลุ่มตัวอย่าง ได้แก่ แบบคำขอจดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าผ่าน จำนวน 200 รายการคำขอ และแบบคำขอจดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าไม่ผ่าน จำนวน 200 รายการคำขอ รวมทั้งสิ้น 400 รายการคำขอ ซึ่งทำการประเมินผลการตรวจสอบใหม่อีกครั้งโดยพนักงานเจ้าหน้าที่ผู้เชี่ยวชาญ จำนวน 3 ท่าน ประเมินอย่างเป็นอิสระต่อกัน หากผลการประเมินผ่านทั้ง 3 ท่าน จะถือว่าเป็นแบบคำขอจดแจ้งเครื่องสำอางที่พนักงานเจ้าหน้าที่ผู้เชี่ยวชาญพิจารณาว่าผ่าน

ในการเลือกกลุ่มตัวอย่างสำหรับการวิจัยครั้งนี้ใช้วิธีการสุ่มตัวอย่างแบบง่าย (simple random sampling: SRS)

2. ประชากรและกลุ่มตัวอย่างสำหรับการศึกษาความพึงพอใจของผู้ใช้งานที่มีต่อโปรแกรม

ประชากร ได้แก่ บุคลากรตำแหน่งเภสัชกรในกลุ่มงานคุ้มครองผู้บริโภคและเภสัชสาธารณสุข สำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา จำนวน 16 ท่าน

กลุ่มตัวอย่าง ได้แก่ บุคลากรตำแหน่งเภสัชกรในกลุ่มงานคุ้มครองผู้บริโภคและเภสัชสาธารณสุข สำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา จำนวน 10 ท่าน ที่ปฏิบัติงานหรือเคยปฏิบัติงานเกี่ยวข้องกับการตรวจสอบผลิตภัณฑ์เครื่องสำอางก่อนออกสู่ท้องตลาด หรือการตรวจสอบผลิตภัณฑ์เครื่องสำอางหลังออกสู่ท้องตลาด ในการเลือกกลุ่มตัวอย่างใช้วิธีการเลือกตัวอย่างแบบเจาะจง (purposive sampling)

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ ประกอบด้วย โปรแกรมสำหรับตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ และแบบประเมินความพึงพอใจของผู้ใช้งานโปรแกรม

1. ภาษาและเครื่องมือที่ใช้ในการพัฒนาโปรแกรม

1.1 ภาษาคอมพิวเตอร์

- 1) ภาษาพีเอชพี (PHP)
- 2) ภาษาไพธอน (Python)
- 3) ภาษาเอชทีเอ็มแอล (HyperText Markup Language: HTML)
- 4) ภาษาเอสคิวแอล (Structured Query Language: SQL)

1.2 โปรแกรมคอมพิวเตอร์

- 1) โปรแกรม Text-editor: Visual Studio code
- 2) โปรแกรมจัดการฐานข้อมูล MySQL: phpMyAdmin

1.3 เครื่องมืออื่น ๆ

- 1) เครื่องมือที่ใช้ในการตัดคำภาษาไทย PyThaiNLP v3.0.8
- 2) หลักเกณฑ์การพิจารณาการจดแจ้งเครื่องสำอาง (ฉบับปรับปรุง 2564)
- 3) ระบบจดแจ้งเครื่องสำอางอัตโนมัติ e-Submission

2. แบบประเมินความพึงพอใจของผู้ใช้งานโปรแกรม

สรุปผลการวิจัย

การศึกษาวิจัยการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ ผู้วิจัยขอสรุปผลการวิจัย ดังนี้

1. ผลการพัฒนาโปรแกรมสำหรับตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ

ผู้วิจัยได้ออกแบบโครงสร้างฐานข้อมูลด้วย entity relationship diagram (ER diagram) ซึ่งเป็นแบบจำลองความสัมพันธ์ระหว่างข้อมูล (entity relationship model) ทำการออกแบบและ

จัดทำฐานข้อมูลที่เกี่ยวข้องด้วยภาษา SQL ใช้ระบบจัดการฐานข้อมูล MySQL และใช้โปรแกรม phpMyAdmin เป็นเครื่องมือช่วยในการจัดการฐานข้อมูล พัฒนาโปรแกรมเชื่อมโยงกับฐานข้อมูล ด้วยภาษา PHP เป็นหลัก โดยพัฒนาร่วมกับภาษา HTML เพื่อช่วยให้โปรแกรมมีการตอบสนองกับ ผู้ใช้งานได้ดียิ่งขึ้น และสามารถเรียกใช้งานได้จากเว็บเบราว์เซอร์ได้ ในการเปรียบเทียบสายอักขระ ด้วยวิธีการหาระยะทางเลเวนชเตย์นพัฒนาด้วยภาษา PHP ส่วนขั้นตอนการตัดคำภาษาไทยใช้ Library PythaiNLP ซึ่งพัฒนาด้วยภาษา Python

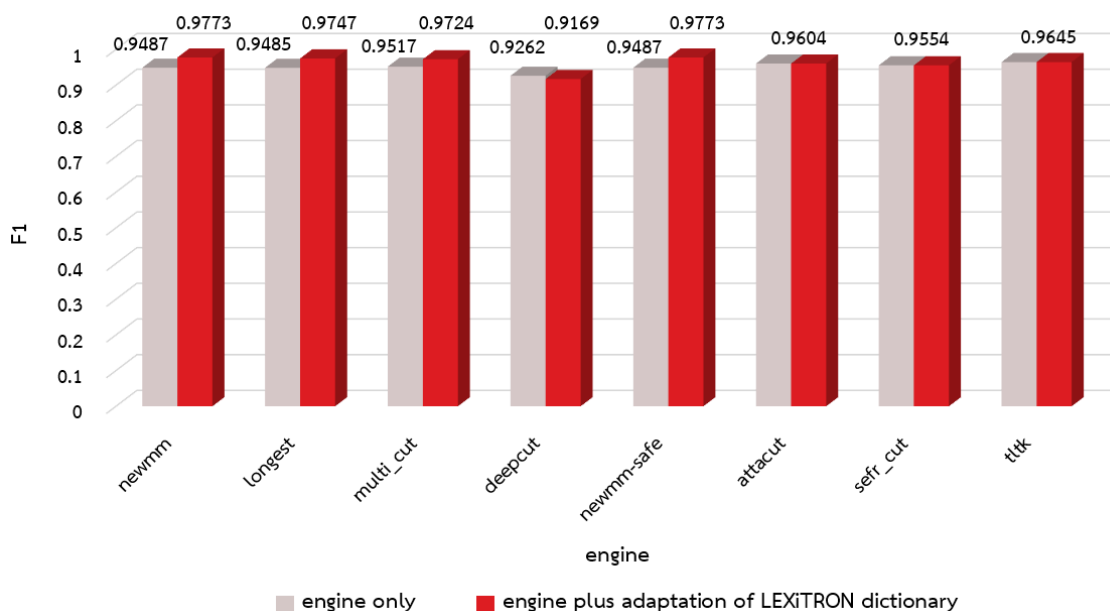
จากผลการศึกษาเมื่อใช้เทคนิคเอ็นแกรม (N-gram) ในการกำหนดคุณลักษณะของคำ สามารถสรุปได้ว่า คุณลักษณะของคำที่ทำให้การเปรียบเทียบสายอักขระให้มีประสิทธิภาพสูงสุดและ เหมาะสมกับการนำไปใช้งานวิจัยนี้ที่สุดคือ คุณลักษณะของคำแบบ trigram

จากผลการประเมินประสิทธิภาพของโปรแกรมในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ โดยใช้ เครื่องมือตัดคำภาษาไทยเพียงอย่างเดียว เมื่อพิจารณาจากค่า F1 พบว่าที่ค่า threshold 85 การใช้ เครื่องมือ attacut ในการตัดคำช่วยให้การทำงานของโปรแกรมมีประสิทธิภาพสูงสุด รองลงมาคือ tltk และ sefr_cut ตามลำดับ และที่ค่า threshold 90 การใช้เครื่องมือ tltk ในการตัดคำจะช่วยให้ การทำงานของโปรแกรมมีประสิทธิภาพสูงสุด รองลงมาคือ attacut และ sefr_cut ตามลำดับ

สำหรับการใช้เครื่องมือตัดคำต่าง ๆ ร่วมกับการใช้ adaptation of LEXITRON dictionary ในการตัดคำภาษาไทย เมื่อพิจารณาจากค่า F1 พบว่าที่ค่า threshold 85 และ 90 การใช้เครื่องมือ newmm และ newmm-safe ร่วมกับการใช้พจนานุกรม ช่วยในการทำงานของโปรแกรมมี ประสิทธิภาพสูงสุด รองลงมาคือ longest และ multi_cut ตามลำดับ และที่ค่า threshold 90 การใช้เครื่องมือ newmm และ newmm-safe ร่วมกับการใช้พจนานุกรม ช่วยในการทำงานของ โปรแกรมมีประสิทธิภาพสูงสุด รองลงมาคือ longest และ multi_cut ตามลำดับ

การเปรียบเทียบประสิทธิภาพของโปรแกรมในการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และ ส่วนประกอบในตำรับเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ ผลของการตัดคำภาษาไทย ระหว่างการใช้เครื่องมือตัดคำเพียงอย่างเดียว กับการใช้เครื่องมือตัดคำร่วมกับ adaptation of LEXITRON dictionary ที่ค่า threshold 90 พบว่าการใช้เครื่องมือตัดคำ newmm, longest, multi_cut และ newmm-safe ร่วมกับ adaptation of LEXITRON dictionary ช่วยเพิ่ม ประสิทธิภาพการทำงานของโปรแกรม โดยการใช้เครื่องมือตัดคำ newmm และ newmm-safe ร่วมกับ adaptation of LEXITRON dictionary ช่วยทำให้โปรแกรมทำงานได้มีประสิทธิภาพมากที่สุด ส่วนการใช้เครื่องมือตัดคำ attacut, sefr_cut และ tltk ร่วมกับ adaptation of LEXITRON

dictionary ไม่ได้มีผลทำให้ประสิทธิภาพการทำงานของโปรแกรมเปลี่ยนแปลงเมื่อเทียบกับการใช้เครื่องมือตัดคำเพียงอย่างเดียว



รูปที่ 16 ผลการเปรียบเทียบประสิทธิภาพของโปรแกรมระหว่างการใช้เครื่องมือตัดคำเพียงอย่างเดียวกับการใช้เครื่องมือตัดคำร่วมกับการใช้ adaptation of LEXiTRON dictionary ที่ threshold 90

สำหรับงานวิจัยนี้การกำหนดค่า threshold ที่ 90 และใช้เครื่องมือ newmm หรือ newmm-safe ร่วมกับการใช้ adaptation of LEXiTRON dictionary จะเหมาะสมกับการใช้งานช่วยให้การตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งและการพิจารณาคำขอจดแจ้งเครื่องสำอางมีประสิทธิภาพมากที่สุด

ผลการศึกษานี้ให้เห็นว่าโปรแกรมที่พัฒนาขึ้นสามารถตรวจสอบคำห้ามใช้และคำที่สามารถนำมาใช้เป็นชื่อผลิตภัณฑ์เครื่องสำอางได้โดยต้องมีสารที่สอดคล้องในสูตรตำรับได้ตามวัตถุประสงค์การประมวลผลการตรวจสอบมีความถูกต้องแม่นยำ เป็นไปในทิศทางเดียวกัน สามารถนำไปใช้เพื่อช่วยเจ้าหน้าที่ในการตรวจสอบข้อมูลในแบบคำขอจดแจ้งเครื่องสำอางเบื้องต้นได้ และเป็นแนวทางในการพัฒนาสำหรับใช้ในองค์กรอื่นๆต่อไป

2. ผลการประเมินความพึงพอใจของผู้ใช้งานที่มีต่อโปรแกรมสำหรับตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ

ผลการประเมินความพึงพอใจของผู้ใช้งานโปรแกรม โดยเภสัชกรในกลุ่มงานคุ้มครองผู้บริโภค และเภสัชสาธารณสุข สำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา จำนวน 10 ท่าน ที่ปฏิบัติงาน

หรือเคยปฏิบัติงานเกี่ยวข้องกับการตรวจสอบผลิตภัณฑ์เครื่องสำอางก่อนออกสู่ท้องตลาด หรือการตรวจสอบผลิตภัณฑ์เครื่องสำอางหลังออกสู่ท้องตลาด ใน 3 ด้าน ได้แก่ 1) ด้านตรงตามความต้องการของผู้ใช้งานโปรแกรม มีความพึงพอใจอยู่ในระดับ มากที่สุด มีค่าเฉลี่ยเท่ากับ 4.80 2) ด้านประสิทธิภาพของโปรแกรม มีความพึงพอใจอยู่ในระดับ มากที่สุด มีค่าเฉลี่ยเท่ากับ 4.67 และ 3) ด้านความสะดวกและง่ายต่อการใช้งาน ความพึงพอใจอยู่ในระดับ มากที่สุด มีค่าเฉลี่ยเท่ากับ 4.55 สรุปได้ว่า ความพึงพอใจของผู้ใช้โปรแกรมสำหรับการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับเครื่องสำอางซึ่งทั้ง 3 ด้านที่กล่าวมาอยู่ใน ระดับมากที่สุด

อภิปรายผล

จากการศึกษาเมื่อกำหนดค่า threshold ที่ 80, 85 และ 90 พบการเกิดผลลบลง (false negative) โดยมีสาเหตุการเกิดผลลบลง ได้แก่ 1) ผลการตัดคำภาษาไทยได้คำห้ามใช้ หรือคำที่ใกล้เคียงคำห้ามใช้ เช่น คำว่า “เลสเซอร์” กับ “เลเซอร์” ซึ่งมีความคล้ายคลึงกัน 87.5%, คำว่า “ดีเอส” เป็นคำทับศัพท์ของการนำด้วยภาษาอังกฤษมาใช้ เมื่อผ่านการตัดคำแล้วได้คำว่า “ดี” และ “เอส” ซึ่งคำว่า “ดี” เป็นคำห้ามใช้ในภาษาไทย 2) ผลการตัดคำภาษาไทยได้คำที่สามารถนำไปใช้ในชื่อการค้าและชื่อเครื่องสำอางได้โดยต้องมีสารที่สอดคล้องในสูตรตำรับ เช่น คำว่า “ซิกเนเจอร์” ผ่านขั้นตอนการตัดคำแล้วได้คำว่า “ซิก” และ “เนเจอร์” ซึ่งคำว่า “เนเจอร์” เป็นคำที่สามารถนำไปใช้ในชื่อการค้าและชื่อเครื่องสำอางได้โดยต้องมีสารที่สอดคล้องในสูตรตำรับ เมื่อไม่พบสารในตำรับ จึงทำให้ผลการตรวจสอบคลาดเคลื่อนจากการตรวจสอบเป็นไม่ผ่านการตรวจสอบ 3) คำภาษาอังกฤษที่เป็นคำพ้องรูปกับคำห้ามใช้ เช่น การใช้ข้อความ “EAU DE PARFUM” หรือ “EAU DE TOILETTE” เป็นชื่อการค้าหรือชื่อเครื่องสำอาง ซึ่งข้อความดังกล่าวสื่อถึงประเภทของน้ำหอม พนักงานเจ้าหน้าที่ได้ตรวจสอบข้อความดังกล่าวในเรื่องความสอดคล้องและความเหมาะสมในการใช้เป็นชื่อการค้าหรือชื่อเครื่องสำอาง และประเมินผลเป็นผ่านการตรวจสอบ แต่การตรวจสอบด้วยโปรแกรมในขั้นตอนการตัดคำจะได้คำว่า “DE” ซึ่งเป็นคำห้ามใช้ในภาษาอังกฤษ จึงทำให้โปรแกรมประเมินผลเป็นไม่ผ่านการตรวจสอบ

นอกจากนี้ยังพบการเกิดผลบวกลง (false positive) ซึ่งสาเหตุการเกิดผลบวกลง ได้แก่ การใช้คำที่สามารถนำไปใช้ในชื่อการค้าและชื่อเครื่องสำอางได้โดยต้องมีสารที่สอดคล้องในสูตรตำรับ ซึ่งคำภาษาไทยบางคำเขียนติดกันโดยไม่มีเว้นวรรค เช่น “ฮันนี่โกลด์” เมื่อพนักงานเจ้าหน้าที่ทำการตรวจสอบพบว่าข้อความดังกล่าวประกอบด้วยคำ จำนวน 2 คำ คือ “ฮันนี่” และ “โกลด์” ซึ่งทั้งสองคำเป็นคำที่สามารถใช้เป็นชื่อการค้าหรือชื่อเครื่องสำอางได้ โดยต้องมีสารที่สอดคล้องในสูตรตำรับ แต่ในสูตรตำรับไม่พบสารที่มีความสอดคล้องกับคำข้างต้น พนักงานเจ้าหน้าที่จึงประเมินผลการพิจารณาเป็นไม่ผ่านการตรวจสอบ แต่การตรวจสอบด้วยโปรแกรม การใช้เครื่องมือตัดคำภาษาไทยบางอย่าง

ไม่สามารถตัดคำภาษาไทยดังกล่าวได้ และได้นำคำว่า “ฮันนี่โกลด์” ไปเปรียบเทียบกับคำที่มีในฐานข้อมูล แต่ไม่พบว่ามีความคล้ายคลึงกับคำว่า “ฮันนี่” และ “โกลด์” เลย จึงไม่มีการไปเทียบหาสารที่เกี่ยวข้องในสูตรตำรับ ทำให้โปรแกรมแสดงผลการพิจารณาเป็นผ่านการตรวจสอบ

ในการกำหนดค่า threshold ความคล้ายคลึงกันของคำต้องเป็นค่าที่เหมาะสม หากกำหนดค่า threshold ที่สูงเกินไป ทำให้การเปรียบเทียบสายอักขระแบบประมาณ (approximate string matching) มีความใกล้เคียงกับการเปรียบเทียบสายอักขระที่เหมือนกันทุกประการ (exact string matching) ส่งผลให้ไม่ได้ประโยชน์จากการเปรียบเทียบสายอักขระโดยการหาระยะทางเลเวนชเตย์น ยกตัวอย่าง กรณีชื่อผลิตภัณฑ์เครื่องสำอางมีคำว่า “ALGOHOL” เป็นส่วนหนึ่งของชื่อ เมื่อเปรียบเทียบกับคำในฐานข้อมูลพบว่าคล้ายคลึงกับคำว่า “ALCOHOL” ซึ่งเป็นคำที่สามารถใช้เป็นชื่อผลิตภัณฑ์เครื่องสำอางได้โดยต้องมีสารที่เกี่ยวข้องในสูตรตำรับ ทั้งสองคำต่างกันที่การสะกดคำด้วยตัวอักษร G แทน C ซึ่งอาจเป็นการสะกดผิดโดยที่ไม่ได้ตั้งใจ หรือเป็นความตั้งใจของผู้ประกอบการเพื่อให้ชื่อผลิตภัณฑ์เครื่องสำอางมีความเป็นเอกลักษณ์ โดดเด่น แตกต่างจากผลิตภัณฑ์อื่น จากการเปรียบเทียบสายอักขระพบว่า ทั้งสองคำมีความยาวสายอักขระเท่ากับ 7 เมื่อเปรียบเทียบสายอักขระโดยการหาระยะทางเลเวนชเตย์น แทนที่ G ด้วย C 1 ตำแหน่ง ทั้งสองคำจะมีความคล้ายคลึงกันเท่ากับ 86% ซึ่งหากกำหนดค่า threshold ที่ 85 จะพบว่าทั้งสองคำมีความคล้ายคลึงกัน จากนั้นโปรแกรมจะไปทำการเทียบหาสารในตำรับที่เกี่ยวข้องกับชื่อ ALCOHOL ต่อไป แต่หากกำหนดค่า threshold ที่ 90 โปรแกรมจะไม่สามารถตรวจสอบได้ว่า ALGOHOL มีความคล้ายคลึงกับคำว่า ALCOHOL ในฐานข้อมูล ทำให้ไม่สามารถไปเทียบหาสารในตำรับที่เกี่ยวข้องกับชื่อ ALCOHOL ได้ ส่งผลให้เกิดความคลาดเคลื่อนของการพิจารณาค่าขอจดแจ้งเครื่องสำอางได้

ส่วนการกำหนดค่า threshold ที่ต่ำเกินไป ทำให้การเปรียบเทียบสายอักขระตรวจพบคำที่มีความคล้ายคลึงกันกับคำในฐานข้อมูลจำนวนมาก หากมีความคล้ายคลึงกับคำห้ามใช้ จะทำให้แสดงผลการพิจารณาเป็นไม่ผ่านการตรวจสอบ หรือหากพบว่ามีความคล้ายคลึงกับคำที่สามารถใช้เป็นชื่อผลิตภัณฑ์เครื่องสำอางได้ โดยต้องมีสารที่สอดคล้องในสูตรตำรับ โปรแกรมจะไปทำการเทียบหาสารในตำรับที่เกี่ยวข้องกับชื่อนั้นๆ หากไม่พบสารในตำรับ จะแสดงผลการพิจารณาเป็นไม่ผ่านการตรวจสอบ ซึ่งอาจทำให้ผลการพิจารณาค่าขอจดแจ้งเครื่องสำอางคลาดเคลื่อนได้เช่นกัน

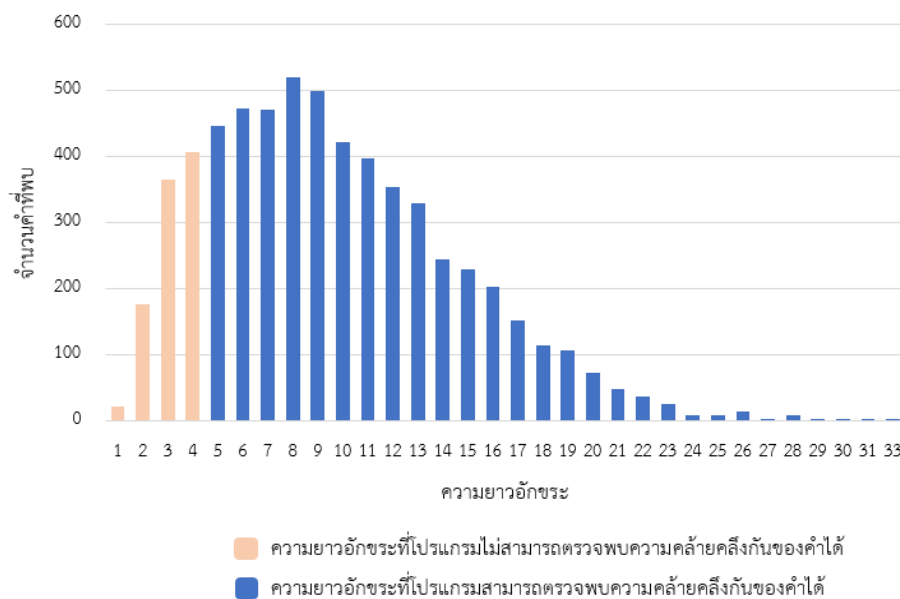
จากการเปรียบเทียบสายอักขระโดยการหาระยะทางเลเวนชเตย์น เมื่อทำการตัดออก แทรกหรือแทนที่อักขระ 1 ครั้ง จะมีค่าเท่ากับ 1 และเมื่อนำไปคำนวณหาความคล้ายคลึงกันของคำ ซึ่งจากสูตรการคำนวณจะเป็นการหาระหว่างระยะทางเลเวนชเตย์นต่อความยาวสูงสุดของสายอักขระใดสายอักขระหนึ่งจากสายอักขระสองชุด ดังนั้นหากคำมีความยาวของสายอักขระที่สั้น เมื่อมีการ

เปลี่ยนแปลงเพียง 1 ตำแหน่ง ในการเปรียบเทียบสายอักขระจะมีผลทำให้ความคล้ายคลึงกันของคำลดลงมากกว่าคำที่มีสายอักขระยาวกว่า

คำที่ใช้ในการเปรียบเทียบ	คำในฐานข้อมูล	จำนวนอักขระของสายอักขระที่ยาวที่สุด	ค่าความคล้ายคลึงกันของคำ (%)
W H I T E R N I N G	W H I T E N I N G	10	90
แ อ ล ก อ ฮ ย ร ๋	แ อ ล ก อ ฮ ย ล ๋	9	89
C O L L A G E L	C O L L A G E N	8	88
A L G O H O L	A L C O H O L	7	86
L O V E	C L O V E	5	80

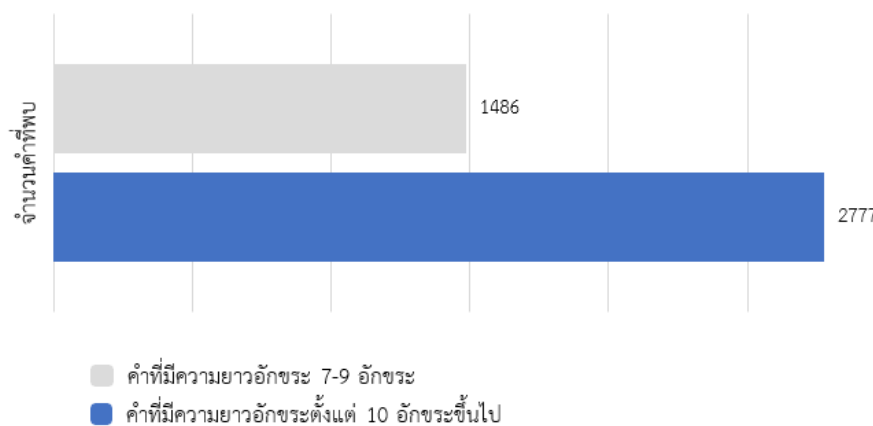
รูปที่ 17 ตัวอย่างการเปรียบเทียบสายอักขระและค่าความคล้ายคลึงกันของคำ

จะเห็นว่าปัจจัยที่มีผลต่อการเปรียบเทียบสายอักขระแบบประมาณอีกประการหนึ่งที่สำคัญคือ ความยาวสายอักขระ คำแต่ละคำมีความยาวสายอักขระที่แตกต่างกัน เมื่อพิจารณาความยาวอักขระของคำที่พบในงานวิจัยนี้ พบว่ามีความยาวอักขระตั้งแต่ 1 จนถึง 33 อักขระ และจากข้อมูลที่น่าเสนอข้างต้นแสดงให้เห็นว่าการกำหนดค่า threshold ที่ 80 คำที่มีความยาวในช่วง 5-6 อักขระเมื่อมีการเปลี่ยนแปลง 1 ตำแหน่ง โปรแกรมสามารถตรวจพบความคล้ายคลึงกันของคำได้ เช่นเดียวกับการกำหนดค่า threshold ที่ 85 และ 90 ส่งผลให้โปรแกรมสามารถตรวจพบความคล้ายคลึงกันของคำที่มีความยาวในช่วง 7-9 อักขระ และตั้งแต่ 10 อักขระขึ้นไป ตามลำดับ ส่วนคำที่มีความยาวอักขระในช่วง 1-4 อักขระ โปรแกรมไม่สามารถตรวจพบความคล้ายคลึงกันของคำได้



รูปที่ 18 ข้อมูลความยาวอักขระของคำที่พบในงานวิจัย

สำหรับงานวิจัยนี้คำที่พบส่วนใหญ่เป็นคำที่มีความยาวอักขระตั้งแต่ 10 อักขระขึ้นไป มีจำนวน 2,777 คำ ดังนั้นผลการเปรียบเทียบประสิทธิภาพของโปรแกรม จึงพบว่าที่ค่า threshold 90 โปรแกรมมีประสิทธิภาพสูงสุดในการเปรียบเทียบสายอักขระ โดยมีค่า F1 สูงสุดเมื่อเทียบกับที่ค่า threshold 85 และ 80



รูปที่ 19 ข้อมูลคำที่พบในงานวิจัย

แม้ว่าการตรวจสอบชื่อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางโดยใช้โปรแกรมฯ จะให้ผลลัพธ์ที่ดี แต่ยังไม่สามารถใช้ทดแทนเจ้าหน้าที่ได้อย่างสมบูรณ์ เนื่องจากอาจมีข้อมูลชื่อการค้า ชื่อเครื่องสำอาง หรือส่วนประกอบในตำรับที่เป็นข้อมูลใหม่หรือข้อมูลที่มีความซับซ้อน เช่น คำพ้องรูป คำพ้องเสียง โปรแกรมอาจตรวจสอบคลาดเคลื่อนได้ ซึ่งยังคงต้องให้พนักงานเจ้าหน้าที่ต้องพิจารณาซ้ำอีกครั้งหนึ่ง แต่โปรแกรมฯ สามารถชี้เป้าแบบคำขอจดแจ้งที่มีความเสี่ยงในการใช้ชื่อการค้าหรือชื่อเครื่องสำอางที่ไม่เป็นไปตามหลักเกณฑ์การพิจารณาการจดแจ้งเครื่องสำอาง ช่วยให้พนักงานเจ้าหน้าที่ตรวจสอบได้อย่างถูกต้องและรวดเร็วมากยิ่งขึ้น

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

คำห้ามใช้และคำที่สามารถใช้เป็นชื่อผลิตภัณฑ์เครื่องสำอางได้ในฐานข้อมูลส่วนหนึ่งมีรูปแบบเป็นคำทับศัพท์ภาษาอังกฤษ ซึ่งวิธีการสะกดคำทับศัพท์ภาษาอังกฤษไม่มีรูปแบบที่ตายตัว จึงควรมีนักภาษาศาสตร์เข้ามาช่วยในการตรวจสอบความถูกต้องของคำก่อนเก็บค่าลงในฐานข้อมูล เพื่อช่วยให้ผลการตรวจสอบและพิจารณาคำขอจดแจ้งเครื่องสำอางมีความสมบูรณ์มากยิ่งขึ้น

การศึกษาครั้งนี้เก็บรวบรวมแบบคำขอจดแจ้งเครื่องสำอางที่ยื่นเข้าระบบ e-Submission ของจังหวัดพระนครศรีอยุธยาเพียงแห่งเดียว และมีขนาดตัวอย่างแบบคำขอจดแจ้งเครื่องสำอางน้อย ซึ่งอาจมีความหลากหลายไม่เพียงพอที่จะเป็นตัวแทนของประชากรแบบคำขอจดแจ้งเครื่องสำอางทั้งประเทศได้ ดังนั้นก่อนนำโปรแกรมไปใช้จึงควรมีการทดสอบในตัวอย่างแบบคำขอจดแจ้งเครื่องสำอาง

ในจำนวนที่มากขึ้นและทดสอบในวงกว้าง ซึ่งอาจพบประเด็นที่ต้องปรับปรุงเพิ่มเติมหรือทำให้ทราบประสิทธิภาพที่แท้จริงของโปรแกรมได้

2. ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

เพื่อเพิ่มความสะดวกในการใช้งานและช่วยให้โปรแกรมทำงานได้อย่างมีประสิทธิภาพมากขึ้น ครอบคลุมการเปรียบเทียบสายอักขระของคำที่มีความยาวอักขระที่แตกต่างกัน ควรปรับรูปแบบของโปรแกรมให้สามารถปรับเปลี่ยนค่า threshold ความคล้ายคลึงกันของคำได้โดยง่าย หรือให้มีการตรวจสอบความยาวของคำก่อน จากนั้นจึงเลือกค่า threshold ที่เหมาะสมเพื่อทำการเปรียบเทียบสายอักขระต่อไป



รายการอ้างอิง



1. มาร์เก็ตเธียร์. (2565). **ตลาดความงาม 2565 มีแวกลับมารุ่ง Beauty Tech คือตัวหนุน กรณีศึกษาอริอัล**. เข้าถึงเมื่อ 21 กันยายน. เข้าถึงได้จาก <https://marketeeronline.co/archives/266777>.
2. กรมเจรจาการค้าระหว่างประเทศ. (2564). **เครื่องสำอางไทยฮอต คนรักสวยรักงามแห่ซื้อปด้นไทยขึ้นแทนส่งออกอันดับ 2 เอเชีย อันดับ 10 ของโลก เน้นใช้ FTA เพิ่มโอกาสขยายตลาด**. เข้าถึงเมื่อ 21 กันยายน. เข้าถึงได้จาก <https://www.dtn.go.th/th/news/categories/5cff753c1ac9ee073b7bd1c5>.
3. "พระราชบัญญัติเครื่องสำอาง พ.ศ. 2558." (2558). ราชกิจจานุเบกษา. เล่ม 132, ตอนที่ 86 ก (8 กันยายน): 5-25.
4. เกษรา จันทรวงศ์ไพศาล และชิตชนก เรือนก้อน. (2564). "สถานการณ์การรับจดแจ้งเครื่องสำอางด้วยระบบ e-Submission และการกำกับดูแลของสำนักงานคณะกรรมการอาหารและยา ปงบประมาณ พ.ศ. 2559-2562." **วารสารอาหารและยา** 28, 1 (มกราคม-เมษายน): 60-71.
5. สำนักงานคณะกรรมการอาหารและยา. สำนักควบคุมเครื่องสำอางและวัตถุอันตราย. (2559). "คู่มือจดแจ้งเครื่องสำอางระบบอัตโนมัติ." สิงหาคม.
6. สำนักงานคณะกรรมการอาหารและยา. สำนักควบคุมเครื่องสำอางและวัตถุอันตราย. (2564). "หลักเกณฑ์การพิจารณาการจดแจ้งเครื่องสำอาง (ฉบับปรับปรุง 2564)." 2 ธันวาคม.
7. สรรพโชค สิงหสุวรรณ. (2565). **Data Warehouse ต่างจาก Database อย่างไร**. เข้าถึงเมื่อ 20 เมษายน. เข้าถึงได้จาก https://race.nstru.ac.th/home_ex/blog/topic/show/6160.
8. กิตติ ภัคดีวัฒนกุล และทวีศักดิ์ กาญจนสุวรรณ. (2544). **สร้างระบบสารสนเทศบนเว็บด้วย FrontPage 2002**. กรุงเทพฯ: เคทีพี คอมพ์แอนด์ คอนซัลท์.
9. อากาศ ธาตุโลหะ. (2547). **ทรัพยากรสารสนเทศเพื่อการค้นคว้า**. ชลบุรี: พี.เค. กราฟฟิคพริ้นต์.
10. โอภาส เอี่ยมสิริวงศ์. (2558). **ระบบฐานข้อมูล Database Systems**. กรุงเทพฯ: ซีเอ็ดยูเคชั่น.
11. เพียรทิพย์ ศรีสุวรรณ. (2564). "แบบจำลองข้อมูล (Data Model)." เอกสารประกอบการสอน สาขาเทคโนโลยีสารสนเทศและการสื่อสาร คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา

12. มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา. กลุ่มภารกิจบริการและพัฒนาสื่อการเรียนการสอน สำนักคอมพิวเตอร์. **ความสัมพันธ์ (Relationships) ของระบบฐานข้อมูล**. เข้าถึงเมื่อ 20 เมษายน. เข้าถึงได้จาก https://cst.tsu.ac.th/courseonline/course/computer_it/lesson1/lesson1.html
13. บัณฑิต อูราธรรมกุล และ กานดา รุณนะพงศา. (2549). "การตัดคำภาษาไทยด้วยวิธีปรับปรุงกฎและพจนานุกรมแบบใหม่." **JCSSE 2006**; 34-40.
14. ศิริวรรณ ชัยวิฑูรณกุล, คำรัส วงศ์สว่าง และชมทิพย์ พรพนมชัย. (2546). "การปรับปรุงวิธีการค้นหาสายอักขระโดยใช้วิธีการตัดแบ่งและการทำแฮชชิ่ง." **วิทยานิพนธ์ปริญญาโท สาขาวิชาวิทยาการคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยมหิดล**.
15. Gomaa, Wael H., and Aly A. Fahmy. (2013). "A survey of text similarity approaches." **international journal of Computer Applications** 68, 13 (April): 13-18.
16. นวรัตน์ คำเสียง. (2554). "การพัฒนาเครื่องมือเพื่อการวิเคราะห์ความสนใจร่วมกันในงานวิจัยสำหรับนักวิจัย." **วิทยานิพนธ์ปริญญาโท สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย**.
17. Levenshtein, Vladimir I. (1996). "Binary codes capable of correcting deletions, insertions, and reversals." **Soviet physics doklady** 10, 8 (February); 707-710.
18. Behara, Krishna N.S., Ashish Bhaskar, and Edward Chung. (2020). "A novel approach for the structural comparison of origin-destination matrices: Levenshtein distance." **Transportation Research Part C: Emerging Technologies** 111, (February): 513-530.
19. Christanti, Viny M., and Dali S. Naga. (2018). "Fast and accurate spelling correction using trie and Damerau-Levenshtein distance bigram." **Telkonnika** 16, 2 (April): 827-833.
20. Yulianto, Muhamad M., Riza Arifudin, and Alamsyah Alamsyah. (2018). "Autocomplete and spell checking Levenshtein distance algorithm to getting Text Suggest Error Data Searching in Library." **Scientific Journal of Informatics** 5, 1 (May): 67-75.
21. Haldar, Rishin., and Debajyoti Mukhopadhyay. (2011). "Levenshtein distance technique in dictionary lookup methods: An improved approach." **arXiv preprint arXiv:1101.1232**.

22. Nagowah, Leckraj. et al. (2016). "Automatically Fixing Syntax Errors Using the Levenshtein Distance." **ADBU Journal of Engineering Technology** 5, 2 (December).
23. จักรินทร์ สันติรัตนภักดี และศุภกฤษฎี นวัตกรรมกุล. (2564). "การออกแบบและพัฒนาระบบการจำแนกข้อร้องเรียนรถโดยสารสาธารณะเพื่อติดแท็กปัญหาการให้บริการ". **JEDT** 9, 2 (July-December): 77-91.
24. จิตติมา ผลึกเพชร. (2560). "การประยุกต์ใช้อัลกอริทึมเลเวนชเตย์นเพื่อหาชื่อพ้องรูปและชื่อคล้ายของพืชสมุนไพร." วิทยานิพนธ์ปริญญาโทมหาบัณฑิต สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร.
25. นิคม ถนอมเสียง. (2561). "การคำนวณขนาดตัวอย่าง Sample Size Determination." เอกสารประกอบการสอน สาขาวิชาวิทยาการระบาดและชีวสถิติ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น.
26. Aiumsiriwong, O. (2012). **System analysis and design**. Bangkok: SE-EDUCATION Public Company Limited.
27. Supaartagorn, C. (2018). **Web application development with PHP and MySQL+AJAX jQuery**. 2nd ed. Bangkok: Simplify Press.
28. พิชิต วิจิตรบุญรักษ์. (2554). **HMTL: ภาษาเขียนเว็บ**. เข้าถึงเมื่อ 10 มีนาคม 2565. เข้าถึงได้จาก https://www.bu.ac.th/knowledgecenter/executive_journal/july_sep_11/pdf/aw32.pdf
29. วรณพงษ์ ภัททิย์ไพบูลย์ และคณะ. (2565). **PyThaiNLP/pythainlp: PyThaiNLP v3.0.8 Released!**. เข้าถึงเมื่อ 12 มีนาคม 2565. เข้าถึงได้จาก <https://zenodo.org/record/6554863#.Yx6xTXZBxPY>
30. วิกานดา ผาพันธ์ และอัญชญา พิมพ์พิศาล. (2563). "การพยากรณ์ทิศทางของราคาหุ้นรายวันจากข้อความข่าวภาษาไทย โดยใช้วิธีการประมวลผลภาษาธรรมชาติ." **วารสารวิทยาศาสตร์ประยุกต์** 19, 1 (June): 59-79.
31. ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ. (2538). **พจนานุกรมอิเล็กทรอนิกส์ ไทย-อังกฤษ เล็กจิตรอน**. เข้าถึงเมื่อ 15 มีนาคม 2565. เข้าถึงได้จาก <http://www.sansarn.com/lexto/license-lexitron.php>
32. Parikh, Nish., Gyanit Singh, and Neel Sundaresan. (2013). "Query suggestion with large scale data." **Handbook of Statistics** 31, 493-518.

33. Kumar, Gulshan. (2014). "Evaluation metrics for intrusion detection systems-a study." **Evaluation** 2, 11 (November): 11-7.
34. บุญชม ศรีสะอาด. (2556). **การวิจัยเบื้องต้น ฉบับปรับปรุงใหม่**. พิมพ์ครั้งที่ 9. กรุงเทพฯ: สุวีริยาสาส์น.
35. อัครพล เอกวงศ์อนันต์. (2548). "การระบุค่าไทยและค่าทับศัพท์ด้วยแบบจำลองเอ็นแกรม." วิทยานิพนธ์ปริญญาโทมหาบัณฑิต สาขาวิชาอักษรศาสตร์ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย





ภาคผนวก ก

ข้อมูลคำที่ห้ามนำมาใช้เป็นชื่อหรือส่วนหนึ่งของชื่อการค้าหรือชื่อเครื่องสำอาง

ลำดับ	ชื่อห้ามใช้	ภาษา	ลำดับ	ชื่อห้ามใช้	ภาษา
1	ACHE	en	26	ANTI FUNGUS	en
2	ACNE CONTROL	en	27	ANTI GRAVITY	en
3	ACNE OUT	en	28	ANTI INFLAME	en
4	ACNE THERAPY	en	29	ANTI INFLAMMATION	en
5	ACNE TREATMENT	en	30	ANTI PIGMENT	en
6	A-CURE	en	31	ANTI PIGMENTATION	en
7	ACUTE	en	32	ANTI RASH	en
8	ADRENALIN	en	33	ANTI SUNBURN	en
9	ADRENALINE	en	34	ANTI SWEAT	en
10	AFTERBITE	en	35	ANTIBALD	en
11	AGAINST GRAVITY	en	36	ANTI-BODY	en
12	ALCOHOL BALL	en	37	ANTI-CHAFE	en
13	ALCOHOL PAD	en	38	ANTI-CHAFING	en
14	ALLERGIC	en	39	ANTI-DOTE	en
15	ALLERGY	en	40	ANTI-GEN	en
16	ANALGESIC	en	41	ANTI-MELANIN	en
17	ANDROGEN	en	42	ANTISEPTIC	en
18	ANESTHETIC	en	43	ANTI-SWEAT	en
19	ANTI ALLERGIC	en	44	AROCA	en
20	ANTI ALLERGY	en	45	AROKA	en
21	ANTI ANALGESIC	en	46	ATOCONTROL	en
22	ANTI BUG	en	47	ATOPIC	en
23	ANTI FRECKLE	en	48	ATOPIC CONTROL	en
24	ANTI FUNGAL	en	49	ATOPY	en
25	ANTI FUNGI	en	50	AYURVEDA	en

ลำดับ	ชื่อห้ามใช้	ภาษา	ลำดับ	ชื่อห้ามใช้	ภาษา
51	AYURVEDIC	en	78	CANDIDIASIS	en
52	B TOX	en	79	CELL BRIGHT	en
53	BFGF	en	80	CELL DETOX	en
54	BIG BOOB	en	81	CELL LIFT	en
55	BASIC FIBROBLAST GROWTH FACTOR	en	82	CELLULITE CONTROL	en
56	BITE RELIEF	en	83	CELLULITE FIGHTER	en
57	BOTOX	en	84	CHAFE	en
58	BOTOXIA	en	85	CHAFED	en
59	BOTOXIN	en	86	CHAFFING	en
60	BOTOX-LIKE	en	87	CHAFING	en
61	BOTOX-LIKED	en	88	CORONA VIRUS	en
62	BREAST ENHANCER	en	89	COVID	en
63	BREAST ENLARGEMENT	en	90	COVID-19	en
64	BTOX	en	91	DIFFLAM	en
65	B-TOX	en	92	DOG	en
66	BTX	en	93	DRY MOUTH	en
67	BUG	en	94	DYSMENORRHEA	en
68	BUG CONTROL	en	95	ECZEMA	en
69	BUGGIE	en	96	EDEMA	en
70	BUGGY	en	97	EF-G	en
71	BUG'N	en	98	EGF	en
72	BUGS	en	99	EGFR	en
73	BUGS AWAY	en	100	EJECTION	en
74	BURN RELIEF	en	101	ELECTROPHORESIS	en
75	BUST ENHANCER	en	102	ENLARGE	en
76	BUST UP	en	103	EPIDERMAL GROWTH FACTOR	en
77	CALLUS	en	104	ESTROGEN	en

ลำดับ	ชื่อห้ามใช้	ภาษา	ลำดับ	ชื่อห้ามใช้	ภาษา
105	FACE OFF	en	132	HORMONE	en
106	FACE TRANSPLANT	en	133	HYPOALLERGENIC	en
107	FAT REDUCE	en	134	IGE	en
108	FAT REDUCING	en	135	IGF-1	en
109	FIT	en	136	IMMUNE	en
110	FRECKLE	en	137	IMMUNIZATION	en
111	FUNGAL	en	138	IMMUNIZE	en
112	FUNGI	en	139	IMMUNOLOGY	en
113	FUNGUS	en	140	IMPLANT	en
114	GALVANIC	en	141	INCREASE BLOOD CIRCULATION	en
115	GARGLE	en	142	INFECT	en
116	GERM	en	143	INFLAMMATORY	en
117	GF	en	144	INFLAME	en
118	GINGIVITIS	en	145	INFLAMMATION	en
119	GMP	en	146	INFLAMMATORY	en
120	GROW	en	147	INHALE	en
121	GROWING	en	148	INHALER	en
122	GROWTH	en	149	INHIBIT	en
123	GROWTH FACTOR	en	150	INHIBIT ENZYME	en
124	GUASA	en	151	INJECT	en
125	GYNECOLOGICAL	en	152	INJECTION	en
126	HAIR GROW	en	153	INSECT	en
127	HAIR GROWTH	en	154	INSECT REPELLANT	en
128	HAIR ROOT	en	155	INSULIN-LIKE GROWTH FACTOR-1	en
129	HAIRY	en	156	INTENSE PULSE LIGHT	en
130	HEALER	en	157	IONTO	en
131	HIFU	en	158	IONTOPHORESIS	en

ลำดับ	ชื่อห้ามใช้	ภาษา	ลำดับ	ชื่อห้ามใช้	ภาษา
159	IPL	en	187	MICROORGANISM	en
160	IRRITANT	en	188	MORPHINE	en
161	IRRITATE	en	189	MOSQUI	en
162	IRRITATION	en	190	MOSQUITO	en
163	JOINT LUBRICANT	en	191	MUSCLE	en
164	LASER	en	192	NAIL GROWTH	en
165	LAZER	en	193	NAPPY	en
166	LICE	en	194	NAPPY RASH	en
167	LIPOFAT	en	195	NARCOTIC	en
168	LIPOLYSIS	en	196	NICOTIANA	en
169	LOUSE	en	197	NORMAL SALINE	en
170	LUBRICANT	en	198	NORMAL SALINE SOLUTION	en
171	MEDICAL	en	199	NSS	en
172	MEDICATE	en	200	OEDEMA	en
173	MEDICATED	en	201	OPIUM	en
174	MEDICATION	en	202	ORTHO	en
175	MESO	en	203	OSTEOARTHRITIS	en
176	MESO CARE	en	204	PAIN	en
177	MESO THERAPY	en	205	PAPAVER	en
178	MESO WHITE	en	206	PENIS	en
179	MESOFAT	en	207	PET	en
180	MESOHAIR	en	208	PHARMACOPOEIA	en
181	METABOLISM	en	209	PHEROMONE	en
182	MEZO CARE	en	210	PHYTO STEM CELL	en
183	MEZO THERAPY	en	211	PHYTOESTROGEN	en
184	MEZOCARE	en	212	PIGMENT REDUCING	en
185	MEZOTHERAPY	en	213	PLANT STEM CELL	en
186	MICRO NEEDLE	en	214	PLASMA	en

ลำดับ	ชื่อห้ามใช้	ภาษา	ลำดับ	ชื่อห้ามใช้	ภาษา
215	PREDNISOLONE	en	241	STEM CELL	en
216	PROGESTERONE	en	242	STEROID	en
217	Pseudomonas aeruginosa	en	243	STIMULATE	en
218	PSORIASIS	en	244	STIMULATE HAIR GROWTH	en
219	RASH	en	245	STIMULATING	en
220	REDNESS RELIEF	en	246	STIMULATING HAIR GROWTH	en
221	REDUCE PIGMENT	en	247	STIMULATION	en
222	REFIT	en	248	STIMULATOR	en
223	REMODELER	en	249	STING FREE	en
224	RESTRUCTURE	en	250	STOP ACNE	en
225	RHEUMATOID	en	251	STOP AGE	en
226	RHEUMATOID ARTHRITIS	en	252	STOP AGING	en
227	ROOT HAIR	en	253	STOP HAIR FALL	en
228	ROOT TREATMENT	en	254	SUN BURN	en
229	RUKSA	en	255	SYNDROME	en
230	SCAR ERASER	en	256	TDD	en
231	SCAR REMOVER	en	257	TDS	en
232	SCAR TREATMENT	en	258	TESTOSTERONE	en
233	SEDUCE	en	259	THERAPEUTIC	en
234	SEX	en	260	THROAT	en
235	SKIN BLEACH	en	261	THYROID	en
236	SKIN BLEACHING	en	262	TOXIC	en
237	SLENDER	en	263	TOXIN	en
238	SLIMMING	en	264	TRANS DERMAL	en
239	SPERM	en	265	TRANSDERMAL DRUG DELIVERY	en
240	SPICULE	en	266	TRANSDERMAL SYSTEM	en

ลำดับ	ชื่อห้ามใช้	ภาษา	ลำดับ	ชื่อห้ามใช้	ภาษา
267	TRANSDERMAL DRUG SYSTEM	en	295	เสริมทรวงอก	th
268	VAGINA	en	296	เสริมนม	th
269	VAGINITIS	en	297	เสริมหน้าอก	th
270	VARGINAL	en	298	เสริมอาหาร	th
271	VEIN	en	299	เห็ดขี้ควาย	th
272	VET	en	300	เอ็นยึด	th
273	VETERINARY	en	301	แก้แพ้	th
274	VIBRATE	en	302	แก้กระ	th
275	VIBRATER	en	303	แก้คัน	th
276	VIBRATING	en	304	แก้ฝ้า	th
277	VIRUS	en	305	แก้สิว	th
278	VULVITIS	en	306	แช่ผ้า	th
279	WART	en	307	แต้มสิ่ว	th
280	WRINKLE ERASER	en	308	แผล	th
281	แผลเป็น	th	309	แผลเป็น	th
282	เกลื้อน	th	310	แมลง	th
283	เคลียร์กระ	th	311	ไดอะเบติก	th
284	เคลียร์ฝ้า	th	312	ไทรอยด์	th
285	เคลียร์สิ่ว	th	313	ไล่	th
286	เจลกระ	th	314	ไล่แมลง	th
287	เจลฝ้า	th	315	ไล่ยุง	th
288	เจลสิ่ว	th	316	ไวรัส	th
289	เบาหวาน	th	317	กระ	th
290	เมโส	th	318	กระตุ้น	th
291	เลเซอร์	th	319	กระฝ้า	th
292	เสถียร	th	320	กันผมร่วง	th
293	เสถียรเซลล์	th	321	กันยุง	th
294	เส้นเลือด	th	322	กัวซา	th

ลำดับ	ชื่อห้ามใช้	ภาษา	ลำดับ	ชื่อห้ามใช้	ภาษา
323	ขจัดกระ	th	352	น้ำมันยา	th
324	ขจัดฝ้า	th	353	บรรเทา	th
325	ขจัดสิว	th	354	บรรเทาการอักเสบ	th
326	ขยี้กระ	th	355	บรรเทาอักเสบ	th
327	ขยี้ฝ้า	th	356	บรรเทาอาการอักเสบ	th
328	ขยี้สิ่ว	th	357	บำบัด	th
329	ข้อเข้า	th	358	ปลุกผม	th
330	ข้อเสื่อม	th	359	ปลุกหนวด	th
331	ขัดกระ	th	360	ปวดข้อ	th
332	ขัดฝ้า	th	361	ผมร่วง	th
333	ขัดสิ่ว	th	362	ผิวขาว	th
334	ครีมกระ	th	363	ฝ้า	th
335	ครีมฝ้า	th	364	ผื่น	th
336	ครีมสิ่ว	th	365	พญานู	th
337	คลายกล้ามเนื้อ	th	366	พลังแรด	th
338	คลายปวด	th	367	พลังช้าง	th
339	ฆ่าเชื้อ	th	368	พลังม้า	th
340	ฆ่าแบคทีเรีย	th	369	พิษ	th
341	ฆ่าไวรัส	th	370	พิษงู	th
342	ซีพจร	th	371	พิษผึ้ง	th
343	ดูดกระ	th	372	ฟันปลอม	th
344	ดูดฝ้า	th	373	ฟาร์มาโคเปีย	th
345	ดูดสิ่ว	th	374	พีโรโมน	th
346	ตกกระ	th	375	ยับยั้ง	th
347	ต้านพิษ	th	376	ยา	th
348	ทาสิว	th	377	ยาดม	th
349	น้ำมันเขียว	th	378	ยานวด	th
350	น้ำมันเหลือง	th	379	ยาหม่อง	th
351	น้ำมันกฤษา	th	380	รอยแผล	th

ลำดับ	ชื่อห้ามใช้	ภาษา	ลำดับ	ชื่อห้ามใช้	ภาษา
381	ระงับเห็ือ	th	410	ลูกปะคบ	th
382	รักษา	th	411	ว่านดอกทอง	th
383	รักษากระ	th	412	ว่านมหาเสน่ห์	th
384	รักษาฝ้า	th	413	ว่านรุกราคะ	th
385	รักษาสิว	th	414	ว่านสาวหลง	th
386	รุกราคะ	th	415	สเต็ม	th
387	ราคะ	th	416	สเต็มเซลล์	th
388	ลดเซลลูโลส	th	417	สเปิร์ม	th
389	ลดกระ	th	418	สลาย	th
390	ลดกระฝ้า	th	419	สลายไขมัน	th
391	ลดการอักเสบ	th	420	สลายเซลลูโลส	th
392	ลดบวม	th	421	สลายกระ	th
393	ลดปวด	th	422	สลายผิวสัมผัส	th
394	ลดฝ้า	th	423	สลายฝ้า	th
395	ลดฝ้ากระ	th	424	สาวหลง	th
396	ลดรอยกระฝ้า	th	425	สิิว	th
397	ลดสิิว	th	426	สิิวฝ้า	th
398	ลดสิิวเสี้ยน	th	427	สีสั๊ก	th
399	ลดอักเสบ	th	428	หญัาริแพร่	th
400	ลดอาการ	th	429	หน้าขาว	th
401	ลบกระ	th	430	หยุดสิิว	th
402	ลบฝ้า	th	431	หล่อลื่น	th
403	ลบรอยแผลเป็น	th	432	หลัง	th
404	ลบสิิว	th	433	อักเสบ	th
405	ลอก	th	434	อายุรเวท	th
406	ลอกกระ	th	435	อาหาร	th
407	ลอกฝ้า	th	436	อาหารเสริม	th
408	ลอกสิิว	th	437	ฮิยู่่ม	th
409	ลูกประคบ	th	438	น้ำมนต์	th

ลำดับ	ชื่อห้ามใช้	ภาษา	ลำดับ	ชื่อห้ามใช้	ภาษา
439	REMEDIES	en	468	ANTICOVID	en
440	MEDICI	en	469	ANTICOVID-19	en
441	ALCOHOL 70%	en	470	ANTICO	en
442	ALCOHOL 80%	en	471	พุทธร	th
443	ALCOHOL 90%	en	472	พุทธิ	th
444	ACNE CLEAR	en	473	ครีมพุงสวย	th
445	AFTER EFFECT	en	474	ขาวเฟือก	th
446	BURN	en	475	เซรั่มหน้าเด็ก	th
447	DOODEE	en	476	ดูดี	th
448	FOLLICLE	en	477	ต้องมนต์	th
449	G1	en	478	ทองคำ	th
450	IMPULSE	en	479	เนื้อทอง	th
451	LONG HAIR	en	480	แอลกอฮอล์70%	th
452	PRE CARE	en	481	แอลกอฮอล์80%	th
453	RE HAIR	en	482	แอลกอฮอล์90%	th
454	RECOVERY CARE	en	483	อโรคา	th
455	RELAXING	en	484	หม่อง	th
456	RICHMONEY	en	485	น้ำมันหม่อง	th
457	SPEED WHITE	en			
458	TENSOR	en			
459	D	en			
460	DE	en			
461	SAFE	en			
462	BUTTERFLY MELASMA	en			
463	AGE RETURN	en			
464	CELL	en			
465	SUN BLOCK	en			
466	UV BLOCK	en			
467	WATER PROOF	en			

ภาคผนวก ข

คำที่สามารถนำมาใช้เป็นส่วนหนึ่งของชื่อผลิตภัณฑ์เครื่องสำอางได้ โดยคำนั้นจะต้องมีความสอดคล้องกับส่วนประกอบในตำรับของเครื่องสำอาง

ลำดับ	ชื่อกลุ่ม	คำที่สามารถใช้แทนได้
1	CAVIAR	คาเวียร์
2	PEARL	ไข่มุก
3	GINSENG	โสม
5	VITAMIN C	VIT C, วิตามินซี, วิตามินซี
7	GOLD	โกลด์
8	PURE GOLD	24K GOLD
9	WHITE	WHITENING, BRIGHTENING, LIGHTENING, BRIGHT, ไวท์เทนนิ่ง, ไบรท์เทนนิ่ง, ไลท์เทนนิ่ง, ไบรท์, ไวท์
10	ACNE	ANTI ACNE, BLACKHEAD
11	ANTI-CELLULITE	FIRMING
12	FIRM	FIRMING, UPLIFTING, LIFTING, LIFT
15	PORE TIGHTENING	PORE MINIMIZER, PORE MINIMIZER, PORELESS
17	SUNSCREEN	UV SUNSCREEN, UV PROTECTOR, UV PROTECTION, SUN PROTECTION, กันแดด
20	ANTIDANDRUFF	ขจัดรังแค
27	TEA TREE OIL	TEA TREE, ที ทรี ออยล์
28	MILK	MILK PROTEIN
29	RICE	RICE MILK, JASMINE RICE, BLACK JASMINE RICE, น้่านมข้าว, ข้าวหอมมะลิ, ข้าวหอมนิล, ข้าว
30	LEMONGRASS	ตะไคร้
31	VITAMIN E	VIT E, วิตามินอี, วิตามินอี
32	COCONUT OIL	COCONUT, น้ำมันมะพร้าว
33	CARROT	แครอท
34	COLLAGEN	คอลลาเจน
35	RETINOL	เรตินอล

ลำดับ	ชื่อกลุ่ม	คำที่สามารถใช้แทนได้
36	ALBUMIN	EGG PROTEIN
37	ALOE VERA	ALOE, ว่านหางจระเข้
39	TURMERIC	ขมิ้น, ขมิ้นชัน
40	GLUTATHIONE	กลูต้าไธโอน
41	HYALURON	HYALURONIC, HYALURONIC ACID, ไฮยาลูรอน
42	THANAKA	ทานาคา
43	HONEY	BEE, น้ำผึ้ง
44	BIRD'S NEST	BIRD NEST, รังนก
45	SILK	ไหม
46	KAFFIR LIME	BERGAMOT, มะกรูด
47	YOGHURT	YOGURT, โยเกิร์ต
48	WILD TURMERIC	ว่านนางคำ
49	GARCINIA	ส้มแขก
51	MAKAMPOM	มะขามป้อม
52	SNAIL	สเนล
57	CUCUMBER	แตงกวา
58	LAKOOCHA	มะหาด
59	TAMARIND	มะขาม
60	CORDYCEPS	ถั่งเช่า
61	POMEGRANATE	ทับทิม
62	PUMPKIN	ฟักทอง
64	GUAVA LEAF	ใบฝรั่ง
65	CLOVE	กานพลู
67	PERILLA OIL	Perilla Seed Oil, น้ำมันเมล็ดงาอ่อน
76	VETIVER	หญ้าแฝก
77	CARDAMOM	กระวาน
78	LINGZHI	เห็ดหลินจือ
79	GINGER	ขิง
81	LEMON	เลมอน

ลำดับ	ชื่อกลุ่ม	คำที่สามารถใช้แทนได้
82	PASSION FRUIT	เสาวรส
83	MARIGOLD	ดาวเรือง
89	MICHELIA	จำปี
90	CENTELLA ASIATICA	CENTELLA, GOTU KOLA, ใบบัวบก
91	KARIYAT	ฟ้าทะลายโจร
92	RADISH	หัวไชเท้า
93	PLAI	ไพล
95	GOJI BERRY	GOJI
96	SAFFLOWER	ดอกคำฝอย
98	BUTTERFLY PEA	อัญชัน
101	THANTAWAN	ทานตะวัน
105	LESSER GALANGAL	กระชาย
107	APPLE	แอปเปิ้ล
108	SUGAR	น้ำตาล
109	MORINGA OIL	น้ำมันมะรุม
112	COFFEE	กาแฟ
116	CHILLI	พริก
117	SANDALWOOD	ไม้จันทน์
119	SESAME	งาดำ
122	SACHA INCHI OIL	น้ำมันถั่วอินคา, น้ำมันถั่วดาวอินคา
123	NONI	ยอ
127	YLANG YLANG	กระดังงา
129	MANGOSTEEN	MANGOSTEEN PEEL, เปลือกมังคุด, มังคุด
130	GRAPE SEED	เมล็ดองุ่น, สารสกัดเมล็ดองุ่น
131	ARBUTIN	ALPHA ARBUTIN
134	ALCOHOL	แอลกอฮอล์
136	RICE BRAN OIL	น้ำมันรำข้าว
138	HERB	HERBAL, PLANT, NATURE, NATURAL, เฮอร์เบิล, แพลนท์, เนเจอร์, เนเชอรัล

ลำดับ	ชื่อกลุ่ม	คำที่สามารถใช้แทนได้
141	PEACH	PEACHY
146	ROSELLE	กระเจี๊ยบแดง
147	DAFFODIL	ดอกดารารัตน์
152	ROSE	กุหลาบ



ภาคผนวก ค

ชื่อสารป้องกันความเสียหายที่สามารถใช้แทนที่สารที่เป็นส่วนประกอบในตำรับเครื่องสำอางได้

ลำดับ	ชื่อสาร	ชื่อสารป้องกันความเสียหาย
1	NIACINAMIDE	NICOTINAMIDE
2	ETHYL ALCOHOL	ETHANOL
3	GLYCERIN	GLYCEROL
4	HAMAMELIS VIRGINIANA BARK EXTRACT	WITCH HAZEL BARK EXTRACT
5	AMAMELIS VIRGINIANA LEAF EXTRACT	WITCH HAZEL LEAF EXTRACT



ภาคผนวก ง



แบบประเมินความพึงพอใจ

การใช้งานโปรแกรมสำหรับตรวจสอบซื้อการค้า ชื่อเครื่องสำอาง และส่วนประกอบ
ในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ

คำชี้แจง

แบบประเมินนี้มีวัตถุประสงค์เพื่อศึกษาเกี่ยวกับความพึงพอใจของเจ้าหน้าที่ผู้ใช้งานที่มีต่อโปรแกรมสำหรับตรวจสอบซื้อการค้า ชื่อเครื่องสำอาง และส่วนประกอบในตำรับในแบบคำขอจดแจ้งเครื่องสำอางด้วยวิธีการเปรียบเทียบสายอักขระ จึงใคร่ขอให้เจ้าหน้าที่ผู้ใช้งานตอบแบบประเมินนี้ทุกข้อและตอบตามความเป็นจริง เพื่อประโยชน์ในการนำผลประเมินไปใช้ในการพัฒนาโปรแกรมในด้านต่างๆที่เป็นประโยชน์ต่อไป

แบบประเมินแบ่งออกเป็น 3 ส่วน คือ

ส่วนที่ 1 ข้อมูลทั่วไป

ส่วนที่ 2 แบบสอบถามความพึงพอใจของผู้ใช้งานโปรแกรมฯ

ส่วนที่ 3 ข้อเสนอแนะอื่นๆ

โดยใช้แบบสอบถามความพึงพอใจตามแบบมาตราส่วนประมาณค่า (Rating Scale) ของลิเคิร์ท (Likert Scale) และคำถามปลายเปิดสำหรับข้อเสนอแนะเพิ่มเติมของผู้ใช้งานโปรแกรม

การประเมินความพึงพอใจใช้แบบสอบถามในรูปแบบ Likert Scale ซึ่งเป็นเครื่องมือในการวัดเชิงคุณภาพโดยวัดเป็นระดับต่างๆจากสูงถึงต่ำซึ่งในการวิจัยนี้จะใช้ 5 ระดับ โดยกำหนดคะแนนเป็น 1 – 5 คะแนน ดังนั้นช่วงความกว้างของแต่ละชั้นเป็น 0.8 ซึ่งได้จากการคำนวณโดยใช้สูตร

$$\text{ความกว้างของอันตรภาคชั้น} = (\text{คะแนนสูงสุด} - \text{คะแนนต่ำสุด}) / \text{จำนวนชั้น}$$

$$\text{ความกว้างของอันตรภาคชั้น} = (5 - 1) / 5 = 0.8$$

ทำให้ได้เกณฑ์การแปลความหมายดังนี้

4.21 – 5.00 ระดับมากที่สุด

3.41 – 4.20 ระดับมาก

2.61 – 3.40 ระดับปานกลาง

1.81 – 2.60 ระดับน้อย

1.00 – 1.80 ระดับน้อยที่สุด

ส่วนที่ 1 ข้อมูลทั่วไป

โปรดทำเครื่องหมาย ✓ ลงใน ที่อยู่หน้าข้อความที่เกี่ยวข้องกับผู้ประเมิน

- | | | | |
|------------------|--|--------------------------------------|-----------------------------------|
| 1. เพศ | <input type="checkbox"/> ชาย | <input type="checkbox"/> หญิง | |
| 2. อายุ | <input type="checkbox"/> 25 ปี หรือต่ำกว่า | <input type="checkbox"/> 26-30 ปี | <input type="checkbox"/> 31-35 ปี |
| | <input type="checkbox"/> 36-40 ปี | <input type="checkbox"/> 41 ปีขึ้นไป | |
| 3. ระดับการศึกษา | <input type="checkbox"/> ต่ำกว่าปริญญาตรี | <input type="checkbox"/> ปริญญาตรี | |
| | <input type="checkbox"/> ปริญญาโท | <input type="checkbox"/> ปริญญาเอก | |

ส่วนที่ 2 แบบสอบถามความพึงพอใจของผู้ใช้งานโปรแกรมฯ

เป็นแบบประเมินความพึงพอใจในการใช้งานโปรแกรมฯ ซึ่งจะประเมินทั้งหมด 3 ด้าน ดังนี้

1. ด้านตรงตามความต้องการของผู้ใช้งานโปรแกรม
2. ด้านประสิทธิภาพของโปรแกรม
3. ด้านความสะดวกและง่ายต่อการใช้งาน

โปรดทำเครื่องหมาย ✓ ลงในช่องที่ตรงกับความเป็นจริงมากที่สุด

ระดับการประเมินความพึงพอใจ ระดับ 5 หมายถึง มากที่สุด

ระดับ 4 หมายถึง มาก

ระดับ 3 หมายถึง ปานกลาง

ระดับ 2 หมายถึง น้อย

ระดับ 1 หมายถึง น้อยที่สุด

รายการประเมิน		ระดับความพึงพอใจ				
		5	4	3	2	1
ด้านตรงตามความต้องการของผู้ใช้งานโปรแกรม						
1	ความสามารถในการตรวจสอบคำห้ามใช้ในชื่อการค้า และชื่อเครื่องสำอาง					
2	ความสามารถในการตรวจสอบความสอดคล้องระหว่างชื่อการค้า/ ชื่อเครื่องสำอางและส่วนประกอบในตำรับเครื่องสำอาง					
3	โปรแกรมมีความรวดเร็วในการตอบสนองต่อผู้ใช้งาน					
4	โปรแกรมช่วยลดขั้นตอนการทำงานที่เคยปฏิบัติ					
5	โปรแกรมช่วยลดระยะเวลาในการทำงาน					

ด้านประสิทธิภาพของโปรแกรม						
6	ความถูกต้องของผลลัพธ์ที่ได้จากการประมวลผลในโปรแกรม					
7	ความเร็วในการประมวลผลของโปรแกรม					
8	โปรแกรมสามารถแจกแจงรายละเอียดของผลลัพธ์ได้อย่างครบถ้วน					
ด้านความสะดวกและง่ายต่อการใช้งาน						
9	การเข้าใช้โปรแกรมมีความสะดวก ง่ายต่อการใช้งาน					
10	มีการออกแบบให้ใช้งานโปรแกรมได้ง่าย เมนูไม่ซับซ้อน					
11	ความสวยงาม ความทันสมัย และน่าสนใจของหน้าโปรแกรม					
12	ขนาดตัวอักษรและรูปแบบตัวอักษรอ่านได้ง่ายและสวยงาม					

ส่วนที่ 3 ข้อเสนอแนะอื่นๆ

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

นางสาวจิราภรณ์ วงษ์เลิศ
หัวหน้าโครงการวิจัย

ประวัติผู้เขียน

ชื่อ-สกุล	นางสาวจิราภรณ์ วงษ์เลิศ
วัน เดือน ปี เกิด	10 มีนาคม 2536
สถานที่เกิด	กรุงเทพมหานคร
วุฒิการศึกษา	ปริญญาเภสัชศาสตรบัณฑิต สาขาวิชาการบริหารทางเภสัชกรรม
ที่อยู่ปัจจุบัน	สำนักงานสาธารณสุขจังหวัดพระนครศรีอยุธยา เลขที่ 130 ถนนอุทุมทอง ตำบลหอรตันไชย อำเภอพระนครศรีอยุธยา จังหวัดพระนครศรีอยุธยา

