Journal of Applied Statistics and Information Technology

Vol 8 | No 1 | 2023

(January - June 2023)

NIDA WISDOM for Change
คณะสถิติประยุกต์
Graduate School of Applied Statistics

ISSN : 2465-4949 (Online)

บทบรรณาธิการ

วารสารสถิติประยุกต์และเทคโนโลยีสารสนเทศ เป็นวารสารวิชาการที่จัดทำโดยคณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์ มีจุดประสงค์เพื่อส่งเสริมงานวิจัยในด้านสถิติประยุกต์ เทคโนโลยีสารสนเทศและศาสตร์ที่เกี่ยวข้อง โดยการเผยแพร่ผลงานวิจัยที่มีคุณภาพจากนักวิชาการ นักวิจัยหรือนิสิตนักศึกษาโดยทั่วไปในรูปของวารสารแบบออนไลน์ สำหรับฉบับที่ 1 ปี 2566 ประกอบด้วยบทความวิจัยจำนวน 3 บทความ กองบรรณาธิการวารสารสถิติประยุกต์และเทคโนโลยีสารสนเทศหวังเป็นอย่างยิ่งว่าบทความวิจัยที่ตีพิมพ์ในวารสารฉบับนี้จะเป็นประโยชน์ต่อผู้อ่านที่จะนำไปประยุกต์ใช้หรือทำวิจัยต่อไป

กองบรรณาธิการ

รองศาสตราจารย์ ดร.สุรพงค์ เอื้อวัฒนามงคล **บรรณาธิการ**

คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

ผู้ช่วยศาสตราจารย์ ดร. ปรีชา วิจิตรธรรมรส **รองบรรณาธิการ**

คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

ศาสตราจารย์ ดร.สำรวม จงเจริญ **กรรมการ**

คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

ศาสตราจารย์ ดร. ชิดชนก เหลือสินทรัพย์ **กรรมการ**

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา **กรรมการ**

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

รองศาสตราจารย์ ดร.เยาวดี เต็มธนาภัทร์ **กรรมการ**

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

รองศาสตราจารย์ ดร. ฐิติวดี ชัยวัฒน์ **กรรมการ**

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

รองศาสตราจารย์ ดร.วราภรณ์ จิรชีพพัฒนา **กรรมการ**

คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

คณะกรรมการบริหารวารสาร

ผู้ช่วยศาสตราจารย์ ดร. ปราโมทย์ ลือนาม	ประธานคณะกรรมการ
อาจารย์ ดร.ศิวิกา ดุษฎีโหนด	กรรมการ
ผู้ช่วยศาสตราจารย์ ภัทราวดี ธนวงศ์สุวรรณ	กรรมการ
รองศาสตราจารย์ ดร.สุรพงศ์ เอื้อวัฒนามงคล	กรรมการ
ผู้ช่วยศาสตราจารย์ ดร.ธนาสัย สุคนธ์พันธุ์	กรรมการ
อาจารย์ ดร.ธนชาตย์ ฤทธิ์บำรุง	กรรมการ
นางสุภาพร หงษาวงษ์	กรรมการและเลขานุการ

สารบัญ

บทความวิจัย **หน้า**

Influent Factors for the failure of IT Project Implementation
And the Engineering Mindset of Project Managers
Kittisak Umaji and Worapat Paireekreng 1-19

The Power Garima-Rayleigh distribution: properties and applications
Issaraporn Thiamsorn, Ekapak Tanprayoon and Sirinapa Aryuyuen 20-31

Business Failure Prediction by Hybrid Data Mining Approach:
A Case of Thailand Agribusiness
Jeerawadee Pumjaroen 35-48

ข้อความที่ปรากฏในบทความที่ตีพิมพ์ในวารสารวิชาการฉบับนี้ถือเป็นความคิดเห็นส่วนบุคคลของผู้เขียนแต่ละท่าน ความผิดพลาดของข้อความและผลที่อาจเกิดจากนำข้อความเหล่านั้นไปใช้ผู้เขียนบทความจะเป็นผู้รับผิดชอบแต่เพียงผู้เดียว

ปัจจัยที่มีผลต่อความล้มเหลวในการดำเนินโครงการทางด้านเทคโนโลยีสารสนเทศและกรอบแนวคิดเชิงวิศวกรรมของผู้จัดการโครงการ

กิตติศักดิ์ อุมาจิ[1,*] และ วรภัทร ไพรีเกรง[2]

Received: 11 October 2022; Revised: 3 January 2023; Accepted: 3 January 2023

บทคัดย่อ

จากแนวโน้มของการลงทุนทางด้านเทคโนโลยีสารสนเทศที่สูงขึ้นในช่วงปีที่ผ่านมาทำให้เกิดโครงการทางด้านสารสนเทศขึ้นเป็นจำนวนมาก แต่จากสถิติพบว่ามีมากถึง 2 ใน 3 ของที่ไม่สามารถดำเนินการตามแผนงานที่วางไว้ได้ การวิจัยครั้งนี้มีจึงวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีผลต่อความล้มเหลวของการดำเนินการโครงการทางด้านเทคโนโลยีสารสนเทศ และศึกษาความสัมพันธ์ระหว่างผลลัพธ์ในการดำเนินโครงการทางด้านเทคโนโลยีสารสนเทศกับ กรอบแนวคิดเชิงวิศวกรรมของผู้จัดการโครงการ รวมไปถึงการศึกษาปัจจัยที่ส่งผลต่อการพัฒนากรอบแนวคิดเชิงวิศวกรรมของผู้จัดการโครงการ โดยใช้กลุ่มตัวอย่างผู้จัดการโครงการทางด้านเทคโนโลยีสารสนเทศในประเทศไทยจำนวน 214 คน โดยใช้การเลือกกลุ่มตัวอย่างแบบเจาะจง เครื่องมือที่ใช้ในการวิจัย ได้แก่ แบบสอบถามที่ผ่านกระบวนการ IOC และ สถิติที่ใช้ในการวิจัย ได้แก่ ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน โดยใช้ค่าสัมประสิทธิ์อัลฟ่าครอนบาชในการทดสอบความเชื่อมั่นของแบบสอบถาม

ผลการวิจัยพบว่าการบริหารขอบเขตโครงการและการบริหารการสื่อสารในโครงการเป็น 2 กลุ่มหลักที่ส่งผลต่อความล้มเหลวของโครงการ โดยมีเปอร์เซ็นต์รวมกันถึง 58.76% นอกจากนี้ยังพบว่า กรอบแนวคิดเชิงวิศวกรรมของผู้จัดการโครงการ ซึ่งประกอบไปด้วยแนวคิดในการปรับตัว แนวคิดในการเติบโต และ แนวคิดในการเรียนรู้ เป็นหนึ่งในทักษะที่สำคัญที่ส่งผลต่อความสำเร็จหรือล้มเหลวในการดำเนินโครงการ และประสบการณ์โดยตรงในการบริหารโครงการของผู้จัดการโครงการน่าจะเป็นหนึ่งในปัจจัยสำคัญที่มีส่วนในการพัฒนาแนวคิดเชิงวิศวกรรมสำหรับผู้จัดการโครงการ

คำสำคัญ: กรอบแนวคิดเชิงวิศวกรรม, ผู้จัดการโครงการ, เทคโนโลยีสารสนเทศ, ความล้มเหลวของโครงการ

[*] *Corresponding Author Email : 637191110004@dpu.ac.th*

[1,2] วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์

Influent Factors for the failure of IT Project Implementation And the Engineering Mindset of Project Managers

Kittisak Umaji[1,*] and Worapat Paireekreng [2]

Received: 11 October 2022; Revised: 3 January 2023; Accepted: 3 January 2023

Abstract

Due to the trend of higher investment in information technology in the recent years. As a result, there are many information technology projects occurred. However, the statistics show that more than two-thirds of the projects were unable to implement based on plan. The purposes of this research were to study the factors affecting the failure of the IT project implementation. Also, to study the relationship between the IT project implementation outcomes and the Engineering Mindset skills of project managers. These includes the study of factors affecting the development of the Project Manager's Engineering Mindset. The samples of this survey were chosen from 214 project managers who worked in Thailand and involved in information technology projects by specific purpose technique. In the study, a questionnaire that passed the IOC reviewing process was used to survey data. The statistics used in the research were the mean average method and standard deviation (S.D.) In addition, Cronbach's alpha coefficient was used to test the confidence of the questionnaire.

The research findings showed that project scope management and project communication management were the two main factors that affected project failure with a combined percentage of 58.76%. The Engineering Mindset of project managers, including Adaptive Mindset, Growth Mindset and Learning Mindset, is one of the key factors that affect the success or failure of a project. It was also found that direct experience in project management is likely to be one of the key factors in the development of an engineering mindset for project managers.

Keywords: Engineering Mindset, Project Manager, Information Technology, Project Failure

[*] *Corresponding Author Email : 637191110004@dpu.ac.th*

[1,2] College of Innovative Technology and Engineering, Dhurakij Pundit University

1. บทนำ

1.1 ที่มาและความสำคัญของงานวิจัย

เทคโนโลยีสารสนเทศเป็นปัจจัยที่เข้ามามีบทบาทอย่างมากในโลกยุคปัจจุบัน ทั้งในแง่ของการอำนวยความสะดวกในชีวิตประจำวันและการแข่งขันเพื่อสร้างโอกาสทางธุรกิจ ดังนั้นการดำเนินโครงการทางด้านเทคโนโลยีสารสนเทศจึงเป็นสิ่งสำคัญในการขับเคลื่อนอนาคต จากการสำรวจของ Gartner (2022) พบว่าการลงทุนทางด้านเทคโนโลยีของบริษัทขนาดใหญ่ทั่วโลกในปี พ.ศ. 2565 เพิ่มขึ้นจากปี พ.ศ. 2564 ประมาณ 132,688 ล้านดอลล่าสหรัฐ และมีแนวโน้มที่จะลงทุนเพิ่มขึ้นอีกประมาณ 274,929 ล้านดอลล่าสหรัฐในปี พ.ศ. 2566 ซึ่งเป็นสิ่งที่แสดงให้เห็นว่าโครงการทางด้านเทคโนโลยีสารสนเทศเป็นสิ่งที่ได้รับความสำคัญและมีแนวโน้มที่จะเกิดขึ้นอีกเป็นจำนวนมาก แต่จากผลการสำรวจของ The Standish Group (2021) พบว่าผลการดำเนินโครงการทางด้านเทคโนโลยีสารสนเทศทั่วโลกในปี พ.ศ.2564 มีอัตราความสำเร็จตรงตามแผนงานและงบประมาณที่วางไว้เพียง 31% เท่านั้น ซึ่งเป็นเพียง 1 ใน 3 ของการทำโครงการทั้งหมด อีก 69% หรือประมาณ 2 ใน 3 คือโครงการที่ล้มเหลว หรือต้องใช้เวลาและทรัพยากรมากกว่าที่ได้วางแผนไว้ และจากการศึกษาถึงสาเหตุของความล้มเหลวของโครงการ พบว่าทุกสาเหตุมีความเกี่ยวข้องกับการบริหารจัดการโครงการทั้งทางตรงและทางอ้อม ดังนั้นการที่สามารถบริหารโครงการได้อย่างมีประสิทธิภาพจึงเป็นปัจจัยสำคัญที่ช่วยให้โครงการประสบความสำเร็จและบรรลุตามวัตถุประสงค์ของโครงการที่วางไว้

จากที่ได้กล่าวมาข้างต้น ผู้จัดการโครงการ (Project Manager) จึงนับเป็นบุคคลที่มีผลโดยตรงต่อความสำเร็จหรือล้มเหลวของโครงการ ความสามารถในการจัดการ วางแผน ควบคุม รวมถึงทักษะในการแก้ปัญหาของผู้จัดการโครงการจึงเป็นปัจจัยสำคัญที่ส่งผลต่อผลลัพธ์ของโครงการ

กรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) เป็นหนึ่งทักษะที่ได้รับการยอมรับอย่างกว้างขวางในการแก้ปัญหาว่าเป็นกระบวนการคิดที่สามารถวางแผนงานและแก้ปัญหาได้อย่างมีประสิทธิภาพ ดังนั้น หากผู้จัดการโครงการมีทักษะในการบริหารโครงการและมีกรอบแนวคิดเชิงวิศวกรรมที่ดี ควรจะส่งผลต่อความสำเร็จของโครงการในอัตราที่สูงขึ้น สามารถวางแผนงาน แก้ไขปัญหาที่พบในโครงการได้อย่างมีประสิทธิภาพและเป็นระบบ ในขณะเดียวกันผู้จัดการโครงการที่ขาดกรอบแนวคิดเชิงวิศวกรรม อาจจะส่งผลให้ประสบปัญหาในการดำเนินโครงการ การแก้ปัญหาซึ่งอาจจะส่งผลให้โครงการเกิดความล้มเหลวได้ จากสมมุติฐานนี้ จึงได้ทำการศึกษาถึงความสัมพันธ์ระหว่างผลลัพธ์ในการดำเนินโครงการกับกรอบแนวคิดเชิงวิศวกรรมของผู้จัดการโครงการ โดยแสดงให้เห็นถึงความแตกต่างของกรอบแนวคิดเชิงวิศวกรรมในผู้จัดการโครงการที่มีผลลัพธ์อัตราความสำเร็จของโครงการแตกต่างกัน รวมถึงศึกษาปัจจัยอื่นๆ ที่อาจส่งผลต่อการพัฒนาความสามารถและทักษะของกรอบแนวคิดเชิงวิศวกรรม

1.2 วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีผลต่อความล้มเหลวของการดำเนินการโครงการทางด้านเทคโนโลยีสารสนเทศ และศึกษาความสัมพันธ์ระหว่างผลลัพธ์ในการดำเนินโครงการทางด้านเทคโนโลยีสารสนเทศกับกรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) ของผู้จัดการโครงการ รวมไปถึงการศึกษาปัจจัยที่ส่งผลต่อการพัฒนากรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) ของผู้จัดการโครงการ นอกจากนี้ผลจากการศึกษายังสามารถนำไปประยุกต์ใช้ในการคัดเลือกหรือมอบหมายงานให้กับผู้จัดการโครงการเพื่อเพิ่มอัตราผลสำเร็จของโครงการทางด้านเทคโนโลยีสารสนเทศต่อไป

2. งานวิจัยที่เกี่ยวข้อง

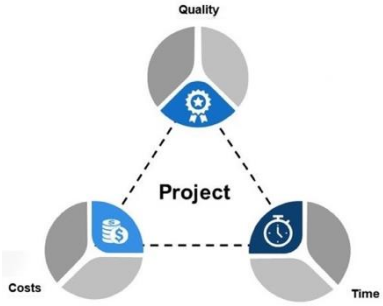
2.1 โครงการเทคโนโลยีสารสนเทศ (Information Technology Project)

โครงการเทคโนโลยีสารสนเทศ (Information Technology Project) (ADHIATMA และ LEGOWO, 2018) เป็นโครงการประเภทหนึ่งที่เกี่ยวข้องกับโครงสร้างพื้นฐานด้านระบบสารสนเทศหรือระบบคอมพิวเตอร์ ตัวอย่างของ

โครงการทางด้านสารสนเทศ ได้แก่ การพัฒนาเว็บ การพัฒนาซอฟต์แวร์ การพัฒนาแอพพลิเคชั่นบนอุปกรณ์พกพา การกำหนดค่าเครือข่าย การใช้ซอฟต์แวร์ การติดตั้งฮาร์ดแวร์ การจัดการฐานข้อมูล รวมถึงระบบสำรองข้อมูลและกู้คืนภัยพิบัติ (Backup and Disaster Recovery)

Tohidi, H. (2011) ได้ทำการศึกษาเกี่ยวกับการดำเนินโครงการทางด้านเทคโนโลยีสารสนเทศ พบว่า เทคโนโลยีเป็นพื้นฐานของปัจจัยความสำเร็จขององค์กรที่ช่วยเพิ่มศักยภาพการแข่งขัน เพิ่มประสิทธิภาพและประสิทธิผลของการทำงาน อีกทั้งยังเป็นการเพิ่มความท้ายทายในการทำงานให้กับบุคลากรในองค์กรอีกด้วย

เนื่องจากโครงการเทคโนโลยีสารสนเทศถือเป็นโครงการประเภทหนึ่ง จึงเป็นกิจกรรมกิจกรรมที่เกี่ยวข้องกับการใช้ทรัพยากรต่างๆ เพื่อนำมาลงทุนสร้างผลงานที่ก่อให้เกิดประโยชน์ต่อกลุ่มเป้าหมาย โดยกิจกรรมดังกล่าวจะต้องเป็นหน่วยอิสระที่สามารถทำการวิเคราะห์ วางแผน และบริหารได้ นอกจากนั้นจะต้องมีวัตถุประสงค์ที่ชัดเจน มีกำหนดเวลาเริ่มต้นและสิ้นสุดที่แน่ชัด การดำเนินงานจะต้องอยู่ภายใต้งบประมาณที่ได้ตั้งไว้ และได้ผลงานที่มีคุณภาพตามเกณฑ์ที่กำหนด (Akampurira, 2013)



รูปภาพที่ 1 ความสัมพันธ์ขององค์ประกอบที่มีผลต่อการดำเนินโครงการ

2.2 องค์ประกอบการบริหารโครงการ

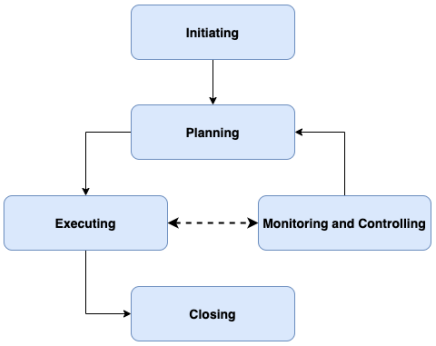
ผู้จัดการโครงการ (Project Manager) (PMI, 2022) คือ บุคคลที่ได้รับมอบหมายให้เป็นรับผิดชอบในการบรรลุวัตถุประสงค์ของโครงการ โดยการตั้งเป้าหมายของโครงการและใช้ทักษะและความเชี่ยวชาญเพื่อสร้างแรงบันดาลใจให้รู้สึกถึงจุดประสงค์ร่วมกันของสมาชิกโครงการ ที่มีหน้าที่รับผิดชอบในการบริหารจัดการ วางแผนงาน มอบหมายงาน ผลักดันควบคุม ติดตามผลและแก้ไขปัญหาต่างๆในโครงการ เป็นผู้ที่เห็นภาพรวมทั้งหมดของโครงการ รวมถึงเป็นผู้รับผิดชอบต่อผลงานทั้งหมดที่เกิดขึ้นจากการดำเนินงานในโครงการ (Gasemagha และ Kowang, 2021)

ผู้จัดการโครงการจะต้องมีทักษะด้านการสื่อสารระหว่างผู้มีส่วนได้ส่วนเสียของโครงการทั้งหมด ได้แก่ ผู้สนับสนุน ผู้ที่จะใช้ประโยชน์จากผลงานของโครงการ ผู้ควบคุมทรัพยากร และสมาชิกในทีมโครงการ โดยจะต้องมีชุดเครื่องมือและเทคนิคที่เหมาะสมในการแก้ไขปัญหา และจะต้องเป็นผู้ที่พัฒนาทักษะของตนเองและสมาชิกในโครงการผ่านการทบทวนบทเรียนเมื่อโครงการเสร็จสิ้น (PMBOK Guide, 2017)

Blaskovics (2016) ได้ทำการศึกษาความสำคัญของผู้จัดการโครงการที่ส่งผลต่อความสำเร็จของโครงการ พบว่า บุคลิก ทัศนคติ และความเป็นผู้นำของผู้จัดการโครงการมีผลอย่างมากต่อความสำเร็จของโครงการและการยอมรับของลูกค้า แต่โครงการแต่ละโครงการอาจต้องการผู้จัดการโครงการที่มีบุคลิกและทัศนคติที่แตกต่างกันออกไป นอกจากนี้ Patanakul (2011) ก็ได้ทำการศึกษาผลกระทบเชิงประจักษ์ของแนวทางที่ใช้ในการมอบหมายผู้จัดการโครงการในโครงการ โดยเน้นที่ประสิทธิภาพในการจัดการหลายโครงการในเวลาเดียวกัน ผลการวิจัยยืนยันถึงความสำคัญของการมอบหมายผู้จัดการโครงการที่เหมาะสมสำหรับโครงการสามารถเพิ่มประสิทธิภาพในด้านการพัฒนาโครงการ การเรียนรู้ ความสำเร็จของโครงการ และความพึงพอใจต่อทุกฝ่ายที่เกี่ยวข้อง

การบริหารโครงการ (Project Management) (PMI, 2021) เป็นการนำกระบวนการ วิธีการ ทักษะ ความรู้ และประสบการณ์ไปประยุกต์ใช้เพื่อให้บรรลุวัตถุประสงค์เฉพาะของโครงการตามเกณฑ์การยอมรับโครงการภายใน

ขอบเขตงานที่กำหนด กระบวนการจัดการโครงการแบ่งออกเป็น 5 ขั้นตอน คือ การเริ่มต้น (Initiating) การวางแผน (Planning) การดำเนินการ (Executing) การตรวจสอบและควบคุม (Monitoring and Controlling) และ การปิดโครงการ (Closing)



รูปภาพที่ 2 กระบวนการดำเนินโครงการตามมาตรฐานของ PMI

สถาบัน Project Management Institute (PMI) ได้นิยามกำหนดกรอบความรู้ในการบริหารโครงการออกเป็น 10 เรื่อง (Knowledge Area) ดังนี้

(1) การบริหารการบูรณาการโครงการ (Project Integration Management)
(2) การบริหารขอบเขตโครงการ (Project Scope Management)
(3) การบริหารเวลาโครงการ (Project Time Management)
(4) การบริหารค่าใช้จ่ายโครงการ (Project Cost Management)
(5) การบริหารคุณภาพโครงการ (Project Quality Management)
(6) การบริหารทรัพยากรของโครงการ (Project Resource Management)
(7) การบริหารการสื่อสารในโครงการ (Project Communications Management)
(8) การบริหารความเสี่ยงโครงการ (Project Risk Management)
(9) การบริหารการจัดซื้อจัดจ้างของโครงการ (Project Procurement Management
(10) การบริหารผู้มีส่วนได้ส่วนเสียของโครงการ (Project Stakeholders Management)

กรอบความรู้ในการจัดการโครงการทั้ง 10 เรื่องนี้ ครอบคลุมกระบวนการจัดการโครงการทั้ง 5 ขั้นตอน ตามกระบวนการดำเนินโครงการตามมาตรฐานของ PMI (รูปภาพที่ 2) ผู้จัดการโครงการจะต้องมีความเข้าใจ ตระหนัก และสามารถนำไปประยุกต์ใช้ได้อย่างมีประสิทธิภาพ โดยที่กรอบความรู้ในการจัดการโครงการเหล่านี้จะเป็นส่วนสำคัญที่ช่วยผู้จัดการโครงการในการประเมินและรับมือความเสี่ยงที่อาจส่งผลต่อผลลัพธ์ของการดำเนินโครงการ และยังเป็นเครื่องมือช่วยให้ผู้จัดการโครงการสามารถตัดสินใจเชิงรุกได้อย่างมีประสิทธิภาพ ดังนั้น เมื่อผู้จัดการโครงการสามารถเข้าใจแนวคิดของกรอบความรู้ในการจัดการโครงการก็จะช่วยให้สามารถควบคุมปัจจัยต่าง ๆที่เกี่ยวข้องกับการดำเนินโครงการ และเพิ่มอัตราความสำเร็จของโครงการให้สูงขึ้น

ดังนั้น ในงานวิจัยนี้จะใช้กรอบความรู้ในการจัดการโครงการทั้ง 10 เรื่อง (Knowledge Area) ข้างต้น เป็นแนวทางในการจัดหมวดหมู่ของปัญหาที่พบ เพื่อให้สอดคล้องกับมาตรฐานของการบริหารโครงการในระดับสากล

2.3 คำนิยามความล้มเหลวของโครงการ (Project Failure Definition)

ความล้มเหลวของโครงการ คือ โครงการที่ไม่สามารถปฏิบัติหรือดำเนินการได้ตามแผนที่วางไว้ ในโครงการทางด้านเทคโนโลยีสารสนเทศได้กำหนดความล้มเหลวในการส่งมอบโครงการคือ (Becker, 2017) งบประมาณ

ค่าใช้จ่ายเกินกว่าที่กำหนด การส่งมอบโครงการเกินจากระยะเวลาที่กำหนด การบกพร่องในความสามารถและคุณภาพที่ส่งมอบ และ ความล้มเหลวในการดำเนินโครงการหรือยกเลิกโครงการ

McManus และ Wood-Harper (2008) ได้ทำการศึกษาปัจจัยที่ทำให้โครงการล้มเหลว โดยได้กำหนดความล้มเหลวไว้เป็น 3 มิติ คือ เวลา (ความเร่งด่วนทางธุรกิจ) ต้นทุน (งบประมาณ) และคุณภาพ (ฟังก์ชันหรือความสามารถของผลิตภัณฑ์) ซึ่งจากผลการศึกษาพบสาเหตุที่ทำให้โครงการล้มเหลวมากถึง 51 สาเหตุ โดยที่สามารถจัดกลุ่มของสาเหตุได้เป็น 3 กลุ่ม คือ ปัจจัยทางด้านธุรกิจ 19.6% ปัจจัยทางด้านการบริหารโครงการ 53% และปัจจัยทางด้านเทคนิค 27.4% ซึ่งจะเห็นได้ว่าการบริหารโครงการเป็นปัจจัยสำคัญที่ส่งผลอย่างมากต่อผลลัพธ์ของโครงการ

Pinto และ Mantel (2015) ได้ทำการศึกษาปัจจัยที่มีผลให้โครงการล้มเหลว โดยการกำหนดความล้มเหลวนอกเหนือจากเรื่อง เวลา งบประมาณ และคุณภาพแล้ว ยังมีการพิจารณาครอบคลุมถึงเรื่องกระบวนการดำเนินการ การรักษาความสัมพันธ์ในการทำงานร่วมกับกลุ่มต่าง ๆ ทั้งภายในและภายนอกโครงการ การเป็นที่รู้จักของโครงการ ประโยชน์ที่ได้รับจากโครงการ รวมไปถึงความพึงพอใจของผู้ใช้บริการอีกด้วย ซึ่งจากการศึกษาได้ทำการกำหนด 10 ปัจจัยสำคัญที่มีผลต่อการดำเนินโครงการไว้ดังนี้

(1) การกำหนดเป้าหมายและทิศทางของโครงการอย่างชัดเจน
(2) การสนับสนุนจากผู้บริหารระดับสูง
(3) การวางแผนรายละเอียดขั้นตอนการดำเนินการสำหรับการดำเนินโครงการ
(4) การสื่อสาร การให้คำปรึกษาและการใช้งานแก่ลูกค้า การรับฟังความต้องการจากทุกฝ่ายที่เกี่ยวข้อง
(5) การสรรหาบุคลากร การคัดเลือก และการฝึกอบรมตามความจำเป็น
(6) งานด้านเทคนิค ความพร้อมใช้งานของเทคโนโลยีที่จำเป็นและความเชี่ยวชาญของบุคลากรเพื่อบรรลุขั้นตอนการดำเนินการทางเทคนิคเฉพาะ
(7) การยอมรับของลูกค้า
(8) การตรวจสอบและควบคุมโครงการที่ครอบคลุมทุกขั้นตอนของกระบวนการดำเนินการ
(9) การสื่อสารกับผู้มีบทบาทต่าง ๆในการดำเนินโครงการ
(10) การแก้ไขปัญหา ความสามารถในการจัดการกับเหตุการณ์ที่ไม่คาดคิดและกระทบกับแผนงานที่วางไว้

2.4 กรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset)

กรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) เป็นวิธีการคิดที่รวมกระบวนการแก้ปัญหาและการออกแบบเข้ากับทักษะชีวิตที่ส่งเสริมการปฏิสัมพันธ์กับผู้อื่น มีการกำหนดปัญหาและจัดหาวิธีแก้ปัญหาที่เหมาะสม ออกแบบและพัฒนาผลิตภัณฑ์ที่เป็นนวัตกรรม แสวงหา คิด และอยากรู้อยากเห็นอยู่เสมอ มีความมุ่งมั่นและเชื่อมโยงกับการเป้าหมายที่กำหนดไว้ (Caluori, 2014) แนวคิดนี้จะเป็นการย่อยปัญหาให้สามารถจับต้องได้ เพื่อให้เข้าใจและมีมุมมองต่อปัญหาที่ชัดเจนมากขึ้น เมื่อมองเห็นปัญหาที่ถูกต้องก็จะสามารถแก้ไขปัญหาได้อย่างตรงประเด็น นอกจากนี้ยังจะทำให้เห็นโอกาสที่เกิดขึ้นในปัญหาเหล่านั้น เพื่อให้เกิดการพัฒนากระบวนการทำงานให้มีประสิทธิภาพมากขึ้น (Miller, 2018) โดยสามารถสรุปความสำคัญของกรอบแนวคิดเชิงวิศวกรรม (Hope, 2022) คือ สิ่งเชื่อมโยงแนวคิดที่ซับซ้อนเข้ากับประสบการณ์ในโลกแห่งความเป็นจริงและนำไปสู่การเรียนรู้ที่ขึ้นกับตัวบุคคล ช่วยให้เกิดความคิดสร้างสรรค์ นำเสนอแนวคิดที่ดีที่สุดโดยไม่มีข้อจำกัด เรียนรู้ที่จะสื่อสารและทำงานร่วมกับผู้อื่นไปพร้อมกับการแก้ปัญหาที่เกิดขึ้น และเป็นทักษะที่สามารถปรับเปลี่ยนสิ่งรอบตัวให้เกิดประโยชน์ได้

โดยที่กรอบแนวคิดเชิงวิศวกรรมได้มีการจัดกลุ่มแนวคิดย่อย (Umaji & Paireekreng, 2021) เป็น 3 กลุ่ม คือ

(1) แนวคิดในการปรับตัว (Adaptive Mindset) คือ แนวคิดที่ประเมินข้อเท็จจริงและสถานการณ์ หรือสิ่งแวดล้อมในปัจจุบัน และปรับเปลี่ยนอย่างเหมาะสมเพื่อให้สามารถแก้ปัญหาได้ในทุกสถานการณ์
(2) แนวคิดในการเติบโต (Growth Mindset) คือ แนวคิดซึ่งบุคคลเชื่อว่าความสามารถพัฒนา สติปัญญา และความสามารถของตนเองได้

(3) แนวคิดในการเรียนรู้ (Learning Mindset) คือ แนวคิดที่จูงใจให้เปิดรับประสบการณ์ใหม่ ๆ มีความเชื่อว่าสามารถพัฒนาได้ด้วยการเรียนรู้และเติบโตจากประสบการณ์

Jamieson และ Donald (2020) ศึกษาความสัมพันธ์ระหว่างทักษะความเป็นผู้นำ (Leadership) และกรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) พบว่า ผู้ที่มีกรอบแนวคิดเชิงวิศวกรรมสูง มีความสามารถทางด้านเทคนิค การสื่อสาร และการทำงานเป็นทีม มีศักยภาพที่จะพัฒนาความสามารถระดับความเป็นผู้นำในองค์กรและสังคมมากกว่าแนวทางทางวิศวกรรมแบบดั้งเดิม สอดคล้องกับงานวิจัยของ Pia Lappalainen (2009) ที่สนับสนุนเรื่องการสื่อสาร ความร่วมมือ การทำงานเป็นทีมและการเรียนรู้ เป็นทักษะสำคัญของกรอบแนวคิดเชิงวิศวกรรม และงานวิจัยของ Katz (1955) และ Mumford และคณะ (2000) ที่กล่าวถึงความสัมพันธ์ระหว่างทักษะความเป็นผู้นำและกรอบแนวคิดเชิงวิศวกรรมที่เป็นทักษะจำเป็นสำหรับการเป็นผู้นำ

Lottero และ Lachapelle (2020) ได้ทำการวิจัยเปรียบเทียบกรอบแนวคิดเชิงวิศวกรรมของเด็กนักเรียนชั้นประถมศึกษาปีที่ 5 อายุระหว่าง 10 ถึง 11 ปีในประเทศสหรัฐอเมริกา โดยใช้กลุ่มตัวอย่าง 2,086 คน พบว่า 3 ใน 4 ของกลุ่มนักเรียนที่ได้รับการศึกษาเพิ่มเติมทางด้านแนวคิดเชิงวิศวกรรมตอบแบบสอบถามที่แสดงให้เห็นถึงคะแนนทางด้านกรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) และความคิดแบบเติบโต (Growth Mindset) ที่สูงขึ้นอย่างเห็นได้ชัด มีทักษะในการเรียน การเข้าสังคม การสื่อสาร และการแก้ปัญหาที่ดีขึ้น นอกจากนี้ยังสามารถรับมือและเรียนรู้จากความล้มเหลวได้ดีกว่ากลุ่มนักเรียนทั่วไป

Cunningham (2017) ได้ศึกษาการนำกรอบแนวคิดเชิงวิศวกรรมไปประยุกต์สอดแทรกในการเรียนการสอนทางด้านวิทยาศาสตร์ (STEM) พบว่านักเรียนสามารถนำกรอบแนวคิดเชิงวิศวกรรมไปใช้ในการแก้ไขปัญหาทางวิทยาศาสตร์รวมถึงปัญหาที่มีความซับซ้อนได้ดีขึ้น และสามารถวิเคราะห์ปัญหาและหาวิธีในการค้นหาคำตอบได้อย่างตรงประเด็นและมีความหมายมากขึ้น

2.5 กรอบแนวคิดการวิจัยและสมมติฐานการวิจัย

จากการทบทวนวรรณกรรมที่เกี่ยวของกับกรอบแนวคิดเชิงวิศวกรรมพบว่า การศึกษากลุ่มตัวอย่างที่มีหรือได้รับการฝึกฝนทักษะทางด้านกรอบแนวคิดเชิงวิศวกรรมจะเป็นผู้ที่มีความสามารถและทักษะในการแก้ปัญหา การเรียนรู้ พัฒนาตนเอง รวมถึงทักษะการสื่อสารและการเข้าสังคมที่ดีกว่ากลุ่มตัวอย่างทั่วไป แต่จากงานวิจัยที่ศึกษานั้นจะเป็นลักษณะของการศึกษากลุ่มเป้าหมายที่ใช้ทักษะในการแก้ปัญหาเฉพาะเรื่อง เช่น กลุ่มผู้นำที่ศึกษาในแง่มุมของการสื่อสารและการทำงานเป็นทีม หรือ กลุ่มนักเรียนที่ศึกษาในด้านของการแก้ปัญหาทางวิทยาศาสตร์ เป็นต้น

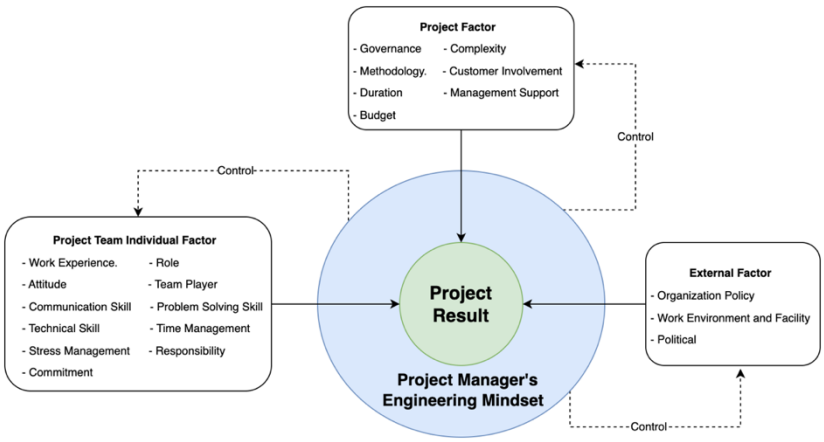
ผู้จัดการโครงการ (Project Manager) เป็นผู้ที่มีบทบาทสำคัญอย่างยิ่ง ที่จะส่งผลทั้งโดยตรงและโดยอ้อมต่อความสำเร็จหรือล้มเหลวของโครงการ โดยจะต้องเป็นผู้ที่จะต้องประยุกต์ใช้กรอบความรู้ในการจัดการโครงการทั้ง 10 เรื่อง (Knowledge Area) รวมถึงจะต้องเป็นผู้ความคุมและจัดการโครงการทั้ง 5 ขั้นตอน โดยจะต้องอาศัยทั้ง Soft skills และ Hard Skills เพื่อให้สามารถดำเนินโครงการได้อย่างมีประสิทธิภาพตามเป้าหมายที่วางไว้

ดังนั้น จึงเป็นประเด็นที่น่าสนใจในการศึกษากรอบแนวคิดเชิงวิศวกรรมของผู้จัดการโครงการ ที่ส่งผลต่อความสำเร็จหรือล้มเหลวของโครงการ โดยตั้งสมมุติฐานได้ว่า ผู้จัดการโครงการ (Project Manager) ซึ่งเป็นผู้ควบคุมการดำเนินโครงการและมีหน้าที่โดยตรงในการจัดการควบคุมปัจจัยต่างๆ ที่ส่งผลต่อผลลัพธ์ของโครงการ (Project Control & Monitoring and Risk Management) หากผู้จัดการโครงการเป็นผู้ที่มีกรอบแนวคิดเชิงวิศวกรรม (Project Manager's Engineering Mindset) ที่ดี จะส่งผลให้การบริหารจัดการและแก้ปัญหาในโครงการมีแนวโน้มที่จะประสบผลสำเร็จในอัตราที่สูงขึ้นเมื่อเทียบกับผู้จัดการโครงการโดยทั่วไป

ในการดำเนินโครงการยังมีปัจจัยอื่นๆที่ส่งผลต่อผลลัพธ์การดำเนินงาน ไม่ว่าจะเป็นปัจจัยส่วนบุคคลของสมาชิกในโครงการ เช่น ประสบการณ์ ความสามารถของบุคลากร ทัศนคติ ปัจจัยที่เกี่ยวของกับโครงการ เช่น วิธีการดำเนินโครงการ (Agile, Water Fall ฯลฯ) ความซับซ้อน ระยะเวลา และ งบประมาณของโครงการ และ ปัจจัยภายนอก

อื่นๆ เช่น นโยบายขององค์กร สภาพแวดล้อมในการทำงาน และ การเมืองภายนอก เป็นต้น ดังนั้นการศึกษาปัจจัยต่างๆ ที่มีผลต่อความล้มเหลวในการดำเนินโครงการจึงเป็นอีกหัวข้อสำคัญที่จะช่วยให้ผู้จัดการโครงการเกิดความตระหนัก และวางแผนในการจัดการ หาวิธีการรับมือ และแก้ปัญหาได้ล่วงหน้า

จากปัจจัยต่างๆที่กล่าวมาข้างต้น สามารถสร้างแผนภาพกรอบการวิจัยได้ตามภาพที่ 3 ดังนี้



รูปภาพที่ 3 กรอบงานวิจัยปัจจัยที่มีผลต่อผลลัพธ์ของโครงการ

3. ระเบียบวิธีวิจัย

3.1 เครื่องมือการวิจัย

ในการดำเนินการวิจัย ได้ใช้แบบสอบถามเป็นเครื่องมือในการเก็บข้อมูล โดยแบบสอบถามจะแบ่งออกเป็น 3 ส่วน ดังนี้

(1) ข้อมูลส่วนบุคคลของผู้ตอบแบบสอบถาม โดยเป็นลักษณะคำถามแบบสำรวจรายการและระบุคำตอบ

(2) ข้อมูลการดำเนินโครงการย้อนหลัง 3 ปี ของผู้ตอบแบบสอบถาม โดยเป็นการให้ผู้ตอบแบบสอบถามระบุคำตอบตามข้อเท็จจริง ข้อมูลจะแบ่งออกเป็น 4 ส่วน คือ

- จำนวนโครงการที่ทำทั้งหมด
- จำนวนโครงการที่ดำเนินการสำเร็จตามแผน
- จำนวนโครงการที่ดำเนินการสำเร็จแต่ไม่เป็นไปตามแผน
- จำนวนโครงการที่ดำเนินการไม่สำเร็จ

(3) สำรวจแนวความคิดของผู้จัดการโครงการ โดยเป็นลักษณะคำถามแบบสอบถามแบบมาตราส่วน 5 ระดับ ซึ่งแบ่งออกเป็น 2 ส่วน ดังนี้

- กรอบแนวคิดของการเป็นผู้จัดการโครงการ
- กรอบแนวคิดแบบวิศวกรรม (Engineering Mindset) โดยได้รับอิทธิพลมาจากงานวิจัยของ Dr. George D. Ricco's (2017) ซึ่งมีคำถามที่ใช้ในการวิจัยดังนี้

ตารางที่ 1 คำถามในงานวิจัยจากงานวิจัยของ Dr. George D. Ricco's

ข้อ	เชิงบวก (P) หรือ เชิงลบ (N)	คำถาม
1	N	You have a certain amount of intelligence, and you really can't do much to change it คุณคิดว่าคุณมีสติปัญญาพอสมควรและไม่สามารถเปลี่ยนแปลงอะไรได้มากนัก
2	N	Your intelligence is something about you that you can't change very much. คุณคิดว่าความฉลาดของคุณเป็นสิ่งที่ไม่สามารถเปลี่ยนแปลงได้มากนัก
3	P	No matter who you are, you can significantly change your intelligence level ไม่ว่าคุณจะเป็นใครคุณสามารถเปลี่ยนระดับสติปัญญาของคุณได้อย่างมาก
4	N	To be honest, you can't really change how intelligent you are คุณคิดว่าคุณไม่สามารถเปลี่ยนแปลงความฉลาดของคุณได้
5	P	You can always substantially change how intelligent you are คุณคิดว่าคุณสามารถเปลี่ยนแปลงความฉลาดของคุณได้เสมอ
6	N	You can learn new things, but you can't really change your basic intelligence คุณสามารถเรียนรู้สิ่งใหม่ ๆ ได้ แต่คุณไม่สามารถเปลี่ยนแปลงสติปัญญาพื้นฐานของคุณได้
7	P	No matter how much intelligence you have, you can always change it quite a bit คุณคิดว่าไม่ว่าคุณจะมีสติปัญญามากแค่ไหนคุณก็สามารถเปลี่ยนแปลงได้เสมอ
8	P	You can change even your basic intelligence level considerably คุณสามารถเปลี่ยนแปลงได้แม้กระทั่งระดับสติปัญญาพื้นฐานของคุณ
9	N	You have a certain amount of talent, and you can't really do much to change it คุณมีความสามารถในระดับหนึ่งและไม่สามารถเปลี่ยนแปลงอะไรได้มากนัก
10	N	Your talent in an area is something about you that you can't change very much ความสามารถของคุณเป็นสิ่งที่คุณไม่สามารถเปลี่ยนแปลงได้มากนัก
11	P	No matter who you are, you can significantly change your level of talent ไม่ว่าคุณจะเป็นใครคุณสามารถเปลี่ยนระดับความสามารถได้อย่างมาก
12	N	To be honest, you can't really change how much talent you have คุณคิดว่าคุณไม่สามารถเปลี่ยนแปลงความสามารถของคุณได้
13	P	You can always substantially change how much talent you have คุณสามารถเปลี่ยนแปลงความสามารถของคุณได้อย่างมาก
14	N	You can learn new things, but you can't really change your basic level of talent คุณสามารถเรียนรู้สิ่งใหม่ ๆ ได้ แต่คุณไม่สามารถเปลี่ยนระดับความสามารถพื้นฐานของคุณได้
15	P	No matter how much talent you have, you can always change it quite a bit ไม่ว่าคุณจะมีความสามารถมากแค่ไหนคุณก็สามารถเปลี่ยนแปลงได้เสมอ
16	P	You can change even your basic level of talent considerably คุณสามารถเปลี่ยนระดับความสามารถได้อย่างมาก

โดยผู้วิจัยได้นำคำถามในงานวิจัยมาทำการแบ่งประเภทย่อยของกรอบแนวคิดเชิงวิศวกรรมออกเป็น 3 ประเภท คือ แนวคิดในการปรับตัว (Adaptive Mindset) แนวคิดในการเติบโต (Growth Mindset) และ แนวคิดในการเรียนรู้ (Learning Mindset) ตามที่แสดงในตารางที่ 2 ดังนี้

ตารางที่ 2 การแบ่งประเภทข้อคำถาม

กลุ่มคำถามตามประเภทย่อยของแนวคิดเชิงวิศวกรรม	ข้อคำถาม
แนวคิดในการปรับตัว (Adaptive Mindset)	1,2,3,11,13,16
แนวคิดในการเติบโต (Growth Mindset)	4,5,7,8,9,12
แนวคิดในการเรียนรู้ (Learning Mindset)	6,10,14,15

3.2 กลุ่มเป้าหมาย

(1) ประชากรที่ใช้ศึกษา คือ ผู้จัดการโครงการทางด้านเทคโนโลยีสารสนเทศในประเทศไทย

(2) กลุ่มตัวอย่าง คือ ผู้จัดการโครงการทางด้านเทคโนโลยีสารสนเทศในประเทศไทย จำนวน 214 คน โดยใช้การเลือกกลุ่มตัวอย่างแบบเจาะจง (Purposive Sampling)

3.3 ขั้นตอนการดำเนินการวิจัย

(1) กำหนดตัวแปรที่ต้องการศึกษา ได้แก่ ปัจจัยที่มีผลต่อผลลัพธ์ในการดำเนินโครงการ

(2) กำหนดข้อมูลและตัวชี้วัดจากตัวแปรที่ศึกษา โดยจะต้องระบุข้อมูลและลักษณะของข้อมูลที่ต้องการว่ามีลักษณะอย่างไร ให้สอดคล้องกับวัตถุประสงค์ หรือปัญหาและขอบเขตของการวิจัย

(3) กำหนดแหล่งข้อมูลว่าต้องการข้อมูลหรือรวบรวมข้อมูลมาจากแหล่งข้อมูลใด

(4) ออกแบบแบบสอบถาม และวางแผนวิธีการเก็บรวบรวมข้อมูล

(5) นำแบบสอบถามที่ออกแบบไว้ เข้าสู่กระบวนการ Index of Item Objective Congruence (IOC) โดยผ่านการพิจารณาและตรวจสอบจากผู้ทรงคุณวุฒิ 3 ท่าน

(6) ดำเนินการเก็บข้อมูลตามแผนการเก็บข้อมูลที่วางไว้ โดยใช้ Google Form เป็นเครื่องมือในการเก็บข้อมูล

(7) สรุปผลการวิจัย

3.4 สถิติที่ใช้ในการวิจัย

การวิเคราะห์ข้อมูลในการวิจัยนี้เป็นการวิเคราะห์ข้อมูลเชิงปริมาณโดยใช้สถิติพรรณนา ได้แก่ ค่าเฉลี่ย (Mean Average) และส่วนเบี่ยงเบนมาตรฐาน (S.D.) โดยนำผลที่ได้มาเปรียบเทียบกับปัจจัยต่าง ๆที่มีผลต่อผลลัพธ์ของการดำเนินโครงการ นอกจากนี้ยังมีการใช้ค่าสัมประสิทธิ์อัลฟ่าครอนบาช (Cronbach's alpha) ในการทดสอบความเชื่อมั่นของแบบสอบถาม โดยค่าสัมประสิทธิ์อัลฟ่าครอนบาชมีสูตรในการคำนวณตามสมการ (1) ดังนี้

$$\alpha = \frac{N\bar{c}}{\bar{v} + (N-1)\bar{c}} \tag{1}$$

โดยที่ : N = *จำนวนข้อของแบบสอบถาม*, $\bar{c}$ = *ความแปรปรวนร่วมเฉลี่ย*, $\bar{v}$ = *ความแปรปรวนเฉลี่ย*

การประเมินความเที่ยงตรงของสัมประสิทธิ์อัลฟ่าครอนบาช ได้มีการพิจารณาจากเกณฑ์การประเมินความเที่ยงตรง ดังนี้ (ศิริชัย กาญจนวาสี, 2544)

ตารางที่ 3 การแปลความหมายระดับความเที่ยงตรงของค่าสัมประสิทธิ์แอลฟา

ค่าสัมประสิทธิ์แอลฟา (α)	การแปลความหมายระดับความเที่ยงตรง
มากกว่า 0.9	ดีมาก
มากกว่า 0.8	ดี
มากกว่า 0.7	พอใช้
มากกว่า 0.6	ค่อนข้างพอใช้
มากกว่า 0.5	ต่ำ
น้อยกว่าหรือเท่ากับ 0.5	ไม่สามารถรับได้

ซึ่งการประเมินความเที่ยงตรงของสัมประสิทธิ์อัลฟ่าครอนบาชจากผลการสำรวจโดยแบ่งตามกลุ่มคำถามประเภทย่อยของแนวคิดเชิงวิศวกรรม (Engineering Mindset) ได้ผลตามที่แสดงในตารางที่ 4 ดังนี้

ตารางที่ 4 ค่าสัมประสิทธิ์แอลฟาของกลุ่มคำถามย่อยของแนวคิดเชิงวิศวกรรม

กลุ่มคำถามตามประเภทย่อยของแนวคิดเชิงวิศวกรรม	ค่าสัมประสิทธิ์แอลฟา (α)
แนวคิดในการปรับตัว (Adaptive Mindset)	0.780
แนวคิดในการเติบโต (Growth Mindset)	0.763
แนวคิดในการเรียนรู้ (Learning Mindset)	0.841

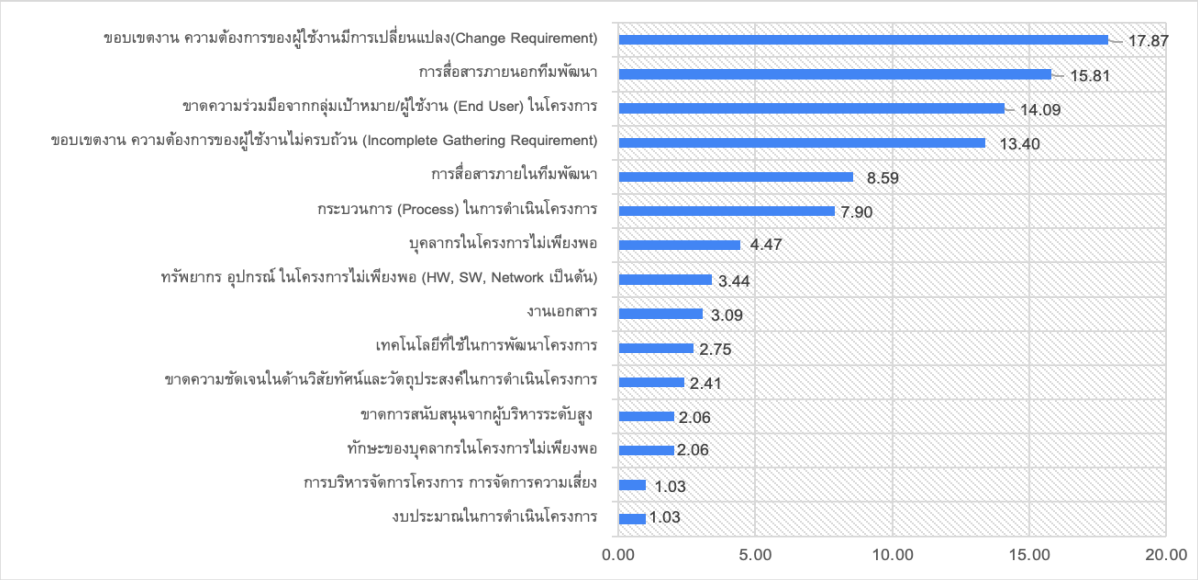
4. ผลการวิจัย

4.1 ผลการศึกษาปัจจัยที่มีผลต่อความล้มเหลวของการดำเนินการโครงการทางด้านเทคโนโลยีสารสนเทศ

จากผลการสำรวจพบว่า ปัจจัยที่มีผลต่อความล้มเหลวของโครงการ มีความสอดคล้องอย่างชัดเจนกับงานวิจัยของ Pinto และ Mantel ที่สำรวจในปี ค.ศ. 2015 โดยผลจากการสำรวจในงานวิจัยชิ้นนี้ ได้แบ่งย่อยกลุ่มของปัญหาออกเป็น 15 กลุ่มปัญหา ตามตารางที่ 5 ดังนี้

ตารางที่ 5 ปัจจัยที่มีผลทำให้การดำเนินโครงล้มเหลว

ลำดับ	ปัญหาทางด้าน	จำนวน	เปอร์เซ็นต์
1	ขอบเขตงาน ความต้องการของผู้ใช้งานมีการเปลี่ยนแปลง (Change Requirement)	52	17.87%
2	การสื่อสารภายนอกทีมพัฒนา	46	15.81%
3	ขาดความร่วมมือจากกลุ่มเป้าหมาย/ผู้ใช้งาน (End User) ในโครงการ	41	14.09%
4	ขอบเขตงาน ความต้องการของผู้ใช้งานไม่ครบถ้วน (Incomplete Gathering Requirement)	39	13.40%
5	การสื่อสารภายในทีมพัฒนา	25	8.59%
6	กระบวนการ (Process) ในการดำเนินโครงการ	23	7.90%
7	บุคลากรในโครงการไม่เพียงพอ	13	4.47%
8	ทรัพยากร อุปกรณ์ ในโครงการไม่เพียงพอ (HW, SW, Network เป็นต้น)	10	3.44%
9	งานเอกสาร	9	3.09%
10	เทคโนโลยีที่ใช้ในการพัฒนาโครงการ	8	2.75%
11	ขาดความชัดเจนในด้านวิสัยทัศน์และวัตถุประสงค์ในการดำเนินโครงการ	7	2.41%
12	ขาดการสนับสนุนจากผู้บริหารระดับสูง	6	2.06%
13	ทักษะของบุคลากรในโครงการไม่เพียงพอ	6	2.06%
14	งบประมาณในการดำเนินโครงการ	3	1.03%
15	การบริหารจัดการโครงการ การจัดการความเสี่ยง	3	1.03%
	รวม	**291**	**100%**



รูปภาพที่ 4 กราฟแสดงสัดส่วนเปอร์เซ็นต์ของปัญหาที่พบ

จากปัจจัยที่มีผลทำให้การดำเนินโครงล้มเหลวในตารางที่ 5 เมื่อนำมาจัดหมวดหมู่ตามกรอบความรู้ในการบริหารโครงการ (Knowledge Area) ที่กำหนดโดย PMI จะสามารถจัดโดยเรียงลำดับตามสัดส่วนเปอร์เซ็นต์รวมจากมากไปน้อยได้ตามตารางที่ 6 ดังนี้

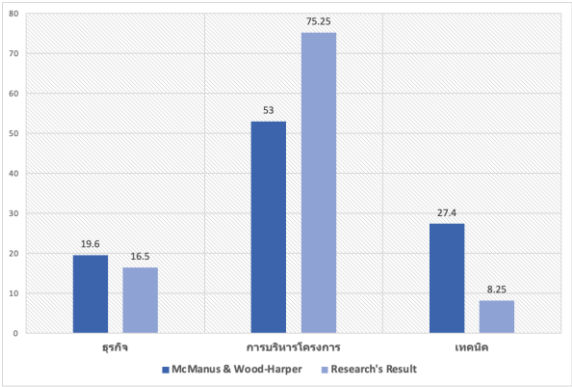
ตารางที่ 6 การจัดหมวดหมู่ตามกรอบความรู้ในการบริหารโครงการ

กรอบความรู้ในการบริหารโครงการ	**ปัญหา**	**เปอร์เซ็นต์รวม**
การบริหารขอบเขตโครงการ (Project Scope Management)	● ขอบเขตงาน ความต้องการของผู้ใช้งานมีการเปลี่ยนแปลง (Change Requirement) ● ขอบเขตงาน ความต้องการของผู้ใช้งานไม่ครบถ้วน (Incomplete Gathering Requirement)	31.27 %
การบริหารการสื่อสารในโครงการ (Project Communications Management)	● การสื่อสารภายนอกทีมพัฒนา ● การสื่อสารภายในทีมพัฒนา ● งานเอกสาร	27.49 %
การบริหารผู้มีส่วนได้ส่วนเสียของโครงการ (Project Stakeholders Management)	● ขาดความร่วมมือจากกลุ่มเป้าหมาย/ผู้ใช้งาน (End User) ในโครงการ ● ขาดการสนับสนุนจากผู้บริหารระดับสูง	16.15 %
การบริหารทรัพยากรของโครงการ (Project Resource Management)	● บุคลากรในโครงการไม่เพียงพอ ● ทรัพยากร อุปกรณ์ ในโครงการไม่เพียงพอ ● เทคโนโลยีที่ใช้ในการพัฒนาโครงการ ● ทักษะของบุคลากรในโครงการไม่เพียงพอ	12.71 %
การบริหารการบูรณาการโครงการ (Project Integration Management)	● กระบวนการ (Process) ในการดำเนินโครงการ ● ขาดความชัดเจนในด้านวิสัยทัศน์และวัตถุประสงค์ในการดำเนินโครงการ	10.31 %
การบริหารค่าใช้จ่ายโครงการ (Project Cost Management)	● งบประมาณในการดำเนินโครงการ	1.03 %
การบริหารความเสี่ยงโครงการ (Project Risk Management)	● การบริหารจัดการโครงการ การจัดการความเสี่ยง	1.03 %

เมื่อนำผลจากการสำรวจในตารางที่ 5 มาเปรียบเทียบกับผลการสำรวจและจัดกลุ่มของ McManus และ Wood-Harper ในปี 2008 สามารถสรุปข้อมูลได้ดังนี้

ตารางที่ 7 การเปรียบเทียบกับผลการวิจัยของ McManus และ Wood-Harper

ปัจจัยที่ทำให้โครงการล้มเหลว	McManus และ Wood-Harper	ผลการสำรวจของงานวิจัย
ทางด้านธุรกิจ	19.60%	รวม 16.50% • ขาดความร่วมมือจากกลุ่มเป้าหมาย/ผู้ใช้งาน (End User) ในโครงการ (14.09%) • ขาดความชัดเจนในด้านวิสัยทัศน์และวัตถุประสงค์ในการดำเนินโครงการ (2.41%)
ทางด้านการบริหารโครงการ	53%	รวม 75.25% • ขอบเขตงาน ความต้องการของผู้ใช้งานมีการเปลี่ยนแปลง (Change Requirement) (17.87%) • การสื่อสารภายนอกทีมพัฒนา (15.81%) • ขอบเขตงาน ความต้องการของผู้ใช้งานไม่ครบถ้วน (Incomplete Gathering Requirement) (13.40%) • การสื่อสารภายในทีมพัฒนา (8.59%) • กระบวนการ (Process) ในการดำเนินโครงการ (7.90%) • บุคลากรในโครงการไม่เพียงพอ (4.47%) • งานเอกสาร (3.09%) • ขาดการสนับสนุนจากผู้บริหารระดับสูง (2.06%) • งบประมาณในการดำเนินโครงการ (1.03%) • การบริหารจัดการโครงการ การจัดการความเสี่ยง (1.03%)
ทางด้านเทคนิค	27.4%	รวม 8.25% • ทรัพยากร อุปกรณ์ ในโครงการไม่เพียงพอ (HW, SW, Network เป็นต้น) (3.44%) • เทคโนโลยีที่ใช้ในการพัฒนาโครงการ (2.75%) • ทักษะของบุคลากรในโครงการไม่เพียงพอ (2.06%)



รูปภาพที่ 5 กราฟแสดงการเปรียบเทียบกับผลการวิจัยของ McManus และ Wood-Harper

จากการเปรียบเทียบกับผลการวิจัยของ McManus และ Wood-Harper พบว่า ปัจจัยทางด้านการบริหารโครงการยังคงเป็นสาเหตุหลักที่ทำให้เกิดความล้มเหลวในการดำเนินโครงการ โดยในผลสำรวจในงานวิจัยนี้ ความล้มเหลวที่มีผลมาจากปัจจัยทางด้านการบริหารโครงการมีมากถึง 75.25% ซึ่งมากกว่าผลการสำรวจของ McManus และ Wood-Harper ถึง 22.25% ในขณะที่ความล้มเหลวที่มีผลมาจากปัจจัยทางด้านธุรกิจและเทคนิคมีแนวโน้มที่จะลดลง โดยเฉพาะอย่างยิ่งปัจจัยทางด้านเทคนิคที่ลดลงถึง 18.90% ซึ่งน่าจะเป็นผลมาจากการที่เทคโนโลยีมีการพัฒนาปรับเปลี่ยนอย่างรวดเร็วในระหว่างปี ค.ศ. 2008 ถึงปัจจุบัน ไม่ว่าจะเป็นเรื่องของระบบ Cloud Infrastructure ที่ช่วยลดปัญหาทางด้านของฮาร์ดแวร์ รวมถึงการที่สถาบันการศึกษาต่างๆได้ให้ความสำคัญในการผลิตทรัพยากรบุคคลออกมารองรับความต้องการของตลาดแรงงานได้อย่างต่อเนื่องและมีประสิทธิภาพ

ปัจจัยที่มีผลต่อความล้มเหลวของโครงการที่พบนั้น นอกจากจะสามารถแก้ไขด้วยการใช้กรอบความรู้ในการบริหารโครงการ (Knowledge Area) ตามมาตรฐานของ PMI แล้ว ยังมีมาตรฐานอื่นๆ ที่สามารถนำมาเป็นแนวทางในการปรับปรุงประสิทธิภาพของการบริหารโครงการ หนึ่งในมาตรฐานที่เป็นที่ยอมรับอย่างกว้างขวางในอุตสาหกรรมทางด้านเทคโนโลยีสาสนเทศ ได้แก่ มาตรฐานการบริหารงานบริการด้านสารสนเทศ (ITIL : Information Technology Infrastructure Library) โดยในตารางที่ 8 ได้สรุปหัวข้อแนวทางการแก้ปัญหา (Practices Area) ตามมาตรฐาน ITIL ไว้ดังนี้

ตารางที่ 8 การจัดปัจจัยปัญหากับหัวข้อแนวทางการแก้ปัญหา (Practices Area) ตามมาตรฐาน ITIL

กรอบความรู้ใน การบริหารโครงการ (PMI)	**ปัญหา**	**หัวข้อตามมาตรฐาน ITIL (Practices Area)**
การบริหารขอบเขตโครงการ (Project Scope Management)	• ขอบเขตงาน ความต้องการของผู้ใช้งานมีการเปลี่ยนแปลง (Change Requirement) • ขอบเขตงาน ความต้องการของผู้ใช้งานไม่ครบถ้วน (Incomplete Gathering Requirement)	• Change control • Business analysis • Release management

ตารางที่ 8 การจัดปัจจัยปัญหากับหัวข้อแนวทางการแก้ปัญหา (Practices Area) ตามมาตรฐาน ITIL (ต่อ)

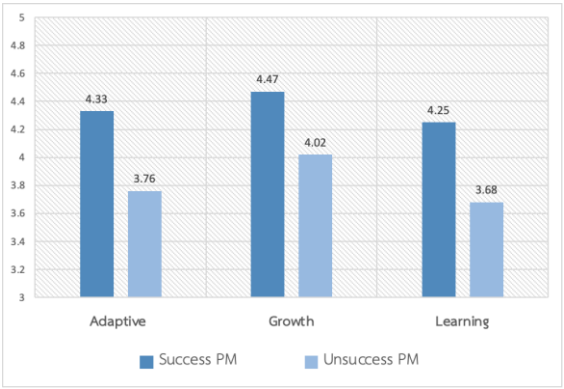
กรอบความรู้ใน การบริหารโครงการ (PMI)	**ปัญหา**	**หัวข้อตามมาตรฐาน ITIL (Practices Area)**
การบริหารการสื่อสารในโครงการ (Project Communications Management)	• การสื่อสารภายนอกทีมพัฒนา • การสื่อสารภายในทีมพัฒนา • งานเอกสาร	• Relationship management • Knowledge management
การบริหารผู้มีส่วนได้ส่วนเสียของโครงการ (Project Stakeholders Management)	• ขาดความร่วมมือจากกลุ่มเป้าหมาย/ผู้ใช้งาน (End User) ในโครงการ • ขาดการสนับสนุนจากผู้บริหารระดับสูง	• Relationship management
การบริหารทรัพยากรของโครงการ (Project Resource Management)	• บุคลากรในโครงการไม่เพียงพอ • ทรัพยากร อุปกรณ์ ในโครงการไม่เพียงพอ • เทคโนโลยีที่ใช้ในการพัฒนาโครงการ • ทักษะของบุคลากรในโครงการไม่เพียงพอ	• Workforce and talent management • Architecture management • Capacity and performance management • Software development and management
การบริหารการบูรณาการโครงการ (Project Integration Management)	• กระบวนการ (Process) ในการดำเนินโครงการ • ขาดความชัดเจนในด้านวิสัยทัศน์และวัตถุประสงค์ในการดำเนินโครงการ	• Project management • Strategy management
การบริหารค่าใช้จ่ายโครงการ (Project Cost Management)	• งบประมาณในการดำเนินโครงการ	• Service financial management
การบริหารความเสี่ยงโครงการ (Project Risk Management)	• การบริหารจัดการโครงการการจัดการความเสี่ยง	• Risk management

4.2 ผลการศึกษาความสัมพันธ์ระหว่างผลลัพธ์ในการดำเนินโครงการทางด้านเทคโนโลยีสารสนเทศกับทักษะการคิดเชิงวิศวกรรม (Engineering Mindset) ของผู้จัดการโครงการ

จากผลการสำรวจ เมื่อนำผลจากแบบสอบถามมาคำนวณหาค่าเฉลี่ย (Mean Average) และส่วนเบี่ยงเบนมาตรฐาน (S.D.) เพื่อทำการเปรียบเทียบในมิติของกรอบแนวคิด (Mindset) ระหว่างผู้จัดการโครงการที่มีอัตราความสำเร็จในการดำเนินโครงการสูง (Success PM - ดำเนินโครงการได้ตรงตามแผนงาน ระยะเวลา งบประมาณ และคุณภาพ) และแนวคิดของผู้จัดการโครงการที่มีอัตราความสำเร็จในการดำเนินโครงการต่ำ (Unsuccess PM - ดำเนินโครงการได้ไม่ตรงตามแผนงาน ระยะเวลา งบประมาณ หรือ คุณภาพ) ได้ตามตารางที่ 9 ดังนี้

ตารางที่ 9 การเปรียบเทียบค่าเฉลี่ยคะแนนกรอบแนวคิดเชิงวิศวกรรมระหว่างกลุ่มของผู้จัดการโครงการ

ประเภทย่อยของกรอบแนวคิดเชิงวิศวกรรม	Success PM		Unsuccess PM	
	Mean	S.D.	Mean	S.D.
แนวคิดในการปรับตัว (Adaptive Mindset)	4.33	0.77	3.76	0.87
แนวคิดในการเติบโต (Growth Mindset)	4.47	0.71	4.02	0.73
แนวคิดในการเรียนรู้ (Learning Mindset)	4.25	0.84	3.68	0.78



รูปภาพที่ 6 การเปรียบเทียบค่าเฉลี่ยคะแนนกรอบแนวคิดเชิงวิศวกรรมระหว่างกลุ่มของผู้จัดการโครงการ

จากข้อมูลการเปรียบเทียบค่าเฉลี่ยของกรอบแนวคิดเชิงวิศวกรรมของผู้จัดการโครงการในตารางที่ 9 (รูปภาพที่ 6) พบว่า คะแนนเฉลี่ยของกลุ่มของผู้จัดการโครงการที่มีอัตราความสำเร็จในการดำเนินโครงการสูง (Success PM) จะสูงกว่าคะแนนเฉลี่ยของกลุ่มของผู้จัดการโครงการที่มีอัตราความสำเร็จในการดำเนินโครงการต่ำ (UnSuccess PM) ในทั้ง 3 กลุ่มแนวคิด (แนวคิดในการปรับตัว การเติบโต และการเรียนรู้) อย่างมีนัยสำคัญ โดยที่ในแต่ละกลุ่มจะส่งผลต่อผลลัพธ์ของการบริหารโครงการ ดังนี้

1. ผู้จัดการโครงการที่มีกรอบแนวคิดในการปรับตัว (Adaptive Mindset) ที่ดี จะสามารถปรับเปลี่ยนแผน เครื่องมือ หรือวิธีการในการดำเนินงานได้อย่างเหมาะสมตามสถานการณ์ที่เปลี่ยนไป ตามสภาพความเป็นจริง ณ ขณะนั้น มีความยืดหยุ่นในการทำงาน จึงสามารถรับมือและแก้ปัญหาที่เกิดขึ้นในโครงการได้อย่างรวดเร็วและมีประสิทธิภาพ
2. ผู้จัดการโครงการที่มีกรอบแนวคิดในการเติบโต (Growth Mindset) ที่ดี จะช่วยให้การพัฒนาขีดความสามารถของผู้จัดการโครงการเป็นไปได้อย่างรวดเร็ว ซึ่งจะส่งผลโดยตรงต่อแนวคิดและวิธีการบริหารโครงการ การแก้ไขปัญหาได้อย่างมีประสิทธิภาพมากขึ้น

3. ผู้จัดการโครงการที่มีกรอบแนวคิดในการเรียนรู้ (Learning Mindset) ที่ดี จะเป็นผู้ที่เปิดรับการเรียนรู้เรื่องใหม่ ๆ จากช่องทางต่าง ๆ รวมถึงการเรียนรู้จากประสบการณ์ที่ผ่านมาทั้งที่ประสบความสำเร็จและล้มเหลว เพื่อนำมาประยุกต์ใช้ให้เกิดประโยชน์สูงสุดในการบริหารโครงการ

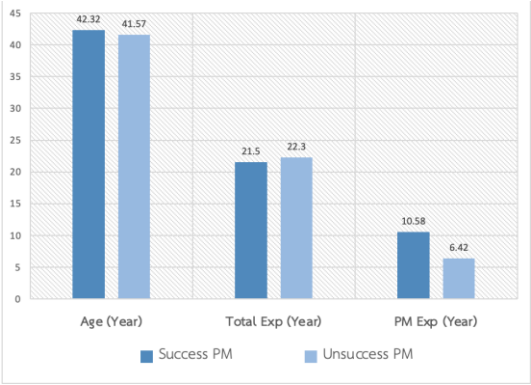
ซึ่งจากข้อมูลแสดงให้เห็นว่าผู้จัดการโครงการที่มีกรอบแนวคิดในการปรับตัว การเติบโต และการเรียนรู้ที่ดี จะส่งผลต่อภาพรวมของกรอบแนวคิดเชิงวิศวกรรมที่ดี และมีแนวโน้มที่จะประสบผลสำเร็จในการจัดการและบริหารโครงการมากขึ้น

4.3 ผลการศึกษาปัจจัยที่ส่งผลต่อผลลัพธ์ของการดำเนินโครงการอันเป็นปัจจัยจากผู้จัดการโครงการ

นอกเหนือจากการศึกษาปัจจัยทางด้านแนวคิด (Mindset) ในงานวิจัยนี้ยังได้ศึกษาถึงปัจจัยอื่น ๆที่อาจจะส่งผลต่อผลลัพธ์ของการดำเนินโครงการ ได้แก่ ปัจจัยทางด้านอายุ ประสบการณ์การทำงานโดยรวม และประสบการณ์ในการเป็นผู้จัดการโครงการ โดยจากข้อมูลการสำรวจได้ผลดังนี้

ตารางที่ 10 การเปรียบเทียบค่าเฉลี่ยปัจจัยอื่น ๆ ระหว่างกลุ่มของผู้จัดการโครงการ

ปัจจัย	Success PM	Unsuccess PM
อายุของผู้จัดการโครงการ (ปี)	42.32	41.57
ประสบการณ์การทำงานโดยรวม (ปี)	21.5	22.3
ประสบการณ์ในการเป็นผู้จัดการโครงการ (ปี)	10.58	6.42



รูปภาพที่ 7 การเปรียบเทียบค่าเฉลี่ยปัจจัยอื่น ๆ ระหว่างกลุ่มของผู้จัดการโครงการ

จากข้อมูลการเปรียบเทียบค่าเฉลี่ยของปัจจัยอื่น ๆของผู้จัดการโครงการในตารางที่ 10 (รูปภาพที่ 7) พบว่าอายุ และประสบการณ์โดยรวมของผู้จัดการโครงการอยู่ในช่วงที่ใกล้เคียงกัน จึงไม่ควรจะมีผลต่อผลลัพธ์ของโครงการมากนัก ในขณะที่ประสบการณ์ในการบริหารโครงการมีความแตกต่างกันอย่างเห็นได้ชัด โดยจากข้อมูลสามารถสรุปได้ว่า ผู้จัดการโครงการที่มีประสบการณ์ในการบริหารโครงการมากกว่า มีแนวโน้มในการบริหารโครงการสำเร็จในอัตราที่สูงกว่าผู้จัดการโครงการที่มีประสบการณ์น้อยกว่า

ผลจากการสำรวจพบว่ากรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) ของผู้จัดการโครงการ ซึ่งประกอบไปด้วยแนวคิดในการปรับตัว (Adaptive Mindset) แนวคิดในการเติบโต (Growth Mindset) และ แนวคิดในการเรียนรู้ (Learning Mindset) เป็นหนึ่งในทักษะที่สำคัญที่ส่งผลต่อความสำเร็จหรือล้มเหลวในการดำเนินโครงการ นอกจากนี้ยังพบว่าประสบการณ์โดยตรงในการบริหารโครงการที่ผ่านประสบการณ์การเรียนรู้และการแก้ปัญหา น่าจะเป็นหนึ่งในปัจจัยสำคัญที่มีส่วนในการพัฒนากรอบแนวคิดเชิงวิศวกรรมสำหรับผู้จัดการโครงการอีกด้วย

ดังนั้นในการดำเนินโครงการทางด้านเทคโนโลยีสารสนเทศที่มีขนาดใหญ่ มีความซับซ้อนหรือความเสี่ยงสูง ควรเลือกผู้จัดการโครงการที่มีจำนวนปีประสบการณ์ในการบริหารโครงการโดยตรงสูง ประกอบกับอัตราความสำเร็จของการดำเนินโครงการที่ผ่านมาของผู้จัดการโครงการเป็นเงื่อนไขแรก เนื่องจากเป็นผู้ที่มีแนวโน้มที่จะมีแนวคิดเชิงวิศวกรรมที่ดี ควรจะสามารถวางแผน บริหาร และรับมือกับปัญหาต่าง ๆ ที่จะผ่านเข้ามาในโครงการได้ดี โดยอาจจะมีข้อมูลอื่น ๆ ประกอบในการคัดเลือก เช่น ลักษณะ ขนาดหรือความซับซ้อนของโครงการที่เคยบริหาร เป็นต้น เพื่อเป็นการลดโอกาสสำหรับความล้มเหลวของโครงการ

5. สรุปผลการวิจัย

5.1 อภิปรายผลการวิจัย

ในขณะที่โลกมีการเปลี่ยนแปลงอยู่ตลอดเวลา โดยเฉพาะอย่างยิ่งในการพัฒนาโครงการทางด้านเทคโนโลยีสารสนเทศซึ่งมีความซับซ้อนและปัจจัยมากมายทั้งภายในและภายนอกที่จะส่งผลให้โครงการไม่สามารถเป็นไปได้ตามแผนที่วางไว้ ผู้จัดการโครงการจึงเป็นบุคคลที่มีบทบาทและสำคัญอย่างยิ่งที่จะส่งผลโดยตรงต่อความสำเร็จหรือล้มเหลวของโครงการ นอกจากทักษะในการบริหารจัดการโครงการที่ดีแล้ว ผู้จัดการโครงการมีความจำเป็นอย่างยิ่งที่จะต้องมีทักษะพิเศษที่จะต้องพร้อมรับมือทุกการเปลี่ยนแปลง พัฒนาและก้าวข้ามขีดจำกัดของตนเอง รวมถึงเรียนรู้เพื่อให้เท่าทันกับโลกที่เปลี่ยนแปลงไป รูปแบบที่เป็นแนวปฏิบัติที่ดีในวันนี้จะต้องมีการปรับปรุงและพัฒนาอยู่อย่างสม่ำเสมอเพื่อให้พร้อมรับมือทุกสภาพของปัญหาที่อาจเกิดขึ้นได้ในระหว่างการดำเนินโครงการ

จากผลการวิจัยพบว่ากรอบความรู้ในการบริหารโครงการที่กำหนดขึ้นโดย PMI ยังคงเป็นเครื่องมือสำคัญในการช่วยให้โครงการประสบความสำเร็จโดยมีแนวโน้มที่จะประสบความสำเร็จมากขึ้นเมื่อนำมาประยุกต์ใช้ร่วมกันกับกรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) โดยเน้นไปที่ 3 ทักษะหลัก ได้แก่ แนวคิดในการปรับตัว (Adaptive Mindset) เพื่อให้สามารถรับมือกับความเปลี่ยนแปลง หรือปัญหาต่างๆที่เกิดขึ้นได้ แนวคิดในการเติบโต (Growth Mindset) เพื่อให้ไม่เกิดข้อจำกัดในการดำเนินงานหรือการแก้ปัญหา และ แนวคิดในการเรียนรู้ (Learning Mindset) เพื่อค้นหาแนวทางหรือรูปแบบใหม่ ๆ ที่เหมาะสมต่อการการดำเนินงาน รวมถึงการเรียนรู้ ถอดบทเรียนจากข้อผิดพลาดต่างๆที่เกิดขึ้น เพื่อเป็นการพัฒนาให้เกิดการปรับปรุงการทำงานที่มีประสิทธิภาพต่อไป

นอกจากนี้ยังพบว่า การบริหารขอบเขตโครงการ (Scope Management) และการบริหารการสื่อสารในโครงการ (Communication Management) เป็น 2 กลุ่มหลักที่ส่งผลต่อความล้มเหลวของโครงการ โดยมีเปอร์เซ็นต์รวมกันถึง 58.76% โดยจาก 10 กลุ่มกรอบความรู้ในการบริหารโครงการ (Knowledge Area) ที่กำหนดโดย PMI มีถึง 7 กลุ่มที่ส่งผลโดยตรงต่อความล้มเหลวของโครงการ (ตารางที่ 6) มีเพียง 3 กลุ่มเท่านั้นที่ไม่ได้ส่งผลโดยตรง ได้แก่ การบริหารเวลาโครงการ (Time Management) การบริหารคุณภาพโครงการ (Quality Management) และ การบริหารการจัดซื้อจัดจ้างของโครงการ (Procurement Management) แต่ก็มีความเป็นไปได้ว่าทั้ง 3 กลุ่มนี้จะส่งผลกระทบทางอ้อมต่อความสำเร็จและล้มเหลวของโครงการเช่นกัน ทั้งนี้ ในปัจจุบันได้มีกรอบการทำงาน (Framework) ซึ่งเป็นเครื่องมือในการบริหารจัดการโครงการที่เป็นที่นิยมอยู่เป็นจำนวนมาก เช่น มาตรฐาน CMMI, Information technology infrastructure library (ITIL) หรือ ISO 29110 ซึ่งมาตรฐานและเครื่องมือเหล่านี้จะช่วยเป็นแนวทางในการลดความเสี่ยงและข้อผิดพลาดในการทำเนินโครงการได้เป็นอย่างดี จึงควรจะพิจารณานำไปใช้ในการบริหารโครงการ

การศึกษาความสัมพันธ์ระหว่างผลลัพธ์ในการดำเนินโครงการทางด้านเทคโนโลยีสารสนเทศกับกรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) ของผู้จัดการโครงการ พบว่ากรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) ของผู้จัดการโครงการ ซึ่งประกอบไปด้วยแนวคิดในการปรับตัว (Adaptive Mindset) แนวคิดในการเติบโต (Growth Mindset) และ แนวคิดในการเรียนรู้ (Learning Mindset) เป็นหนึ่งในทักษะที่สำคัญที่ส่งผลต่อความสำเร็จหรือล้มเหลวในการดำเนินโครงการ นอกจากนี้ยังพบว่าประสบการณ์โดยตรงในการบริหารโครงการน่าจะเป็นหนึ่งในปัจจัยสำคัญที่มีส่วนในการพัฒนากรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) สำหรับผู้จัดการโครงการ ในขณะที่อายุหรือ

ประสบการณ์โดยรวมของผู้จัดการโครงการไม่น่าจะเป็นปัจจัยที่มีความสำคัญต่อการพัฒนากรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) สำหรับผู้จัดการโครงการมากนัก

จากผลการเปรียบเทียบค่าเฉลี่ยทางด้านกรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) จะเห็นได้อย่างชัดเจนว่ากลุ่มของผู้จัดการโครงการที่มีค่าเฉลี่ยสูง มีโอกาสและแนวโน้มที่จะช่วยให้โครงการประสบความสำเร็จได้มากกว่าเมื่อเทียบกับกลุ่มของผู้จัดการโครงการที่มีค่าเฉลี่ยต่ำ ซึ่งสอดคล้องกับผลการวิจัยของ Jamieson และ Donald ที่ศึกษาความสัมพันธ์ระหว่างกรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) และด้านความเป็นผู้นำ (Leadership) ผลจากการศึกษาพบว่า ผู้นำที่มีกรอบแนวคิดเชิงวิศวกรรม จะมีโอกาสที่จะก้าวขึ้นไปเป็นผู้นำในระดับสังคมหรือองค์กรมากกว่าผู้นำแบบดั้งเดิม และเป็นไปในทิศทางเดียวกันกับงานวิจัยของ Lottero และ Lachapelle (2020) ที่พบว่าผู้ที่มีกรอบแนวคิดเชิงวิศวกรรม ที่ดี จะมีแนวคิดแบบเติบโต (Growth) ที่ดี และสามารถปรับตัว (Adaptive) เพื่อรับมือและแก้ปัญหาได้ อีกทั้งยังสามารถเรียนรู้ (Learning) จากความผิดพลาดและนำมาใช้เป็นแนวทางในการพัฒนาตนเองในอนาคตได้ รวมถึงความสามารถในการวิเคราะห์ ค้นหาคำตอบ และวิธีการแก้ไขได้อย่างมีประสิทธิภาพ ซึ่งสอดคล้องกับการศึกษาของ Cunningham ด้วยเช่นกัน ดังนั้น จากผลการศึกษาจึงสามารถแสดงให้เห็นได้ว่า การที่จะเป็นผู้จัดการโครงการทางด้านเทคโนโลยีสารสนเทศที่ประสบความสำเร็จได้นั้น ควรจะมีทักษะหลักๆ ที่สำคัญ ได้แก่ แนวคิดในการปรับตัว (Adaptive Mindset) แนวคิดในการเติบโต (Growth Mindset) และ แนวคิดในการเรียนรู้ (Learning Mindset) ซึ่งเป็นองค์ประกอบของกรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) นั่นเอง

5.2 ข้อเสนอแนะสำหรับงานวิจัยต่อเนื่อง

การบริหารโครงการเป็นทั้งศาสตร์และศิลป์ ซึ่งมีปัจจัยมากมายที่เกี่ยวข้อง ดังนั้น จึงยังคงมีอีกหลายปัจจัยที่ควรจะนำมาพิจารณาประกอบเป็นปัจจัยที่อาจจะส่งผลต่อผลการดำเนินโครงการ เช่น ระดับการศึกษา หรือ ใบรับรองความสามารถ (Certificate) ของผู้จัดการโครงการ หรืออาจจะเป็นปัจจัยอื่นๆ เช่น ขนาด ขอบเขตและความซับซ้อนของโครงการ หรือ ระเบียบวิธี (Methodology) ในการดำเนินโครงการ เช่น Agile Waterfall หรือ Hybrid เป็นต้น

นอกจากนี้ การพัฒนากรอบกระบวนการ (Framework) ในการประเมินความสามารถทางด้านกรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) ของผู้จัดการโครงการเพื่อเป็นเครื่องมือในการมอบหมายโครงการ หรือการศึกษาเพื่อหาปัจจัยหรือวิธีการที่จะพัฒนากรอบแนวคิดเชิงวิศวกรรม (Engineering Mindset) ให้สามารถพัฒนาได้อย่างรวดเร็วยิ่งขึ้น การนำกรอบแนวคิดเชิงวิศวกรรมไปประยุกต์ใช้ในรูปแบบต่างๆ หรือการสร้างแนวปฏิบัติของกรอบแนวคิดเชิงวิศวกรรม ก็ถือเป็นหัวข้อที่น่าสนใจที่จะนำมาศึกษาเพิ่มเติมในอนาคต เพื่อเป็นการต่อยอด เพิ่มประสิทธิภาพของผู้จัดการโครงการ อีกทั้งยังเป็นการเพิ่มอัตราความสำเร็จในการบริหารโครงการอีกด้วย

บรรณานุกรม

ศิริชัย กาญจนวาสี. (2544). **การเลือกใช้สถิติที่เหมาะสมสำหรับการวิจัย**. (พิมพ์ครั้งที่ 4). กรุงเทพฯ: บุญศิริการพิมพ์.

ADHIATMA, D., & LEGOWO, N. (2018). 4 DOMAINS OF COBIT FRAMEWORK VERSION 5.0 THAT CAN BE USED TO EVALUATING INFORMATION TECHNOLOGY PROJECT MANAGEMENT GOVERNANCE IN THE IT FUNCTION. **Journal of Theoretical and Applied Information Technology**, 96(23).

Akampurira (2013), **Project Planning and Management: An Aspect of Development**. Hamburge: Anchor Academic Publishing.

Becker, D. (2017). **Predicting Outcomes for Big Data Projects: Big Data Project Dynamics (BDPD)**. 2017 IEEE International Conference on Big Data (BIGDATA).

Blaskovics, B. (2016). The impact of project manager on project success—The case of ICT sector. **Society and Economy**, 38(2), 261-281.

Caluori, N. (2014). **Mindset and motivation**. West Point. Retrieved August 4, 2022, https://www.westpoint.edu/sites/default/files/inline-images/centers_research/center_for_teching_excellence/PDFs/mtp_project_papers/Caluori_14.pdf.

Cunningham, C. M. (2017). **Engineering in elementary STEM education: Curriculum design, instruction, learning, and assessment**. Teachers College Press.

Gartner (2022). **Gartner Forecasts Worldwide IT Spending to Grow 3% in 2022**. Gartner. Retrieved August 4, 2022, https://www.gartner.com/en/newsroom/press-releases/2022-06-14-gartner-forecasts-worldwide-it-spending-to-grow-3-percent-in-2022#:~:text=Worldwide%20IT%20spending%20is%20projected,latest%20forecast%20by%20Gartner%2C%20Inc.

Gasemagha, A. A., & Kowang, T. O. (2021). Project manager role in project management success. **International Journal of Academic Research in Business and Social Sciences**, 11(13), 1345-1355.

Hope (2022). **Engineering Mindset.** Center for Work Life. Retrieved Aug 4, 2022 from https://centerforworklife.com /emotional-intelligenceself-improvement/engineering-mindset.

Jamieson, M., & Donald, J. (2020). **Building the engineering mindset: Developing leadership and management competencies in the engineering curriculum**. Proceedings of the Canadian Engineering Education Association (CEEA).

Katz, R. (1955). **Skills of an effective administrator**. Harvard BusinessReview.

Lottero‐Perdue, P. S., & Lachapelle, C. P. (2020). Engineering mindsets and learning outcomes in elementary school. **Journal of Engineering Education**, 109(4), 640-664.

McManus, J., & Wood-Harper, T. (2008). **A study in project failure**. British Computer Society.

Mumford, M. D., Zaccaro, S. J., Harding, F. D., Jacobs, T. O., & Fleishman, E. A. (2000). Leadership skills for a changing world: Solving complex social problems. **The Leadership Quarterly**, 11(1), 11-35.

Patanakul, P. (2011). Project manager assignment and its impact on multiple project management effectiveness: An empirical study of an IT organization. **Engineering management journal**, 23(4), 14-23.

Pia Lappalainen (2009) Communication as part of the engineering skills set, **European Journal of Engineering Education**, 34:2, 123-129, DOI: 10.1080/03043790902752038.

Pinto, J. K., & Mantel, S. J. (1990). The causes of project failure. **IEEE transactions on engineering management**, 37(4), 269-276.

PMBOK Guide. (2017). **A guide to the project management body of knowledge**. (6th ed.). Project Management Institute.

Project Academy. (2018). **What is Engineering Mind-set thinking?**. Project Academy. Retrieved August 4, 2022, http://www.projectacademy.org/Documents/Engineering-Mind-set-summary-11.pdf.

Project Management Institute. (2021). **Project Management Body of Knowledge**. (7th ed.). Project Management Institute.

Project Management Institute. **Who are Project Managers?**. PMI. Retrieved August 4, 2022 from https://www.pmi.org /about /learn-about-pmi/who-are-project-managers/.

Ricco, G. D., Girtz, S., & Silliman, S. E. (2017). **Exploring Engineering Mindset**. In 2017 ASEE Annual Conference & Exposition.

The Standish Group. (n.d.). **Benchmarks and Assessments**. Standish. Retrieved August 4, 2022 from https://www.standish group.com/benchmark.

Tohidi, H. (2011). Human Resources Management main role in Information Technology project management. **Procedia Computer Science**, 3, 925-929.

Umaji, K., & Paireekreng, W. (2021). **A Study of Project Manager Skills for Enterprise Data Platform Project**. In 2021 2nd International Conference on Big Data Analytics and Practices (IBDAP), 84-89.

การแจกแจงการิมากำลัง-เรย์ลี: คุณสมบัติและการประยุกต์ใช้

อิศราภรณ์ เทียมศร[1], เอกภักดิ์ ตันประยูร[2] และ ศิรินภา อายุยืน[3,*]

Received: 14 November 2022; Revised: 15 January 2023; Accepted: 24 January 2023

บทคัดย่อ

งานวิจัยนี้ ได้นำเสนอตัวแบบช่วงชีวิตที่อยู่ในวงศ์ของตัวแบบวางนัยการิมากำลัง ซึ่งเรียกว่าการแจกแจงการิมากำลัง-เรย์ลี มีการศึกษาคุณสมบัติทางสถิติของการแจกแจงนี้ ได้แก่ ฟังก์ชันควอนไทล์ ฟังก์ชันอยู่รอด โมเมนต์ และสถิติอันดับ มีการประมาณค่าพารามิเตอร์ของการแจกแจกแจงการิมากำลัง-เรย์ลีด้วยวิธีภาวะน่าจะเป็นสูงสุด รวมทั้งมีการประยุกต์ใช้การแจกแจงการิมากำลัง-เรย์ลีกับข้อมูลจริง ซึ่งเป็นข้อมูลช่วงชีวิต จำนวน 2 ชุด ผลการศึกษาพบว่าการแจกแจงการิมากำลัง-เรย์ลี เป็นตัวแบบที่อธิบายลักษณะการแจกแจงความน่าจะเป็นของค่าสังเกตได้ใกล้เคียงกับข้อมูลจริง นั่นคือการแจกแจงนี้มีความยืดหยุ่นและครอบคลุมกับข้อมูลช่วงชีวิตมากกว่าการแจกแจงอื่น ๆ ที่เกี่ยวข้อง เช่น การแจกแจงเรย์ลี การแจกแจงการิมากำลัง-ลินด์เลย์ และการแจกแจงลินด์เลย์

คำสำคัญ: การแจกแจงการิมากำลัง-เรย์ลี, ข้อมูลช่วงชีวิต, การประมาณค่าภาวะน่าจะเป็นสูงสุด, ฟังก์ชันอยู่รอด

[*] *Corresponding Author Email : sirinapa_a@rmutt.ac.th*

[1,2,3] ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

Power Garima-Rayleigh distribution: properties and applications

Issaraporn Thiamsorn[1], Ekapak Tanprayoon[2] and Sirinapa Aryuyuen[3,*]

Received: 14 November 2022; Revised: 15 January 2023; Accepted: 24 January 2023

Abstract

This paper considers a new lifetime model of the class of power Garima generalized models, called the power Garima-Rayleigh (PG-R) distribution. Some statistical properties are provided, including quantile function, hazard function, moments, and order statistics. The maximum likelihood estimation is used to estimate the parameters of the PG-R distribution. In addition, we apply the PG-R distribution with two real-life data sets. The study found that the PG-R distribution is a model that describes the probability distribution of observation values close to the real data. The proposed distribution is more flexible and comprehensive with lifetime data than other related distributions, such as the Rayleigh, power Garima-Lindley, and Lindley distributions.

Keywords: Power Garima-Rayleigh distribution, lifetime data, maximum likelihood estimation, hazard function

[*] *Corresponding Author Email :* sirinapa_a@rmutt.ac.th

[1,2,3] Department of Mathematics and Computer Science, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi

1. Introduction

Many new statistical distributions have been derived using the commonly known distributions through different types of transformations, compounding, or mixing. Flexible distributions are obtained from extending the classical distributions by introducing one or more additional parameter(s) to the baseline distribution. Many generalized families of distributions have been proposed and studied for modelling data in many applied areas such as economics, engineering, biological studies, environmental sciences, medical sciences, and finance (Gupta et al., 1998). Some of the generated (G) families are: the beta-G (Eugene et al., 2002), gamma-G (Zografos & Balakrishnan, 2009; Ristic & Balakrishnan, 2012), transformed-transformer (T-X; Alzaatreh et al., 2013), Weibull-G (Bourguignon et al., 2014), exponentiated half-logistic-G (Cordeiro et al., 2014), half logistic-G family (Cordeiro et al., 2016; Soliman et al., 2017), exponentiated Weibull-G (Hassan & Elgarhy, 2016), generalized transmuted-G (Nofal et al., 2017), Gompertz-G (Alizadeh et al., 2017), beta Weibull-G (Yousof et al., 2017), new Weibull-G (Ahmad et al., 2018), and transmuted Gompertz-G (Reyad et al., 2018).

Recently, Aryuyuen et al. (2021) proposed a power Garima-G (PG-G) family of distributions, which is obtained by using the concept of T-X family by Alzaatreh et al. (2013) when a generator T distributed as the power Garima (PG) distribution. The cumulative density function (cdf) and the probability density function (pdf) of a random variable X distributed as the PG-G family are represented, respectively.

$$\mathrm{F}(x;\alpha,\beta,\lambda)=1-\left[1+\frac{\beta}{2+\beta}\left(\frac{\mathrm{G}^{\alpha}(x;\xi)}{1-\mathrm{G}^{\alpha}(x;\xi)}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{\mathrm{G}^{\alpha}(x;\xi)}{1-\mathrm{G}^{\alpha}(x;\xi)}\right)^{\lambda}\right\},\ x>0 \tag{1}$$

and

$$\mathrm{f}(x;\alpha,\beta,\lambda)=\frac{\alpha\beta\lambda\,\mathrm{g}(x;\xi)\mathrm{G}^{\alpha-1}(x;\xi)}{(2+\beta)\left[1-\mathrm{G}^{\alpha}(x;\xi)\right]^{2}}\left[(1+\beta)+\beta\left(\frac{\mathrm{G}^{\alpha}(x;\xi)}{1-\mathrm{G}^{\alpha}(x;\xi)}\right)^{\lambda}\right]$$
$$\times\left(\frac{\mathrm{G}^{\alpha}(\mathrm{x};\xi)}{1-\mathrm{G}^{\alpha}(\mathrm{x};\xi)}\right)^{\lambda-1}\exp\left\{-\beta\left(\frac{\mathrm{G}^{\alpha}(\mathrm{x};\xi)}{1-\mathrm{G}^{\alpha}(\mathrm{x};\xi)}\right)^{\lambda}\right\}, \tag{2}$$

where $\alpha,\beta,\lambda>0,$ and $\mathrm{G}(x;\xi)$ is the cdf of any existing (baseline) distribution with a parameter vector ξ.

In practice, the PG-G family of distributions: (a) is the primary method of introducing an additional parameter(s) to generate an extended version of the baseline distribution; (b) improve the characteristics of the traditional distributions; (c) make the kurtosis more flexible compared to the baseline distribution; (d) generated distributions that have the pdf with various shapes, i.e., symmetric, right-skewed, left-skewed, and reversed-J shaped; (e) define special models with all types of hazard rate function; (f) define special models having a closed-form for cdf, survival function as well as hazard rate function; (g) provide consistently better fits than other generated distributions having the same or higher number of parameters. Thus, we are interested in constructing a new distribution to analyse lifetime data using the PG-G family of distributions using some baseline distribution. This study uses a Rayleigh distribution as the baseline distribution (Aryuyuen et al., 2021).

Rayleigh (1880) derived the Rayleigh distribution; it was introduced in connection with a problem in the field of acoustics. The cdf and pdf of the Rayleigh distribution are given by

$$\mathrm{G}(x;\sigma)=1-e^{-x^{2}/(2\sigma^{2})}\ \text{and}\ \mathrm{g}(x;\sigma)=\frac{1}{\sigma^{2}}xe^{-x^{2}/(2\sigma^{2})}\ \text{for}\ x>0\ \text{and}\ \sigma>0. \tag{3}$$

In nature, physical phenomena in many fields of science (for example, noise theory, lethality, radar return, and others) have amplitude distributions that the Rayleigh density function can characterize. Since then, extensive work has occurred related to this distribution in different areas of science and technology. It has some relations with well-known distributions like the Weibull, chi-square, or extreme value distributions. The hazard function of the Rayleigh distribution is an increasing function of time (Rayleigh, 1880; Dey et al., 2014; Al-Babtain, 2020). Statistical inference of one parameter Rayleigh distribution, such as parameter estimation, predictions, and testing, have been extensively studied by several authors (See Johnson et al., 1994; Abd-Elfattah et al., 2006; Dey & Das, 2007; Dey, 2009; Al-Babtain, 2020).

This paper proposes a new lifetime model of the class of power Garima generalized models called the power Garima-Rayleigh distribution. Some statistical properties of the proposed distribution are provided. The algorithm for generating a random number from the proposed distribution is shown. We estimate the model

parameters using the maximum likelihood (ML) method. Moreover, applications are illustrated. Finally, we offer some concluding remarks.

2. Methods

The methods of this research are:

2.1 The cdf and pdf of the proposed distribution and its mathematical properties are studied.

2.2 The appropriate parameter estimation method for the proposed distribution using the ML method is provided.

2.3 Application of the proposed distribution to real data sets has been studied by comparing it to some existing distributions using the criteria of Akaike information criterion (AIC), corrected Akaike information criterion (CAIC), and Kolmogorov-Smirnov (D*) statistics.

3. Results and Discussion

This section proposes a new distribution, namely the power Garima-Rayleigh distribution. We provided some statistical properties, parameter estimation, and applications of the proposed distribution. The remaining part of this article is presented as follows:

3.1 The power Garima-Rayleigh distribution

Let X be a random variable with the cdf of the PG-G family of distributions as in (1), when $G(x;\sigma)$ is the cdf of the Rayleigh distribution as (3). Then the cdf of X is

$$F(x;\boldsymbol{\theta}) = 1 - \left[1 + \frac{\beta}{2+\beta}\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right\},\ x>0 \tag{4}$$

where $\boldsymbol{\theta} = (\alpha,\beta,\lambda,\sigma)^T$ and $\alpha,\beta,\lambda,\sigma > 0.$ Its corresponding pdf is

$$f(x;\boldsymbol{\theta}) = \frac{\alpha\beta\lambda x e^{-x^2/(2\sigma^2)}}{(2+\beta)\sigma^2}\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha-1}}{\left(1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}\right)^2}\left[(1+\beta)+\beta\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]$$
$$\times\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda-1}\exp\left\{-\beta\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right\}. \tag{5}$$

We have a random variable X with the cdf as (4) and pdf as (5) distributed as the power Garima-Rayleigh (PG-R) distribution with parameters α, β, λ and σ, denoted by $X \sim \text{PG-R}(\alpha,\beta,\lambda,\sigma)$. Some pdf plots of PG-R distribution with different values of α, β, λ and σ are shown in Figure1.

3.2 Some statistical properties

We derive statistical properties of the PG-R distribution, including survival and hazard functions, moments, quantile function, skewness and kurtosis, algorithm for generating random variables, and order statistics.

3.2.1 Survival and hazard functions

The survival function $S(x)$, is the probability that a subject survives longer than time x, $S(x) = 1 - F(x).$ The survival function of the PG-R distribution is

$$S(x;\boldsymbol{\theta}) = \left[1 + \frac{\beta}{2+\beta}\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right\},\ x>0. \tag{6}$$

From the expression of the hazard function $\mathrm{h}(x)$, $\mathrm{h}(x)=\mathrm{f}(x)/\mathrm{S}(x)$, we have the hazard function of the PG-R distribution, that is

$$\mathrm{h}(x;\boldsymbol{\theta})=\frac{\alpha\beta\lambda x e^{-x^2/(2\sigma^2)}}{(2+\beta)\sigma^2}\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha-1}}{\left(1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}\right)^2}\left[(1+\beta)+\beta\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]$$
$$\times\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda-1}\left[1+\frac{\beta}{2+\beta}\left(\frac{\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]^{-1}. \qquad (7)$$

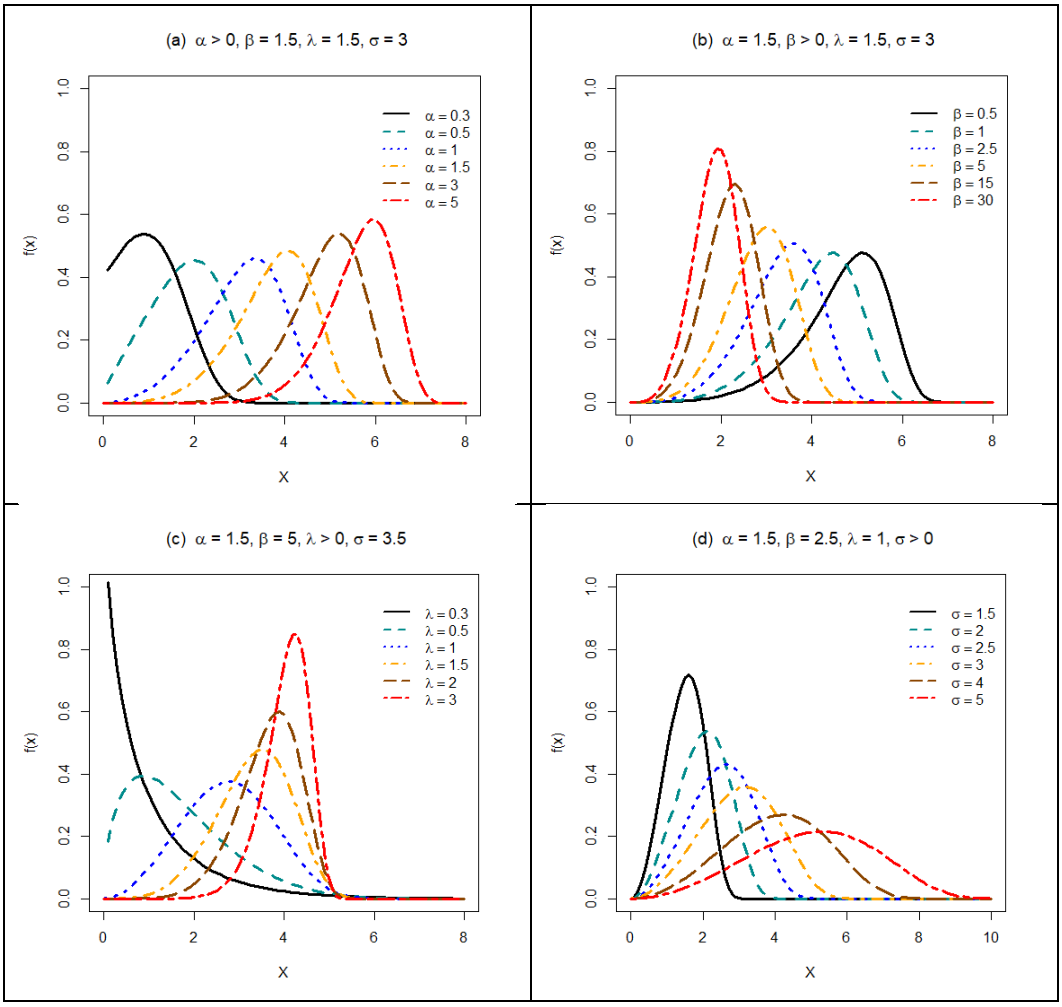


Figure 1 Some pdf plots of the PG-R distribution with different values of α, β, λ and σ.

3.2.2 Moments

The r^{th} moments of the PG-G family of distributions (Aryuyuen et al., 2021) is

$$\mathrm{E}(X^r)=\frac{\alpha\beta\lambda}{2+\beta}\sum_{i,j=0}^{\infty}\sum_{r=0}^{\infty}\phi_r\left[(1+\beta)\delta_1(i,j)+\beta\delta_2(i,j)\right],\ \text{for}\ \phi_r=\int_{-\infty}^{\infty}x^r b_{i,j}^{\alpha}(x;\xi)\,dx,\ \text{where}$$

$$b_{i,j}^{\alpha}(x;\xi)=\frac{g(x;\xi)}{\left[1-G^{\alpha}(x;\xi)\right]G(x;\xi)}\left(\frac{G^{\alpha}(x;\xi)}{1-G^{\alpha}(x;\xi)}\right)^{j}\ \text{for}\ i,j=0,1,2,\ldots\ \text{and}$$

$\delta_k(i,j) = \dfrac{\Gamma\left[\lambda(i+k)+1\right]}{\Gamma(j+1)\Gamma\left[\lambda(i+k)-j+1\right]}\left(\dfrac{\beta^i(-1)^{i+j}}{i!}\right)^j$ for $k = 1,2$ and $\Gamma(.)$ is a gamma function.

From the cdf and pdf in (3), the r^{th} moments of the PG-R distribution is

$$\mathrm{E}(X^r) = \frac{\alpha\beta\lambda}{(2+\beta)}\sum_{i,j=0}^{\infty}\sum_{r=0}^{\infty}\phi_r\left[(1+\beta)\delta_1(i,j)+\beta\delta_2(i,j)\right], \text{ where} \tag{8}$$

$$\phi_r = \frac{1}{\sigma^2}\int_{-\infty}^{\infty}\left\{\frac{x^{r+1}e^{-x^2/(2\sigma^2)}}{(1-e^{-x^2/(2\sigma^2)})\left[1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}\right]}\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^j\right\}dx\,.$$

From the moment of the PG-R distribution, we have the mean and variance are respectively

$$\mathrm{E}(X) = \frac{\alpha\beta\lambda}{2+\beta}\sum_{i,j=0}^{\infty}\sum_{r=0}^{\infty}\phi_1\left[(1+\beta)\delta_1(i,j)+\beta\delta_2(i,j)\right], \text{ and}$$

$$V(X) = \frac{\alpha\beta\lambda}{(2+\beta)}\left\{\sum_{i,j=0}^{\infty}\sum_{r=0}^{\infty}\phi_2\left[(1+\beta)\delta_1(i,j)+\beta\delta_2(i,j)\right]-\frac{\alpha\beta\lambda}{2+\beta}\left(\sum_{i,j=0}^{\infty}\sum_{r=0}^{\infty}\phi_1\left[(1+\beta)\delta_1(i,j)+\beta\delta_2(i,j)\right]\right)^2\right\}.$$

3.2.3 Quantile function and related measurements

Let X be a random variable distributed as the PG-R distribution with cdf as (4). Theoretically, $\mathrm{F}(x;\alpha,\beta,\lambda,\sigma) = U$, where U is a uniform random variable on the interval (0,1). Then the quantile function (qf) of the PG-R distribution is obtained by inverting equation (4) as follows:

$$\mathrm{Q}(u;\alpha,\beta,\lambda,\sigma) = \sigma\left\{-2\log\left[1-\left(1+\left[-\frac{1}{\beta}\left(\mathrm{W}_{-1}\left[-(1-u)(2+\beta)\exp\left(-2-\beta\right)\right]+\beta+2\right)\right]^{-1/\lambda}\right)^{-1/\alpha}\right]\right\}^{1/2} \tag{9}$$

where $\mathrm{W}_{-1}[\cdot]$ denotes the negative branch of the `LambertW` function (see Corless et al., 1996; Veberic, 2012), i.e., $\mathrm{W}_{-1}[z]\exp\left\{\mathrm{W}_{-1}[z]\right\} = z$, where z is a complex number.

Median: In particular, the median can be derived from (9) be setting $u = 0.5$. Then, the median is given by

$$\text{Median} = \sigma\left\{-2\log\left[1-\left(1+\left[-\frac{1}{\beta}\left(\mathrm{W}_{-1}\left[-\left(1+\frac{\beta}{2}\right)\exp\left(-2-\beta\right)\right]+\beta+2\right)\right]^{-1/\lambda}\right)^{-1/\alpha}\right]\right\}^{1/2}.$$

Algorithm for generating variate: We can generate a random variable of X_i from the PG-R distribution using the following algorithm.

Step 1: Generate U_i from the uniform distribution on the interval (0, 1) for $i = 1,2,...,n$.

Step 2: Generate X_i by using the qf of the PG-R distribution in (9), $X_i = \mathrm{Q}(u_i;\alpha,\beta,\lambda,\sigma)$.

For $\mathrm{W}_{-1}[\cdot]$ in equation (9), it computes with `lambertWm1` function in the **LamW** package (Adler, 2015) in contribution package in R (R Core Team, 2022).

Since the moments of the PG-R distribution is not a close form. Thus, the expression of mean, standard deviation, skewness, and kurtosis is in the close form, too. However, these values are shown in terms of the simulation study. Some simulations are performed for illustration purposes of how the skewness of the PG-R distribution with different values of each parameter that is appeared in Figure1. Random samples with size 1,000 are generated 10,000 times, that is, $X_i = \mathrm{Q}(u_i;\alpha,\beta,\lambda,\sigma)$. The results, including sample mean, variance, skewness value (SV), and kurtosis value (KV), are shown in Table 1. Skewness can be used as a measure of the symmetry of distribution for SV = 0. It is defined as a symmetrical distribution (left-skewed for SV < 0 and right-skewed for SV > 0). Meanwhile, the KV is often compared to the kurtosis of the normal distribution that is KV = 3. If the KV

> 3, the dataset has heavier tails than a normal distribution. Moreover, the dataset has lighter tails than a normal distribution for KV < 3.

We can consider the effects of the shape parameters on the pdf-plotted of PG-R distribution in Figure 1, and the values of skewness and kurtosis values in Table 1, which the results are follows. For fixed other parameters but α changed, the distribution shapes are right-skewed and left-skewed when $\alpha < 1$ and $\alpha \geq 1$, respectively (see Figure 1 (a)). For fixed other parameters but β changed, the distribution shapes are left-skewed, and is close to symmetric when β increase (see Figure 1 (b)). Figure 1 (c) shows that the various shape when different values of λ (for fixed other parameters); the distribution is decreasing function when $\lambda \leq 0.5$, but it is unimodal distribution when $\lambda \geq 0.5$; for $0.5 \leq \lambda < 1$ the distribution is right-skewed; for $\lambda \geq 1$ the distribution is left-skewed. Figure 1 (d) shows that the values for σ increase and the distribution is flatten but is still left-skewed.

Table 1 Results of the average values of the sample mean, variance, skewness, and kurtosis of PG-R distribution.

Figure 1	α	β	λ	σ	sample mean	variance	SV	KV
(a) $\alpha \uparrow$	*5*	1.5	1.5	3	5.6669	0.5127	-0.6228	3.3300
	3	1.5	1.5	3	4.8972	0.5995	-0.6045	3.2596
	1.5	1.5	1.5	3	3.7586	0.7113	-0.5175	3.0176
	1	1.5	1.5	3	3.0525	0.7416	-0.4039	2.7756
	0.5	1.5	1.5	3	1.8474	0.6360	-0.0525	2.3674
	0.3	1.5	1.5	3	1.0642	0.4123	0.3676	2.4358
(b) $\beta \uparrow$	1.5	*30*	1.5	3	1.8720	0.2296	-0.1768	2.7719
	1.5	*15*	1.5	3	2.2005	0.3124	-0.2073	2.7692
	1.5	*5*	1.5	3	2.8518	0.4901	-0.3089	2.8009
	1.5	*2.5*	1.5	3	3.3514	0.6198	-0.4168	2.8867
	1.5	*1*	1.5	3	4.0974	0.7743	-0.6058	3.1648
	1.5	*0.5*	1.5	3	4.6923	0.8586	-0.7637	3.5045
(c) $\lambda \uparrow$	1.5	5	*3*	3.5	4.0460	0.2569	-0.7412	3.7268
	1.5	5	*2*	3.5	3.6649	0.4671	-0.5139	3.1472
	1.5	5	*1.5*	3.5	3.3270	0.6666	-0.3089	2.7987
	1.5	5	*1*	3.5	2.7644	0.9777	0.0555	2.5559
	1.5	5	*0.5*	3.5	1.7010	1.2950	0.9253	3.6533
	1.5	5	*0.3*	3.5	1.0197	1.2068	1.9518	7.7386
(d) $\sigma \uparrow$	1.5	2.5	1	*5*	5.0469	2.9459	-0.0967	2.4988
	1.5	2.5	1	*4*	4.0365	1.8857	-0.0961	2.4986
	1.5	2.5	1	*3*	3.0278	1.0605	-0.0959	2.4981
	1.5	2.5	1	*2.5*	2.5227	0.7364	-0.0965	2.4979
	1.5	2.5	1	*2*	2.0185	0.4715	-0.0961	2.4971
	1.5	2.5	1	*1.5*	1.5141	0.2650	-0.0969	2.4976

3.2.4 Order statistics

Let $X_{(1)},\ldots,X_{(n)}$ denote the order statistics of a random sample X_i for $i = 1,2,...,n$ from the PG-R distribution with the pdf in equation (5) and cdf in equation (4); then the pdf of $X_{(j)}$ for $j = 1,2,...,n$ is

$$f_{X_{(j)}}(x) = \frac{\Gamma(n+1)\alpha\beta\lambda x(1-e^{-x^2/(2\sigma^2)})^{\alpha-1}e^{-x^2/(2\sigma^2)}}{\Gamma(j)\Gamma(n-j+1)\sigma^2(2+\beta)\left[1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}\right]^2}\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda-1}$$

$$\times\left[(1+\beta)+\beta\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right\}$$

$$\times\left(1-\left[1+\frac{\beta}{2+\beta}\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right\}\right)^{j-1}$$

$$\times\left(\left[1+\frac{\beta}{2+\beta}\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right]\exp\left\{-\beta\left(\frac{(1-e^{-x^2/(2\sigma^2)})^{\alpha}}{1-(1-e^{-x^2/(2\sigma^2)})^{\alpha}}\right)^{\lambda}\right\}\right)^{n-j}.$$

3.3 Parameter estimation

The ML estimates of the unknown parameters for the PG-R distribution are determined based on complete samples. Let $X_1,\ldots,X_n$ be a random sample of size n from the PG-R distribution with set of parameters $\boldsymbol{\theta}=(\alpha,\beta,\lambda,\sigma)^T$. The log-likelihood function of $\boldsymbol{\theta}$ can be expressed as

$$\ell(x_i;\boldsymbol{\theta}) = n\log(\alpha)+n\log(\beta)+n\log\lambda-n\log(2+\beta)-2n\log(\sigma)+\sum_{i=1}^{n}\log x_i-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2$$
$$+(\alpha\lambda-1)\sum_{i=1}^{n}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]-(\lambda+1)\sum_{i=1}^{n}\log\left(1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right)$$
$$+\sum_{i=1}^{n}\log\left[1+\beta+\beta\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]-\beta\sum_{i=1}^{n}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}.$$

The elements of the score function $\mathrm{U}(\boldsymbol{\theta})=(\mathrm{U}_{\alpha},\mathrm{U}_{\beta},\mathrm{U}_{\lambda},\mathrm{U}_{\sigma})$ are given by

$$\mathrm{U}_{\alpha}=\frac{n}{\alpha}+\lambda\sum_{i=1}^{n}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]$$
$$+(\lambda+1)\sum_{i=1}^{n}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]\right)-\beta\lambda\sum_{i=1}^{n}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha\lambda}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]}{\left[1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right]^{\lambda+1}}\right)$$
$$+\beta\lambda\sum_{i=1}^{n}\left[1+\beta+\beta\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]^{-1}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha\lambda}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]}{\left[1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right]^{\lambda+1}}\right),$$

$$\mathrm{U}_{\beta}=\frac{n}{\beta}-\frac{n}{2+\beta}$$
$$+\sum_{i=1}^{n}\left[(1+\beta)+\beta\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]^{-1}\left[1+\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]-\sum_{i=1}^{n}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda},$$

$$\mathrm{U}_{\lambda}=\frac{n}{\lambda}+\alpha\sum_{i=1}^{n}\log\left[1-e^{-x_i^2/(2\sigma^2)}\right]-\sum_{i=1}^{n}\log\left(1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right)$$
$$+\beta\sum_{i=1}^{n}\left[1+\beta+\beta\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]^{-1}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\log\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)$$
$$-\beta\sum_{i=1}^{n}\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\log\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right),$$

$$U_\sigma = -\frac{2n}{\sigma}+\frac{1}{\sigma^3}\sum_{i=1}^{n}x_i^2-(\alpha\lambda-1)\sum_{i=1}^{n}\left(\frac{x_i^2e^{-x_i^2/(2\sigma^2)}}{1-e^{-x_i^2/(2\sigma^2)}}\right)$$
$$+\frac{\alpha(\lambda+1)}{\sigma^3}\sum_{i=1}^{n}\left(\frac{x_i^2\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha-1}e^{-x_i^2/(2\sigma^2)}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)+\frac{\alpha\beta\lambda}{\sigma^3}\sum_{i=1}^{n}\frac{x_i^2e^{-x_i^2/(2\sigma^2)}\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha\lambda-1}}{\left(1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right)^{\lambda+1}}$$
$$-\frac{\alpha\beta\lambda}{\sigma^3}\sum_{i=1}^{n}x_i^2\left[1+\beta+\beta\left(\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}{1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}}\right)^{\lambda}\right]^{-1}\frac{\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha\lambda-1}e^{-x_i^2/(2\sigma^2)}}{\left(1-\left[1-e^{-x_i^2/(2\sigma^2)}\right]^{\alpha}\right)^{\lambda+1}}.$$

Then the ML estimates of the parameters α, β, λ and σ are obtained by setting the last two equations to be zero and solving them. Clearly, it is difficult to solve them, therefore computer packages. In this study, the `nlm` function in the **stats** package and contribution package in R (R Core Team, 2022) is used to find the ML estimates.

3.4 Application study

In this section, we demonstrate the flexibility and the potentiality of the PG-R distribution through two real data sets, which data set are:

Data I: The data set is gauge lengths of 10 mm (Kundu & Raqab, 2009). This data set consists of 63 observations: 1.901, 2.132, 2.203, 2.228, 2.257, 2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614, 2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977, 2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332, 3.346, 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886, 3.971, 4.024, 4.027, 4.225, 4.395, 5.020.

Data II: The second data is the strength data set originally reported by Badar and Priest in 1982 (see Dey, 2014). It describes the strength measured in GPA for single-carbon fibers and impregnated 1000-carbon fiber tows. This data set consists of 62 observations: 0.5620, 0.5640, 0.7290, 0.8020, 0.9500, 1.0530, 1.1110, 1.1940, 1.2080, 1.2160, 1.2470, 1.2560, 1.2710, 1.2777, 1.3050, 1.3130, 1.3900, 1.4290, 1.4740, 1.4900, 1.5030, 1.5200, 1.5220, 1.5240, 1.5510, 1.6090, 1.6320, 1.6320, 1.6760, 1.6840, 1.6850, 1.7280, 1.7400, 1.7640, 1.7850, 1.8040, 1.8160, 1.8240, 1.8360, 1.8790.

We use these real data sets to illustrate the importance and flexibility of the PG-R distribution and compare it with some lifetime distributions which related the PG-R distribution and the PG-G family of distributions with four-parameter, such as the Rayleigh (R), power Garima-Lindley (PG-L; Aryuyuen et al., 2021), and Lindley (L) distributions. All the model parameters are estimated by the ML method for the proposed model. These applications will be used to determine the estimated parameters of each distribution. The value of ML estimates of each distribution is obtained using the `nlm` function in the stats package and contribution package in R (R Core Team, 2022). The results are presented in Tables 2-3. In each application, we compare the fit of the PG-G distribution with the fit of the corresponding baseline distribution (R distribution), the PG-G family with a four-parameter distribution (PG-L distribution), and the L distribution. The criteria of the Akaike information criterion (AIC), the corrected Akaike information criterion (CAIC), and Kolmogorov-Smirnov (D*) statistics are used. The model gives the smallest values of AIC, CAIC, and D*. Therefore, it is the best model for fitting data.

Table 2 ML estimates of the model parameters and goodness of measures for their estimates for Data I.

Distributions	ML estimates	AIC	CAIC	D* (p-value)
PG-R	$\hat{\alpha}=40.3156,\ \hat{\beta}=2.3421,$ $\hat{\lambda}=0.3076,\ \hat{\sigma}=1.3118$	121.19	121.88	0.0770 (0.8492)
R	$\hat{\sigma}=2.2067$	189.04	189.11	0.3607 (<0.0001)
PG-L	$\hat{\alpha}=6.5264,\ \hat{\beta}=216.0021,$ $\hat{\lambda}=0.7011,\ \hat{\tau}=0.2743$	129.65	130.34	0.0810 (0.8031)
L	$\hat{\tau}=0.5392$	244.72	244.78	0.4308 (<0.0001)

Table 3 ML estimates of the model parameters and goodness of measures for their estimates for Data II.

Distribution	ML estimates	AIC	CAIC	D* (p-value)
PG-R	$\hat{\alpha} = 2.5351,\ \hat{\beta} = 114.5274,$ $\hat{\lambda} = 0.7918,\ \hat{\sigma} = 4.1840$	95.81	86.51	0.0490 (0.9984)
R	$\hat{\sigma} = 1.2464$	121.60	121.67	0.2551 (<0.0001)
PG-L	$\hat{\alpha} = 1.4544,\ \hat{\beta} = 21.2383,$ $\hat{\lambda} = 2.1026,\ \hat{\tau} = 0.4522$	95.98	96.68	0.0559 (0.9902)
L	$\hat{\tau} = 0.9015$	176.00	176.06	0.5454 (<0.0001)

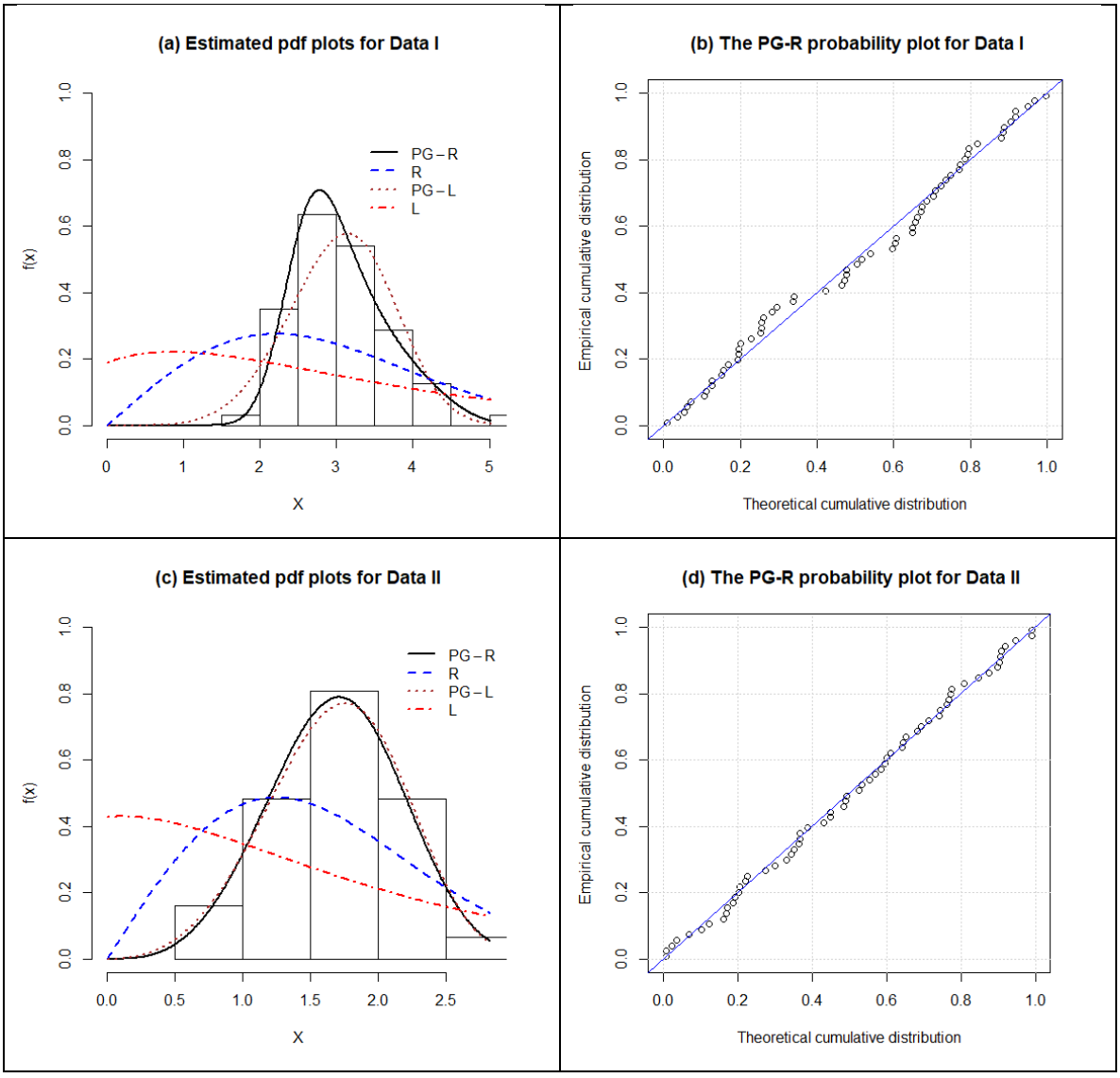


Figure 2 Plots of the estimated pdf for each distribution and the PG-R probability plot for the data sets.

Figure 2 shows the plots of the estimated pdf of the fitted models for the data sets. The numerical values of the AIC, CAIC, and D* statistics are listed in Tables 2-3. The PG-R model gives the lowest values for the AIC, CAIC, and D* statistics for the data set among the fitted model. The PG-R distribution could fit the data better than the PG-L, R, and L distributions. A density plot compares the fitted densities of the models with the empirical histogram of the observed data (Figure 2).

4. Conclusion

This paper proposes a new four-parameter distribution, namely the PG-R distribution. The PG-R distribution is motivated by the wide use of the Rayleigh distribution in practice and the fact that generalization provides more flexibility to analyze positive real-life data. We derive explicit expressions for the quantile function and order statistics. We also provide a foundation for some mathematical properties of this distribution, including the derivation of the hazard rate function, moments, quantile function, generating numbers, and order statistics. The model parameters are estimated by maximum likelihood. Applying the PG-R distribution to a real data set indicates that the new distribution outperforms several distributions, including the PG-L, Rayleigh, and Lindley distributions.

Acknowledgement

The authors gratefully acknowledge participating in the Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi, Thailand. We also thank those who could not be mentioned here for their kindness and encouragement. Moreover, the authors thank the anonymous reviewers for their comments and suggestions.

References

Abd Elfattah, A. M., Hassan, A. S., & Ziedan, D. M. (2006). Efficiency of maximum likelihood estimators under different censored sampling schemes for Rayleigh distribution. **Interstat**, 1, 1-16.

Adler A. (2015). **lamW: Lambert-W Function**. CRAN.R-project. Retrieved November 1, 2022 from https://CRAN.R-project.org/package=lamW.

Ahmad, Z., Elgarhy, M., & Hamedani, G. G. (2018). A new Weibull-X family of distributions: properties, characterizations and applications. **Journal of Statistical Distributions and Applications**, 5(1), 1-18.

Al-Babtain, A. A. (2020). A new extended Rayleigh distribution. **Journal of King Saud University-Science**, 32(5), 2576-2581.

Alizadeh, M., Cordeiro, G. M., Pinho, L. G. B., & Ghosh, I. (2017). The Gompertz-G family of distributions. **Journal of Statistical Theory and Practice**, 11(1), 179-207.

Alzaatreh, A., Lee, C., & Famoye, F. (2013). A new method for generating families of continuous distributions. **Metron**, 71(1), 63-79.

Aryuyuen, S., Bodhisuwan, W., & Ngamkham, T. (2021). Power Garima-generated family of distributions: properties and application. **Lobachevskii Journal of Mathematics**, 42(2), 287-299.

Bourguignon, M., Silva, R. B., & Cordeiro, G. M. (2014). The Weibull-G family of probability distributions. **Journal of data science**, 12(1), 53-68.

Cordeiro, G. M., Alizadeh, M., & Diniz Marinho, P. R. (2016). The type I half-logistic family of distributions. **Journal of Statistical Computation and Simulation**, 86(4), 707-728.

Cordeiro, G. M., Alizadeh, M., & Ortega, E. M. (2014). The exponentiated half-logistic family of distributions: Properties and applications. **Journal of Probability and Statistics**.

Corless, R. M., Gonnet, G. H., Hare, D. E., Jeffrey, D. J., & Knuth, D. E. (1996). On the LambertW function. **Advances in Computational mathematics**, 5(1), 329-359.

Dey, S. (2009). Comparison of Bayes estimators of the parameter and reliability function for Rayleigh distribution under different loss functions. **Malaysian Journal of Mathematical Sciences**, 3, 247-264.

Dey, S., & Das, M. K. (2007). A note on prediction interval for a Rayleigh distribution: Bayesian approach. **American Journal of Mathematical and Management Sciences**, 27(1-2), 43-48.

Dey, S., Dey, T., & Kundu, D. (2014). Two-parameter Rayleigh distribution: different methods of estimation. **American Journal of Mathematical and Management Sciences**, 33(1), 55-74.

Eugene, N., Lee, C., & Famoye, F. (2002). Beta-normal distribution and its applications. **Communications in Statistics-Theory and methods**, 31(4), 497-512.

Gupta, R.C., Gupta P.L., Gupta R.D. (1998). Modeling failure time data by Lehman alternatives. **Communications in Statistics-Theory and methods**, 27(4), 887-904.

Hassan, A. S., & Elgarhy, M. (2016). A new family of exponentiated Weibull - generated distributions. **International Journal of Mathematics and Its Applications**, 4(1-D), 135-148.

Johnson, N.L., Kotz, S. & Balakrishnan, N. (1994). **Continuous Univariate Distribution**. (2nd ed.). John Wiley & Sons.

Kundu, D., & Raqab, M. Z. (2009). Estimation of R= P (Y< X) for three-parameter Weibull distribution. **Statistics & Probability Letters**, 79(17), 1839-1846.

Nofal, Z. M., Afify, A. Z., Yousof, H. M., & Cordeiro, G. M. (2017). The generalized transmuted-G family of distributions. **Communications in Statistics-Theory and Methods**, 46(8), 4119-4136.

R Core Team (2022). **R: A language and environment for statistical computing**. R-project. Retrieved November 1, 2022 from https://www.R-project.org/.

Rayleigh, L. (1880). XII. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. **The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science**, 10(60), 73-78.

Reyad, H., Jamal, F., Othman, S., & Hamedani, G. G. (2018). The transmuted Gompertz-G family of distributions: properties and applications. **Tbilisi Mathematical Journal**, 11(3), 47-67.

Ristić, M. M., & Balakrishnan, N. (2012). The gamma-exponentiated exponential distribution. **Journal of statistical computation and simulation**, 82(8), 1191-1206.

Soliman, A. H., Elgarhy, M. A. E., & Shakil, M. (2017). Type II half logistic family of distributions with applications. **Pakistan Journal of Statistics and Operation Research**, 245-264.

Veberič, D. (2012). Lambert W function for applications in physics. **Computer Physics Communications**, 183(12), 2622-2628.

Yousof, H. M., Rasekhi, M., Afify, A. Z., Ghosh, I., Alizadeh, M., & Hamedani, G. G. (2017). The beta Weibull-G family of distributions: theory, characterizations and applications and applications. **Pakistan Journal of Statistics**, 33(2), 95-116.

Zografos, K., & Balakrishnan, N. (2009). On families of beta-and generalized gamma-generated distributions and associated inference. **Statistical methodology**, 6(4), 344-362.

Business Failure Prediction by Hybrid Data Mining Approach: A Case of Thailand Agribusiness

Jeerawadee Pumjaroen[1,*]

Received: 29 January 2023; Revised: 30 April 2023; Accepted: 16 May 2023

Abstract

The failure of a business could significantly impact private companies, the government, and the whole economy. Therefore, predicting business failure is always one major research problem in business and economics. Many methods, such as theoretical models, statistical models, and data mining techniques, were applied to predict business failures. This research developed a business failure prediction model to classify failed and non-failed companies from one to three years before the failure by a hybrid data mining technique. The interest of this research is to integrate clustering and classification techniques to predict business failure, which can be beneficial for further research related to business failure prediction or early warning models. The study involved 3,118 agribusiness companies that submitted their financial statements from 2016 to 2020 in Thailand. Based on the data of financial statements, a single classifier, including decision tree (DT), logistic regression (LR), and neural network (NN), was compared with a hybrid data mining technique—clustering and classification. The results showed that applying the hybrid method, k-mean and DT, helped to improve the business failure prediction performance.

Keywords: data mining, business failure, prediction, early warning model, classification

[*] *Corresponding Author Email : jeerawadee_p@rmutt.ac.th*
[1] Applied Statistics, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi (RMUTT)

1. Introduction

Business failure prediction has been a topic of interest and tends to gain more attention because of its importance to many parties—the private companies, the government, and the whole economy (Tsai, 2014; Wanke, Barros, & Faria, 2015). The failure of a business can cause social costs and the economy involving entrepreneurs, workers, lenders, suppliers, and clients, which are related to increased debt. Rising household debt possibly negatively affected the nation's economy (Pumjaroen, 2019; Pumjaroen, Vichitthamaros, & Sethapramote, 2020).

Both private and government sectors will benefit if the prediction is reliable (Geng, Bose, & Chen, 2015; Klepáč & Hampel, 2017; Lin & McClean, 2001; Pumjaroen & Sethapramote, 2023). Since business failure does not suddenly happen just in one day, the loss is a gradual process that evolves over a considerable period (Altman, 1968; Lincoln, 1984; Webb, 2003). Hence, the firms and the policymakers notice the early sign of business failure; they could initiate remedial measures to avoid deterioration before the collapse. The firm could make a strategic plan to reform itself; the investors have information to better manage their investment portfolios by reallocating the stocks' funds and avoiding companies that are about to fail. Financial sectors related to investment and lending would better assess the default risk Gepp, Kumar, and Bhattacharya (2010); (Tsai, 2014). The government policymaker will have a longer time to find a strategy for healing the economy. Therefore, accurate business failure prediction models would help the economy stabilize and increase for all involved.

Many empirical studies apply business failure as the outcome of financial distress (Geng et al., 2015). Since if the firm's financial health is weakened, it deals with financial distress, developing into a financial crisis and ending as a failure. Because financial analysis informs the company's standing, it helps identify scenarios leading to its financial health and management's quality decisions (Klepáč & Hampel, 2017).

There are many aliases in the area of business failure prediction, such as financial distress prediction, bankruptcy prediction, and business failure prediction. However, they are all related to the model aiming to predict the failure of a business before it actually happens. Some models apply only the available financial information associated with the company. Many studies include other attributes, such as industrial and economic indicators (Geng et al., 2015).

Since agribusiness is one of the driving sectors of Thailand's economy, this research aimed to develop its business failure prediction model to classify failed and non-failed companies. Data mining techniques have become popular alternatives in business failure prediction (Li, Sun, & Wu, 2010). Most of the business failure predictions in Thailand were conducted by a single classifier (Buanak, 2016; Narungsri, 2005). However, many recent studies indicate that a hybrid method, which combines multiple approaches, outperforms a single process (Alapati & Sindhu, 2016; Jader & Aminifar, 2022; Shalaby, Belal, & Omar, 2021). Moreover, some related studies of business failure prediction found that a hybrid method helps improve the model performance over using only a single process (Hsieh, 2005; West, Dellana, & Qian, 2005). Hence, this research focused on comparing the prediction performance of single and hybrid classified methods with different previous years' early warnings before the business failure. The interest of this research is to integrate the clustering and classification techniques to predict Thailand's agribusiness failure, which can be beneficial for further research related to the model of business failure prediction or early warning model. The result showed whether applying the clustering technique before classification helped improve Thailand's agribusiness business failure prediction.

The remainder of the article is organized as follows—Section Two reviews related studies on a business failure description and prediction model using data mining techniques. The research steps and empirical data collected for agribusiness failure prediction are described in Section Three. The results are discussed in Section Four. Finally, Section Five presents the conclusions of the empirical results and their implications.

2. Conceptual Background

2.1 Description of Business Failure

Many terms are used to describe when firms face financial difficulties, such as financial distress, business bankruptcy, firm default, and business failure (Geng et al., 2015). As for business failure, Dimitras, Zanakis, and Zopounidis (1996) defined it as a situation in which the firms could not continually operate.

The financial statement has long been applied to predict business failure. Since it is the accounting report assessing a firm's information focusing on the financial situation and the results of its operations, the financial statement helps provide useful information for making economic decisions. Altman (1968) was among the first researchers who used financial ratios to predict corporate bankruptcy.

In Thailand, a registered juristic person must report a financial statement to the Department of Business Development (DBD), Ministry of Commerce every year (Development, 2020). If the firm does not submit the financial report for consecutively three years, the officer will label its status as an unoccupied firm. In case the company cannot operate the business, it must report to the department for liquidation.

As for those mentioned above, the firm that registers the liquidation or does not report the financial statements for three consecutive years will be defined as a business failure for this research.

2.2 Business Failure Prediction Models

The business failure prediction model could be classified into three groups: statistical models (such as discriminant and regression analysis), data mining techniques (such as neural networks and decision trees), and theoretical models (such as expert evaluation and market risk models) (Klepáč & Hampel, 2017). Nowadays, many studies have applied hybrid methods to enhance prediction accuracy (Alapati & Sindhu, 2016; Klepáč & Hampel, 2017; Tsai & Chen, 2010). The procedure combines more than one model from statistical and machine learning techniques to improve prediction performance. Recent studies of business failure prediction have also used a hybrid method and confirmed that combining multiple approaches outperforms a single process (Hsieh, 2005; West et al., 2005). More specifically, pre-classification by clustering technique before classification helps improve predictive performance (Hsieh, 2005; Tsai, 2014).

2.2.1 Classification

Classification is considered a supervised technique since it learns by a set of examples, acting as a supervisor. The supervisor has the expertise of the environment, represented by a group of input-output examples. As for classifying unknown patterns, the process tries to generate a classifier or a model from a training sample set for each class during the learning task. The learning process tries to map between the input-output examples and correctly label the training set. Consequently, the model obtained from the training process is applied to classify an unknown pattern into each class. The classification techniques used frequently for business failure are decision tree (DT), logistic regression (LR), and neural network (NN) (Kumar & Ravi, 2007; Tsai, 2014).

The first statistical model for business failure prediction was published in the 1960s by Beaver (1966), which was a univariate model. As for the multivariate technique, Altman (1968) proposed the application of discriminant analysis. However, the method was popular in business failure prediction before the 1980s (Geng et al., 2015) because of its stick assumptions, such as linear separability, multivariate normality, and equal covariances. LR was suggested in the business failure area to overcome the discriminant limitations. Martin (1977) pioneered the logistic model to forecast bank failure. LR is mainly applied to forecast binary or multi-class response attributes. The model cannot be generated by linear regression directly since the dependent variable is a nominal type. Instead of predicting a point estimate of the event, it predicts the odds of the event. The logistic model has been widely used for classification problems; however, its restrictive assumptions make data mining techniques more popular in real-world prediction problems (Olson, Delen, & Meng, 2012).

Data mining techniques have influenced the area of business failure prediction since the 1990s. NN and DT are the popular technique for this area (Klepáč & Hampel, 2017).

NN is a nonlinear mathematical approach to the data mining technique. The method comprises a group of neural nodes linking with the weighted nodes. Each node can simulate a neuron of creatures, and the connection among these nodes is equal to the synaptic that connects among the neurons. Generally, neural networks consist of input, hidden, and output layers. Each layer is connected, of which input units will be connected to hidden units, and the hidden units will later be joined to the output units.

DT is a non-parametric prediction method as a recursive partitioning using a divide and conquers technique to distinguish instances. Its structure comprises a root node, branches, and leaf nodes. Each internal node denotes a test on an attribute, each branch represents the outcome of a test, and each leaf node holds a class label. The topmost node in the tree is the root node. The decision tree aims to divide the unknown data recursively until every data belongs to a specific class. The algorithm for the decision tree is generally implemented in two phases. Tree-building is the first phase; in that phase, the tree is divided until all the data have its class in a top-down fashion. The second phase is tree pruning, where predictions and accuracy are improved bottom-up.

2.2.2 Clustering

Clustering is an unsupervised technique that differs from classification. The approach aims to arrange a given data of unlabelled input patterns into meaningful clusters based on a measure of similarity (Jain, Murty, & Flynn, 1999). Pattern clustering results in some well-separated groups in the feature space, which summarise and visualize data in the given collection. Algorithms in the clustering technique can be grouped into two categories-- hierarchical and non-hierarchical (or partitional) clustering. The hierarchical clustering algorithm generates a hierarchy of clusters, in which a distinct singleton cluster will be combined one by one until satisfied some thresholds. The result will produce a series of arborescence partitions. The work of Beranová, Basovníková, and Martinovičová (2013) applied this method to cluster agricultural enterprises. Some might prefer partitional clustering, especially in business problems (Dissayarungkun, 2021; Olafsson, Li, & Wu, 2008). K-means is one of the well-known partitional clustering algorithms since it procedures with a simple algorithm but gives efficient clustering results (Tsai & Chen, 2010). The method aims to partition the data set into k clusters in which each observation belongs to the cluster with the nearest mean (centroid). The algorithm computes centroids and repeats until the optimal centroid is found. The process starts with selecting the k point as the initial centroids. Each observation in the data set will be assigned to the closest centroid and then recomputed the new centroids of each

cluster. The process reassigns each observation regarding their nearest new centroids. Repeatedly the process until some convergence criterion is satisfied or the assignment is not changed.

3. Research Methodology

The research follows the important steps of the CRoss Industry Standard Process for Data Mining (CRISP-DM) (Shearer, 2000), which is the standard data mining process. This study focuses on data understanding, preparation, modeling, and evaluation. RapidMiner Studio Version 9.10 was used for the data analysis processing.

3.1 Data Understanding and Preparation

The research focused on predicting agribusiness failure in Thailand based on the annual financial statement collected from the Department of Business Development database. Regarding The Thailand Standard Industrial Classification (TSIC), agribusiness is classified in A--Agriculture, Hunting, Forestry, and Fishing. This research studied 3,118 companies that submitted their financial statements from 2016 to 2020 (The recent financial statement data collected was in 2019). The input variables were related to 12 kinds of financial statements and firm information attributes, including business size, company age, two consecutive years of profit or loss, recently registered capital, working capital, corporate income tax, revenue, authorized capital stock, total assets, current asset, liability, and operating profit margin.

3.2 Number of Failed Business Status

Since the research aimed to predict business failure in 2020, the target attribute was a nominal type--the "Failure" or "Non-Failure" status in 2020. There were 100 failed companies in the study, including those registered for liquidation or labeled as unoccupied firms. The rest of them are labeled as non-failure firms, including 3,018 companies.

3.3 The Time Span of the Data Set

The previous studies in Thailand used the financial statements data obtained during the last three years before the failure, which achieved the accuracy of training data of about 70% to 80%. Buanak (2016) forecasted bankruptcies of small and medium-sized enterprises (SMEs) in Thailand ahead of one to three years before failure happened. An imbalanced sample of 604 firms from 2008 to 2015, including 71 failed companies and 532 active enterprises, was applied in the training model. The study compared the accuracy of Altman, Zmijewski, and discriminant models. The finding showed that discriminant analysis gave the most accuracy of training data with 77.45%,75.80%, and 81.16% for one, two, and three years ahead. Narungsri (2005) applied logistic regression to predict financial failure for small and medium companies from one to three years ahead. The study estimated the model by the financial statement from 1999 to 2001. The sample size was 671 firms containing 321 failed companies and 350 active firms. The selected model was applied to predict the business failure of testing data, including 55 companies for each group of failed and non-failed. The result showed that the one-year-ahead prediction was better than two and three years ahead, which achieved the accuracy of training data at 72.70%, 69.27%, 67.79%, and for testing data at 71.82%, 68.87%, and 66.98% consecutively.

This research also implemented the model based on the financial statement obtained for three previous years before the companies were labeled as business failures in 2020. For example, in an early warning for one year (t-1), the study used financial data by 2019 to predict whether the firm was marked as a failure in 2020. As for the two-year prediction (t-2) of 2020, the financial statement before 2018 was applied to the model. The model used only the financial information before 2017 to predict a three-year earlier failure (t-3) in 2020.

3.4 Data Preparation

As the initial data set, the business status in 2020 and the financial statement were combined. The missing values were remedied by interpolation. Stratified random sampling was applied to partition the data into training and testing data sets. The training data set was applied to generate the learning models, while the testing data set was used to test the models' predictive ability. The random partitions of the training and testing data sets might impact the analysis; hence, we conduct both single and multiple splits.

Regarding the single split, to avoid undertraining or overtraining, we applied the training ratios at 0.6, 0.7, and 0.8. As for the training data set, it worked with a balanced sample size. For instance, of 0.6 split ratios, this training data set contained 120 observations, including 60 failed firms and 60 non-failed firms, since the observations labeled as failed firms totally in the data set were 100 firms. As for the testing data set, the number of data was 2,998 firms—40 failed firms and 2,958 non-failed firms, which the data did not use for generating the model.

The research also conducted multiple splits or cross-validation, a resampling procedure, to evaluate the models with ten folds. Cross-validation aims to reduce the bias associated with the random sampling of the training

and holdout data samples. The analysis randomly selected 100 failed firms and 100 non-failed firms regarding a balanced training data set. This input data was partitioned into ten subsets of equal size at 20 observations. A single subset was retained as the test data set (20 observations) of those ten subsets. The remaining nine subsets worked as a training data set (180 observations). The cross-validation process was then iterated ten times, with each of the ten subsets used exactly once as the test data. The ten results from the ten iterations were combined to produce the estimation.

This training data set of this study might be quite small for general problems using the data mining technique; however, as for the limitation of data in the fields of business failure prediction, this training sample size was acceptable. As seen in many works related to business failure prediction, such as the work of Geng et al. (2015); Klepáč and Hampel (2017), who applied data mining models such as LR, NN, and DT, also used around the same size of this research for the training data set.

3.5 Modelling

In the first step of data modeling, we generated single-classified models. LR, NN, and DT, well-known and widely used classification methods in business failure prediction, were conducted. Regarding the RapidMiner operator, as for LR, the logistic regression operator was used. This operator applied an S-shaped curve formed by the logit transformation representing the probability of an event—Failure Non-Failur. The decision tree operator was applied for DT, in which the algorithm collects nodes to identify a decision on values affiliation to a class of business status. Each node represents a splitting rule for one variable. The new nodes were repeatedly created until meeting the stopping criterion. A prediction of a business status was identified based on the majority of Examples that reached this leaf during generation. The Neural Net operator was conducted for NN, which learns a model through a feed-forward neural network trained by a multi-layer perceptron. Hereafter, the outperform model will be selected regarding the predictive performance (Figure 1).

Secondly, the research conducted the hybrid classifier by combining the clustering method in the selected classifier to improve business failure prediction (Figure 2). The k-mean clustering is one of which the well-known clustering in business fields (Tsai & Chen, 2010) was applied for this task. Then we compared the predictive performance of the selected model from the first step and the hybrid model from the second step.

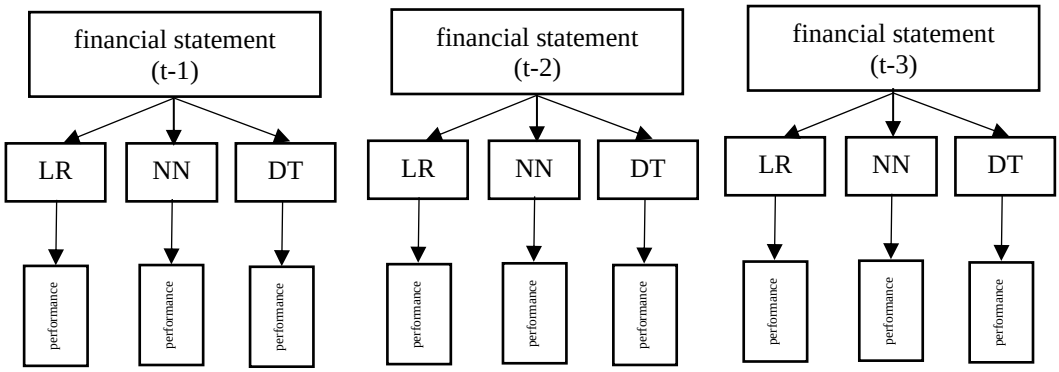


Figure 1: A single method for business failure prediction

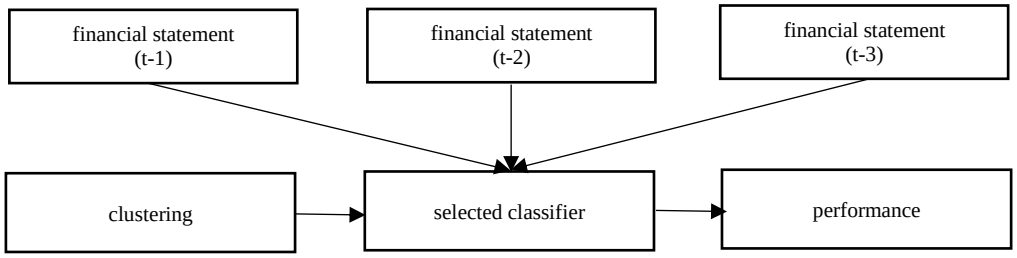


Figure 2: A hybrid method of clustering combined with classification for business failure

3.6 Evaluation

The research evaluated the model performance in terms of accuracy, recall, and precision, the common evaluation metrics of machine learning. (Davis & Goadrich, 2006; Geng et al., 2015).

The accuracy is the ratio of correctly predicted observations, including failure and non-failure, to the total number of observations. Whereas recall is the ratio of correctly predicted failure observations to the total number of actual failure observations, precision is the ratio of correctly predicted failure observations to the total number of predicted failure observations. The confusion matrix displays the classification prediction performance in Table 1.

Table 1: Confusion matrix for business failure prediction

		Actual	
		T (failure)	F (non-failure)
Predict	T (failure)	TP	FP
	F(non-failure)	FN	TN

$$\text{Accuracy}=\frac{TP+TN}{TP+FP+FN+TN}$$

$$\text{Recall}=\frac{TP}{TP+FN}$$

$$\text{Precision}=\frac{TP}{TP+FP}$$

4. Results and Discussion

The predictive performance of LR, DT, and NN are shown in Tables 2 and 3. The tables display the accuracy, recall, and precision based on the training ratios of 0.6, 0.7, 0.8, and ten-fold cross-validation of the early warning for t-1, t-2, and t-3.

Table 2 shows the single classifier performances of LR, NN, and DT to predict business failure. The results varied according to the performance evaluation criterion, the training ratios or the cross-validation, the early warning period, and the training or testing data set. For example, the accuracy of cross-validation for t-1 in the training data set, the result of LR was outstanding, gaining at 87.67%; however, in the testing data, NN was outperformed at 70.00%. Hence, the research considered the number of models achieving better performance in the last section of Table 2. However, please note that an accuracy rate is generally applied to evaluate the correctly classified sample number to the total sample number of a balanced data set. However, the data set of business failure is highly class-imbalanced; the accuracy rate measure cannot evaluate model performance for the minority class well (Sun, Li, Fujita, Fu, & Ai, 2020). The purpose of the business failure prediction is to capture the failure of a business (Klepáč & Hampel, 2017), which is the minority class of the data set. Hence, this research focused on the model with the highest recall and precision rates because they concentrate on correctly predicted failure business.

Figure 3 shows the recall and precision of each single technique, whereas the last row of Table 2 focuses on the number of models gaining better recall and precision. The results indicated that DT achieved better

performance in both training and testing data sets (except for t-2, DT was equal to LR for training, and DT was not different from NN for the testing data set). Many research suggested applying black-box algorithms such as NN and the support vector machines method (SVM) to predict business failure; however, those techniques are less comprehensible by human users than DT (Olson et al., 2012). Our results also were consistent with Klepáč and Hampel (2017). They indicated that DT offers better accuracy for agriculture companies in the EU than LR and SVM. In addition, the study of Olson et al. (2012) related to bankruptcy prediction found that DT was relatively more accurate than NN and SVM.

Since this research aimed to improve business failure prediction by applying the hybrid classifier--segmentation before classification, the k-mean clustering was applied. K-mean separated the data into subgroups based on the company's age and the two consecutive years of profit or loss before conducting the DT classification method to improve the business failure prediction performance. However, k-mean is the non-hierarchical method; identifying the number of groups or k is needed before using the technique. Therefore, the research applied the hierarchical method with agglomerative algorithms to roughly set the number of clusters. The study decided to apply k=3 for the k-mean clustering. In addition, as the suggestion by Hair, Black, Babin, and Anderson (2014), the cluster solution was validated by considering the variable not included in the cluster analysis—business size (small, medium, and large) to ensure the practical significance of the cluster solutions. The results showed the relevant reasons. For example, the largest group of the clustering analysis trended to be the companies that had not conducted business for a long time and had two consecutive years of loss; most of the companies in the group were small-size businesses.

Table 3 displays the comparison of the predictive performance of DT and hybrid DT to predict business failure. For the same reason mentioned above, this part also considered only recall and precision. As for the training data set, clustering before classification helped improve classification performance for t-1 and t-2; however, the performance did not differ for t-3. As for the testing data set, clustering before classifying helped improve classification performance for t-1 and t-3, but as for t-2, it gained lower performance (The last row of Table 3 and Figure 4). Moreover, the result found that long-time forecasting seemed to decrease the model prediction performance, which was according to the previous studies of Klepáč and Hampel (2017); Narungsri (2005).

Table 2: Model performance of DT, LR, NN

	Splits	Training data									Testing data								
		t-1			t-2			t-3			t-1			t-2			t-3		
		LR	NN	DT	LR	NN	DT	LR	NN	DT	LR	NN	DT	LR	NN	DT	LR	NN	DT
Accuracy	0.6	**100.00**	88.33	88.33	**95.83**	90.83	88.33	**90.83**	86.67	85.83	58.30	**67.20**	54.37	49.72	**55.33**	50.12	58.54	**77.23**	50.12
	0.7	**97.14**	90.00	83.57	**88.57**	86.43	77.14	**83.57**	**83.57**	72.86	64.71	**76.04**	50.48	**63.32**	54.01	44.49	63.96	**75.72**	37.11
	0.8	85.00	**86.25**	83.75	81.25	**84.38**	80.00	80.63	**82.50**	67.50	65.06	**76.12**	55.93	**66.19**	62.02	55.93	**70.51**	**70.51**	29.33
	Cross-validation	**87.67**	82.11	83.44	**81.56**	79.28	79.72	79.67	78.83	**80.06**	64.50	**70.00**	67.00	61.00	**66.00**	65.00	**68.50**	68.00	65.00
Recall	0.6	**100.00**	86.67	**100.00**	96.67	95.00	**98.33**	91.67	81.67	**98.33**	57.50	60.00	**87.50**	55.00	**77.50**	75.00	52.50	55.00	**72.50**
	0.7	98.57	82.86	**100.00**	87.14	95.71	**100.00**	84.29	78.57	**100.00**	60.00	46.67	**73.33**	50.00	83.33	**93.33**	50.00	46.67	**93.33**
	0.8	83.75	82.50	**98.75**	82.50	93.75	**96.25**	77.50	83.75	**100.00**	50.00	45.00	**85.00**	55.00	80.00	**95.00**	50.00	55.00	**95.00**
	Cross-validation	87.44	85.56	**97.00**	85.00	84.22	**96.11**	85.22	84.56	**95.44**	71.00	76.00	**79.00**	68.00	74.00	**81.00**	75.00	75.00	**82.00**
Precision	0.6	**100.00**	89.66	81.08	**95.08**	87.69	81.94	90.16	**90.74**	78.67	4.37	5.76	**5.84**	3.49	**5.35**	4.67	4.05	**7.64**	4.53
	0.7	95.83	**96.67**	75.27	**89.71**	80.72	68.63	83.10	**87.30**	64.81	5.36	**6.31**	4.61	4.37	**5.56**	5.14	4.45	**6.22**	4.56
	0.8	85.90	**89.19**	75.96	**80.49**	78.95	72.64	**82.67**	81.71	60.61	4.59	**6.12**	5.88	5.16	6.43	**6.48**	5.43	**5.91**	4.14
	Cross-validation	**87.83**	80.04	76.31	**79.52**	76.64	72.38	**76.70**	75.87	72.98	62.83	**67.86**	63.71	59.65	**63.79**	61.36	**66.37**	65.79	61.19
number of models gaining better performance	Accuracy	**3**	1	0	**3**	1	0	**2**	**2**	1	0	**4**	0	**2**	**2**	0	2	**3**	0
	Recall	1	0	**4**	0	0	**4**	0	0	**4**	0	0	**4**	0	1	**3**	0	0	**4**
	Precision	**2**	**2**	0	**4**	0	0	**2**	**2**	0	0	**3**	1	0	**3**	1	1	**3**	0
	Recall and Precision	3	2	**4**	**4**	0	**4**	2	2	**4**	0	3	**5**	0	**4**	**4**	1	3	**4**

Note: The highlight is the better performanc

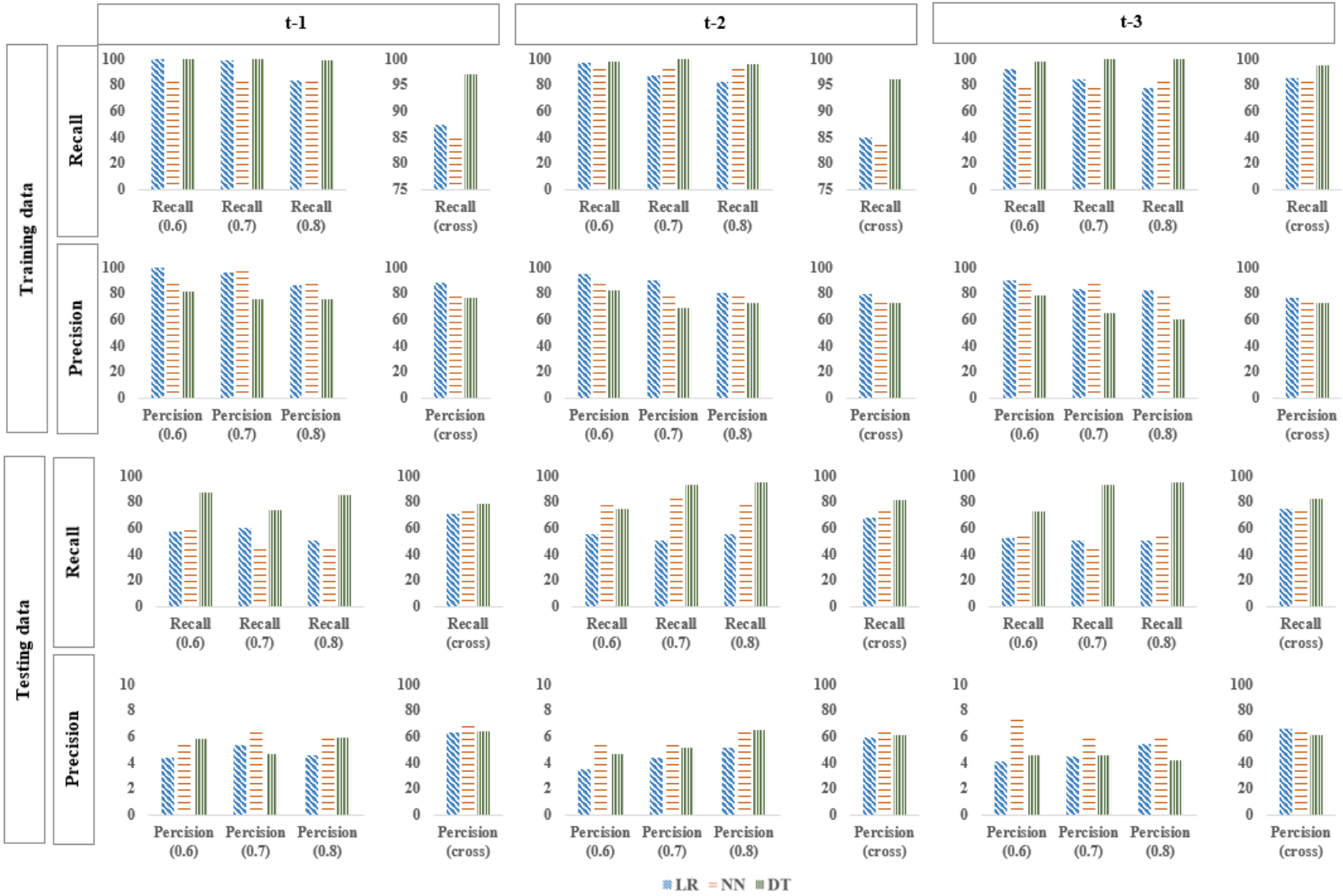


Figure 3: Recall and precision of DT, LR, NN
Note: () Training ratio or cross-validation

Table 3: Comparing the predictive performance of the single classifier and hybrid classifier of DT

	Splits	t-1		t-2		t-3		t-1		t-2		t-3	
		Single	Hybrid	Single	Hybrid	Single	Hybrid	Single	Hybrid	Single	Hybrid	Single	Hybrid
Accuracy	0.6	88.33	**89.17**	**88.33**	86.67	**85.83**	82.50	54.37	**57.74**	50.12	**54.45**	**50.12**	47.15
	0.7	83.57	**88.03**	77.14	**83.80**	72.86	**78.87**	50.48	**53.96**	44.49	**53.10**	37.11	**44.22**
	0.8	83.75	**88.13**	80.00	**83.75**	67.50	**79.38**	55.93	**60.03**	55.93	**56.66**	29.33	**58.91**
	Cross-validation	83.44	**88.22**	79.72	**86.44**	**80.06**	77.28	67.00	**68.50**	65.00	**70.50**	**65.00**	57.50
Recall	0.6	**100.00**	**100.00**	98.33	**100.00**	98.33	**100.00**	**87.50**	85.00	75.00	**87.50**	72.50	**85.00**
	0.7	**100.00**	**100.00**	**100.00**	97.18	**100.00**	98.59	73.33	**89.66**	**93.33**	75.86	**93.33**	86.21
	0.8	**98.75**	**98.75**	96.25	**98.75**	**100.00**	91.25	**85.00**	**85.00**	**95.00**	80.00	**95.00**	70.00
	Cross-validation	97.00	**97.44**	**96.11**	94.22	95.44	**99.89**	**79.00**	**79.00**	**81.00**	77.00	**82.00**	**82.00**
Precision	0.6	81.08	**82.19**	**81.94**	78.95	**78.67**	74.07	5.84	**6.13**	4.67	**5.85**	4.53	**4.95**
	0.7	75.27	**80.68**	68.63	**76.67**	64.81	**70.71**	4.61	**5.74**	**5.14**	4.86	4.56	**4.61**
	0.8	75.96	**81.44**	72.64	**75.96**	60.61	**73.74**	5.88	**6.46**	**6.48**	5.67	4.14	**5.30**
	Cross-validation	76.31	**82.27**	72.38	**81.54**	**72.98**	68.78	63.71	**65.29**	61.36	**68.14**	**61.19**	55.03
number of models gaining better performance	Accuracy	0	**4**	1	**3**	2	2	0	**4**	0	**4**	2	2
	Recall	3	**4**	2	2	2	2	3	3	**3**	1	**3**	2
	Precision	0	**4**	1	**3**	2	2	0	**4**	2	2	1	**4**
	Recall and Precision	3	**8**	3	**5**	4	4	3	**7**	**5**	3	4	**6**

Note: The highlight is the better performance

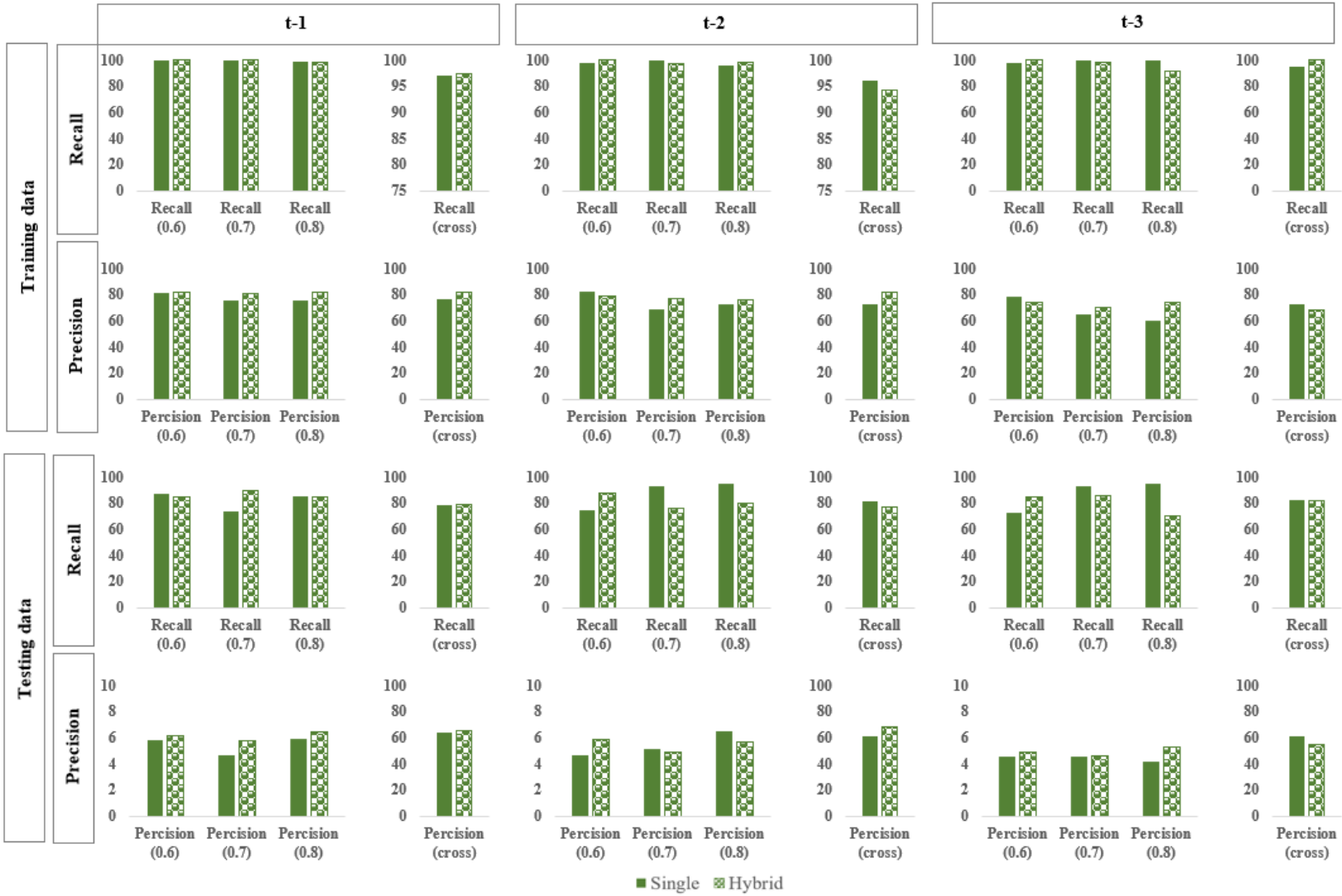


Figure 4: Recall and precision of the single classifier and hybrid classifier of DT
Note: () Training raio or cross-validation

5. Conclusions

Business failure can significantly impact many parties—the private companies, the government, and the whole economy. Therefore, predicting business failure is always one major research problem in business and economics. This research aimed to improve business failure prediction for Thailand agribusiness using hybrid classifiers. This case studied 3,118 companies submitting their financial statements from 2016 to 2020, and there were 100 failed companies in the study, including those registered for liquidation or labeled as unoccupied firms.

As for the single classifier, focusing on recall and precision, the empirical result indicated that DT performs better than LR and NN. This result was consistent with Klepáč and Hampel (2017), who showed DT offers better accuracy for agriculture companies in the EU than LR and the SVM. Moreover, Olson et al. (2012), whose study related to bankruptcy prediction, suggest that DT was relatively more accurate than NN and SVM. Although much research concludes NN obtained more accuracy than DT, it is less comprehensible by human users than DT (Olson et al., 2012).

Since the research focused on comparing the prediction performance of single and hybrid classified methods, segmentation and classification were applied to improve the business failure prediction. The k-mean clustering technique was applied to separate the data into three groups before using the DT classification method. The results of the hybrid classifier of DT were not consensus for all t-1, t-2, and t-3. As for the testing data set, the hybrid helped improve classification performance for t-1 and t-3, but for t-2, it gained lower performance than a single DT. However, when considering the overall of all periods of t-1, t-2, and t-3, The results indicated that hybrid classifiers helped improve classification performance. Clustering before classification was beneficially supported by the studies of Alapati and Sindhu (2016) and Tsai (2014). Nevertheless, the results indicated that increasing the previous year of early warning before the failure will decrease the accuracy of the business failure prediction model, which is consistent with earlier studies by Klepáč and Hampel (2017) and Narungsri (2005).

In conclusion, in this case, the results suggest applying the hybrid method of clustering and classification to the work of business failure prediction. Even though, in numerous instances, clustering can help improve classification performance; however, it does not always improve the quality of classification (Piernik & Morzy, 2021).

For future work, I suggest considering these issues. First, more rule nodes were desired to improve the NN performance; however, the compensation of this task is the more complex model. Second, several clustering techniques can be pre-processed before classification for comparisons. Third, other advanced classification methods could be regarded as the hybrid method in the same way. Forth, different data sets with higher failed firms should be applied for this work.

References

Alapati, K. Y., & Sindhu, K. (2016). Combining Clustering with Classification: A Technique to Improve Classification Accuracy. **International Journal of Computer Science Engineering (IJCSE)**, 5(6).

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. **The journal of finance**, 23(4), 589-609.

Beaver, W. H. (1966). Financial ratios as predictors of failure. **Journal of accounting research**, 71-111.

Beranová, M., Basovníková, M., & Martinovičová, D. (2013). Clustering of agricultural enterprises. **Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis**, 61(2), 289-296.

Buanak, S. (2016). **A Bankruptcy Test for Small and Medium-Size Enterprises (SMEs) in Thailand**. Thammasat University, Thailand.

Davis, J., & Goadrich, M. (2006). **The relationship between Precision-Recall and ROC curves**. Paper presented at the Proceedings of the 23rd international conference on Machine learning.

Development, D. o. B. (2020). **Manual for Submitting Financial Statements**. DBD. Retrieved from https://www.dbd.go.th/download/document_file/balance/manual_finacial2563.pdf.

Dimitras, A. I., Zanakis, S. H., & Zopounidis, C. (1996). A survey of business failures with an emphasis on prediction methods and industrial applications. **European Journal of Operational Research**, 90(3), 487-513.

Dissayarungkun, P. (2021). Cluster Analysis of Small Enterprises According to a Risk of Business. **Journal of Applied Statistics and Information Technology**, 6(2), 1-16.

Geng, R., Bose, I., & Chen, X. (2015). Prediction of financial distress: An empirical study of listed Chinese companies using data mining. **European Journal of Operational Research**, 241(1), 236-247.

Gepp, A., Kumar, K., & Bhattacharya, S. (2010). Business failure prediction using decision trees. **Journal of forecasting**, 29(6), 536-555.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). Multivariate data analysis: Pearson new international edition. **Essex: Pearson Education Limited**, 1(2).

Hsieh, N.-C. (2005). Hybrid mining approach in the design of credit scoring models. **Expert Systems with Applications**, 28(4), 655-665.

Jader, R., & Aminifar, S. (2022). Predictive Model for Diagnosis of Gestational Diabetes in the Kurdistan Region by a Combination of Clustering and Classification Algorithms: An Ensemble Approach. **Applied Computational Intelligence and Soft Computing**, 2022.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. **ACM computing surveys (CSUR)**, 31(3), 264-323.

Klepáč, V., & Hampel, D. (2017). Predicting financial distress of agriculture companies in EU. **Agricultural Economics-Zemedelska Ekonomika**.

Kumar, P. R., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques–A review. **European Journal of Operational Research**, 180(1), 1-28.

Li, H., Sun, J., & Wu, J. (2010). Predicting business failure using classification and regression tree: An empirical comparison with popular classical statistical methods and top classification mining methods. **Expert Systems with Applications**, 37(8), 5895-5904.

Lin, F. Y., & McClean, S. (2001). **A data mining approach to the prediction of corporate failure**. In Applications and Innovations in Intelligent Systems VIII, 93-106.

Lincoln, M. (1984). An empirical study of the usefulness of accounting ratios to describe levels of insolvency risk. **Journal of banking & finance**, 8(2), 321-340.

Martin, D. (1977). Early warning of bank failure: A logit regression approach. **Journal of banking & finance**, 1(3), 249-276.

Narungsri, K. (2005). **Financial Failure Prediction for Small and Medium Enterprises in Thailand**. Suranaree University of Technology, Thailand.

Olafsson, S., Li, X., & Wu, S. (2008). Operations research and data mining. **European Journal of Operational Research**, 187(3), 1429-1448.

Olson, D. L., Delen, D., & Meng, Y. (2012). Comparative analysis of data mining methods for bankruptcy prediction. **Decision Support Systems**, 52(2), 464-473.

Piernik, M., & Morzy, T. (2021). A study on using data clustering for feature extraction to improve the quality of classification. **Knowledge and Information Systems**, 63(7), 1771-1805.

Pumjaroen, J. (2019). **Early warning system for real economy: a case study of Thailand**. (Doctoral Dissertation). National Institute of Development Administration,

Pumjaroen, J., & Sethapramote, Y. (2023). Evaluating The Robustness And Implementing Double One-Sided Hodrick-Prescott For Cycle Extraction. **ABAC Journal**, 43(2), 143-160.

Pumjaroen, J., Vichitthamaros, P., & Sethapramote, Y. (2020). Forecasting Economic Cycle with a Structural Equation Model: Evidence from Thailand. **International Journal of Economics and Financial Issues**, 10(3), 47-57.

Shalaby, M., Belal, N. A., & Omar, Y. (2021). Data clustering improves siamese neural networks classification of Parkinson's disease. **Complexity**, 2021, 1-9.

Shearer, C. (2000). The CRISP-DM model: the new blueprint for data mining. **Journal of data warehousing**, 5(4), 13-22.

Sun, J., Li, H., Fujita, H., Fu, B., & Ai, W. (2020). Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting. **Information Fusion**, 54, 128-144.

Tsai, C.-F. (2014). Combining cluster analysis with classifier ensembles to predict financial distress. **Information Fusion**, 16, 46-58.

Tsai, C.-F., & Chen, M.-L. (2010). Credit rating by hybrid machine learning techniques. **Applied soft computing**, 10(2), 374-380.

Wanke, P., Barros, C. P., & Faria, J. R. (2015). Financial distress drivers in Brazilian banks: A dynamic slacks approach. **European Journal of Operational Research**, 240(1), 258-268.

Webb, A. R. (2003). **Statistical pattern recognition**. (2nd ed.). John Wiley & Sons.

West, D., Dellana, S., & Qian, J. (2005). Neural network ensemble strategies for financial decision applications. **Computers & operations research**, 32(10), 2543-2559.

ติดต่อหรือส่งบทความ

บรรณาธิการวารสารสถิติประยุกต์และเทคโนโลยีสารสนเทศ
คณะสถิติประยุกต์
สถาบันบัณฑิตพัฒนบริหารศาสตร์
148 ถ. เสรีไทย แขวงคลองจั่น
เขตบางกะปิ กทม. 10240

โทรศัพท์ : 02-727-3038
โทรสาร : 02-374-3061
อีเมลล์ : asit-journal@as.nida.ac.th
Website : https://ph02.tci-thaijo.org/index.php/asit-journal/index

NIDA WISDOM for Change
คณะสถิติประยุกต์
Graduate School of Applied Statistics

ISSN : 2465-4949 (Online)