บทที่ 2
ทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้อง

งานวิจัยเรื่อง การพัฒนาตัวแบบการทำนายความสำเร็จของสตาร์ตอัป จากประสิทธิผลการปรากฏตัวทางดิจิทัลโดยใช้การเรียนรู้ของเครื่อง ได้ทำการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องเพื่อใช้เป็นแนวทางในการทำการวิจัย ดังนี้

2.1 ทบทวนวรรณกรรม
- 2.1.1 การคาดการณ์ความสำเร็จของสตาร์ตอัป
- 2.1.2 ความสำคัญของการระดมทุนสำหรับธุรกิจสตาร์ทอัพ
- 2.1.3 การปรากฏตัวทางดิจิทัล (Digital Presence)
- 2.1.4 ปัจจัยที่ใช้ในการทำนายความสำเร็จของสตาร์ตอัป
- 2.1.5 การใช้ประโยชน์จาก Digital Marketing เพื่อความสำเร็จของ Startup
- 2.1.6 การใช้การเรียนรู้ของเครื่อง (Machine learning) ในการทำนาย
- 2.1.7 การทำนายโดยใช้สถิติ กับการใช้การเรียนรู้ของเครื่อง
- 2.1.8 โมเดลการเรียนรู้ของเครื่องที่ใช้ในการทำนายความสำเร็จของสตาร์ตอัป

2.2 งานวิจัยที่เกี่ยวข้อง

2.3 กรอบแนวคิดการวิจัย

โดยมีรายละเอียดดังต่อไปนี้

2.1 ทบทวนวรรณกรรม

2.1.1 การคาดการณ์ความสำเร็จของสตาร์ตอัป

การทำนายความสำเร็จของสตาร์ตอัปได้รับความสนใจอย่างมากในงานวิจัยช่วงหลังๆ มานี้ งานวิจัยหลายชิ้นได้สำรวจการทำนายความสำเร็จของสตาร์ตอัปโดยใช้การเรียนรู้ของเครื่องในการพัฒนาแบบจำลองที่สามารถทำนายผลลัพธ์ในอนาคตของบริษัทที่เพิ่งก่อตั้งขึ้น (Shi et al., 2024) การมีอยู่ของฐานข้อมูลสาธารณะที่ให้ข้อมูลหลากหลายเกี่ยวกับสตาร์ตอัปได้ขับเคลื่อนการพัฒนาแบบจำลองในการทำนาย ซึ่งแต่ละโมเดลนั้นมีวิธีการและการนิยามเกณฑ์ความสำเร็จที่แตกต่างกัน (Zadehnoori, 2023) นักวิจัยได้สำรวจแนวทาง Hybrid Intelligence โดยการผสานความเชี่ยวชาญของมนุษย์กับความสามารถของเครื่องจักร เพื่อปรับปรุงการทำนายในบริบทที่ไม่แน่นอน อาทิ ความสำเร็จของสตาร์ตอัปในระยะแรกเริ่ม (Dellermann et al., 2017)

เทคนิคการเรียนรู้ของเครื่องได้ถูกนำมาใช้ในการสร้างโมเดลทำนายที่สามารถระบุปัจจัยที่มีอิทธิพลต่อความสำเร็จของสตาร์ตอัป ช่วยให้ผู้ประกอบการสามารถตัดสินใจได้อย่างมีข้อมูล (Asmoro et al., 2018) งานวิจัยยังได้มุ่งเน้นไปที่การใช้ปัญญาประดิษฐ์ (Artificial Intelligence: AI) ในการทำนายความสำเร็จของสตาร์ตอัป แก้ไขปัญหาการทำนายผลลัพธ์ของการประกอบการท่ามกลางความไม่แน่นอน (Morande et al., 2023) นอกจากนี้ งานวิจัยยังได้เน้นถึงศักยภาพของการเรียนรู้ของเครื่องในการให้ข้อมูลเชิงลึกที่เกี่ยวกับการทำนายความสำเร็จของสตาร์ตอัป (Pranav et al., 2024)

ปัจจัยภายในที่เป็นข้อมูลพื้นฐานของธุรกิจที่สำคัญ อาทิ ประสบการณ์ของผู้ก่อตั้ง ความมั่นใจในตนเอง ความสามารถในการจัดการองค์กร และทักษะของทีม ได้รับการยอมรับว่าเป็นตัวกำหนดที่สำคัญต่อความสำเร็จของสตาร์ตอัป (Lovrincevic, 2022) นอกจากนี้ การใช้ปัญญาประดิษฐ์ในการทำนายความสำเร็จของสตาร์ตอัปยังได้รับการเน้นย้ำ โดยเฉพาะในบริบทของการลดอคติในการทำนายความสำเร็จสำหรับ Venture Capitalists (Te et al., 2023) การเสนอแนวคิดในการระบุปัจจัยความสำเร็จที่สำคัญสำหรับสตาร์ตอัปผ่านแบบจำลองการวิเคราะห์รู้สึก (Sentiment Analysis) และการทำเหมืองข้อมูล (Text Data Mining) ได้ถูกเสนอเพื่อช่วยในการสร้างแบบจำลองธุรกิจที่ประสบความสำเร็จและมีกำไรในระยะยาว (Asgari et al., 2022)

จากความสำคัญของการคาดการณ์ความสำเร็จของสตาร์ตอัปที่กล่าวไป จึงเป็นที่น่าสนใจว่าปัจจัยใดบ้างที่เป็นตัวบ่งชี้หรือทำนายความสำเร็จเหล่านั้นได้ และในยุคดิจิทัลปัจจุบัน การปรากฏตัวทางดิจิทัล (Digital Presence) ถือเป็นหนึ่งในปัจจัยที่มีบทบาทสำคัญอย่างยิ่ง ที่เราจะพิจารณาในหัวข้อถัดไป

2.1.2 ความสำคัญของการระดมทุนสำหรับธุรกิจสตาร์ทอัพ

ธุรกิจสตาร์ตอัปเปรียบเสมือนเมล็ดพันธุ์แห่งนวัตกรรม ซึ่งมีศักยภาพในการเติบโตเป็นธุรกิจที่ประสบความสำเร็จ สามารถปฏิวัติอุตสาหกรรม และกำหนดทิศทางของอนาคต อย่างไรก็ตาม อาทิเดียวกับเมล็ดพันธุ์ทั่วไป สตาร์ตอัปต้องการการบ่มเพาะและทรัพยากรในการเจริญเติบโต และการระดมทุนมีบทบาทสำคัญในการมอบปัจจัยหล่อเลี้ยงที่จำเป็นต่อความสำเร็จ

การระดมทุนที่เพียงพอมีความสำคัญอย่างยิ่งสำหรับสตาร์ตอัป เนื่องจากการระดมทุนจะมอบทรัพยากรทางการเงินที่จำเป็นสำหรับการนำแนวคิดไปสู่ความเป็นจริง และการฝ่าฟันอุปสรรคในระยะเริ่มต้นของการพัฒนา การระดมทุนทำหน้าที่เป็นตัวเร่งปฏิกิริยา ทำให้สตาร์ตอัปสามารถแปลงแนวคิดให้เป็นผลิตภัณฑ์หรือบริการที่จับต้องได้ เข้าถึงตลาดเป้าหมาย และสร้างรากฐานที่มั่นคงสำหรับการเติบโตอย่างยั่งยืน นอกจากการมอบทรัพยากรทางการเงินแล้ว การระดมทุนยังทำหน้าที่เป็นสัญญาณแห่งความไว้วางใจและความมั่นใจในสตาร์ตอัป ซึ่งดึงดูดการลงทุนและลูกค้าเพิ่มเติม (Pertsiya, 2023)

ธุรกิจสตาร์ตอัปมีความต้องการด้านการระดมทุนที่หลากหลาย และเหตุผลในการแสวงหาการสนับสนุนทางการเงินนั้นแตกต่างกันไปขึ้นอยู่กับระยะของการพัฒนา อุตสาหกรรม และเป้าหมายเฉพาะ ต่อไปนี้คือเหตุผลหลักบางประการที่การระดมทุนมีความจำเป็นสำหรับธุรกิจสตาร์ตอัป

- ครอบคลุมค่าใช้จ่ายเริ่มต้น
  การเริ่มต้นธุรกิจเกี่ยวข้องกับค่าใช้จ่ายที่หลากหลาย และการมีเงินทุนที่เพียงพอเป็นสิ่งสำคัญในการครอบคลุมค่าใช้จ่ายเหล่านี้ ค่าใช้จ่ายเริ่มต้นอาจรวมถึงค่าธรรมเนียมทางกฎหมาย ใบอนุญาต การจดทะเบียน การวิจัยตลาด การสร้างแบรนด์ และการพัฒนาเว็บไซต์ หากไม่มีเงินทุนเพียงพอ สตาร์ทอัพจำนวนมากจะไม่สามารถเริ่มต้นธุรกิจได้
- รักษาสภาพคล่องทางการเงินที่เพียงพอ
  การรักษาสภาพคล่องทางการเงินที่เพียงพอเป็นสิ่งจำเป็นสำหรับสุขภาพทางการเงินของบริษัท การระดมทุนสามารถเป็นเกราะป้องกันของสตาร์ตอัป เพื่อให้แน่ใจว่าสตาร์ทอัพมีกระแสเงินสดเพียงพอที่จะครอบคลุมค่าใช้จ่ายในแต่ละวัน จัดการค่าใช้จ่ายที่ไม่คาดคิด และเชื่อมช่องว่างระหว่างการชำระเงินให้กับซัพพลายเออร์และการรับเงินจากลูกค้า การระดมทุนยังสามารถช่วยจัดการช่องว่างกระแสเงินสดที่เกิดจากการชำระเงินล่าช้าหรือการลดลงตามฤดูกาล ซึ่งอาจก่อให้เกิดปัญหาร้ายแรงสำหรับสตาร์ทอัพ
- การขยายธุรกิจ
  เมื่อธุรกิจเติบโตขึ้น การระดมทุนจะมีความสำคัญสำหรับการขยายธุรกิจ การระดมทุนช่วยให้บริษัทสามารถขยายการดำเนินงาน เข้าสู่ตลาดใหม่ และคว้าโอกาสในการเติบโต ซึ่งอาจเกี่ยวข้องกับการจ้างบุคลากรใหม่ การขยายผลิตภัณฑ์และบริการที่นำเสนอ หรือการย้ายไปยังสถานที่ที่ใหญ่ขึ้น
- การอัปเกรดอุปกรณ์และเทคโนโลยี
  การลงทุนในอุปกรณ์และเทคโนโลยีใหม่สามารถเพิ่มประสิทธิภาพและผลผลิตของสตาร์ทอัพได้อย่างมาก การระดมทุนช่วยให้สตาร์ทอัพสามารถจัดหาเครื่องมือ เครื่องจักร และซอฟต์แวร์ที่จำเป็นเพื่อปรับปรุงการดำเนินงาน ปรับปรุงคุณภาพผลิตภัณฑ์ และรักษาความสามารถในการแข่งขัน
- การจ้างงานและการลงทุนในบุคลากร
  การสร้างทีมที่แข็งแกร่งเป็นสิ่งสำคัญสำหรับความสำเร็จของสตาร์ทอัพ การระดมทุนช่วยให้สตาร์ทอัพสามารถดึงดูดและรักษาบุคลากรที่มีความสามารถสูง โดยเสนอเงินเดือน ผลประโยชน์ และโอกาสในการพัฒนาวิชาชีพที่แข่งขันได้

- การตลาดและการโฆษณา
  การเข้าถึงกลุ่มเป้าหมายและการสร้างการรับรู้ถึงแบรนด์เป็นสิ่งจำเป็นสำหรับสตาร์ทอัพในการสร้างแรงผลักดันในตลาด การระดมทุนสามารถสนับสนุนแคมเปญการตลาดและการโฆษณา รวมถึงการตลาดดิจิทัล การโฆษณาบนโซเชียลมีเดีย และความพยายามด้านการประชาสัมพันธ์
- การวิจัยและพัฒนา
  การลงทุนในการวิจัยและพัฒนา (R&D) เป็นสิ่งสำคัญสำหรับสตาร์ทอัพในการสร้างสรรค์นวัตกรรมและก้าวนำหน้า การระดมทุนสามารถสนับสนุนกิจกรรม R&D ช่วยให้สตาร์ทอัพสำรวจเทคโนโลยีใหม่ พัฒนาผลิตภัณฑ์หรือบริการใหม่ และปรับปรุงข้อเสนอที่มีอยู่
- การปฏิบัติตามมาตรฐานการปฏิบัติตามกฎระเบียบ
  การระดมทุนมีบทบาทสำคัญในการทำให้แน่ใจว่าสตาร์ทอัพปฏิบัติตามมาตรฐานการปฏิบัติตามกฎระเบียบและข้อกำหนดทางกฎหมาย สามารถใช้เพื่อครอบคลุมค่าใช้จ่ายที่เกี่ยวข้องกับการได้รับการรับรอง ใบอนุญาต และการอนุญาต ตลอดจนการดำเนินการตามมาตรการการปฏิบัติตามที่จำเป็น
- การดึงดูดนักลงทุนในอนาคต
  การรักษาเงินทุนยังสามารถช่วยดึงดูดนักลงทุนในอนาคตได้ เมื่อสตาร์ทอัพระดมทุนได้สำเร็จจะแสดงให้เห็นถึงความน่าเชื่อถือและศักยภาพต่อนักลงทุนรายอื่น ทำให้การรักษาเงินทุนในรอบต่อๆ ไปง่ายขึ้น

แม้ว่าการระดมทุนจะสามารถเพิ่มโอกาสในการประสบความสำเร็จได้ แต่สิ่งสำคัญคือต้องยอมรับว่าการระดมทุนนั้นไม่ได้เป็นการรับประกันความสำเร็จ ซึ่งขึ้นอยู่กับปัจจัยต่างๆ อาทิ ประสบการณ์ของผู้ก่อตั้ง และรูปแบบธุรกิจที่แข็งแกร่งมีบทบาทสำคัญ นอกจากนี้ ผู้ก่อตั้งควรตระหนักถึงข้อเสียที่อาจเกิดขึ้นจากการระดมทุน อาทิ การลดสัดส่วนการถือหุ้น และแรงกดดันในการส่งมอบผลลัพธ์ ด้วยการตัดสินใจอย่างรอบคอบและการหาพันธมิตรด้านการระดมทุนที่เหมาะสม และการบริหารทรัพยากรทางการเงินอย่างมีประสิทธิภาพ สตาร์ทอัพจะสามารถเพิ่มโอกาสในการประสบความสำเร็จ และสร้างผลกระทบที่ยั่งยืนในอุตสาหกรรมของตนได้ (*Business Funding – Five Reasons You Need It*, 2024; *Types of Startup Funding: Pre-Seed, VC Financing, and More | Leader Bank*, 2024; Gottfeld, 2025)

2.1.3 การปรากฏตัวทางดิจิทัล (Digital Presence)

การปรากฏตัวทางดิจิทัล หมายถึง การแสดงตัวตนของบุคคลหรือองค์กรในสภาพแวดล้อมออนไลน์ ซึ่งครอบคลุมถึงความสามารถในการมองเห็น อัตลักษณ์ และการมีส่วนร่วมบนแพลตฟอร์มดิจิทัลต่างๆ การปรากฏตัวทางดิจิทัลมีบทบาทสำคัญในการกำหนดการรับรู้และปฏิสัมพันธ์ทั้งในบริบททางวิชาการและธุรกิจ ชี้ให้เห็นว่าการปรากฏตัวทางดิจิทัลของนักวิชาการดิจิทัลสามารถแตกต่างกันอย่างมาก โดยบางคนรักษาขอบเขตแบบดั้งเดิมระหว่างตัวตนออนไลน์และออฟไลน์ ในขณะที่บางคนใช้วิธีการที่ผสมผสานมากขึ้น นำไปสู่การสร้างโปรไฟล์ที่สอดคล้องกันแทนที่จะเป็นตัวตนที่แยกส่วน (Quintas-Mendes & Paiva, 2023)

ในแวดวงธุรกิจ การมีตัวตนออนไลน์ที่แข็งแกร่งเป็นสิ่งจำเป็นสำหรับความสำเร็จ เนื่องจากช่วยเพิ่มการมองเห็นและชื่อเสียง ซึ่งเป็นสิ่งสำคัญสำหรับการสร้างความได้เปรียบในการแข่งขัน (Cioppi et al., 2019) นอกจากนี้ การใช้สื่อสังคมออนไลน์อย่างมีกลยุทธ์ได้รับการพิสูจน์แล้วว่ามีอิทธิพลเชิงบวกต่อความสำเร็จทางการเงินของบริษัท ซึ่งเน้นย้ำถึงความสำคัญของการปรากฏตัวทางดิจิทัลอย่างมีเจตนา (DeLeon & Brown, 2023)

การปรากฏตัวทางดิจิทัลไม่ได้เป็นเพียงแค่การมีตัวตนออนไลน์เท่านั้น แต่ยังเกี่ยวข้องกับการจัดการอัตลักษณ์และการมีส่วนร่วมอย่างแข็งขันเพื่อสร้างความสัมพันธ์ที่มีความหมายและบรรลุผลลัพธ์ที่ต้องการในด้านต่างๆ (Almassry & Eid, 2023; Bruce et al., 2023) จึงมีความน่าสนใจที่จะนำปัจจัยนี้มาเป็นส่วนหนึ่งในการทำนายความสำเร็จของสตาร์ตอัป

2.1.4 ปัจจัยที่ใช้ในการทำนายความสำเร็จของสตาร์ตอัป

ในการทำนายความสำเร็จของสตาร์ตอัป มีปัจจัยสำคัญหลายประการที่ถูกระบุผ่านการวิจัย ปัจจัยอาทิ ประสบการณ์ก่อนหน้าในการเริ่มบริษัท ความมั่นใจในตนเอง การจัดการภายในองค์กรสตาร์ตอัป ความสามารถในระดับทีม การเป็นผู้ประกอบการ ความกล้าเสี่ยง การมุ่งเน้นที่ความสำเร็จ ความน่าเชื่อถือของคุณภาพบริการออนไลน์ และการสร้างเครือข่ายออนไลน์ ได้รับการเน้นย้ำว่าเป็นปัจจัยที่มีผลสำคัญต่อความสำเร็จของสตาร์ตอัป (Fitria & Hakim, 2022; Lovrincevic, 2022; Pham & Pham, 2021) นอกจากนี้ การมีพันธมิตรทางธุรกิจ อายุของบริษัท ความทุ่มเทของพันธมิตรทางธุรกิจ พบว่ามีอิทธิพลต่อความสำเร็จของสตาร์ตอัป (Díaz-Santamaría & Bulchand-Gidumal, 2021; Ковальов et al., 2024)

ยิ่งไปกว่านั้น ปัจจัยอาทิ คุณภาพของทรัพยากรบุคคล ความทุ่มเทของผู้ก่อตั้งสตาร์ตอัป และเจตนารมณ์เชิงกลยุทธ์ของสตาร์ตอัปในการมอบคุณค่าให้กับลูกค้า ถูกเน้นว่าเป็นสิ่งสำคัญสำหรับความสำเร็จของสตาร์ตอัป (Ehrensperger et al., 2020; Sadma, 2021; Zhou, 2023) นอกจากนี้ การเปิดกว้างของทีมต่อทรัพยากรภายนอก ทรัพยากรทางการเงิน เทคโนโลยี และทุนมนุษย์ และเกณฑ์การตัดสินใจอาทิ ภูมิภาค การแข่งขัน โอกาสในการระดมทุน และความเป็นผู้นำ ก็ถูกระบุว่าเป็นสิ่งสำคัญสำหรับความยั่งยืนและความสำเร็จของสตาร์ตอัป (Korpysa et al., 2023; Marullo et al., 2018)

การใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ในการทำนายความสำเร็จของสตาร์ตอัปนั้นได้รับความสนใจอย่างมากในช่วงหลายปีที่ผ่านมา งานวิจัยส่วนใหญ่มุ่งเน้นไปที่การใช้ปัจจัยที่เกี่ยวข้องกับข้อมูลภายในของบริษัท อาทิ ข้อมูลเกี่ยวกับผู้ก่อตั้ง ทีมบริหาร ความสามารถของบริษัทในด้านต่างๆ จำนวนเงินที่ได้รับจากการระดมทุน และข้อมูลทางการเงินอื่นๆ ดังที่กล่าวมาก่อนหน้านี้ อย่างไรก็ตาม ยังไม่มีการศึกษาที่นำปัจจัยด้านการปรากฏตัวทางดิจิทัล (Digital Presence) มาใช้ในการทำนายความสำเร็จของสตาร์ตอัปอย่างเป็นระบบ

การปรากฏตัวทางดิจิทัลนั้นเป็นปัจจัยสำคัญในยุคดิจิทัล ซึ่งครอบคลุมทั้งสื่อที่เป็นเจ้าของ (Owned Media) และสื่อที่ได้รับ (Earned Media) ของสตาร์ตอัป การวิเคราะห์ประสิทธิผลของการปรากฏตัวทางดิจิทัลนี้อาจให้ข้อมูลเชิงลึกที่สำคัญเกี่ยวกับความสามารถของสตาร์ตอัปในการสร้างการ

รับรู้ ดึงดูดลูกค้า และสร้างความน่าเชื่อถือในตลาด ซึ่งล้วนเป็นปัจจัยที่อาจส่งผลต่อความสำเร็จของสตาร์ตอัป

งานวิจัยนี้จึงมีจุดมุ่งหมายที่จะเติมเต็มช่องว่างในองค์ความรู้ปัจจุบัน โดยการนำปัจจัยด้านการปรากฏตัวทางดิจิทัลมาใช้ในการพัฒนาโมเดลทำนายความสำเร็จของสตาร์ตอัป โดยใช้เทคนิคการเรียนรู้ของเครื่อง การศึกษานี้จะช่วยขยายความเข้าใจเกี่ยวกับปัจจัยที่ส่งผลต่อความสำเร็จของสตาร์ตอัป และอาจนำไปสู่การพัฒนาเครื่องมือที่มีประสิทธิภาพมากขึ้นในการประเมินและทำนายศักยภาพของสตาร์ตอัปในอนาคต

จากข้อมูลดังกล่าว เราสามารถแบ่งกลุ่มปัจจัยที่ส่งผลต่อความสำเร็จของสตาร์ตอัปได้เป็นมิติต่างๆ ดังนี้:

1. การปรากฏตัวทางดิจิทัล (Digital Presence Dimension)
    1.1. สื่อที่เป็นเจ้าของ (Owned Media)
    1.2. สื่อที่ได้รับ (Earned Media)
2. ข้อมูลพื้นฐานของธุรกิจ (Business Foundation Dimension)
    2.1. ข้อมูลเกี่ยวกับผู้ก่อตั้ง
    2.2. ทีมบริหาร
    2.3. ความสามารถของบริษัทในด้านต่างๆ
3. การเงิน (Financial Dimension)
    3.1. จำนวนเงินที่ได้รับจากการระดมทุน
    3.2. ข้อมูลทางการเงินอื่นๆ
4. การตลาดและอุตสาหกรรม (Market and Industry Dimension)
    4.1. ขนาดของตลาด
    4.2. การเติบโตของอุตสาหกรรม
    4.3. คู่แข่งในตลาด
5. นวัตกรรมและเทคโนโลยี (Innovation and Technology Dimension)
    5.1. จำนวนสิทธิบัตร
    5.2. การลงทุนในการวิจัยและพัฒนา

ซึ่งยังไม่ปรากฏการศึกษามิติด้านการปรากฏตัวทางดิจิทัลเป็นหลัก ซึ่งเป็นมิติที่ยังไม่ได้รับการศึกษาอย่างเป็นระบบในงานวิจัยก่อนหน้านี้ แต่มีความสำคัญมากขึ้นในยุคดิจิทัลที่สตาร์ตอัปใช้เป็นเครื่องมือสำคัญในการดำเนินกิจการ

2.1.5 การใช้ประโยชน์จาก Digital Marketing เพื่อความสำเร็จของ Startup

สตาร์ตอัปสามารถเพิ่มความสำเร็จของตนได้โดยการใช้ประโยชน์จากการปรากฏตัวทางดิจิทัล (Digital Presence) ผ่านกลยุทธ์และแนวทางต่างๆ งานวิจัยระบุว่าการมีส่วนร่วมอย่างกระตือรือร้นบนแพลตฟอร์ม Social Media สามารถเพิ่มโอกาสความสำเร็จของสตาร์ตอัปและการดึงดูดเงินทุนจากนักลงทุนได้อย่างมาก (Peixoto et al., 2023) การใช้ Social Media ไม่เพียงแต่ช่วยในการสร้างชุมชนแต่ยังช่วยในการเพิ่มประสิทธิภาพการตลาด การติดตามเทรนด์

และการให้บริการลูกค้าที่ดีเยี่ยม การทำงานร่วมกันและการมีส่วนร่วมในระบบนิเวศดิจิทัลมีบทบาทสำคัญในการขับเคลื่อนความสำเร็จของสตาร์ตอัปในยุคดิจิทัล (Joel et al., 2024)

สตาร์ตอัปดิจิทัลสามารถได้รับประโยชน์จากแนวทางเชิงกลยุทธ์ที่รวมถึงความยืดหยุ่นขององค์กร นวัตกรรม และการมีส่วนร่วมของลูกค้าเพื่อให้เกิดการเติบโตที่ยั่งยืนและความได้เปรียบในการแข่งขันในตลาดดิจิทัล (Moradi et al., 2024) โดยการใช้ Digital Marketing เป็นฐานและการทำงานร่วมกับซัพพลายเออร์และผู้จัดจำหน่ายที่เชื่อถือได้ สตาร์ตอัปสามารถเพิ่มการพัฒนาและการเติบโตของตนได้ (Tejaningrum, 2022) นอกจากนี้ การนำกลยุทธ์การตลาดดิจิทัลมาใช้เป็นสิ่งสำคัญสำหรับสตาร์ตอัปในการเพิ่มการรับรู้ของแบรนด์ การเติบโตของลูกค้า และความสำเร็จโดยรวม (“Examining the Role of Advertising to Promote a Startup Business,” 2024)

นอกจากนี้ การใช้ประโยชน์จากโซลูชั่นดิจิทัลล้ำสมัย อาทิ Automation Software, Cloud Computing และ Artificial Intelligence สามารถเพิ่มผลิตภาพ ลดค่าใช้จ่ายในการดำเนินงาน และทำให้การดำเนินงานของสตาร์ตอัปง่ายขึ้น (Muley et al., 2024) การตลาดดิจิทัล Web 3.0 ได้รับการระบุว่าเป็นปัจจัยสำคัญที่ส่งเสริมความสำเร็จและการเติบโตของธุรกิจสตาร์ตอัป (Aladenika et al., 2022) นอกจากนี้ บทบาทของกลยุทธ์ดิจิทัลและความคิดสร้างสรรค์ของมนุษย์ยังมีความสำคัญในการสร้าง Technopreneurship ที่เปลี่ยนแปลงไป เน้นความสำคัญของการตอบสนองต่อความก้าวหน้าทางเทคโนโลยีเพื่อการพัฒนาสตาร์ตอัป (Indrianti et al., 2023)

สรุปแล้ว สตาร์ตอัปสามารถใช้ประโยชน์จากการปรากฏตัวทางดิจิทัล (Digital Presence) ได้โดยการมีส่วนร่วมอย่างกระตือรือร้นบน Social Media การนำกลยุทธ์การตลาดดิจิทัลที่เป็นนวัตกรรมมาใช้ การใช้ประโยชน์จากเทคโนโลยีล้ำสมัย และการส่งเสริมความคิดสร้างสรรค์และการทำงานร่วมกันในระบบนิเวศดิจิทัล โดยการนำแนวทางเหล่านี้มาใช้ สตาร์ตอัปสามารถเพิ่มการมองเห็น ดึงดูดนักลงทุน และขับเคลื่อนการเติบโตที่ยั่งยืนในสภาพแวดล้อมดิจิทัลที่มีการแข่งขันสูง

2.1.6 การใช้การเรียนรู้ของเครื่อง (Machine learning) ในการทำนาย

Machine Learning ได้กลายเป็นเครื่องมือที่ทรงพลังในการทำนายในหลายสาขา มันเกี่ยวข้องกับการเทรนโมเดลการทำนายเพื่อให้สามารถคาดการณ์ได้อย่างแม่นยำโดยอิงจากข้อมูล (Ranjan & Goldsztein, 2022) เทคนิคนี้ถูกนำมาใช้ในหลายๆ ด้าน อาทิ การดูแลสุขภาพ การเงิน การศึกษา และการขนส่ง ยกตัวอย่างอาทิ ในการดูแลสุขภาพ Machine Learning ถูกนำมาใช้ในการทำนายอัตราการเสียชีวิตใน ICU ระยะเวลาการพักรักษาตัว และอัตราความสำเร็จของการกำจัดนิ่วหลังการใช้ Shock Wave Lithotripsy (IWASE et al., 2021; S. W. Yang et al., 2020) ในด้านการเงิน Machine Learning ถูกนำมาใช้ในการทำนายเสถียรภาพทางการเงินทั่วโลกและราคาหุ้น (Shi, 2023; M. Yang, 2023) นอกจากนี้ ในด้านการศึกษา Machine Learning ถูกนำมาใช้ในการทำนายคุณลักษณะการเรียนรู้และผลการเรียนของนักเรียน (Su et al., 2022)

ความยืดหยุ่นและความแม่นยำของโมเดล Machine Learning ได้รับความสนใจเมื่อเปรียบเทียบกับโมเดลสถิติแบบดั้งเดิม โมเดล Machine Learning ขับเคลื่อนด้วยข้อมูลและสามารถปรับตัวกับข้อมูลประเภทต่างๆ ส่งผลให้มีความแม่นยำในการทำนายสูงขึ้น (Lazar et al., 2019) โมเดลเหล่านี้มีบทบาทสำคัญในการทำนายผลลัพธ์ต่างๆ อาทิ การจ้างงาน ผลลัพธ์หลังเกิดโรคหลอด

เลือดสมอง และเหตุการณ์อันตรายในบ่อน้ำมัน (Khan et al., 2020; Research Scholar, Department of Computer Science, Sacred Heart College, Tirupattur, Tamil Nadu, India. et al., 2020; Su et al., 2022)

นอกจากนี้ อัลกอริธึม Machine Learning ถูกนำมาใช้ในการทำนายปรากฏการณ์ที่หลากหลาย ตั้งแต่โปรโตคอลการแจกจ่ายกุญแจควอนตัมไปจนถึงการเกิดมะเร็งกระเพาะอาหารซ้ำหลังการผ่าตัด (R. et al., 2022; Zhou, 2023) ความสามารถในการทำนายของ Machine Learning ถูกนำมาใช้ในการทำนายยอดขาย การใช้โหมดการขนส่ง และการเคลื่อนไหวของราคาหุ้น (Lazar et al., 2019; Mao, 2022; M. Yang, 2023) นอกจากนี้ Machine Learning ยังถูกนำมาใช้ในการทำนายผลลัพธ์ทางคลินิกหลังจากการผ่าตัดอาทิ การผ่าตัดต่อมลูกหมากด้วยหุ่นยนต์และการสลายนิ่วผ่านผิวหนัง (Wong et al., 2019; T. Zhang et al., 2024)

สรุปแล้ว Machine Learning ได้ปฏิวัติแนวทางการทำนายโดยการนำเสนอเครื่องมือที่ทรงพลังในการวิเคราะห์ข้อมูล การระบุรูปแบบ และการคาดการณ์อย่างแม่นยำในหลายสาขา การใช้งานของมันยังคงขยายตัว แสดงถึงศักยภาพในการขับเคลื่อนนวัตกรรมและปรับปรุงกระบวนการตัดสินใจกับข้อมูลที่ไม่เป็นเชิงเส้น (Non-linear) มีความซับซ้อนสูง และมีรูปแบบที่ยากต่อการอธิบายด้วยสถิติแบบดั้งเดิม ทำให้สามารถทำนายผลลัพธ์ในงานที่วิธีการทางสถิติทั่วไปอาจไม่เพียงพอ หรือไม่สามารถทำได้เลย

2.1.7 การทำนายโดยใช้สถิติ กับการใช้การเรียนรู้ของเครื่อง

Machine Learning มีข้อได้เปรียบหลายประการเหนือโมเดลสถิติแบบดั้งเดิมตามที่ได้เน้นในงานวิจัยต่างๆ (Wang et al., 2022) เน้นย้ำว่า วิธีการ Machine Learning เมื่อผสานกับสถิติข้อมูลด้านสุขภาพสามารถดึงข้อมูลที่ซ่อนอยู่จากชุดข้อมูลขนาดใหญ่ได้ดีกว่า แสดงให้เห็นถึงความสามารถในการเรียนรู้และการทำนายทั่วไปที่เหนือกว่าวิธีการสถิติแบบดั้งเดิม

Lyu & Yong (2024) (Lyu & Yong, 2024) ได้กล่าวถึงข้อดีของเทคนิค Machine Learning เหนือวิธีการสถิติแบบดั้งเดิม โดยชี้ให้เห็นว่า Machine Learning มีความเชี่ยวชาญในการแก้ปัญหาที่ซับซ้อนซึ่งยากต่อวิธีการแบบดั้งเดิมอาทิ Simple Linear Regression Machine Learning สามารถผสานข้อมูลเพิ่มเติมได้อย่างมีประสิทธิภาพในต้นทุนการคำนวณที่ต่ำกว่า และปรับโครงสร้างโมเดลเฉพาะสำหรับปัญหาที่กำหนด

นอกจากนี้ Levy & O'Malley (2020) (Levy & O'Malley, 2020) ได้ชี้ให้เห็นว่ามีการศึกษามากมายที่แสดงให้เห็นถึงประสิทธิภาพที่เหนือกว่าของเทคนิค Machine Learning เมื่อเทียบกับโมเดลสถิติแบบดั้งเดิม (Jing et al., 2022) สนับสนุนแนวคิดนี้โดยเสนอว่าอัลกอริทึม Machine Learning มักจะทำงานได้ดีกว่าวิธีการสถิติแบบดั้งเดิม โดยเฉพาะในสถานการณ์ที่มีตัวพยากรณ์จำนวนมากและมีผลปฏิสัมพันธ์ที่สำคัญ

ยิ่งไปกว่านั้น Kwag et al. (2020) (Kwag et al., 2020) พบว่า วิธีการ Machine Learning มักจะมีประสิทธิภาพดีกว่าโมเดลสถิติแบบดั้งเดิมในการศึกษาการประเมินประสิทธิภาพของการเกิดแผ่นดินไหว Wallis et al. (2022) (Wallis et al., 2022) ยังเน้นว่า เทคนิค Machine Learning

อาทิ Artificial Neural Networks และ Support Vector Machines มีประสิทธิภาพดีกว่าโมเดลแบบดั้งเดิมในแง่ของความแม่นยำและมาตรวัดทางสถิติ

สรุปแล้ว งานวรรณกรรมวิจัยแสดงให้เห็นอย่างสม่ำเสมอว่า โมเดล Machine Learning มีข้อได้เปรียบเหนือวิธีการสถิติแบบดั้งเดิมโดยให้ประสิทธิภาพในการทำนายที่ดีกว่า สามารถจัดการกับความสัมพันธ์ที่ซับซ้อนในข้อมูล ผสานข้อมูลเพิ่มเติมได้อย่างมีประสิทธิภาพ และปรับตัวให้เข้ากับโครงสร้างปัญหาเฉพาะ

2.1.8 โมเดลการเรียนรู้ของเครื่องที่ใช้ในการทำนายความสำเร็จของสตาร์ตอัป

ในการพยากรณ์ความสำเร็จของสตาร์ตอัป อัลกอริทึม Machine Learning มีบทบาทสำคัญโดยการวิเคราะห์ปัจจัยต่าง ๆ ที่ส่งผลต่อความสำเร็จของธุรกิจเหล่านี้ หนึ่งในงานวิจัยโดย Bargagli-Stoffi et al. (2021) (Bargagli-Stoffi et al., 2021) ชี้ให้เห็นว่าอัลกอริทึม Supervised Learning (SL) ที่ถูกฝึกด้วยชุดข้อมูลขนาดใหญ่เหมาะสมอย่างยิ่งในการพยากรณ์ความสำเร็จของสตาร์ตอัป โดยเฉพาะเมื่อปัจจัยที่ส่งผลต่อความสำเร็จนั้นไม่ชัดเจนและมีความซับซ้อน ซึ่งบ่งชี้ว่าโมเดล Machine Learning สามารถจัดการกับความไม่แน่นอนและความซับซ้อนที่เกี่ยวข้องกับการพยากรณ์ความสำเร็จของสตาร์ตอัปได้อย่างมีประสิทธิภาพ

นอกจากนี้ Panduri (2023) (Panduri et al., 2023) ยังเน้นถึงการสร้างโมเดล Machine Learning ที่ยั่งยืนโดยใช้ตัวแปรหลายประการในการพยากรณ์ความสำเร็จของธุรกิจสตาร์ตอัป โดยการรวมปัจจัยที่หลากหลายเข้ากับโมเดลการพยากรณ์ อาทิ สภาวะตลาด, มาตรการทางการเงิน และแนวโน้มอุตสาหกรรม โมเดล Machine Learning ที่ยั่งยืนเหล่านี้สามารถให้ข้อมูลเชิงลึกที่มีค่าเกี่ยวกับความน่าจะเป็นของความสำเร็จของสตาร์ตอัป

นอกจากนี้ Allu (2020) (Research Scholar, Department of Computer Science and Engineering, GITAM (Deemed to be University), Visakhapatnam, India and Associate Director in Novartis, Hyderabad, India. et al., 2020) ยังได้กล่าวถึงการนำเทคนิค Machine Learning มาใช้ รวมถึงอัลกอริทึม Random Forest ในการพยากรณ์ความสำเร็จของสตาร์ตอัป ซึ่งช่วยให้นักธุรกิจสามารถตัดสินใจอย่างมีข้อมูล โดยการใช้ประโยชน์จากอัลกอริทึมเหล่านี้ สตาร์ตอัปสามารถประเมินศักยภาพของความสำเร็จและวางกลยุทธ์ให้เหมาะสมเพื่อเพิ่มโอกาสในการเติบโตในตลาดที่มีการแข่งขันสูง

สรุปได้ว่า การเรียนรู้ของเครื่องเป็นเครื่องมือที่ทรงพลังสำหรับการพยากรณ์ความสำเร็จของสตาร์ตอัปโดยการวิเคราะห์ปัจจัยต่าง ๆ การโต้ตอบระหว่างปัจจัย และตัวแปรที่มีผลต่อผลลัพธ์ของธุรกิจเหล่านี้ โดยการใช้ประโยชน์จากอัลกอริทึมที่ซับซ้อนและชุดข้อมูลขนาดใหญ่ โมเดลเหล่านี้สามารถให้ข้อมูลเชิงลึกที่มีค่าต่อผู้ประกอบการ นักลงทุน และผู้มีส่วนได้ส่วนเสีย ช่วยให้พวกเขาสามารถตัดสินใจได้อย่างมีข้อมูลและเพิ่มโอกาสในการประสบความสำเร็จของสตาร์ตอัป

Random Forest

Random Forest เป็นเทคนิคการเรียนรู้แบบกลุ่ม (Ensemble Learning) ที่นิยมใช้อย่างแพร่หลายในงานด้าน Machine Learning มีประสิทธิภาพสูงในการทำงานทั้งด้านการจำแนก

ประเภท (Classification) และการวิเคราะห์ถดถอย (Regression) โดยเทคนิคนี้ถูกนำเสนอครั้งแรกโดย Breiman เมื่อปี ค.ศ. 2001 ซึ่ง Random Forest ทำงานโดยการสร้างต้นไม้ตัดสินใจ (Decision Tree) หลายต้นในขั้นตอนการฝึกโมเดล จากนั้นรวมผลลัพธ์ด้วยการโหวตเลือกคลาสที่ปรากฏมากที่สุด (Mode) ในกรณีการจำแนกประเภท หรือใช้ค่าเฉลี่ยของการทำนายสำหรับงานด้านการถดถอย (Xing et al., 2022) แนวทางแบบกลุ่มนี้ช่วยเสริมประสิทธิภาพในการทำนาย และลดความเสี่ยงจากปัญหาการเกิดการเรียนรู้เกิน (Overfitting) ซึ่งส่งผลให้ Random Forest เหมาะสมอย่างยิ่งกับชุดข้อมูลที่มีตัวแปรจำนวนมาก หรือมีความสัมพันธ์ระหว่างข้อมูลที่ซับซ้อน (Duroux & Scornet, 2018; Ziegler & König, 2014).

ความสามารถรอบด้านของ Random Forest สะท้อนผ่านความสำเร็จในการนำไปประยุกต์ใช้ในหลากหลายสาขา ตั้งแต่การทำนายผลลัพธ์ทางสุขภาพไปจนถึงการประเมินด้านสิ่งแวดล้อม ตัวอย่างเช่น มีรายงานว่าการใช้ Random Forest สามารถทำนายการเสียชีวิตหลังการผ่าตัดของผู้ป่วยสูงอายุได้อย่างแม่นยำ ซึ่งแสดงถึงประสิทธิภาพในการจัดการข้อมูลทางการแพทย์ที่โมเดลแบบดั้งเดิมอาจทำได้ไม่ดีเท่า นอกจากนี้ ในด้านวิทยาศาสตร์สิ่งแวดล้อม Random Forest ยังถูกใช้ในการจำลองรูปแบบเชิงพื้นที่ของเหตุการณ์ไฟป่า โดยแสดงให้เห็นถึงประสิทธิภาพการทำนายที่เหนือกว่าวิธีการถดถอยแบบดั้งเดิม เช่น การถดถอยเชิงเส้นพหุคูณ (Oliveira et al., 2012) สะท้อนให้เห็นถึงความสามารถของ Random Forest ในการวิเคราะห์ความสัมพันธ์ที่ไม่เป็นเชิงเส้น (non-linear relationships) และการมีปฏิสัมพันธ์ระหว่างตัวแปรต่างๆ ซึ่งเป็นเหตุผลสำคัญที่ Random Forest ได้รับความนิยมเพิ่มมากขึ้นในหลายๆ สาขาการวิจัย

นอกจากนี้ ประสิทธิภาพของ Random Forest บางครั้งอาจสูงกว่าวิธีการทั่วไป โดยเฉพาะในกรณีที่โครงสร้างข้อมูลมีความซับซ้อนที่เหมาะกับวิธีการของ Random Forest มากกว่า ตัวอย่างเช่น ในงานวิจัยที่เกี่ยวกับการตรวจจับการฉ้อโกงธุรกรรมการเงินผ่านมือถือ Random Forest แสดงให้เห็นถึงความแม่นยำที่โดดเด่น เมื่อเปรียบเทียบกับการถดถอยโลจิสติก (Logistic Regression) โดยสามารถบรรลุความแม่นยำสูงสุดถึง 98% อย่างไรก็ตาม สิ่งสำคัญที่ควรตระหนักคือ Random Forest ไม่ได้เหนือกว่าวิธีอื่นๆ เสมอไป โดยในบางสถานการณ์ที่มีความสัมพันธ์แบบเชิงเส้นชัดเจน วิธีการเชิงเส้น (Linear Methods) อาจมีประสิทธิภาพดีกว่าหรือเทียบเท่า (P. Liu et al., 2020).

Random Forest ยังขยายขอบเขตการประยุกต์ใช้ไปยังการประมาณค่าความน่าจะเป็น (Probability Estimation) โดยมีเป้าหมายที่จะให้ผลลัพธ์ในรูปแบบเชิงความน่าจะเป็นแทนที่จะเป็นเพียงการทำนายแบบจัดประเภทเท่านั้น ความสามารถนี้มีความสำคัญอย่างยิ่งในสถานการณ์ที่การเข้าใจระดับความแน่นอนของการทำนายมีความสำคัญไม่แพ้ตัวผลลัพธ์ที่ทำนายออกมา เช่น การวินิจฉัยทางการแพทย์หรือการประเมินความเสี่ยง อย่างไรก็ตาม Random Forest จำเป็นต้องมีการปรับแต่งหรือสอบเทียบ (Calibration) อย่างเหมาะสมเพื่อให้ได้ผลลัพธ์ที่แม่นยำสูงสุด

ยิ่งไปกว่านั้นแบบจำลองยังสามารถประยุกต์ใช้กับงานที่ต้องการความรวดเร็วแบบเรียลไทม์ เช่น งานด้านการจำแนกประเภทแพ็กเก็ตข้อมูลบนเครือข่ายความเร็วสูง โดยแสดงให้เห็นถึงความสามารถในการปรับตัวในสภาพแวดล้อมที่ต้องการประมวลผลแบบเรียลไทม์ แม้ว่าจะยังพบความท้าทายในการรักษาความแม่นยำในระดับสูงอยู่บ้าง ซึ่งประเด็นนี้สะท้อนถึงความจำเป็นที่จะต้องรักษาสมดุลระหว่างประสิทธิภาพเชิงการคำนวณ (Computational Efficiency) กับความแม่นยำใน

การจำแนกประเภท เพื่อให้แน่ใจว่าแอปพลิเคชันสามารถใช้ประโยชน์จากจุดแข็งของแบบจำลองได้อย่างเหมาะสมที่สุด (S.-Y. Wang & Wu, 2023)

โดยสรุป Random Forest เป็นวิธีการเรียนรู้แบบกลุ่มที่ได้รับความนิยมเนื่องจากประสิทธิภาพในการแก้ปัญหาซับซ้อนหลากหลายรูปแบบ และยังคงเป็นที่สนใจของการศึกษาวิจัยในสาขาต่าง ๆ อย่างต่อเนื่อง

XGBoost

XGBoost หรือ Extreme Gradient Boosting เป็นอัลกอริทึมการเรียนรู้ของเครื่องแบบกลุ่ม (ensemble) ที่มีประสิทธิภาพสูงและได้รับการยอมรับอย่างกว้างขวางในการประยุกต์ใช้งานด้านการสร้างโมเดลเชิงทำนายหลากหลายประเภท โดย XGBoost พัฒนาต่อยอดจากกรอบแนวคิดพื้นฐานของ Gradient Boosting ด้วยการเพิ่มเทคนิคการควบคุมความซับซ้อน (Regularization) ที่มีประสิทธิภาพยิ่งขึ้น ส่งผลให้สามารถเพิ่มประสิทธิภาพในการทำนายกับข้อมูลใหม่ (Reneralization) และลดความเสี่ยงในการเกิดการเรียนรู้เกินขนาด (Overfitting) ซึ่งเป็นปัญหาทั่วไปในการพัฒนาโมเดลการเรียนรู้ของเครื่อง

จุดเด่นสำคัญของ XGBoost อยู่ที่แนวทางการฝึกโมเดลที่ใช้ต้นไม้ตัดสินใจ (Decision Tree) เป็นตัวเรียนรู้พื้นฐานภายใต้กรอบการทำงานแบบ Boosting โดย Chen และ Guestrin อธิบายว่า XGBoost เป็นระบบการเรียนรู้แบบ Boosting ด้วยต้นไม้ที่มีความสามารถในการขยายขนาด (Scalable) และครบวงจร (End-to-end) ที่มีข้อดีสำคัญทั้งด้านประสิทธิภาพในการคำนวณสูงและความยืดหยุ่นในการประมวลผลข้อมูลประเภทต่างๆ (T. Chen & Guestrin, 2016) อัลกอริทึมนี้ถูกออกแบบมาให้สามารถใช้ทรัพยากรในการฝึกโมเดลและการคำนวณได้อย่างเหมาะสม ทำให้เหมาะสมอย่างยิ่งสำหรับข้อมูลชุดใหญ่ที่มีตัวแปรจำนวนมาก features (Y. Liu et al., 2023; L. Wang et al., 2020).

จากงานวิจัยล่าสุดจำนวนมากแสดงให้เห็นถึงประสิทธิภาพของ XGBoost ในการประยุกต์ใช้งานด้านต่างๆ โดยเฉพาะในวงการแพทย์และการดูแลสุขภาพ เช่น งานวิจวิจัยของ Liu et al. พบว่า XGBoost มีความแม่นยำสูงในการทำนายปัจจัยเสี่ยงการกลับเป็นซ้ำของมะเร็งลำไส้ใหญ่ และให้ผลลัพธ์ดีกว่าโมเดลอย่าง Random Forest และ Support Vector Machine (SVM) ในการจัดการกับข้อมูลจำนวนมาก เช่นเดียวกันกับการศึกษาด้านการทำนายความเสี่ยงของโรคเบาหวานชนิดที่ 2 ซึ่ง XGBoost ให้ผลการทำนายที่มีความแม่นยำสูงกว่าโมเดลการเรียนรู้ของเครื่องแบบดั้งเดิมอย่างมีนัยสำคัญ ทั้งนี้ความสามารถของอัลกอริทึมในการจัดการข้อมูลที่มีความไม่สมดุล (Imbalanced Data) พร้อมคงประสิทธิภาพที่ดี ยังเพิ่มความน่าสนใจในการนำไปใช้งานจริงมากขึ้น (P. Zhang et al., 2022)

ยิ่งไปกว่านั้น การออกแบบเชิงโครงสร้างของ XGBoost ยังช่วยให้สามารถนำไปผสมผสานกับอัลกอริทึมและเทคนิคอื่นๆ ได้อย่างมีประสิทธิภาพ ตัวอย่างเช่น โมเดลลูกผสม (Hybrid Models) ที่รวม XGBoost เข้ากับอัลกอริทึมทางพันธุกรรม (Genetic Algorithms) ซึ่งแสดงให้เห็นถึงศักยภาพในการทำนายที่เพิ่มขึ้นอย่างชัดเจน โดยเฉพาะในการวินิจฉัยข้อผิดพลาดในทางวิศวกรรม (Wu et al., 2021) ความยืดหยุ่นและการปรับตัวที่ดีของ XGBoost ส่งผลให้สามารถนำไปใช้ในหลากหลายสาขา

ตั้งแต่การพยากรณ์ทางการเงินจนถึงการวินิจฉัยทางการแพทย์ แสดงให้เห็นถึงความหลากหลายในการประยุกต์ใช้ของอัลกอริทึม

นอกจากนี้ การพัฒนาเทคนิคด้านการหาค่าพารามิเตอร์ที่เหมาะสม (hyperparameter tuning) ด้วยอัลกอริทึมเชิงเมตาฮิวริสติก (metaheuristic algorithms) ยังช่วยเพิ่มประสิทธิภาพของอัลกอริทึมให้สูงขึ้นไปอีก สามารถบรรลุผลลัพธ์ที่แม่นยำอย่างมากในการแข่งขันด้านการเรียนรู้ของเครื่อง (Gulsun & Aydin, 2024) เช่นเดียวกับที่ Abdulrazzak และคณะได้กล่าวถึงว่า XGBoost ถูกนำไปใช้อย่างมีประสิทธิภาพในแอปพลิเคชันแบบเรียลไทม์ เช่น การตรวจจับโรคดีซ่าน ซึ่งประโยชน์สำคัญอยู่ที่ความเร็วในการฝึกและการทำนายผล (Abdulrazzak et al., 2024)

โดยสรุป XGBoost มีความโดดเด่นในวงการเรียนรู้ของเครื่องจากประสิทธิภาพที่แข็งแกร่ง การคำนวณที่มีประสิทธิภาพสูง และความสามารถในการประยุกต์ใช้งานที่หลากหลาย อัลกอริทึมนี้สามารถจัดการปัญหาทั่วไปในการเรียนรู้ของเครื่อง เช่น ปัญหาการเรียนรู้เกินขนาดและความซับซ้อนเชิงคำนวณได้อย่างดีเยี่ยม ส่งผลให้เป็นเครื่องมือสำคัญสำหรับนักวิทยาศาสตร์ข้อมูลและผู้ปฏิบัติงานที่ต้องการใช้ประโยชน์จากศักยภาพของการเรียนรู้ของเครื่องอย่างเต็มที่

Gradient Boosting

Gradient Boosting เป็นเทคนิคพื้นฐานที่สำคัญในด้านการเรียนรู้ของเครื่อง และเป็นที่ยอมรับอย่างกว้างขวางในด้านประสิทธิภาพของการประยุกต์ใช้งานในการสร้างแบบจำลองเชิงทำนายที่หลากหลาย เทคนิคนี้ทำงานด้วยการรวมกลุ่มของตัวเรียนรู้ที่มีความสามารถจำกัด (Weak Learners) ซึ่งส่วนใหญ่คือต้นไม้ตัดสินใจ (Decision Trees) ที่ได้รับการฝึกฝนอย่างต่อเนื่องเพื่อลดข้อผิดพลาดของการทำนาย โมเดลใหม่แต่ละตัวถูกสร้างขึ้นเพื่อแก้ไขความผิดพลาดของโมเดลก่อนหน้า ซึ่งส่งผลให้ความแม่นยำและความแข็งแกร่งของผลการทำนายโดยรวมดีขึ้น (Natekin & Knoll, 2013) Gradient Boosting ได้แสดงประสิทธิภาพที่ดีในการจัดการปัญหาที่มีความซับซ้อนสูง โดยสามารถจับรูปแบบที่ซับซ้อนซึ่งโมเดลที่เรียบง่ายกว่าอาจมองข้าม (Dorogush et al., 2018; Y. Mao, 2024)

อีกแง่มุมสำคัญของ Gradient Boosting คือการปรับแต่งพารามิเตอร์ (Hyperparameter Tuning) ซึ่งมีผลกระทบต่อประสิทธิภาพของโมเดลอย่างมาก โดยการปรับพารามิเตอร์ที่เหมาะสมมักเกี่ยวข้องกับการหาจุดสมดุลของพารามิเตอร์ เช่น อัตราการเรียนรู้ (Learning Rate) และจำนวนรอบการทำซ้ำ (Iterations) โดยทั่วไปแล้ว อัตราการเรียนรู้ที่ต่ำมักให้ผลลัพธ์ในการทำนายที่ดีขึ้น (Generalization Performance) ในขณะที่อัตราการเรียนรู้ที่สูงเกินไปอาจส่งผลให้เกิดปัญหาการเรียนรู้เกินขนาด (Overfitting) โดยเฉพาะอย่างยิ่งเมื่อไม่ได้ควบคุมจำนวนโมเดลในกลุ่มอย่างเหมาะสม เทคนิคเช่นการตรวจสอบแบบข้ามกลุ่มข้อมูล (Cross-validation) สามารถช่วยในการเลือกพารามิเตอร์ที่ดีที่สุด ช่วยให้โมเดลสามารถทำงานได้ดีในชุดข้อมูลที่หลากหลาย (Bentéjac et al., 2021; Yoon et al., 2023)

แม้จะมีประสิทธิภาพที่สูง Gradient Boosting ยังมีความท้าทายที่ต้องระวัง โดยเฉพาะการเรียนรู้เกินขนาดที่อาจเกิดขึ้นเมื่อโมเดลมีความซับซ้อนสูงจากการทำซ้ำหลายรอบโดยขาดการตรวจสอบอย่างเหมาะสม ดังนั้น งานวิจัยที่กำลังดำเนินอยู่จึงให้ความสำคัญกับการศึกษาเชิงทฤษฎีของ Gradient Boosting เช่น ความสม่ำเสมอของตัวประมาณ (Estimators) ในกรณีพิเศษต่างๆ ซึ่ง

สามารถให้ข้อมูลเชิงลึกและกรอบแนวทางที่ชัดเจนเพื่อจัดการความท้าทายเหล่านี้ได้อย่างมีประสิทธิภาพ (Velthoen et al., 2023)

โดยสรุป Gradient Boosting เป็นเทคนิคที่มีศักยภาพสูงและสามารถปรับใช้ได้อย่างกว้างขวางในวงการเรียนรู้ของเครื่อง ด้วยลักษณะของกระบวนการแก้ไขข้อผิดพลาดอย่างต่อเนื่อง ความสามารถในการปรับแต่งพารามิเตอร์ได้อย่างยืดหยุ่น และความสามารถในการประยุกต์ใช้งานที่หลากหลาย ส่งผลให้ Gradient Boosting กลายเป็นเครื่องมือสำคัญทั้งในการวิจัยทางวิชาการและการประยุกต์ใช้งานจริงในภาคอุตสาหกรรม

AdaBoost

AdaBoost หรือ Adaptive Boosting เป็นหนึ่งในอัลกอริทึมการเรียนรู้แบบกลุ่ม (Ensemble Learning) ที่ริเริ่มโดย Freund และ Schapire โดยมีหลักการสำคัญคือการเพิ่มประสิทธิภาพของตัวจำแนกที่อ่อนแอ (weak classifiers) ผ่านการรวมตัวกันเพื่อสร้างตัวจำแนกที่แข็งแกร่ง (Strong Classifier) หัวใจสำคัญของ AdaBoost คือการเน้นไปที่ตัวอย่างข้อมูลที่ยากต่อการจำแนก ทำให้สามารถปรับปรุงความแม่นยำของการทำนายในงานต่างๆ ได้ดีขึ้น อัลกอริทึมนี้ดำเนินการโดยการปรับน้ำหนักของตัวอย่างที่จำแนกผิดในแต่ละรอบของการเรียนรู้ และค่อยๆ ปรับปรุงโมเดลโดยรวมอย่างต่อเนื่อง ลักษณะการปรับตัวเช่นนี้ทำให้ AdaBoost สามารถรับมือกับสถานการณ์ต่างๆ ได้อย่างมีประสิทธิภาพ ทั้งชุดข้อมูลที่มีความไม่สมดุลและปัญหาการจำแนกประเภทที่หลากหลาย ส่งผลให้เป็นเครื่องมือที่มีความยืดหยุ่นในการเรียนรู้ของเครื่อง

งานวิจัยจำนวนมากได้แสดงให้เห็นถึงประสิทธิภาพของ AdaBoost ในการเพิ่มประสิทธิภาพการจำแนกข้อมูลในหลากหลายสาขา ตัวอย่างเช่น Wei และคณะได้แสดงให้เห็นว่า เมื่อรวม AdaBoost เข้ากับ Support Vector Machines (SVMs) จะสามารถเพิ่มประสิทธิภาพในการรู้จำเป้าหมายในภาพ Synthetic Aperture Radar (SAR) ได้อย่างมีนัยสำคัญ (Wei et al., 2008) ในทำนองเดียวกัน Pan และคณะได้ขยายโมเดล AdaBoost แบบดั้งเดิมจากการจำแนกแบบสองกลุ่ม (Binary Classification) ไปสู่การจำแนกหลายกลุ่ม (Multi-class Classification) ในชื่อ Multi-AdaBoost ซึ่งแสดงศักยภาพในการจัดการงานที่ซับซ้อน เช่น การวิเคราะห์สเปกตรัม (Spectral Analysis) (Pan et al., 2023) ความสามารถในการปรับตัวของ AdaBoost ทำให้มีข้อได้เปรียบทั้งในการจำแนกสองกลุ่มและหลายกลุ่ม และให้ผลลัพธ์ที่มีความแข็งแกร่งในบริบทที่วิธีการดั้งเดิมอาจไม่สามารถทำงานได้อย่างมีประสิทธิภาพ

นอกจากนี้ ฐานทางทฤษฎีของ AdaBoost ยังถูกเสริมด้วยความสามารถในการเลือกคุณลักษณะ (Feature Selection) และเสถียรภาพของโมเดล โดยเฉพาะในกรณีที่มีตัวแปรมิติจำนวนมาก ซึ่งมักเป็นปัญหาที่ท้าทายในสาขาเช่นการวิจัยทางชีวการแพทย์ (Mayr et al., 2018) แอปพลิเคชันล่าสุดในการตรวจสอบแหล่งที่มาของผลิตภัณฑ์ เช่น ชา โดยใช้ Multi-AdaBoost ร่วมกับวิธีวิเคราะห์ขั้นสูง แสดงถึงความสามารถของ AdaBoost ในการใช้งานร่วมกับข้อมูลหลากหลายประเภทและการยกระดับความแม่นยำในการทำนาย (Pan et al., 2023) นอกจากนี้ ยังมีการพัฒนา AdaBoost ร่วมกับเทคนิค Deep Learning เช่น Convolutional Neural Networks (CNN) ซึ่งช่วย

เพิ่มประสิทธิภาพในการทำนายและสามารถจัดการโครงสร้างข้อมูลที่ซับซ้อนได้ดีขึ้น เช่นในงานพยากรณ์ผลผลิตทางการเกษตร (Chandraprabha & Dhanaraj, 2023)

อย่างไรก็ตาม แม้ AdaBoost จะมีจุดแข็ง แต่ก็มีข้อจำกัดเรื่องการคำนวณที่อาจใช้เวลานาน เนื่องจากต้องใช้ตัวจำแนกที่อ่อนแอจำนวนมากเพื่อให้ได้ความแม่นยำที่น่าพอใจ ซึ่งอาจเป็นข้อจำกัดในการนำไปใช้งานจริงในสภาพแวดล้อมที่มีทรัพยากรจำกัด (Tamon & Xiang, 2000) ด้วยเหตุนี้ จึงมีการนำเสนอวิธีการใหม่ๆ เช่น Adaptive Sampling เพื่อบรรเทาภาระในการคำนวณ โดยยังคงรักษาประสิทธิภาพของโมเดลไว้ (Zhou et al., 2020) ในท้ายที่สุด งานวิจัยเกี่ยวกับ AdaBoost ยังแสดงให้เห็นถึงแนวโน้มในการผสานรวมเทคนิคแบบกลุ่มต่างๆ เพื่อใช้ประโยชน์จากจุดแข็งร่วมกัน ซึ่งสะท้อนผ่านการศึกษาโมเดลลูกผสม (Hybrid Models) ที่ผสมผสานหลายอัลกอริทึม boosting เพื่อเพิ่มประสิทธิภาพสูงสุด (Thotad et al., 2023)

สรุปได้ว่า AdaBoost มีความโดดเด่นในกลุ่มการเรียนรู้แบบกลุ่ม ด้วยฐานทฤษฎีที่แข็งแกร่งและความสามารถในการประยุกต์ใช้ที่กว้างขวาง ส่งเสริมประสิทธิภาพการทำงานในหลากหลายแอปพลิเคชันตั้งแต่การประมวลผลภาพจนถึงการวิเคราะห์ความเสี่ยงด้านการเงินและสิ่งแวดล้อม

Multi-Layer Perceptron (MLP)

Multi-Layer Perceptron (MLP) เป็นสถาปัตยกรรมพื้นฐานที่สำคัญในด้านการเรียนรู้ของเครื่อง โดยเฉพาะอย่างยิ่งในกลุ่มการเรียนรู้เชิงลึก (Deep Learning) ซึ่งมีโครงสร้างแบบหลายชั้น ประกอบด้วยชั้นรับข้อมูล (Input Layer) ชั้นแฝงอย่างน้อยหนึ่งชั้น (Hidden Layers) และชั้นแสดงผล (Output Layer) โดยที่แต่ละโหนด (Neuron) ในแต่ละชั้นจะเชื่อมต่อกับทุกโหนดในชั้นถัดไป MLP ได้รับการยอมรับจากความสามารถในการประมาณฟังก์ชันที่มีความไม่เป็นเชิงเส้น (Nonlinear Functions) และสามารถจัดการชุดข้อมูลที่ซับซ้อนได้อย่างมีประสิทธิภาพ

หัวใจสำคัญที่ทำให้ MLP มีประสิทธิภาพคืออัลกอริทึมการแพร่กระจายย้อนกลับ (Back-propagation Algorithm) ซึ่งช่วยให้การฝึกโมเดลมีประสิทธิภาพโดยการลดฟังก์ชันต้นทุน (Cost Function) ผ่านการปรับค่าน้ำหนักระหว่างโหนดต่างๆ ในเครือข่าย กระบวนการนี้ช่วยให้โมเดลสามารถเรียนรู้จากข้อมูลที่ได้รับ โดยทั่วไปแล้ว MLP มีความสามารถในการจำแนกรูปแบบ (Pattern Classification) การประมาณฟังก์ชัน (Function Approximation) และการวิเคราะห์การถดถอย (Regression Analysis) ซึ่งมักให้ผลลัพธ์ที่เหนือกว่าวิธีการแบบดั้งเดิมในหลายด้าน เช่น การสร้างแบบจำลองด้านสิ่งแวดล้อมและการทำนายทางการแพทย์ (Yan et al., 2023; Zhu, 2023)

ในปัจจุบัน มีการนำ MLP ไปใช้งานในสาขาต่างๆ อย่างหลากหลาย ตัวอย่างเช่น ในงานวิศวกรรมสิ่งแวดล้อมมีการใช้ MLP เพื่อทำนายระดับสารมลพิษในอากาศ โดยแสดงให้เห็นถึงศักยภาพในการสร้างแบบจำลองความสัมพันธ์ที่ซับซ้อนโดยไม่จำเป็นต้องอาศัยสมมติฐานเกี่ยวกับการกระจายข้อมูลล่วงหน้า (Ip et al., 2010) ในลักษณะเดียวกัน งานวิจัยพบว่า MLP มีประสิทธิภาพในการทำนายปริมาณรังสีแสงอาทิตย์ ซึ่งมีข้อได้เปรียบเหนือวิธีเชิงประจักษ์แบบดั้งเดิม (Vakili et al., 2015) ในทางการแพทย์ เช่น การทำนายมะเร็งเต้านม MLP แสดงให้เห็นถึงศักยภาพในการวิเคราะห์และจำแนกข้อมูลที่ซับซ้อน ทำให้การวินิจฉัยแม่นยำขึ้น (Al-Shargabi et al., 2019)

การพัฒนาและปรับปรุงสถาปัตยกรรมของ MLP อย่างต่อเนื่องช่วยเพิ่มประสิทธิภาพในการทำงานของโมเดล เทคนิคการผสานรวม MLP กับอัลกอริทึมการปรับแต่งเชิงคำนวณต่างๆ ส่งผลให้เกิดการพัฒนาอย่างมีนัยสำคัญทั้งในด้านความแม่นยำในการทำนายและประสิทธิภาพในการคำนวณ ตัวอย่างเช่น การสร้างแบบจำลองความเสี่ยงในการเกิดดินถล่ม (Landslide Susceptibility Modeling) (S. Xu, 2024) นอกจากนี้ การผสานรวมอัลกอริทึมการเลือกคุณลักษณะเฉพาะเจาะจงสำหรับการประยุกต์ใช้ MLP ยังช่วยยกระดับความแม่นยำในการตรวจจับรูปแบบเฉพาะ เช่น การตรวจจับการโจมตีแบบ Denial-of-service (Jeong et al., 2016) ซึ่งสะท้อนให้เห็นถึงการพัฒนาและวิวัฒนาการที่ต่อเนื่องในการประยุกต์ใช้งาน MLP

โดยสรุป MLP เป็นเครื่องมือที่สำคัญในวงการเรียนรู้ของเครื่อง โดยมีฐานทฤษฎีและการวิจัยเชิงประจักษ์ที่แข็งแกร่ง ความสามารถในการปรับตัวและประสิทธิภาพที่ดีในการประยุกต์ใช้ในด้านต่างๆ เป็นข้อพิสูจน์ถึงความสำคัญทั้งในงานวิจัยทางวิชาการและการใช้งานจริงในภาคอุตสาหกรรม งานวิจัยในอนาคตจะเน้นไปที่การปรับแต่งสถาปัตยกรรม MLP และการผสมผสานกับเทคนิคอื่นๆ ในการจัดการกับปัญหาที่ซับซ้อนยิ่งขึ้นในหลากหลายสาขา

The K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) เป็นวิธีการที่มีชื่อเสียงและได้รับการยอมรับอย่างกว้างขวางในด้านการเรียนรู้ของเครื่อง โดยมักถูกนำไปใช้งานในการจำแนกประเภท (Classification) และการถดถอย (Regression) KNN เป็นวิธีการเรียนรู้แบบไม่ใช้พารามิเตอร์ (Non-parametric) และอิงตามตัวอย่างข้อมูล (Instance-based Learning) ซึ่งหมายความว่าไม่มีข้อสมมติฐานเบื้องต้นเกี่ยวกับการกระจายของข้อมูล แต่จะจำแนกข้อมูลใหม่โดยพิจารณาจากประเภทที่พบมากที่สุดในกลุ่มข้อมูลที่อยู่ใกล้ที่สุดในชุดข้อมูลการฝึก ส่งผลให้อัลกอริทึมใช้หลักการความใกล้เคียงของจุดข้อมูลในการระบุประเภท (Boudjella et al., 2020; Fadlil et al., 2022)

หนึ่งในจุดแข็งที่สำคัญของ KNN คือความเรียบง่ายและง่ายต่อการนำไปปฏิบัติจริง อัลกอริทึมนี้ใช้ทรัพยากรการคำนวณน้อยมากในขั้นตอนการฝึก เนื่องจากเพียงแค่จัดเก็บข้อมูลไว้ อย่างไรก็ตามอาจพบความท้าทายในขั้นตอนการจำแนกประเภท โดยเฉพาะกับชุดข้อมูลขนาดใหญ่ เนื่องจากต้องคำนวณระยะห่างระหว่างจุดข้อมูลใหม่กับทุกจุดในชุดข้อมูล คุณลักษณะดังกล่าวทำให้ KNN มีความเหมาะสมอย่างยิ่งกับชุดข้อมูลขนาดเล็ก แต่สามารถเป็นภาระหนักในกรณีที่มีข้อมูลขนาดใหญ่มาก หากไม่มีการปรับปรุงประสิทธิภาพผ่านวิธีการเช่น การตัดแต่งข้อมูล (Data Pruning) หรือการคำนวณแบบขนาน (Parallel Computing Techniques) (Saadatfar et al., 2020)

KNN ได้รับการประยุกต์ใช้อย่างกว้างขวางในหลากหลายสาขา เช่น ด้านการดูแลสุขภาพที่มีการใช้ KNN ทำนายผลลัพธ์ต่างๆ เช่น อัตราการกลับมารักษาซ้ำของผู้ป่วยในโรงพยาบาลและการวินิจฉัยโรค โดยงานวิจัยแสดงให้เห็นถึงประสิทธิภาพในการพยากรณ์การกลับมารักษาซ้ำของผู้ป่วยโรคหัวใจเต้นผิดจังหวะ (Atrial Fibrillation) และการทำนายโรคหัวใจจากคุณลักษณะของผู้ป่วย ซึ่งแสดงให้เห็นถึงความสามารถที่หลากหลายในด้านการแพทย์ นอกจากนี้ประสิทธิภาพของอัลกอริทึมยังสามารถปรับปรุงเพิ่มเติมได้ด้วยวิธีต่างๆ เช่น เทคนิคการตรวจสอบแบบไขว้ (Cross-validation) ที่ช่วยเพิ่มความแข็งแกร่งและความแม่นยำ (Tembusai et al., 2021)

นอกจากนี้ KNN ยังสามารถผสมผสานกับเทคนิคการเรียนรู้ของเครื่องอื่นๆ เพื่อเพิ่มประสิทธิภาพในการสร้างแบบจำลองเชิงทำนาย รูปแบบต่างๆ ของอัลกอริทึมสามารถนำไปใช้ร่วมกับวิธีการแบบกลุ่ม (Ensemble Methods) หรือกรอบแนวคิดแบบผสม (Hybrid Frameworks) เพื่อเสริมประสิทธิภาพการจำแนกประเภทในสถานการณ์ที่ซับซ้อน ตัวอย่างเช่น การทำนายความเสี่ยงโรคเบาหวาน KNN แสดงประสิทธิภาพในการแข่งขันกับการถดถอยโลจิสติก (Logistic Regression) ได้อย่างมีนัยสำคัญ (P. V. S. Kumar & Kumar, 2023)

สรุปได้ว่าอัลกอริทึม K-Nearest Neighbors ยังคงเป็นเครื่องมือที่สำคัญในชุดเครื่องมือการเรียนรู้ของเครื่อง โดยเฉพาะงานการจำแนกประเภทที่ต้องการข้อมูลเชิงลึก การสร้างต้นแบบอย่างรวดเร็ว และการนำไปใช้งานจริง ความสามารถในการปรับตัวให้เหมาะสมกับชุดข้อมูลที่หลากหลายและประสิทธิภาพที่แข็งแกร่งในการใช้งานต่างๆ ยืนยันถึงความเกี่ยวข้องและความสำคัญในแวดวงนี้อย่างต่อเนื่อง

Logistic Regression

การถดถอยโลจิสติก (Logistic Regression) เป็นวิธีทางสถิติที่นิยมใช้อย่างกว้างขวางในด้านการเรียนรู้ของเครื่องสำหรับปัญหาการจำแนกข้อมูลแบบสองกลุ่ม (Binary Classification) ซึ่งมีจุดเด่นที่ความเรียบง่าย การตีความที่เข้าใจง่าย และประสิทธิภาพในการใช้งานหลากหลายสาขา เช่น ด้านการแพทย์และการเงิน วิธีการนี้สามารถจำลองความสัมพันธ์ระหว่างตัวแปรตามที่มีค่าเป็นสองสถานะกับตัวแปรอิสระตั้งแต่หนึ่งตัวขึ้นไปได้อย่างมีประสิทธิภาพ ทำให้เป็นเครื่องมือที่มีคุณค่าในหลายๆ สาขา ตัวอย่างเช่น งานวิจัยของ Ahmadov และ Boyacı พบว่าการถดถอยโลจิสติกให้ผลลัพธ์ที่ดีเยี่ยมในการวิเคราะห์ความรู้สึก (Sentiment Analysis) โดยให้ค่าความแม่นยำ ความไว และคะแนน F1 สูงสุดเมื่อเทียบกับวิธีการเรียนรู้ของเครื่องอื่นๆ ที่นำมาทดสอบ (Ahmadov & Boyacı, 2022) ผลการศึกษานี้เน้นย้ำถึงความน่าเชื่อถือของการถดถอยโลจิสติกในสถานการณ์ที่ตัวชี้วัดประสิทธิภาพมีความสำคัญ

ยิ่งไปกว่านั้น ความสามารถในการตีความของการถดถอยโลจิสติกมีข้อได้เปรียบอย่างมากในทางคลินิก ซึ่งความเข้าใจในกระบวนการตัดสินใจของโมเดลมีความสำคัญ Chen และ Zhang ระบุว่าความเรียบง่ายและความชัดเจนในการตีความผลลัพธ์ของโมเดลการถดถอยโลจิสติกทำให้โมเดลนี้ได้รับความนิยมในการนำไปประยุกต์ใช้งานจริงในทางคลินิก แม้จะมีรายงานว่าโมเดลการเรียนรู้ของเครื่องที่ซับซ้อนกว่าอาจให้ประสิทธิภาพที่ดีกว่าในบางกรณีก็ตาม (M. Chen & Zhang, 2024) นอกจากนี้ Wang และคณะ ยังเน้นว่าการถดถอยโลจิสติกสามารถใช้เป็นเกณฑ์มาตรฐานเพื่อเปรียบเทียบกับโมเดลขั้นสูงอื่นๆ ซึ่งแสดงให้เห็นถึงความสำคัญของการรักษาความเรียบง่ายในการสร้างแบบจำลองเชิงทำนาย ประเด็นนี้สำคัญมากในสาขาเช่นการแพทย์ที่การตัดสินใจมีผลกระทบสำคัญต่อการดูแลผู้ป่วย (F. Wang & Wu, 2023)

อย่างไรก็ตาม การประเมินบริบทที่เหมาะสมสำหรับการใช้การถดถอยโลจิสติกถือว่ามีความสำคัญ เนื่องจากประสิทธิภาพของวิธีนี้อาจแตกต่างกันมากตามลักษณะของข้อมูลและปัญหาที่กำลังวิเคราะห์ ตัวอย่างเช่น งานวิจวิจัยของ Jacobsen และคณะ ซึ่งเปรียบเทียบการถดถอยโลจิสติกกับเทคนิคการเรียนรู้ของเครื่องอื่นๆ พบว่าในขณะที่ Logistic Regression ให้ผลลัพธ์ที่น่าพอใจ แต่

บางครั้งวิธีเช่น Random Forest และ Gradient Boosting มีประสิทธิภาพในการทำนายสูงกว่าโดยเฉพาะกับข้อมูลที่ซับซ้อน (Jacobsen et al., 2021) ผลการศึกษานี้สอดคล้องกับ Vinter และคณะ ซึ่งพบว่าวิธีการถดถอยโลจิสติกแบบดั้งเดิมไม่ได้มีความเหนือกว่าเทคนิคการเรียนรู้ของเครื่องอื่นในการทำนายผลทางคลินิกอย่างชัดเจน จึงแนะนำให้เลือกโมเดลโดยพิจารณาบริบทเฉพาะที่นำมาวิเคราะห์ (Vinter et al., 2020)

นอกจากนี้ การถดถอยโลจิสติกยังคงมีความสำคัญในงานวิจัยที่เน้นการระบุปัจจัยเสี่ยง ตัวอย่างเช่น งานศึกษาของ Yang และคณะ แสดงให้เห็นการสร้างแบบจำลองที่มีความแม่นยำสูงโดยใช้การถดถอยโลจิสติกแบบพหุตัวแปร ซึ่งเน้นย้ำถึงการใช้งานในประเมินความเสี่ยงทางคลินิก อย่างไรก็ตาม ในกรณีที่เกี่ยวข้องกับการแพร่กระจายของมะเร็งลำไส้ใหญ่ โมเดลการถดถอยโลจิสติกแบบดั้งเดิมอาจมีข้อจำกัดเมื่อเปรียบเทียบกับอัลกอริทึมที่ทันสมัยกว่า ซึ่งชี้ว่าประสิทธิภาพของการถดถอยโลจิสติกอาจถูกจำกัดในสถานการณ์ที่ซับซ้อนเป็นพิเศษ (X. Yang et al., 2023)

สรุปได้ว่า การถดถอยโลจิสติกยังคงเป็นวิธีการสำคัญในด้านการเรียนรู้ของเครื่อง โดยมีความสมดุลระหว่างประสิทธิภาพและการตีความที่ชัดเจน โดยเฉพาะอย่างยิ่งในแวดวงคลินิก อย่างไรก็ตาม ควรพิจารณาข้อจำกัดและลักษณะเฉพาะของชุดข้อมูลที่ใช้ในการวิเคราะห์อย่างระมัดระวัง เมื่อการเรียนรู้ของเครื่องมีการพัฒนาต่อเนื่อง วิธีการดั้งเดิมอย่างการถดถอยโลจิสติกจึงยังเป็นกรอบการทำงานที่แข็งแกร่งซึ่งสามารถพัฒนาและเปรียบเทียบกับอัลกอริทึมที่ซับซ้อนมากขึ้นในอนาคต

Decision Trees

ต้นไม้ตัดสินใจ (Decision Trees) เป็นวิธีการพื้นฐานที่สำคัญในด้านการเรียนรู้ของเครื่อง ซึ่งถูกใช้อย่างแพร่หลายในการจำแนกประเภท (Classification) และการถดถอย (Regression) โมเดลประเภทนี้มีโครงสร้างแบบต้นไม้ ซึ่งจะทำการแบ่งชุดข้อมูลออกเป็นกลุ่มย่อยตามค่าของคุณลักษณะต่างๆ จนกระทั่งได้ผลการทำนายที่โหนดใบ จุดเด่นสำคัญของต้นไม้ตัดสินใจคือความเรียบง่าย ความสามารถในการตีความที่ชัดเจน และการจัดการชุดข้อมูลที่ซับซ้อนได้ดี ซึ่งทำให้อัลกอริทึมนี้ได้รับความนิยมอย่างสูงในวงการเรียนรู้ของเครื่อง

ในปัจจุบัน มีการพัฒนาและปรับปรุงวิธีการต้นไม้ตัดสินใจให้มีประสิทธิภาพมากยิ่งขึ้น ตัวอย่างเช่น Jiang และคณะ (2024) (*Application of Decision Tree and Machine Learning in New Energy Vehicle Maintenance Decision Making*, n.d.; Jiang et al., 2024) ชี้ให้เห็นว่าต้นไม้ตัดสินใจช่วยให้สามารถเรียนรู้กฎการตัดสินใจที่เหมาะสมจากข้อมูลได้ ซึ่งแสดงถึงบทบาทสำคัญในงานเชิงปฏิบัติ เช่น การตัดสินใจในการซ่อมบำรุงรถยนต์พลังงานใหม่ (New Energy Vehicles) นอกจากนี้ นวัตกรรมอย่าง QUBO Decision Tree ที่ใช้เทคนิค annealing ในการแยกข้อมูล แสดงให้เห็นถึงการผลักดันขีดจำกัดของโมเดลแบบดั้งเดิมไปสู่แนวทางใหม่ๆ (Yawata et al., 2022)

ต้นไม้ตัดสินใจถูกนำไปใช้ในหลายสาขา เช่น ในด้านสุขภาพ การศึกษาเกี่ยวกับการทำนายผลลัพธ์การดูแลโรคไวรัส พบว่าต้นไม้ตัดสินใจสามารถให้ผลลัพธ์ที่แม่นยำและเชื่อถือได้ นอกจากนี้ งานวิจัยที่เปรียบเทียบต้นไม้ตัดสินใจกับวิธีการอื่นๆ เช่น การถดถอยโลจิสติก แสดงให้เห็นว่าต้นไม้ตัดสินใจมีประสิทธิภาพในการทำนายที่ดี พร้อมทั้งยังมีความสามารถในการตีความที่สูง ซึ่งมี

ความสำคัญในสถานการณ์ที่จำเป็นต้องเข้าใจกระบวนการตัดสินใจของโมเดล (Setiawan et al., 2024)

ต้นไม้ตัดสินใจได้รับการประยุกต์ใช้อย่างหลากหลายในสาขาการวิจัยต่างๆ เช่น การทำนายปัจจัยเสี่ยงทางการเงิน และการวิเคราะห์ข้อมูลด้านพันธุศาสตร์ โดยงานวิจวิจัยต่างๆ แสดงให้เห็นถึงประสิทธิภาพที่ดีในการประยุกต์ใช้งาน ตั้งแต่การพยากรณ์ไปจนถึงการวิเคราะห์ข้อมูลทางชีวภาพ การวิจัยอย่างต่อเนื่องยังมุ่งเน้นการปรับปรุงประสิทธิภาพของโมเดล เช่น การปรับพารามิเตอร์ (Hyperparameter Tuning) ที่ช่วยเพิ่มประสิทธิภาพของโมเดลอย่างมีนัยสำคัญ (Arista, 2022)

โดยสรุป ต้นไม้ตัดสินใจเป็นองค์ประกอบสำคัญในวงการเรียนรู้ของเครื่อง ซึ่งมีความโดดเด่นในเรื่องของประสิทธิภาพ ความสามารถในการตีความ และความยืดหยุ่นในการประยุกต์ใช้ในหลากหลายสาขา งานวิจัยในอนาคตจะยังคงมุ่งเน้นไปที่การพัฒนาและปรับปรุงอัลกอริทึมนี้ เพื่อให้สามารถจัดการกับความท้าทายที่ซับซ้อนมากขึ้นในอนาคต

Support Vector Machines (SVM)

Support Vector Machines (SVM) เป็นเครื่องมือที่ได้รับการยอมรับอย่างแพร่หลายและมีประสิทธิภาพสูงในด้านการเรียนรู้ของเครื่อง โดยเฉพาะในงานการจำแนกประเภท (Classification) SVM ถูกพัฒนาขึ้นจากหลักการของทฤษฎีการเรียนรู้เชิงสถิติ (Statistical Learning Theory) โดยใช้กลยุทธ์การลดความเสี่ยงเชิงโครงสร้าง (Structural Risk Minimization) ซึ่งช่วยปรับสมดุลระหว่างความซับซ้อนของโมเดลกับข้อผิดพลาดจากการฝึกโมเดล ส่งผลให้มีความสามารถในการทำนายข้อมูลใหม่ได้ดีกว่าวิธีการดั้งเดิม เช่น โครงข่ายประสาทเทียม (Neural Networks) ที่มักใช้วิธีลดความเสี่ยงเชิงประจักษ์ (Empirical Risk Minimization) (He et al., 2019; Xiu-zhi & Ying, 2010; Havlíček et al., 2019) หลักการสำคัญของ SVM คือการแปลงข้อมูลในพื้นที่ป้อนข้อมูล (Input Space) ไปยังพื้นที่ลักษณะที่มีมิติสูงขึ้น (Higher-dimensional Feature Space) ซึ่งช่วยให้สามารถใช้เทคนิค Kernel Trick ได้อย่างมีประสิทธิภาพ ส่งผลให้ SVM สามารถจัดการกับขอบเขตการตัดสินใจที่ไม่เป็นเชิงเส้น (Non-linear Decision Boundaries) ซึ่งอัลกอริทึมอื่นมักจะพบความท้าทายในการคำนวณ (Han & Zhao, 2015; Z. Xu et al., 2015)

กรอบทฤษฎีของ SVM มีพื้นฐานอยู่บนแนวคิดในการเพิ่มระยะห่างระหว่างคลาสต่างๆ ให้สูงที่สุด โดยใช้ไฮเปอร์เพลน (Hyperplanes) ในพื้นที่ลักษณะ วิธีการนี้ช่วยให้ SVM สามารถหาคำตอบที่เหมาะสมที่สุดได้แม้ในชุดข้อมูลที่ซับซ้อนและไม่สามารถแยกเชิงเส้น การพัฒนาเทคนิคเคอร์เนล (Kernel Methods) ได้ขยายขีดความสามารถในการใช้งานของ SVM ในหลากหลายสาขา การที่สามารถปรับแต่งฟังก์ชันเคอร์เนลได้ช่วยให้ SVM สามารถปรับตัวเข้ากับลักษณะเฉพาะของข้อมูลแต่ละชุด ซึ่งช่วยปรับปรุงความแม่นยำในการจำแนกประเภทอย่างมีประสิทธิภาพ ส่งผลให้เหมาะสมกับการนำไปใช้ในงานหลากหลายตั้งแต่การวินิจฉัยทางการแพทย์ จนถึงการวิเคราะห์ภาพ (Wassermann et al., 2010)

ในแง่ของประสิทธิภาพ มีงานวิจัยที่มุ่งเน้นพัฒนากลยุทธ์การคำนวณที่มีประสิทธิภาพสำหรับการฝึกและใช้งาน SVM โดยเฉพาะการพัฒนาอัลกอริทึมแบบขนาน (Parallel Algorithms) และการปรับให้เหมาะสมกับระบบคอมพิวเตอร์สมรรถนะสูง ซึ่งช่วยปรับปรุงความเร็วในการฝึกโมเดลและเพิ่ม

ศักยภาพในการทำงานกับชุดข้อมูลขนาดใหญ่ รวมถึงเทคนิคการหาค่าที่เหมาะสม (Optimization Techniques) (Imbault & Lebart, 2004) ตัวอย่างเช่น การเปลี่ยนไปใช้การเรียนรู้เคอร์เนลหลายตัว (Multiple Kernel Learning) ได้แสดงให้เห็นศักยภาพที่มากขึ้นในการรวมคุณลักษณะหลายประเภทเข้าด้วยกันเพื่อยกระดับประสิทธิภาพในการตัดสินใจของ SVM (Xiao & Liu, 2012)

โดยสรุป Support Vector Machines ถือเป็นวิธีการที่มีรากฐานทางทฤษฎีที่แข็งแกร่งและสามารถประยุกต์ใช้งานได้จริงในหลากหลายด้านของการเรียนรู้ของเครื่อง ความก้าวหน้าในเทคนิคเคอร์เนลและวิธีการคำนวณที่มีประสิทธิภาพ ทำให้ SVM ยังคงมีความสำคัญและมีประสิทธิผลในการจัดการปัญหาการจำแนกประเภทที่หลากหลาย

2.2 งานวิจัยที่เกี่ยวข้อง

การทบทวนวรรณกรรมนี้เกี่ยวกับการประยุกต์ใช้การเรียนรู้ของเครื่อง ในการทำนายความสำเร็จของสตาร์ตอัป ซึ่งเป็นความท้าทายที่มีหลายมิติและดึงดูดความสนใจอย่างมากจากนักลงทุน ผู้กำหนดนโยบาย และผู้ประกอบการ แหล่งข้อมูลต่างๆ เน้นย้ำว่าการทำนายความสำเร็จของสตาร์ตอัปอย่างแม่นยำนั้น จำเป็นต้องสามารถความไม่แน่นอนของสภาพแวดล้อมทางธุรกิจและพิจารณาปัจจัยที่มีอิทธิพลจำนวนมาก

งานวิจัยหลายฉบับเห็นพ้องกันในศักยภาพของการเรียนรู้ของเครื่อง ในการทำนายความสำเร็จของสตาร์ตอัป โดยใช้แบบจำลองวิเคราะห์ชุดข้อมูลขนาดใหญ่เพื่อค้นหาแบบแผนและคาดการณ์ผลลัพธ์ในอนาคต อย่างไรก็ตาม พวกเขาใช้แหล่งข้อมูล เทคนิคการสร้างคุณลักษณะ และกลยุทธ์การเลือกแบบจำลองที่แตกต่างกัน โดยเน้นถึงความสำคัญของการปรับแต่งวิธีการให้เหมาะสมกับงานทำนายที่เฉพาะเจาะจง

Vasquez และคณะ (2023) Piskunova และคณะ (2021) Gangwani และคณะ (2023) Lestari และ Halim (2022) Żbikowski และ Antosiuk (2021) เน้นย้ำถึงความสำคัญของการกำหนดความสำเร็จของสตาร์ตอัปอย่างชัดเจน ในขณะที่บางแห่งกำหนดว่าเป็นการเสนอขายหุ้นเข้าสู่ตลาดหุ้นครั้งแรก (Initial Public Offering: IPO) หรือการควบรวมและซื้อกิจการ (Merger & Acquisition: M&A) ในงานวิจัยของ Gangwani และคณะ (2023) Rodrigues และคณะ (2021) ในขณะที่ Vasquez และคณะ (2023) Piskunova และคณะ (2021) พิจารณาปัจจัยอย่าง อาทิ ความสามารถในการทำกำไร การได้รับเงินทุนเพิ่มเติม หรือการดำเนินธุรกิจต่อไป (Gangwani et al., 2023; Lestari & Halim, 2022; Piskunova et al., 2021; Rodrigues et al., 2021; Vasquez et al., 2023; Żbikowski & Antosiuk, 2021)

Żbikowski และ Antosiuk (2021) ใช้ชุดข้อมูล และแนวทางการวิเคราะห์ที่หลากหลาย จาก Crunchbase ในการสร้างโมเดลที่ปราศจากอคติในการทำนายความสำเร็จทางธุรกิจ โดยเน้นที่ข้อมูลที่มีอยู่เมื่อเริ่มต้นบริษัท ในขณะที่ Piskunova และคณะ (2021) ใช้ข้อมูลจาก Crunchbase โดยวิเคราะห์สตาร์ตอัปจาก 171 ประเทศ และกำหนดความสำเร็จว่าเป็นการควบรวมกิจการหรือ เสนอขายหุ้นเข้าสู่ตลาดหุ้นครั้งแรก ซึ่งใช้แบบจำลองการเรียนรู้ของเครื่องต่างๆ และพบว่าประเภทของอุตสาหกรรม ขนาดทีม จำนวนรอบการระดมทุน และสถานที่ตั้งเป็นปัจจัยที่ทำนายได้อย่างมีนัยสำคัญ

Q. Zhang และคณะ (2017) ตรวจสอบผลกระทบของการมีส่วนร่วมในโซเชียลมีเดียต่อความสำเร็จในการระดมทุนบนแพลตฟอร์ม AngelList ซึ่งบ่งชี้ถึงขอบเขตของแหล่งข้อมูลที่ใช้ในการทำนายความสำเร็จของสตาร์ตอัปที่ขยายกว้างขึ้น (Q. Zhang et al., 2017)

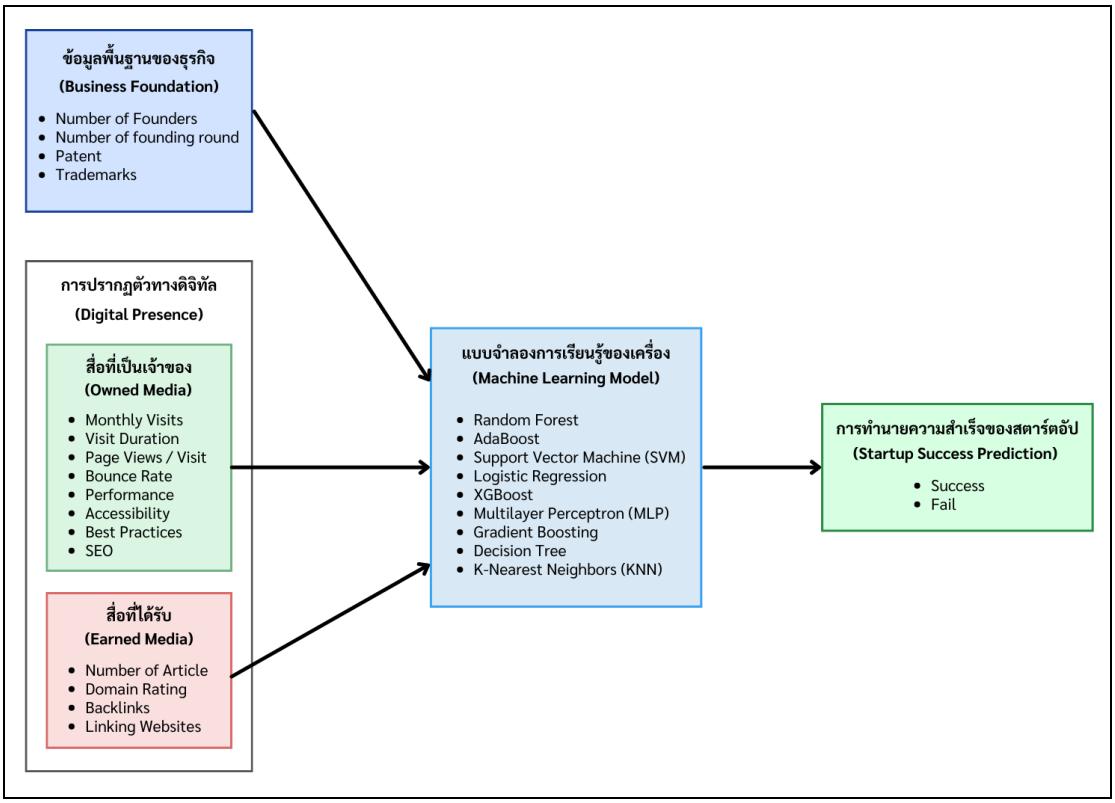
Piskunova และคณะ (2021) เน้นย้ำถึงความสำคัญของการเตรียมข้อมูลที่มีความสมบูรณ์ โดยใช้แนวทางที่ระมัดระวังในการกำจัดข้อมูลที่มีค่าว่าง ซึ่งแตกต่างจากการศึกษาอื่นๆ ที่ใช้เทคนิคการเติมข้อมูลที่ขาดหายไป

Gangwani และคณะ (2023) เน้นถึงการใช้ประโยชน์จากการผสมผสานระหว่างปัจจัยพื้นฐานทางธุรกิจภายในแบบดั้งเดิม และการทำเหมืองข้อมูลจากข้อความ เพื่อสร้างแบบจำลองที่แข็งแกร่งและให้ข้อมูลเชิงลึกมากยิ่งขึ้น

Vasquez และคณะ (2023) สนับสนุนการใช้โมเดลแบบไฮบริดที่รวมโมเดล ML หลายตัวเข้าด้วยกันพร้อมกับกลยุทธ์การตัดสินใจแบบการโหวต (Voting) ซึ่งจะให้ความแม่นยำสูงกว่าโมเดลเดี่ยว พวกเขายังจัดทำรายการ 79 ปัจจัยความสำเร็จ เพื่อเน้นถึงความครอบคลุมของแนวทางของพวกเขา ในขณะที่ Tomy และ Pardede (2018) มุ่งเน้นไปที่ปัจจัยความไม่แน่นอน โดยทำการจำแนกประเภทและใช้การเรียนรู้ของเครื่อง ในการวิเคราะห์ผลกระทบต่อการทำนายความสำเร็จ ซึ่งบ่งชี้ถึงความสำคัญของการรวมการวิเคราะห์ความไม่แน่นอนในกระบวนการทำนาย พวกเขาใช้ชุดข้อมูล ICT ของออสเตรเลีย ฝึกโมเดลด้วยข้อมูล 80% และทดสอบกับข้อมูลที่เหลืออีก 20% ผลการศึกษาพบว่า Naïve Bayes ให้ผลลัพธ์ที่ดีกว่า k-NN และ SVM ในการทำนายความสำเร็จของสตาร์ตอัปตามชุดข้อมูลนี้ (Tomy & Pardede, 2018)

ผู้วิจัยเลือกใช้ข้อมูลพื้นฐานของธุรกิจ (Business Foundation) ในส่วนของปัจจัยภายใน และข้อมูลด้านการปรากฏตัวทางดิจิทัล โดยทำการแบ่งเป็นสองปัจจัยคือ สื่อที่เป็นเจ้าของ (Owned Media) และสื่อที่ได้รับ (Earned Media) โดยใช้เครื่องมือในการวิเคราะห์อย่างเป็นระบบ โดยผู้วิจัยใช้แบบจำลอง 9 แบบ ดังที่ได้กล่าวมาแล้วในบทที่ 1 ซึ่งช่วยให้สามารถเปรียบเทียบประสิทธิภาพของโมเดลในการวิเคราะห์ปัจจัยได้อย่างครอบคลุม และทำการวิเคราะห์สองกรณีในแต่ละแบบจำลอง ได้แก่ การใช้ข้อมูลผสมผสาน (Mixed Features) ระหว่างข้อมูลพื้นฐานของธุรกิจ (Business Foundation) และ การปรากฏตัวทางดิจิทัล (Digital Presence) และการใช้เฉพาะข้อมูลจาก การปรากฏตัวทางดิจิทัล (Digital Presence) ซึ่งจะทำให้สามารถเห็นถึงผลกระทบจากคุณลักษณะแต่ละกลุ่มได้อย่างชัดเจน

2.3 กรอบแนวคิดการวิจัย



รูปที่ 2.1 แผนภาพแสดงกรอบแนวคิดการวิจัย

2.3.1 พื้นฐานทางธุรกิจ (Business Foundation)

คุณลักษณะนี้หมายถึงองค์ประกอบพื้นฐานที่สำคัญต่อการก่อตั้งและดำเนินธุรกิจสตาร์ตอัป ซึ่งประกอบไปด้วย

- จำนวนผู้ก่อตั้ง (Number of Founders) จำนวนผู้ร่วมก่อตั้งธุรกิจ อาจส่งผลต่อความหลากหลายของทักษะและประสบการณ์ รวมถึงการแบ่งปันภาระและความรับผิดชอบในการดำเนินธุรกิจ
- จำนวนรอบการระดมทุน (Number of founding round) จำนวนครั้งที่ธุรกิจสตาร์ตอัปสามารถระดมทุนได้ อาจบ่งชี้ถึงศักยภาพในการเติบโตและความน่าสนใจของธุรกิจต่อนักลงทุน
- สิทธิบัตร (Patent) การมีสิทธิบัตรบ่งบอกถึงนวัตกรรมและความสามารถในการปกป้องทรัพย์สินทางปัญญาของธุรกิจ ซึ่งอาจเป็นปัจจัยสำคัญในการสร้างความได้เปรียบทางการแข่งขัน

- เครื่องหมายการค้า (Trademarks) เครื่องหมายการค้าช่วยสร้างการจดจำแบรนด์และความน่าเชื่อถือของธุรกิจ ซึ่งเป็นสิ่งสำคัญในการสร้างฐานลูกค้าและความภักดี

2.3.2 การปรากฏตัวทางดิจิทัล (Digital Presence)

ในยุคปัจจุบัน การมีตัวตนบนโลกดิจิทัลเป็นสิ่งสำคัญอย่างยิ่งสำหรับธุรกิจสตาร์ตอัป กรอบแนวคิดนี้แบ่งการปรากฏตัวทางดิจิทัลออกเป็น 2 ส่วนหลัก ได้แก่ สื่อที่เป็นเจ้าของ (Owned Media) และสื่อที่ได้รับ (Earned Media)

2.3.2.1 สื่อที่เป็นเจ้าของ (Owned Media)

สื่อที่เป็นเจ้าของหมายถึงช่องทางดิจิทัลที่ธุรกิจสตาร์ตอัปสามารถควบคุมและจัดการได้เอง ซึ่งครอบคลุมถึง

- จำนวนผู้เข้าชมรายเดือน (Monthly Visits) จำนวนครั้งที่มีผู้เข้าชมเว็บไซต์หรือแพลตฟอร์มดิจิทัลของธุรกิจในแต่ละเดือน บ่งบอกถึงความนิยมและความสนใจในธุรกิจ
- ระยะเวลาการเข้าชม (Visit Duration) ระยะเวลาเฉลี่ยที่ผู้เข้าชมอยู่ในเว็บไซต์หรือแพลตฟอร์มดิจิทัลของธุรกิจ สะท้อนถึงความน่าสนใจและคุณภาพของเนื้อหา
- จำนวนหน้าเว็บต่อการเข้าชม (Page Views/Visit) จำนวนหน้าที่ผู้เข้าชมเปิดดูต่อการเข้าชมหนึ่งครั้ง บ่งชี้ถึงความลึกซึ้งในการสำรวจเนื้อหาและข้อมูลของธุรกิจ
- อัตราตีกลับ (Bounce Rate) สัดส่วนของผู้เข้าชมที่ออกจากเว็บไซต์หรือแพลตฟอร์มดิจิทัลทันทีหลังจากเข้าชมหน้าเดียว บ่งบอกถึงความน่าสนใจของหน้าแรกและความเกี่ยวข้องของเนื้อหา
- ประสิทธิภาพ (Performance) ประสิทธิภาพของเว็บไซต์หรือแพลตฟอร์มดิจิทัลในด้านต่างๆ อาทิ ความเร็วในการโหลด ความเสถียร และการทำงานที่ราบรื่น
- การเข้าถึง (Accessibility) ความสามารถในการเข้าถึงเว็บไซต์ หรือเนื้อหาสำหรับผู้ใช้งานทุกกลุ่ม รวมถึงผู้พิการ และเบราว์เซอร์ต่างๆ อย่างทั่วถึง
- แนวทางปฏิบัติที่ดีที่สุด (Best Practices) การปฏิบัติตามแนวทางที่ดีที่สุดในการออกแบบและพัฒนาเว็บไซต์หรือแพลตฟอร์มดิจิทัล เพื่อให้ได้ผลลัพธ์ที่ดีที่สุดในด้านต่างๆ
- การเพิ่มประสิทธิภาพเว็บไซต์เพื่อการค้นหา (SEO) การปรับแต่งเว็บไซต์หรือแพลตฟอร์มดิจิทัลให้เหมาะสมกับเครื่องมือค้นหา เพื่อให้ธุรกิจปรากฏในอันดับต้นๆ ของผลการค้นหา

2.3.2.2 สื่อที่ได้รับ (Earned Media)

สื่อที่ได้รับหมายถึงการกล่าวถึงธุรกิจสตาร์ตอัปในช่องทางดิจิทัลอื่นๆ ที่ธุรกิจไม่ได้เป็นเจ้าของโดยตรง แต่ได้มาจากการสร้างความสัมพันธ์ที่ดีและการสร้างเนื้อหาที่มีคุณค่า ซึ่งครอบคลุมถึง

- จำนวนบทความ หรือข่าวที่กล่าวถึงองค์กร (Number of Article) จำนวนบทความหรือการกล่าวถึงธุรกิจในสื่อออนไลน์ต่างๆ อาทิ ข่าว บทความรีวิว หรือบล็อก

- คะแนนโดเมน (Domain Rating) คะแนนที่บ่งบอกถึงความน่าเชื่อถือ และอิทธิพลของเว็บไซต์ที่กล่าวถึงธุรกิจ
- ลิงก์ย้อนกลับ (Backlinks) จำนวนลิงก์จากเว็บไซต์ภายนอกที่เชื่อมโยงมายังเว็บไซต์ของธุรกิจ บ่งบอกถึงความน่าเชื่อถือ และความนิยมของเว็บไซต์
- เว็บไซต์ที่เชื่อมโยง (Linking Websites) จำนวนโดเมนที่แตกต่างกันที่ลิงก์มายังเว็บไซต์ของธุรกิจ

2.3.3 แบบจำลองการเรียนรู้ของเครื่อง (Machine Learning Model)

กรอบแนวคิดนี้เสนอการนำแบบจำลองการเรียนรู้ของเครื่องหลากหลายประเภทมาใช้ในการวิเคราะห์ข้อมูลปัจจัยพื้นฐานทางธุรกิจและการปรากฏตัวทางดิจิทัล เพื่อทำนายความสำเร็จของธุรกิจสตาร์ตอัป แบบจำลองที่ใช้ในกรอบแนวคิดนี้ ได้แก่

- Random Forest
- AdaBoost
- Support Vector Machine (SVM)
- Logistic Regression
- XGBoost
- Multilayer Perceptron (MLP)
- Gradient Boosting
- Decision Tree
- K-Nearest Neighbors (KNN)

การเลือกใช้แบบจำลองที่เหมาะสมจะขึ้นอยู่กับลักษณะของข้อมูลและผลลัพธ์ในการทำนาย

2.3.4 การทำนายความสำเร็จของสตาร์ตอัป (Startup Success Prediction)

ผลลัพธ์สุดท้ายของกระบวนการวิจัยนี้คือการทำนายว่าธุรกิจสตาร์ตอัปจะประสบความสำเร็จ (Success) หรือล้มเหลว (Fail) โดยอาศัยผลลัพธ์จากการวิเคราะห์ข้อมูลด้วยแบบจำลองการเรียนรู้ของเครื่อง กรอบแนวคิดการวิจัยนี้แสดงให้เห็นถึงความสัมพันธ์ระหว่างปัจจัยพื้นฐานทางธุรกิจ การปรากฏตัวทางดิจิทัล และความสำเร็จของธุรกิจสตาร์ตอัป โดยเสนอการใช้แบบจำลองการเรียนรู้ของเครื่องเป็นเครื่องมือในการวิเคราะห์และทำนายผลลัพธ์