

บทที่ 4

ผลการวิจัยและการอภิปรายผล

จากการวิเคราะห์ชุดข้อมูลด้วยแบบจำลองการเรียนรู้ของเครื่องทั้ง 9 แบบจำลอง เพื่อการวิเคราะห์หาความสำคัญของคุณลักษณะที่มีผลต่อการทำนายความสำเร็จของสตาร์ทอัพ โดยใช้การเรียนรู้ของเครื่อง โดยใช้ข้อมูลแบบผสมผสานจากข้อมูลพื้นฐานทางธุรกิจและข้อมูลประสิทธิภาพการปรากฏตัวทางดิจิทัล ผู้วิจัยได้ผลการวิจัยและอภิปรายผลดังนี้

- 4.1 การวิเคราะห์ความสำคัญของคุณลักษณะ
 - 4.2 การวิเคราะห์อิทธิพลของคุณลักษณะที่มีผลต่อค่าของการทำนาย
 - 4.3 การเปรียบเทียบประสิทธิภาพของแบบจำลอง
 - 4.4 การนำองค์ความรู้ไปใช้งานจริง
- โดยมีรายละเอียดดังนี้

4.1 การวิเคราะห์ความสำคัญของคุณลักษณะ

จากการวิเคราะห์ผลการหาความสำคัญของคุณลักษณะ จากแบบจำลองการเรียนรู้ของเครื่องทั้ง 2 แบบ โดยใช้ 9 แบบจำลอง ในตารางที่ 4.1 สามารถพบว่า ความสม่ำเสมอของความสำคัญของแต่ละคุณลักษณะใน 9 แบบจำลอง สามารถจัดลำดับได้ดังนี้

อันดับที่ 1: EM-Article เป็นคุณลักษณะที่แสดงความสำคัญอย่างสม่ำเสมอและมีค่าสูงในเกือบทุกแบบจำลอง โดยปรากฏในอันดับต้นๆ ของ 6 แบบจำลอง ที่ใช้ข้อมูลแบบผสมผสาน ได้แก่ Random Forest (16.812%), XGBoost (20.257%), Gradient Boosting (29.932%), Logistic Regression (23.982%) และ Decision Tree (30.212%) แม้ค่าเปอร์เซ็นต์จะแตกต่างกัน แต่ความสำคัญยังคงอยู่ในระดับสูงอย่างสม่ำเสมอ ซึ่งสอดคล้องกับการใช้เฉพาะข้อมูลประสิทธิภาพการปรากฏตัวทางดิจิทัล

อันดับที่ 2: OM-Duration แสดงความสม่ำเสมอในการเป็นคุณลักษณะสำคัญโดยปรากฏในทุกแบบจำลองและมักอยู่ในอันดับต้นๆ ใช้ข้อมูลแบบผสมผสาน โดยมีค่าความสำคัญที่ค่อนข้างคงที่ อยู่ในช่วง 6.098-17.333% ในเกือบทุกแบบจำลอง ซึ่งแสดงให้เห็นถึงความน่าเชื่อถือของคุณลักษณะนี้ในการทำนาย ซึ่งสอดคล้องกับการใช้เฉพาะข้อมูลประสิทธิภาพการปรากฏตัวทางดิจิทัล

อันดับที่ 3: EM-Rating แม้จะมีค่าความสำคัญที่แตกต่างกันมากในแต่ละแบบจำลองที่ใช้ข้อมูลแบบผสมผสาน แต่ก็แสดงบทบาทสำคัญอย่างสม่ำเสมอ โดยเฉพาะอย่างยิ่งในแบบจำลอง MLP (30.348%) และ SVM (26.535%) ซึ่งแสดงให้เห็นถึงความสำคัญของปัจจัยนี้ในการวิเคราะห์ ซึ่งสอดคล้องกับการใช้เฉพาะข้อมูลประสิทธิภาพการปรากฏตัวทางดิจิทัล

อันดับที่ 4: OM-Bounce เป็นคุณลักษณะที่แสดงความสม่ำเสมอในระดับปานกลาง โดยมีค่าความสำคัญที่ค่อนข้างคงที่ในหลายแบบจำลอง ยกเว้นในบางแบบจำลอง อาทิ AdaBoost (22.940%)

อันดับที่ 5: BF-Funding แม้จะมีความสำคัญสูงในบางแบบจำลอง ที่ใช้ข้อมูลแบบผสมผสาน อาทิ SVM (31.001%) แต่ความสม่ำเสมอของความสำเร็จในแบบจำลองอื่นๆ ค่อนข้างแปรปรวน ทำให้ความน่าเชื่อถือในภาพรวมน้อยกว่าปัจจัยอื่นๆ ข้างต้น แต่เป็นลำดับที่ 4 สำหรับแบบที่ใช้เฉพาะข้อมูลประสิทธิภาพการปรากฏตัวทางดิจิทัล

คุณลักษณะอื่นๆ อาทิ EM-Backlink, EM-Linking, OM-View, OM-Visit และ OM-Performance แม้จะปรากฏในหลายแบบจำลอง แต่มีความสม่ำเสมอของความสำเร็จค่อนข้างต่ำ และมีค่าความสำคัญที่แตกต่างกันมากในแต่ละแบบจำลอง จึงไม่ถูกจัดอยู่ในอันดับต้นๆ ของการจัดลำดับตามความสม่ำเสมอ การจัดลำดับนี้พิจารณาจากความสำเร็จในการปรากฏเป็นคุณลักษณะสำคัญ และความสม่ำเสมอของค่าความสำคัญในแต่ละแบบจำลอง ซึ่งสะท้อนให้เห็นถึงความน่าเชื่อถือของคุณลักษณะในการนำไปใช้ทำนายผลในสถานการณ์จริง เมื่อเปรียบเทียบระหว่างการใช้ข้อมูลแบบผสมผสาน กับการใช้เฉพาะข้อมูลประสิทธิภาพการปรากฏตัวทางดิจิทัลนั้น พบว่าการเพิ่มขึ้นของเปอร์เซ็นต์ความสำเร็จในปัจจัยที่เหลือในหมวดประสิทธิภาพการปรากฏตัวทางดิจิทัล แสดงให้เห็นว่าปัจจัยเหล่านี้สามารถชดเชยคุณลักษณะด้านข้อมูลพื้นฐานทางธุรกิจที่ถูกนำออกไปได้อย่างมีประสิทธิภาพ

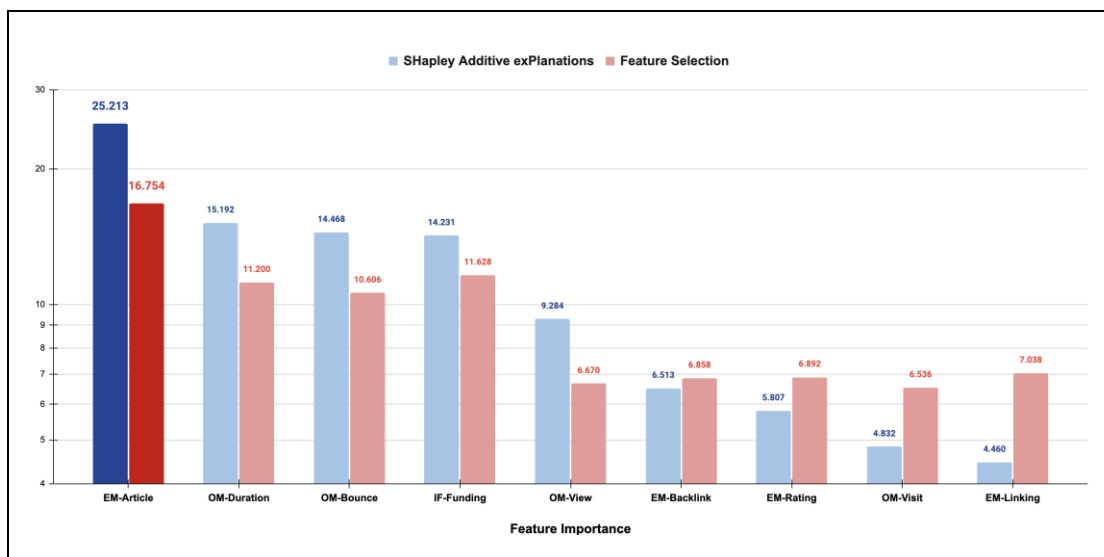
การวิเคราะห์นี้แสดงให้เห็นว่าคุณภาพเนื้อหาของที่มีผู้อ่านกล่าวถึง (EM-Article) ความน่าเชื่อถือของเว็บไซต์ที่กล่าวถึง (EM-Rating) และพฤติกรรมผู้เข้าชมเว็บไซต์ (OM-Duration) เป็นคุณลักษณะที่สำคัญที่สุดในเกือบทุกแบบจำลอง ในขณะที่การระดมทุนและตัวชี้วัดการดำเนินงานอื่นๆ มีความสำคัญรองลงมา

4.2 การวิเคราะห์อิทธิพลของคุณลักษณะที่มีผลต่อค่าของการทำนาย

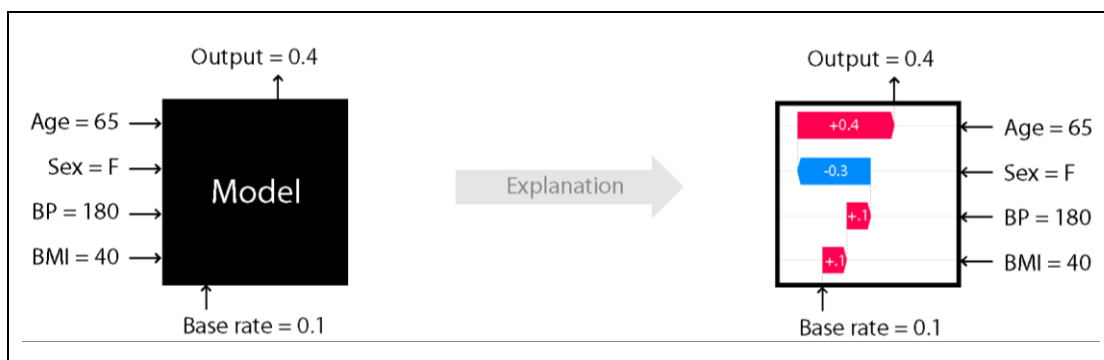
เมื่อเปรียบเทียบลำดับความสำคัญของฟีเจอร์ที่ได้จาก Feature Selection ในแบบจำลอง Random Forest และค่าที่ได้จากวิเคราะห์ด้วย SHAP ในตารางที่ 4.1 และรูปที่ 4.1 พบว่ามีลำดับและค่าที่แตกต่างกัน เป็นผลมาจากความแตกต่างในวิธีการคำนวณ และแนวทางการประเมินความสำคัญของคุณลักษณะ

ตารางที่ 4.1 เปรียบเทียบค่าความสำคัญของฟีเจอร์ที่ได้จาก SHAP และ Feature Selection

Feature	SHapley Additive exPlanations	Feature Selection
EM-Article	25.213 %	16.754 %
OM-Duration	15.192 %	11.200 %
OM-Bounce	14.468 %	10.606 %
BF-Funding	14.231 %	11.628 %
OM-View	9.284 %	6.670 %
EM-Backlink	6.513 %	6.858 %
EM-Rating	5.807 %	6.892 %
OM-Visit	4.832 %	6.536 %
EM-Linking	4.460 %	7.038 %



รูปที่ 4.1 กราฟแสดงค่าความสำคัญของฟีเจอร์ที่ได้จาก SHAP และ Random Forest

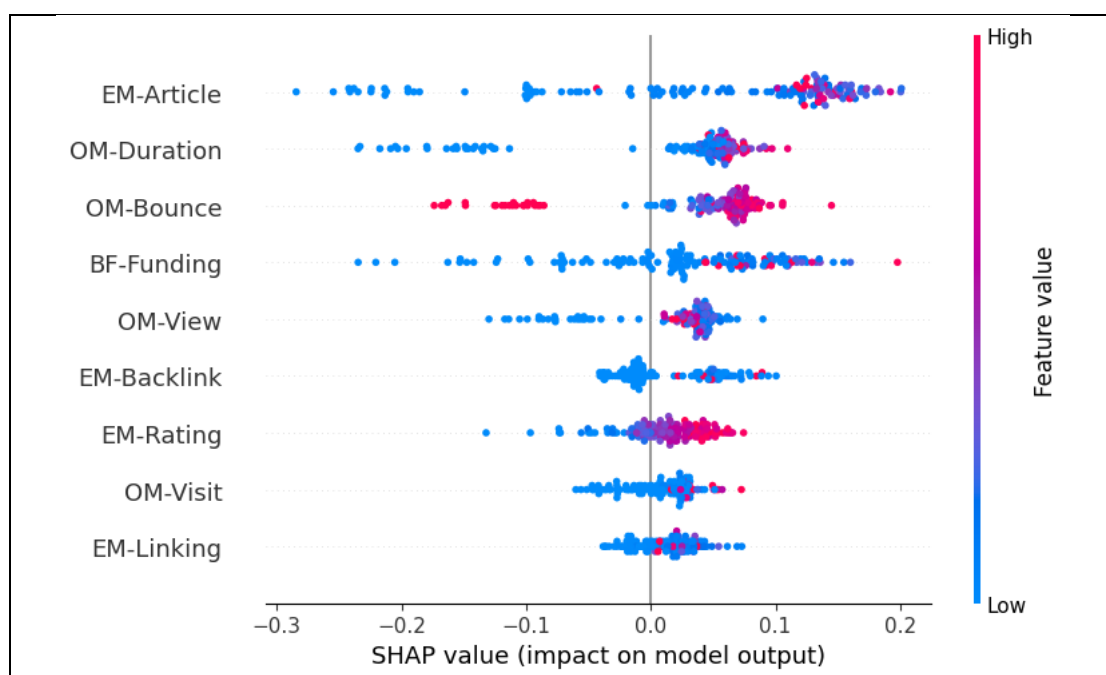


รูปที่ 4.2 แผนภาพแสดงการทำงานของ SHAP (Lundberg & Lee, 2017)

การทำ Feature Selection โดยทั่วไป มักจะทำโดยการดูว่าคุณลักษณะใดมีผลกระทบต่อความแม่นยำของโมเดล (Accuracy) มากที่สุด โดยการวัดความสำคัญของคุณลักษณะจากการที่คุณลักษณะนั้นช่วยลดความผิดพลาด (Error) หรือเพิ่มความแม่นยำ (Accuracy) ในการทำนายของแบบจำลอง ซึ่งไม่ได้คำนึงถึงความสัมพันธ์ระหว่างคุณลักษณะ หรือการมีอิทธิพลร่วมกันของคุณลักษณะต่าง ๆ (Multicollinearity) และไม่สามารถอธิบายได้ว่าแต่ละคุณลักษณะมีผลกระทบต่อค่าที่ได้จากการทำนายมากน้อยเพียงใด ซึ่งเป็นการวัดความสำคัญที่มองแต่ละคุณลักษณะในภาพรวม

ส่วนการวิเคราะห์ SHAP ใช้แนวคิดจากทฤษฎีเกม (Game Theory) เพื่ออธิบายว่าแต่ละคุณลักษณะมีผลกระทบอย่างไรต่อค่าที่ได้จากการทำนายผลของแบบจำลอง โดยการคำนวณค่า

SHAP จะบอกว่าแต่ละคุณลักษณะให้อิทธิพลเพิ่มขึ้นหรือลดลงต่อค่าที่ได้จากการทำนายผลของแบบจำลองจากค่าพื้นฐาน (Base Value) ของชุดข้อมูลในแบบจำลองนั้น ซึ่งช่วยให้สามารถอธิบายการตัดสินใจของแบบจำลองได้ในระดับที่สูงขึ้นได้ โดยช่วยให้สามารถอธิบายผลกระทบของคุณลักษณะในแต่ละชุดข้อมูลได้อย่างละเอียด และเข้าใจถึงการมีอิทธิพลร่วมกันของแต่ละคุณลักษณะ ดังแสดงในรูปที่ 3.4 ซึ่งจะให้ข้อมูลที่มีความหมายเชิงลึกในการอธิบายการตัดสินใจของแบบจำลอง ดังแสดงในรูปที่ 4.3



รูปที่ 4.3 กราฟแสดง SHAP Summary Plot แบบ beeswarm plot

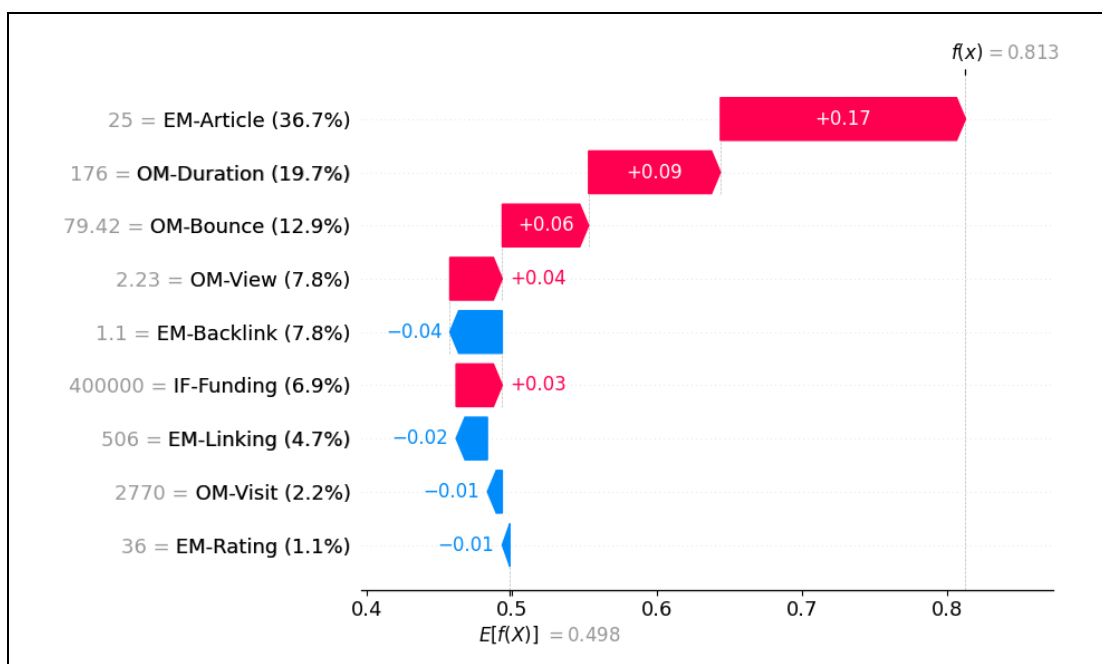
จากการวิเคราะห์ด้วย SHAP แบบ beeswarm plot ในรูปที่ 4.3 นั้น แต่ละจุดแทนข้อมูลหนึ่งชุด โดยตำแหน่งบนแกนนอน แสดงค่า SHAP (ผลกระทบต่อผลลัพธ์ของแบบจำลอง) สีน้ำเงินแทนค่าคุณลักษณะที่มีค่าน้อย สีแดงแทนค่าคุณลักษณะที่มีค่ามาก และด้านขวาหมายถึงการส่งผลบวกต่อค่าที่ทำนาย ด้านซ้ายหมายถึงการส่งผลลบต่อค่าที่ทำนาย การกระจายบนแกนตั้ง แสดงถึงการกระจายตัวของผลกระทบสำหรับคุณลักษณะแต่ละตัว ซึ่งคุณลักษณะส่วนใหญ่มีค่า SHAP กระจุกวัดรอบศูนย์แสดงถึงผลกระทบปานกลาง โดยที่ EM-Article ที่เป็นจำนวนบทความหรือข่าวที่กล่าวถึงองค์กร แสดงการกระจายของค่าผลกระทบกว้างที่สุดบ่งชี้ว่ามีอิทธิพลสูงที่สุด และ OM-Duration ที่เป็นระยะเวลาเฉลี่ยที่ผู้เข้าชมใช้บนเว็บไซต์ในแต่ละครั้งนั้นมีความสำคัญรองลงมา ซึ่งเป็นตัวบ่งชี้ว่าควรให้ความสำคัญกับการสร้างเนื้อหาที่มีคุณภาพเพื่อเพิ่มจำนวนบทความ (Earned Media) และการปรับเนื้อหาในเว็บไซต์เพื่อเพิ่มระยะเวลาในการเข้าชมเว็บไซต์ (Visit Duration) ในนานขึ้น ซึ่งคุณลักษณะ 2 ส่วนนี้ยังมากจะยิ่งส่งผลบวกต่อค่าที่ทำนาย

OM-Bounce ที่เป็นอัตราการเด้งออกของผู้ใช้งาน (Bounce Rate) แสดงถึงเปอร์เซ็นต์ของผู้ที่ออกจากเว็บไซต์หลังจากดูหน้าแรกเพียงหน้าเดียวพบว่ามีความซับซ้อน แม้จะมี Bounce Rate สูง (จุดสีแดง) ในบางกรณีก็ส่งผลบวกต่อค่าที่แบบจำลองทำนาย อาจอธิบายได้ว่ามีผู้ใช้ที่เข้ามาหาข้อมูลเฉพาะเจาะจง เมื่อได้ข้อมูลที่ต้องการก็ออกจากเว็บไซต์ทันที แต่เป็นการใช้งานที่มีประสิทธิภาพ หรือบางหน้าเว็บไซต์อาจเป็น Landing Page ที่ออกแบบมาให้ผู้ใช้ทำ Action เดียว อาทิ การลงทะเบียน หรือการดาวน์โหลด ในทางกลับกัน Bounce Rate ต่ำ (จุดสีน้ำเงิน) บางครั้งกลับส่งผลลบต่อค่าที่แบบจำลองทำนาย อาจเป็นเพราะผู้ใช้อาจหาข้อมูลที่ต้องการไม่พบ จึงต้องเปลี่ยนไปดูหน้าอื่นๆ หลายหน้า หรือการนำทางของเว็บไซต์อาจไม่ดีพอ ทำให้ผู้ใช้ต้องคลิกหลายครั้งเพื่อหาสิ่งที่ต้องการ

จากการวิเคราะห์การกระจายตัวของ "IF-Funding" พบว่าข้อมูลมีการกระจายตัวแบบสมมาตร โดยมีศูนย์กลางอยู่ที่จุดกึ่งกลาง ซึ่งบ่งชี้ว่าปริมาณเงินทุนที่เพิ่มขึ้นไม่ได้เป็นเครื่องรับประกันความสำเร็จเสมอไป โดยอาจมีปัจจัยอื่นๆ ร่วมด้วย อาทิ ประสิทธิภาพในการใช้จ่ายเงินทุน ความเหมาะสมของรูปแบบธุรกิจ และความสามารถในการบริหารเงินทุนของทีม ล้วนมีบทบาทสำคัญในการกำหนดผลลัพธ์ นอกจากนี้ การสังเกตจุดข้อมูลที่อยู่ห่างจากค่าเฉลี่ย (Outliers) ทั้งในทิศทางบวกและลบ แสดงให้เห็นถึงกรณีพิเศษที่เงินทุนจำนวนมากส่งผลต่ออย่างมีนัยสำคัญ ในขณะที่บางกรณีกลับพบผลลัพธ์ในเชิงลบ ซึ่งอาจเป็นผลมาจากการใช้เงินทุนอย่างรวดเร็วเกินความจำเป็น (Burn Rate) การวิเคราะห์นี้เน้นถึงความสำคัญของการพิจารณาปัจจัยที่หลากหลายในการประเมินผลกระทบของเงินทุนต่อความสำเร็จ

คะแนนความน่าเชื่อถือของเว็บไซต์ที่อ้างถึง หรือ EM-Rating แสดงความสัมพันธ์ที่ชัดเจนระหว่างค่าคุณลักษณะและผลกระทบต่อค่าในการทำนาย โดยคะแนนที่สูงกว่า ซึ่งแทนด้วยจุดสีแดง แสดงผลกระทบเชิงบวกอย่างสม่ำเสมอ

เพื่อทำความเข้าใจอิทธิพลของคุณลักษณะที่สำคัญให้มากขึ้น ผู้วิจัยจึงได้ใช้ตัวอย่างชุดข้อมูลที่ 347 มาทำการวิเคราะห์ด้วย SHAP Waterfall Plot ดังแสดงในรูปที่ 4.4



รูปที่ 4.4 กราฟแสดง SHAP Waterfall Plot โดยใช้ข้อมูลแถวที่ 347

SHAP Waterfall Plot ในรูปที่ 4.4 เป็นข้อมูลของชุดข้อมูลลำดับที่ 347 ซึ่งเป็นการแสดงผลของการวิเคราะห์ที่ใช้ SHAP (SHapley Additive exPlanations) เพื่ออธิบายการทำนายของแบบจำลอง Random Forest โดยแสดงผลอิทธิพลของคุณลักษณะต่างๆ ที่มีต่อผลการทำนายค่าผลลัพธ์ของแบบจำลอง ซึ่งค่าของอิทธิพลของคุณลักษณะนั้นๆ จะเป็นการแสดงค่าของอิทธิพลเฉพาะตัวอย่างชุดข้อมูลนั้นๆ เท่านั้น โดยสามารถอธิบายส่วนต่างๆ ของกราฟ ดังต่อไปนี้

1. แกนแนวนอนคือค่า Base Value ($E[f(X)]$)
 - แกนนี้แสดงค่าคาดการณ์พื้นฐานของแบบจำลอง (Base Value) โดยที่ค่า $E[f(X)] = 0.498$
 - ค่านี้จะมีการเปลี่ยนแปลงจากค่าพื้นฐานของแบบจำลอง ผ่านการเพิ่มหรือหักค่าจากฟีเจอร์ต่างๆ ที่มีผลกระทบต่อการทำนาย
2. ค่าผลลัพธ์สุดท้าย $f(x)$
 - ค่า $f(x) = 0.813$ ที่ด้านขวาสุดเป็นค่าผลลัพธ์สุดท้ายที่โมเดลทำนาย หลังจากพิจารณาผลกระทบจากแต่ละฟีเจอร์สำหรับตัวอย่างนั้นๆ ซึ่งเข้าใกล้ 1.0 ถูกต้องตามข้อมูลที่ทำป้ายกำกับไว้ว่าเป็นสตาร์ทอัพที่ประสบความสำเร็จ
 - ค่านี้จะมีการเปลี่ยนแปลงจากค่าพื้นฐานของแบบจำลอง ผ่านการเพิ่ม หรือลดค่าจากคุณลักษณะต่างๆ ที่มีผลกระทบต่อการทำนาย

3. คุณลักษณะต่าง ๆ (Features)

- ทุกบาร์ในกราฟแสดงถึงผลกระทบจากคุณลักษณะต่าง ๆ โดยจะมีการบอกชื่อคุณลักษณะพร้อมกับเปอร์เซ็นต์ของผลกระทบที่คุณลักษณะนั้นมีต่อการทำนาย
- คุณลักษณะที่มีผลกระทบมากที่สุดคือ EM-Article มีอิทธิพล 36.7% ที่เพิ่มค่าได้มากที่สุด 0.17
- คุณลักษณะที่มีผลกระทบน้อยที่สุดคือ EM-Rating มีอิทธิพล 1.1% ซึ่งมีผลกระทบในทางลบ 0.01

4. ทิศทางของบาร์

- บาร์ที่มีสีแดง อาทิ EM-Article และ OM-Duration คือคุณลักษณะที่เพิ่มค่าการทำนายในทางบวก (Positive Impact)
- บาร์ที่มีสีน้ำเงิน อาทิ EM-Backlink และ EM-Linking คือคุณลักษณะที่ลดค่าการทำนาย ส่งผลในทางลบ (Negative Impact)

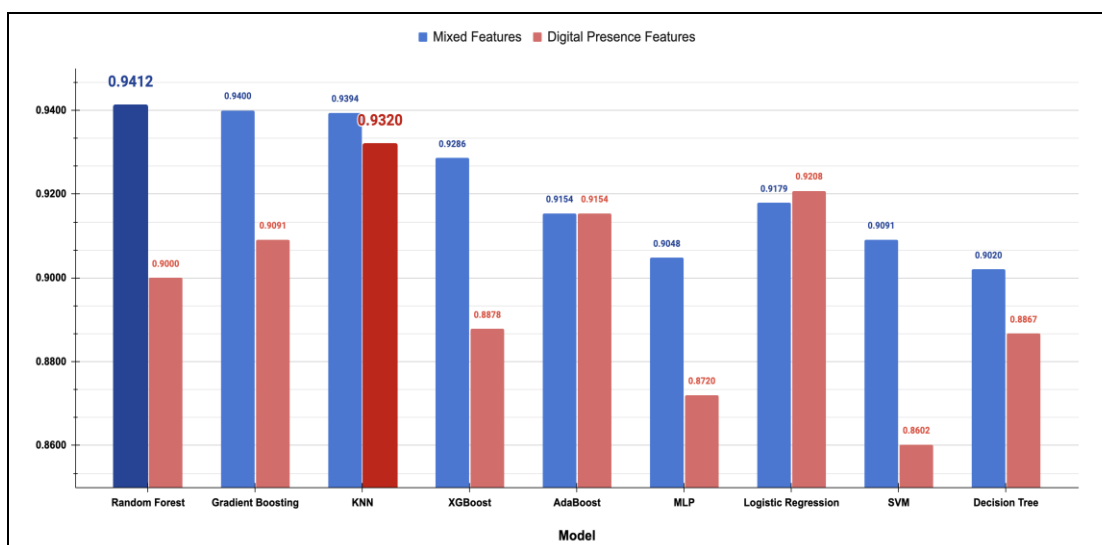
5. การจัดลำดับอิทธิพลของคุณลักษณะ

- คุณลักษณะที่มีผลกระทบสูงสุดจะถูกแสดงอยู่ด้านบนสุด ซึ่งสามารถเห็นได้ว่า EM-Article เป็นคุณลักษณะที่มีอิทธิพลมากที่สุด 36.7% โดยส่งผลในทางบวกต่อค่าของการทำนายถึง 0.17
- คุณลักษณะลำดับถัดไป ได้แก่ OM-Duration ที่มีอิทธิพล 19.7% โดยส่งผลบวกต่อค่าของการทำนาย 0.09 และ OM-Bounce ที่มีอิทธิพล 12.9% โดยส่งผลบวกต่อค่าของการทำนาย 0.06

ขณะที่คุณลักษณะบางตัว อาทิ EM-Backlink EM-Linking OM-Visit และ EM-Rating ส่งผลกระทบในเชิงลบต่อค่าของการทำนาย ซึ่งอาจเป็นเพราะจำนวนผู้เข้าชมเว็บไซต์ และความน่าเชื่อถือของเว็บไซต์ที่อ้างอิงถึงนั้นที่ค่าค่อนข้างต่ำ

4.3 การเปรียบเทียบประสิทธิภาพของแบบจำลอง

การเรียนรู้ของเครื่องในการทำนายความสำเร็จของสตาร์ทอัพ ระหว่างแบบจำลองที่ใช้ข้อมูลแบบผสมผสานระหว่างข้อมูลพื้นฐานทางธุรกิจร่วมกับประสิทธิภาพการปรากฏตัวทางดิจิทัล และแบบจำลองที่ใช้ข้อมูลเฉพาะประสิทธิภาพการปรากฏตัวทางดิจิทัลนั้น สามารถวิเคราะห์ประสิทธิภาพของโมเดลทั้ง 2 แบบ ในตารางที่ 3.9 และตารางที่ 3.10 สามารถเปรียบเทียบประสิทธิภาพของโมเดลทั้ง 9 ดังแสดงในรูปที่ 4.5



รูปที่ 4.5 กราฟแสดงค่า F1-Score ของแบบจำลองที่ใช้ข้อมูลแบบผสมผสานเทียบกับการใช้
เฉพาะ
ข้อมูลประสิทธิภาพการปรากฏตัวทางดิจิทัล

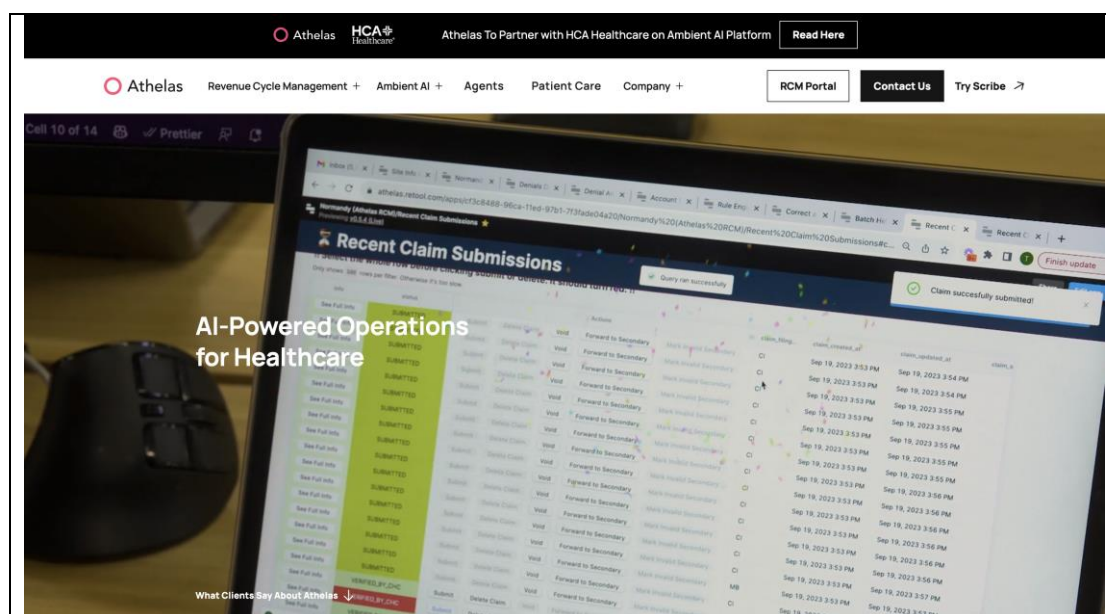
แบบจำลองที่มีประสิทธิภาพสูงที่สุด สำหรับแบบจำลองที่ใช้ข้อมูลแบบผสมผสาน ได้แก่ แบบจำลอง Random Forest มีค่า F1-Score เท่ากับ 0.9412 และค่า AUC เท่ากับ 0.9549 ถือว่าแบบจำลองมีประสิทธิภาพสูง และมีประสิทธิภาพของโมเดลการจำแนกประเภทข้อมูลสูง ในส่วนของแบบจำลองที่ใช้เฉพาะข้อมูลประสิทธิภาพการปรากฏตัวทางดิจิทัลนั้น แบบจำลอง K-Nearest Neighbors มีค่า F1-Score เท่ากับ 0.9320 และ AUC เท่ากับ 0.9174 แม้ว่าค่า AUC จะต่ำกว่าแบบจำลอง Random Forest เล็กน้อย แต่แบบจำลองยังคงแสดงความสามารถที่น่าพอใจ เหมาะสำหรับการใช้ที่ไม่สามารถหาข้อมูลพื้นฐานทางธุรกิจมาใช้ในการวิเคราะห์ได้

4.4 การนำองค์ความรู้ไปใช้งานจริง

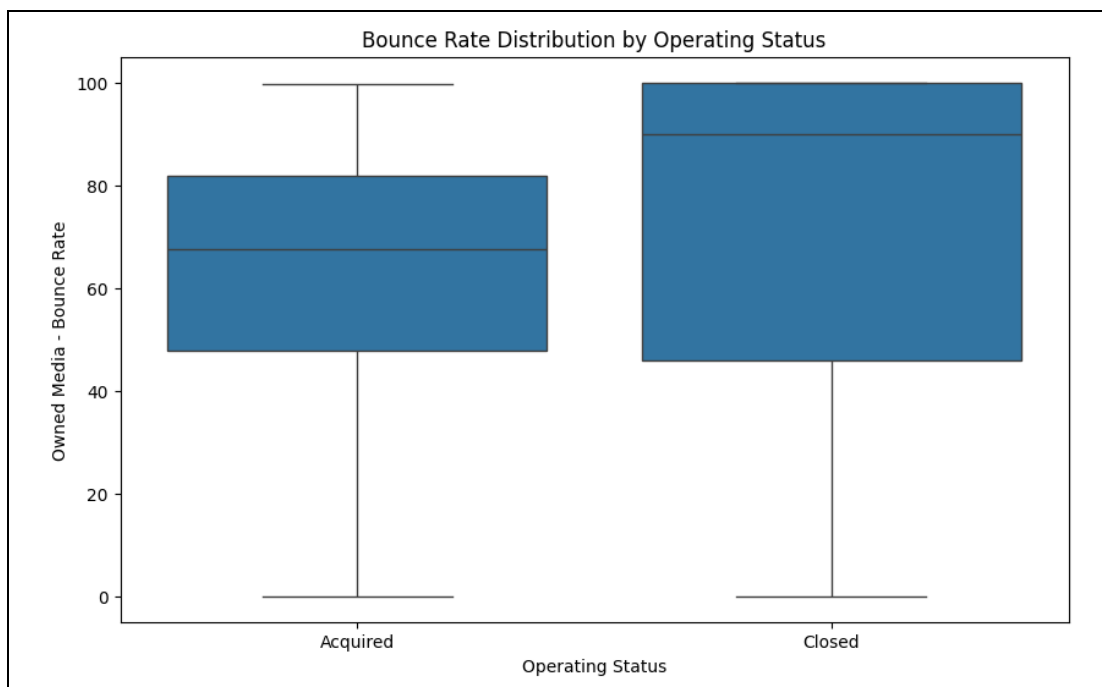
จากงานวิจัยพบว่าคุณลักษณะที่มีอิทธิพลต่อความแม่นยำในการทำนายของแบบจำลอง Random Forest นั้น คือคุณลักษณะด้านการปรากฏตัวทางดิจิทัลมีส่วนสูงถึง 86.19% โดยให้น้ำหนักของคุณลักษณะด้านสื่อที่เป็นเจ้าของ (Owned Media) และสื่อที่ได้รับ (Earned Media) พอๆ กัน และคุณลักษณะที่มีความสำคัญที่สุดคือ จำนวนบทความที่มีผู้อื่นกล่าวถึง ดังนั้นสตาร์ตอัปและผู้ที่เกี่ยวข้องในการสนับสนุนสตาร์ตอัป ควรให้ความสำคัญกับสื่อที่สตาร์ตอัปเป็นเจ้าของ และสื่อที่มีผู้อื่นกล่าวถึง ซึ่งอยู่นอกเหนือการควบคุมของสตาร์ตอัป ซึ่งควรให้ความสำคัญการกลยุทธ์ที่จะทำให้สื่อที่มีความน่าเชื่อถือในอุตสาหกรรมนั้นๆ กล่าวถึงมากที่สุด ก็จะเป็นการเพิ่มโอกาสที่จะประสบความสำเร็จได้ ส่วนคุณลักษณะด้านข้อมูลพื้นฐานทางธุรกิจนั้นมีอิทธิพลเพียง 13.81% โดยคุณลักษณะที่มีความสำคัญ คือจำนวนเงินรวมที่ได้รับจากการระดมทุน ซึ่งอาจจะขัดกับความรู้สึกของ

คนทั่วไป อาจเป็นเพราะว่าคุณลักษณะมีความผันผวนสูง ทำให้แบบจำลองมองว่ามีความสำคัญน้อยกว่า นั่นอาจเป็นเพราะว่าการที่สตาร์ทอัปแต่ละราย มีเป้าหมายและพฤติกรรมการใช้เงินในกิจกรรมต่างๆ ของกิจการที่แตกต่างกัน ซึ่งส่งผลต่อความสำเร็จโดยภาพรวม

นอกจากการให้ความสำคัญกับบทความที่มีผู้อ่านกล่าวถึง ซึ่งเป็นคุณลักษณะที่ส่งผลบวกต่อการทำนายเป็นอันดับ 1 แล้ว จากการวิเคราะห์พบว่าการที่อัตราการตีกลับการเข้าชมเว็บไซต์ (Bounce Rate) สูงสามารถส่งผลทั้งบวกและลบต่อการทำนายของแบบจำลอง ซึ่งเจ้าของและคนออกแบบเว็บไซต์ส่วนใหญ่จะกังวลกับค่าการตีกลับมากจนเกินไป แต่ควรพิจารณาวัตถุประสงค์ในการออกแบบเว็บไซต์ และควรวิเคราะห์ควบคู่กับประสบการณ์ของผู้ใช้งานเว็บไซต์ และเป้าหมายของแต่ละหน้าเว็บไซต์ ดังตัวอย่างในรูปที่ 4.6 เป็นหน้าเว็บไซต์ของ Athelas.com ที่มี Bounce Rate สูงถึง 95.9% แต่ยังคงประสบความสำเร็จ โดยถ้าวิเคราะห์จากการออกแบบเว็บไซต์นั้น ค่อนข้างหาข้อมูลยาก ไม่มีข้อมูลเกี่ยวกับราคา ไม่มีปุ่มลงทะเบียน พื้นที่แสดงภาพเคลื่อนไหวมากเกินไป ทำให้ผู้ใช้งานต้องเสียเวลาเลื่อนลงไปอ่านข้อมูล แต่ถ้าดูสินค้าและบริการของบริษัทนี้ จะพบว่าเป็นบริการสำหรับโรงพยาบาลขนาดใหญ่ ซึ่งทางบริษัทอาจใช้วิธีนำเสนอสินค้าและบริการโดยตรงกับกลุ่มเป้าหมาย (Business to Business: B2B) โดยให้พนักงานขาย ซึ่งมีประสิทธิภาพดีกว่าให้ลูกค้าศึกษาเองจากเว็บไซต์



รูปที่ 4.6 เว็บไซต์ Athelas.com



รูปที่ 4.7 การกระจายตัวของค่าอัตราการตีกลับ (Bounce Rate)

จากรูปที่ 4.7 จะแสดงถึงความแตกต่างของการกระจายตัวของค่าอัตราการตีกลับ ระหว่างกลุ่มที่ประสบความสำเร็จ (Acquired) และปิดกิจการ (Closed) อย่างเด่นชัด แต่เมื่อพิจารณาในรายละเอียด จะสามารถมองเห็นลักษณะร่วมบางประการ กล่าวคือ ถึงแม้ค่ามัธยฐานและช่วงการกระจายตัวของอัตราการตีกลับจะมีความแตกต่างกัน แต่ทั้งสองกลุ่มก็แสดงให้เห็นถึงการกระจายตัวของข้อมูลในช่วงกว้าง ซึ่งสะท้อนให้เห็นว่าอัตราการตีกลับนั้นมีความผันผวนได้มากไม่ว่ากิจการจะอยู่ในสถานะใดก็ตาม ยิ่งไปกว่านั้น การปรากฏของค่าผิดปกติ (Outliers) ในทั้งสองกลุ่ม ชี้ให้เห็นว่ามีบางกรณีที่อัตราการตีกลับเบี่ยงเบนไปจากค่าปกติอย่างมีนัยสำคัญ ซึ่งบ่งชี้ว่ายังมีปัจจัยอื่นๆ ที่นอกเหนือไปจากสถานะการดำเนินงานที่ส่งผลต่ออัตราการตีกลับ

การวิเคราะห์แบบนี้มีประโยชน์อย่างมากสำหรับการบ่งชี้ประสิทธิภาพ เพราะช่วยระบุว่าการวัดผลใดมีอิทธิพลต่อผลลัพธ์ความสำเร็จมากที่สุด ทำให้สามารถเน้นการให้คำปรึกษาและจัดสรรทรัพยากรได้อย่างมีกลยุทธ์มากขึ้น