

บทที่ 1

บทนำ

โครงการเรื่อง การวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิกเพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์ ในบทนี้จะเป็นการกล่าวถึงความสำคัญ และที่มาของปัญหา วัตถุประสงค์ ประโยชน์ที่จะได้รับ ขอบเขตงานของคณะผู้จัดทำและผู้ใช้งานทั่วไปบนเว็บไซต์ เครื่องมือที่ใช้ในการพัฒนาเว็บไซต์ สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล และระยะเวลาในการดำเนินการ รวมถึงนิยามศัพท์เฉพาะของโครงการ ดังนี้

1.1 ความสำคัญและที่มาของปัญหา

มหกรรมกีฬาโอลิมปิกครั้งที่ 1 ในปี ค.ศ. 1896 เป็นกีฬาโอลิมปิกสมัยใหม่ครั้งแรกที่ได้ถูกจัดขึ้นหลังจากหยุดไปเป็นเวลานาน เจ้าภาพอยู่ที่ประเทศต้นกำเนิดของโอลิมปิก กรุงเอเธนส์ ประเทศกรีซ ที่สนามกีฬาโอลิมปิกซึ่งต่อมาได้ถูกทหารฝรั่งเศสทำลายในช่วงสงครามโลก การฟื้นฟูกีฬาโอลิมปิกยุคโบราณได้ดึงดูดนักกีฬาจาก 14 ประเทศ ด้วยผู้แทนนักกีฬากลุ่มใหญ่จากประเทศกรีซ เยอรมนี ฝรั่งเศส และอังกฤษ ในวันที่ 26 ธันวาคม ค.ศ. 1896 นายเจมส์ คอนนอลลี ชาวอเมริกัน ชนะการแข่งขันกระโดดสามเท้า ซึ่งกลายเป็นผู้ชนะเลิศในโอลิมปิกคนแรกในรอบมากกว่า 150 ปี ผู้ชนะได้รับรางวัลเป็นเหรียญเงินและช่อโอลีฟ

โดยในการแข่งขัน Olympic Games 2020 ที่ผ่านมากใน Tokyo นั้นได้นำเทคโนโลยีมาใช้ในการแข่งขันด้วยเนื่องจากสถานะสถานการณ์โลกปัจจุบันนั้นเกิดการแพร่ระบาดของ Covid-19 โดยเทคโนโลยีที่ได้ทำมาใช้นั้นมีหลากหลาย คณะผู้จัดทำจึงเลือกมาสองเทคโนโลยี ที่คิดว่าน่าสนใจได้แก่

The 3D Athlete Tracking เก็บข้อมูลการแข่งขันแบบเรียลไทม์

3DAT หรือ 3D Athlete Tracking เป็นเทคโนโลยีที่จะทำให้การชมกีฬาที่มีความสนุกเพิ่มมากขึ้น โดยเฉพาะการแข่งขันประเภทกรีฑา เพราะนวัตกรรมนี้สามารถบันทึกไต่ระดับความเร็วของนักวิ่งแต่ละลูได้ ผ่านกล้องวิดีโอที่มีระบบ AI เพื่อเก็บสถิตินักกีฬามาวิเคราะห์ได้ทันที และนำข้อมูลเหล่านั้นมาแสดงผลบนจอแบบเรียลไทม์ ทั้งข้อมูลความเร็ว, ไตรกำลังวิ่งนำ, ระยะทางปัจจุบัน หรือระยะทางที่เหลืออยู่ นอกจากนี้ผู้ชมที่บ้านที่มีชุด VR ยังสามารถเชื่อมต่อเพื่อรับชมบรรยากาศการแข่งขันได้แบบเสมือนจริง

The Olympic Broadcasting ถ่ายทอดสดคมชัดระดับ 8K

การถ่ายทอดสดการแข่งขันด้วยความคมชัดระดับ 8K ให้กับผู้ชมภายในประเทศญี่ปุ่น นับเป็นครั้งแรกในประวัติศาสตร์โอลิมปิกอีกด้วย ทีมถ่ายทอดสดจะส่งสัญญาณภาพพร้อมระบบเสียง Immersive Audio ให้ผู้ชมทั่วโลกและในประเทศญี่ปุ่น พร้อมมุมมองแปลกใหม่ที่ไม่เคยเห็นมาก่อน มีการรีเพลย์ภาพ 360 องศา, การถ่ายทอดสดจากกล้อง VR และนำเสนอข้อมูลวิเคราะห์ด้วยระบบประมวลผล AI นับเป็นการปรับตัวในยุค COVID-19 ที่ผู้ชมที่บ้าน สามารถสัมผัสความคมชัดราวกับว่าอยู่ในสนามจริงๆ

ตั้งแต่อดีตที่ผ่านมาผู้คนส่วนใหญ่เริ่มให้ความสำคัญในการเล่นกีฬามากขึ้นและทำให้เกิดการจัดแข่งขันกีฬาประเภทต่างๆ เพิ่มมากขึ้นผู้คนเริ่มให้ความสำคัญในการเล่นกีฬาในระดับอาชีพในการลงแข่งขันในรายการต่างๆ เพื่อชิงเงินรางวัลและชื่อเสียงมากขึ้นและที่สำคัญกว่านั้นคือการที่ได้เป็นหนึ่งในกีฬาที่ตัวเองเชี่ยวชาญ จนทำให้เกิดการแข่งขันงานกีฬาระดับโลกซึ่งประเทศต่าง ๆ จะส่งตัวแทนนักกีฬาที่ได้คัดเลือกเข้าไปแข่งขันกันในระดับโลกเพื่อชิงความเป็นเจ้าเหรียญทอง จากการแข่งขันทั้งหมดนี้มีการแข่งขันมากมายหลายประเภทหลายรายการ นั่นจึงทำให้บุคคลที่สนใจหรือเริ่มสนใจในการเล่นกีฬาไม่สามารถเข้าใจหรือรู้ผลการแข่งขันทั้งหมดว่าแต่ละประเทศที่ส่งนักกีฬาลงแข่งขันทั้งหมดที่ผ่านมาไม่มีประเทศไหนที่ส่งนักกีฬามากที่สุดหรือได้เหรียญมากที่สุดหรือกีฬาประเภทไหนลงแข่งขันมากที่สุด แต่ด้วยจำนวนข้อมูลที่มีปริมาณมหาศาลทำให้ยากต่อ

การนำเสนอให้เข้าใจได้ง่าย ซึ่งถือเป็นข้อจำกัดหากต้องการนำข้อมูลชุดนี้ไปใช้เพื่อเผยแพร่ให้แก่บุคคลภายนอกหรือกลุ่มผู้ใช้ข้อมูลได้รับรู้

จากปัญหาข้างต้นดังนั้นผู้จัดทำโครงการได้ทำการศึกษามหกรรมกีฬาโอลิมปิก 120 ปี เพื่อนำมาเข้าสู่กระบวนการจำแนกตามรายการชุดข้อมูล และประมวลผลของข้อมูลให้เหมาะสมตรงกับความต้องการด้วยการใช้เทคนิคการทำเหมืองข้อมูลในโมเดล Data Classification โดยอาศัยรูปแบบ moving average ในการสร้างความสัมพันธ์ของข้อมูลและนำข้อมูลที่ได้เข้าสู่กระบวนการวิเคราะห์ข้อมูลและนำข้อมูลที่ผ่านการทำเหมืองข้อมูลและการวิเคราะห์ข้อมูลที่อยู่ในรูปแบบ Visualization ที่เสร็จสมบูรณ์แล้วนั้น นำไปเผยแพร่บนเว็บไซต์เพื่อเป็นข้อมูลในการศึกษาและทำความเข้าใจ

1.2 วัตถุประสงค์

เพื่อการวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิก เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์

1.3 ประโยชน์ที่จะได้รับ

ได้ผลการวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิก เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์

1.4 ขอบเขต

1.4.1 ขอบเขตของผู้วิเคราะห์

จัดทำกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM

1) Business Understanding ผู้วิเคราะห์ข้อมูล ได้ทำการหาข้อมูลที่น่าสนใจจากเว็บไซต์ www.kaggle.com โดยข้อมูลที่น่าสนใจและจะนำมาวิเคราะห์นั้นคือข้อมูล Olympic Games, 1986-2021 เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์

2) Data Understanding ผู้วิเคราะห์ข้อมูล ได้ทำการรวบรวมข้อมูลที่ได้มาจากเว็บไซต์ www.kaggle.com และทำการตรวจเช็คข้อมูล สภาพข้อมูล จำนวนข้อมูล

ความผิดพลาดของข้อมูลที่ได้มาและตั้งโจทย์เพื่อหาแนวทางในการวิเคราะห์ข้อมูล และนำข้อมูลที่ได้มาไปเผยแพร่ข้อมูลบนเว็บไซต์

3) Data Preparation ผู้วิเคราะห์ข้อมูล ได้นำข้อมูลที่ได้มานั้นไปเข้ากระบวนการ Data Cleaning เพื่อปรับปรุงข้อมูลที่ได้มาให้อยู่ในรูปแบบที่สมบูรณ์ โดยที่จะตัดข้อมูลที่ไม่ว่าจำเป็นบางส่วนออก เพื่อเหลือข้อมูลที่จำเป็นในการนำไปวิเคราะห์ข้อมูลต่อไป

4) Modeling ผู้วิเคราะห์ข้อมูล ได้เลือกใช้รูปแบบการทำเหมืองข้อมูลแบบ Data Classification เพื่อจัดทำข้อมูลวิเคราะห์ในการพยากรณ์ว่าในการแข่งขันครั้งต่อไปในอีก 4 ปี ในแต่ละประเทศจะได้รับเหรียญรางวัลรวมเท่าไร เหรียญทองเท่าไร เหรียญเงินเท่าไร เหรียญทองแดงเท่าไร และการพยากรณ์หาเปอร์เซ็นต์ความน่าจะเป็นของแต่ละประเทศที่จะเป็นเจ้าเหรียญทอง และการพยากรณ์หาเปอร์เซ็นต์ความน่าจะเป็นในการชนะเลิศของกีฬาแต่ละประเภทของแต่ละประเทศ โดยจะใช้ทฤษฎี Time-Series Forecasting และทฤษฎีสมการต่าง ๆ ในการตรวจสอบความถูกต้องของผู้วิเคราะห์ใช้รูปแบบโมเดล Decision Trees และ Naive Bayes เพื่อตรวจสอบความแม่นยำของข้อมูลและจัดกลุ่มข้อมูลที่มีความสัมพันธ์กัน เพื่อใช้ในการพยากรณ์หาความน่าจะเป็นต่าง ๆ

5) Evaluation ผู้วิเคราะห์ข้อมูล ได้ผลการวิเคราะห์ข้อมูลด้วยเทคนิคทาง Data Mining จากการสร้างโมเดล Decision Trees ทำให้การวัดผลของข้อมูลและการสรุปผลข้อมูลตรงกับวัตถุประสงค์ที่ตั้งไว้

6) Deployment ผู้วิเคราะห์ข้อมูล ได้นำข้อมูลที่ได้วิเคราะห์มานั้นมาแสดงผลในรูปแบบ Visualization ที่จัดขึ้นมาจากนั้นแล้วจึงนำไปแสดงผล บน Web Browser เพื่อเผยแพร่ข้อมูลที่ได้มาให้ผู้ที่สนใจได้ศึกษาและเรียนรู้ในผลรับของการแข่งขันกีฬาโอลิมปิกที่จัดขึ้นมา

1.4.3 ผู้เยี่ยมชมเว็บไซต์

- 1) สามารถดูข้อมูลของการแข่งขันกีฬาโอลิมปิก
- 2) สามารถดูข้อมูลการพยากรณ์ได้
- 3) สามารถดูข้อมูลสารสนเทศ ในลักษณะ รูปแบบแผนภูมิได้
- 4) สามารถเปรียบเทียบข้อมูลตารางจำนวนเหรียญทองของแต่ละปีได้

1.4.4 ผู้ดูแลระบบ

- 1) สามารถเผยแพร่ข้อมูลได้
- 2) สามารถอัปเดตข้อมูลการแข่งขันกีฬาโอลิมปิกเข้าไปในระบบได้
- 3) ผู้ดูแลระบบ สามารถเพิ่ม ลบ แก้ไขข้อมูลการแข่งขันกีฬาได้

1.5 เครื่องมือที่ใช้ในการพัฒนาโปรแกรม

1.5.1 Hardware

- 1) คอมพิวเตอร์ตั้งโต๊ะ PC
 - CPU: AMD Ryzen 5 2600
 - Mainboard: ASRock b450m
 - Memory: Ram 16 GB
- 2) Notebook Acer Nitro 5

1.5.2 Software

- 1) โปรแกรม Tableau Public ใช้ในการแสดงผลข้อมูลในรูปแบบของภาพ
- 2) โปรแกรม Microsoft office ใช้สำหรับการทำรายงานและพาวเวอร์พ้อยในการนำเสนอ
- 3) โปรแกรม RapidMiner Studio ใช้สำหรับการทำเทคนิคเหมืองข้อมูล
- 4) โปรแกรม SubLime ใช้ในการเขียนโค้ด Web app
- 5) โปรแกรม Xampp 5.5.38 ใช้ในการจำลอง Web server
- 6) โปรแกรม FileZilla 3.40.0 โปรแกรมที่ใช้ติดต่อกับ FTP Server เพื่ออัปโหลด (Upload) ไฟล์ (File)
- 7) ชุดคำสั่ง HTML ,PHP , Bootstrap 5 , Javascript ใช้สำหรับการเขียนเว็บไซต์เพื่อเผยแพร่

1.6 สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล

1.6.1 สถานที่ใช้ในการดำเนินการศึกษา

1) มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาภาคพายัพเชียงใหม่ 128 ถนนห้วยแก้ว ตำบลช้างเผือก อำเภอเมือง จังหวัดเชียงใหม่ 50300

2) แหล่งรวบรวมข้อมูล

- เว็บไซต์ www.kaggle.com

1.7 ระยะเวลาในการดำเนินการ

แผนการดำเนินการ	2564				
	พ.ย	ธ.ค	ม.ค	ก.พ	มี.ค
1. ศึกษาและกำหนดความต้องการ	→				
2. วิเคราะห์ออกแบบระบบและสร้างฐานข้อมูล	→	→			
3. เขียนและทดสอบโปรแกรมเขียนและทดสอบโปรแกรม		→	→		
4. ติดตั้ง ทดสอบ และปรับปรุงระบบ				→	
5. ตรวจสอบระบบโดยรวม				→	
6. ประเมินการใช้งานระบบ				→	
7. จัดทำคู่มือการใช้งาน				→	→
8. จัดทำเอกสารประกอบโครงการ	→	→	→	→	→

1.8 นิยามศัพท์เฉพาะ

1.8.1 การวิเคราะห์ข้อมูล หมายความว่า เป็นการจัดระเบียบแยกแยะส่วนต่าง ๆ ของหลักฐาน หรือข้อมูลที่ได้ออกเป็นหมวดหมู่ เพื่อหาคำตอบตามความมุ่งหมาย และตามสมมติฐานที่ได้กำหนดไว้

1.8.2 เว็บไซต์ หมายความว่า เว็บเพจที่จัดทำขึ้น เพื่อนำเสนอข้อมูลต่าง ๆ ผ่านทางคอมพิวเตอร์และอินเทอร์เน็ต โดยจะมีหน้าเว็บเพจหลาย ๆ หน้า que เชื่อมโยงเข้ากับไฮเปอร์ลิงก์ เพื่อให้สามารถเปิดไปยังหน้าเพจต่าง ๆ ได้อย่างง่ายดายและถูกจัดเก็บไว้ใน www. (เว็ลด์ไวด์เว็บ)

1.8.3 เหมือนข้อมูล หมายความว่า การวิเคราะห์ข้อมูลจากข้อมูลจำนวนมาก เพื่อหาความสัมพันธ์ของข้อมูลที่เกี่ยวข้อง โดยทำการจำแนกประเภท รูปแบบ เชื่อมโยงข้อมูลที่มีความสัมพันธ์กัน และหาความน่าจะเป็นที่จะเกิดขึ้น เพื่อให้ได้องค์ความรู้ใหม่ ที่สามารถนำไปใช้ประกอบการตัดสินใจในด้านต่าง ๆ

1.9 บทสรุป

จากบทนำที่ได้กล่าวมาในข้างต้นทั้งหมดนั้น ผู้วิเคราะห์ข้อมูลได้สังเกตเห็นความสำคัญของการนำเทคโนโลยีมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิกเพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์ ด้วยกระบวนการวิเคราะห์ข้อมูลได้ทำการจำแนกข้อมูลออกเป็นประเภทต่าง ๆ ด้วยเทคนิคการทำเหมืองข้อมูลในรูปแบบ Classification โดยอาศัยการเรียนรู้ข้อมูลเก่าในช่วงเวลาที่ผ่านมาและจัดทำโมเดลในรูปแบบต่าง ๆ Decision tree และ Naive Bayes ในการตรวจสอบความแม่นยำของข้อมูลและใช้การพยากรณ์และทฤษฎีสมการต่าง ๆ เข้ามาช่วยในการหาความจะเป็นและการพยากรณ์จากที่ได้ตั้งขอบเขตขึ้น และจัดทำรูปแบบภูมิขึ้นเพื่อนำไปทำการนำเสนอและเผยแพร่ผ่านทางเว็บไซต์ต่อไป