บทที่ 3

วิธีการดำเนินงานโครงงาน

ในการพัฒนาโครงงานการต้องมีการวิเคราะห์และออกแบบข้อมูลให้ออกมาสอดคล้องกับความต้องการของผู้ใช้เพื่อให้เกิดประโยชน์สูงสุด และทำให้โครงการสำเร็จลุล่วง ผู้วิเคราะห์จึงได้วิเคราะห์และออกแบบข้อมูลใหม่โดยใช้เครื่องมือประกอบไปด้วย โดยการใช้กระบวนการพัฒนาฐานข้อมูล CRISP-DM สร้างแบบฟอร์มข้อมูลและปรับโครงสร้างข้อมูลใหม่

3.1 กระบวนการ CRISP-DM

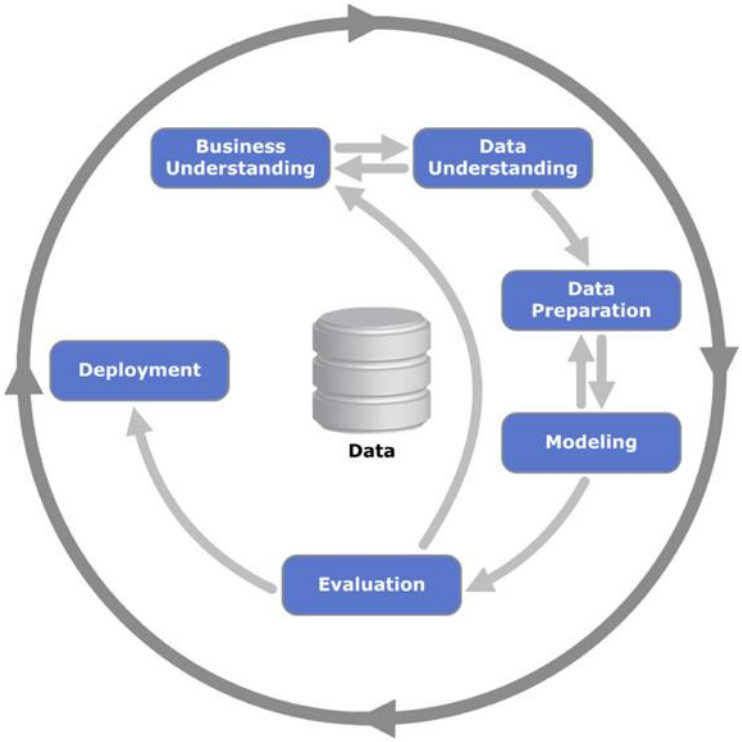
3.2 กระบวนการพยากรณ์

3.3 การออกแบบและพัฒนาเว็บไซต์

3.4 สรุปผล

3.1 กระบวนการ CRISP-DM

กระบวนการ CRISP-DM จะประกอบด้วย 6 ขั้นตอน ในแต่ละขั้นตอนนั้นจะเป็นขั้นตอนที่ต่อเนื่องกันกับขั้นตอนถัดไป ซึ่งจะแสดงด้วยลูกศรที่เชื่อมระหว่างกล่องสี่เหลี่ยมแต่ละกล่อง ตัวอย่างเช่นเมื่อได้ผลลัพธ์จากขั้นตอนการเตรียมข้อมูล (Data Preparation) แล้วจะนำไปสร้างโมเดลจำแนกประเภทข้อมูลในขั้น (Modeling) และหลังจากนั้นอาจจะย้อนกลับมาเปลี่ยนแปลงข้อมูลให้ถูกต้องมากขึ้นเพื่อเพิ่มความถูกต้องของข้อมูลให้มีความถูกต้องมากขึ้น เป็นต้น



ภาพที่ 3.1 รูปภาพประกอบกระบวนการ CRISP–DM

กระบวนการ CRISP–DM ประกอบด้วย 6 ขั้นตอน

1) การทำความเข้าใจธุรกิจ (Business Understanding)

ขั้นตอนแรกมุ่งไปที่การทำความเข้าใจธุรกิจปัญหาและวัตถุประสงค์ของโครงการ จากนั้นแปลงปัญหาให้อยู่ในรูปของโจทย์สำหรับการวิเคราะห์ข้อมูล และวางแผนการดำเนินงานเบื้องต้น

2) การทำความเข้าใจข้อมูล (Data Understanding)

ขั้นตอนนี้เริ่มต้นด้วยการรวบรวมข้อมูล จากนั้นทำความเข้าใจ ตรวจสอบคุณภาพของข้อมูลและเลือกข้อมูลที่เก็บรวบรวมมาและเลือกใช้ข้อมูลที่ได้มาในการวิเคราะห์

3) การเตรียมข้อมูล (Data Preparation)

การเตรียมข้อมูล เป็นขั้นที่จะทำให้ข้อมูลดิบที่รวบรวมมา กลายเป็นข้อมูลสมบูรณ์ที่พร้อมจะเข้าสู่โมเดลในขั้นตอนที่ 4 เช่น การสร้างตาราง การลบข้อมูลที่ไม่ต้องการออก การแปลงข้อมูลให้อยู่ในรูปแบบที่ต้องการ

4) การสร้างโมเดล (Modeling)

ในขั้นตอนนี้ จะเลือกและทดสองสร้างโมเดลหลาย ๆ แบบขึ้นมาที่จะสามารถแก้ไขปัญหาที่ต้องการได้ จากนั้นค่อย ๆ ปรับค่าพารามิเตอร์ในแต่ละโมเดล เพื่อให้ได้โมเดลที่เหมาะสมที่สุดมาใช้ในการแก้ไขปัญหา

5) การวัดประสิทธิภาพของโมเดล (Evaluation)

การวัดประสิทธิภาพของโมเดลที่ได้จากขั้นตอนที่ 4 เพื่อวัดว่าโมเดลมีประสิทธิภาพเพียงพอต่อการนำไปใช้งานแล้วหรือไม่ ซึ่งโมเดลแต่ละประเภทก็จะมีตัววัดประสิทธิภาพที่แตกต่างกันออกไป

6) การนำโมเดลไปใช้งานจริง (Deployment)

การนำโมเดลที่เหมาะสมที่สุดไปใช้งานจริง เพื่อวิเคราะห์และแก้ปัญหาที่ต้องการ

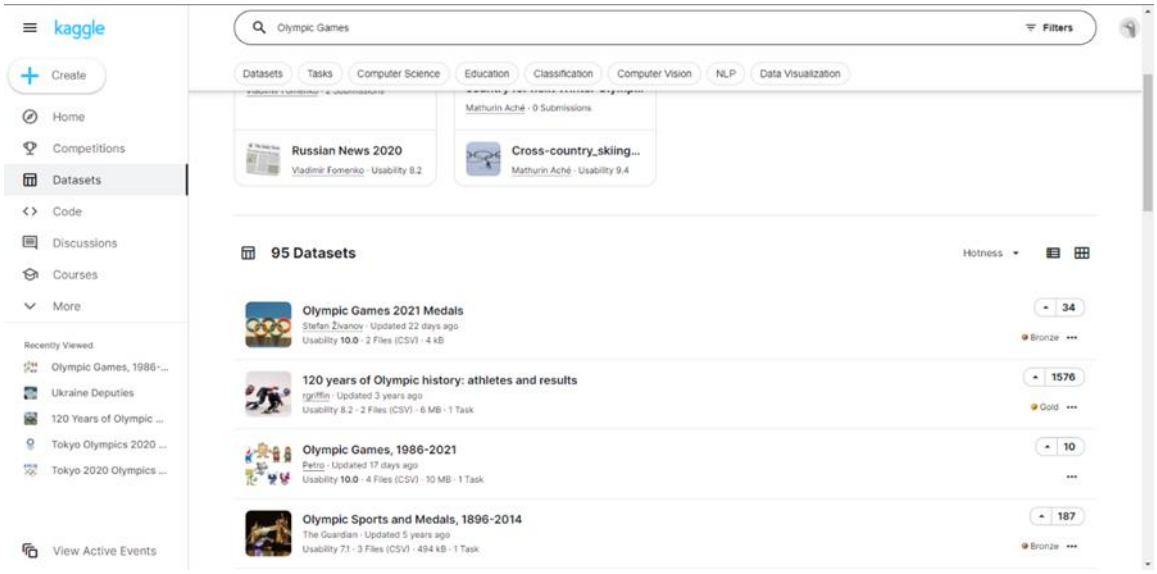
3.1.1 การทำความเข้าใจธุรกิจ (Business Understanding)

ผู้วิเคราะห์มุ่งเน้นไปที่การทำความเข้าใจกระบวนการทางข้อมูลและระบบวิเคราะห์ข้อมูลกับปัญหา เพื่อที่จะนำผลการวิเคราะห์ข้อมูลไปใช้ให้เกิดประโยชน์ โดยที่ผู้วิเคราะห์สนใจในข้อมูลเกมการแข่งขัน Olympic Games, 1986–2021 เพื่อจะนำไปวิเคราะห์ข้อมูลและนำผลข้อมูลไปใช้ให้เกิดประโยชน์ เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์

3.1.2 การทำความเข้าใจข้อมูล (Data Understanding)

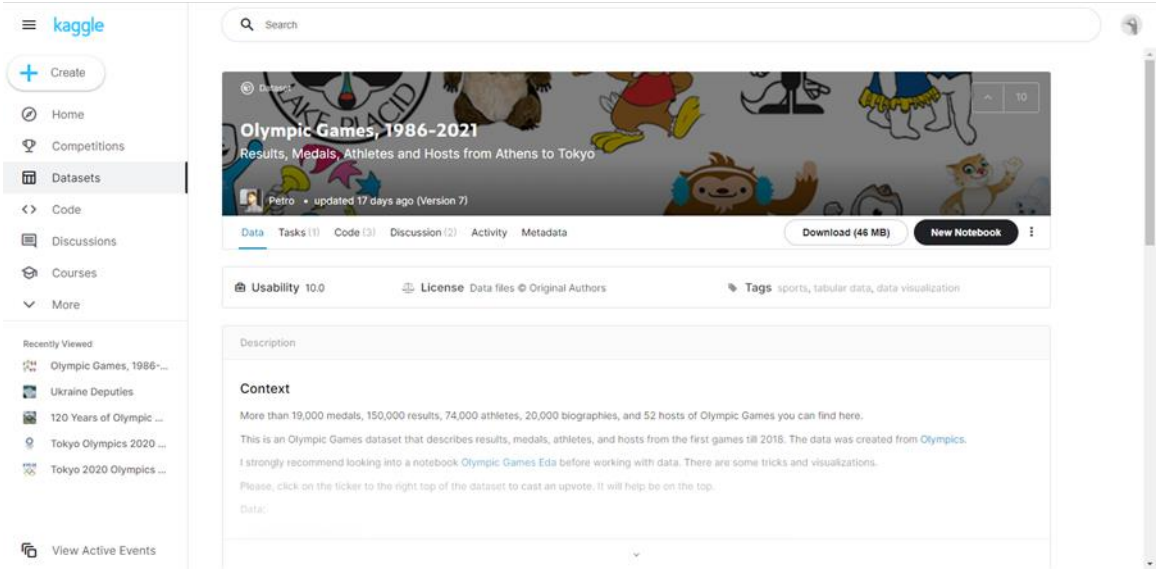
ผู้วิเคราะห์ข้อมูล ได้เริ่มทำการรวบรวมข้อมูลจากเว็บไซต์ www.kaggle.com และโดยที่ได้เลือกชุดข้อมูล Olympic Games, 1986–2021 จากนั้นได้นำชุดข้อมูลที่ได้เลือกมาทำการตรวจเช็คข้อมูล สภาพข้อมูล จำนวนข้อมูล ความผิดพลาดของข้อมูลที่ได้มาและตั้งโจทย์เพื่อหาแนวทางในการวิเคราะห์ข้อมูลและนำข้อมูลที่ได้มาไปเผยแพร่ข้อมูลบนเว็บไซต์

1) ผู้วิเคราะห์ข้อมูล ได้เริ่มทำการรวบรวมข้อมูลจากเว็บไซต์ www.kaggle.com โดยที่กำหนดการค้นหา Olympic Games



ภาพที่ 3.2 ข้อมูลจากจากเว็บไซต์ www.kaggle.com

2) ผู้วิเคราะห์ข้อมูล ได้เลือกชุดข้อมูล Olympic Games, 1986–2021 จะเป็นชุดข้อมูลที่จะนำมาทำการวิเคราะห์

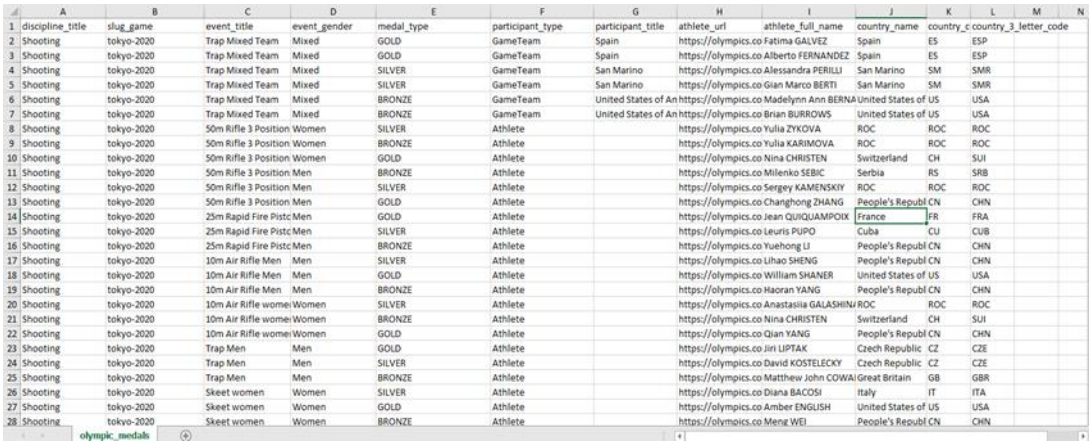


ภาพที่ 3.3 ชุดข้อมูล Olympic Games, 1986–2021

3) ชุดข้อมูล Olympic Games, 1986–2021 ที่ได้เลือกมานำเข้าสู่กระบวนการ ตรวจเช็คข้อมูล สภาพข้อมูล จำนวนข้อมูล ความผิดพลาดของข้อมูล โดยจากผลตรวจเช็คมีไฟล์ชุดข้อมูลทั้งหมด 4 ไฟล์ ได้แก่ olympic_athletes, olympic_hosts, olympic_medals, olympic_results โดยรวมของข้อมูลอยู่ในสภาพที่ดี

โดยทางผู้วิเคราะห์ได้เลือกชุดข้อมูลที่เหมาะสมมันการดำเนินงานทั้งสิ้น 1 ชุด ดังนี้

ชุดข้อมูลที่ 1 ที่ได้เลือกมา คือ olympic_medals โดยชุดข้อมูลนี้จะประกอบไปด้วย ประเภทกีฬา, ชื่อรายการแข่งขัน, หัวข้อการแข่งขันกีฬา, ประเภทของเพศในการลงแข่ง, ประเภทเหรียญรางวัล, ประเภทการแข่งขัน, รายชื่อนักกีฬาประเภททีม, URL ของนักกีฬา, ชื่อนักกีฬา, ชื่อประเทศ, ชื่อตัวย่อของประเทศ, ชื่อตัวย่อของประเทศแบบ 3 ตัวอักษร, คะแนนการแข่งขัน, ประเภทคะแนน



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	discipline_title	slug_game	event_title	event_gender	medal_type	participant_type	participant_title	athlete_url	athlete_full_name	country_name	country_c	country_3_letter_code		
2	Shooting	tokyo-2020	Trap Mixed Team	Mixed	GOLD	GameTeam	Spain	https://olympics.co	Fatima GALVEZ	Spain	ES	ESP		
3	Shooting	tokyo-2020	Trap Mixed Team	Mixed	GOLD	GameTeam	Spain	https://olympics.co	Alberto FERNANDEZ	Spain	ES	ESP		
4	Shooting	tokyo-2020	Trap Mixed Team	Mixed	SILVER	GameTeam	San Marino	https://olympics.co	Alessandra PERILLI	San Marino	SM	SMR		
5	Shooting	tokyo-2020	Trap Mixed Team	Mixed	SILVER	GameTeam	San Marino	https://olympics.co	Gian Marco BERTI	San Marino	SM	SMR		
6	Shooting	tokyo-2020	Trap Mixed Team	Mixed	BRONZE	GameTeam	United States of An	https://olympics.co	Madelynn Ann BERNA	United States of	US	USA		
7	Shooting	tokyo-2020	Trap Mixed Team	Mixed	BRONZE	GameTeam	United States of An	https://olympics.co	Brian BURROWS	United States of	US	USA		
8	Shooting	tokyo-2020	50m Rifle 3 Position	Women	SILVER	Athlete		https://olympics.co	Yulia ZYKOVA	ROC	ROC	ROC		
9	Shooting	tokyo-2020	50m Rifle 3 Position	Women	BRONZE	Athlete		https://olympics.co	Yulia KARIMOVA	ROC	ROC	ROC		
10	Shooting	tokyo-2020	50m Rifle 3 Position	Women	GOLD	Athlete		https://olympics.co	Nina CHRISTEN	Switzerland	CH	SUI		
11	Shooting	tokyo-2020	50m Rifle 3 Position	Men	BRONZE	Athlete		https://olympics.co	Milenko SEBIC	Serbia	RS	SRB		
12	Shooting	tokyo-2020	50m Rifle 3 Position	Men	SILVER	Athlete		https://olympics.co	Sergey KAMENSKIY	ROC	ROC	ROC		
13	Shooting	tokyo-2020	50m Rifle 3 Position	Men	GOLD	Athlete		https://olympics.co	Changhong ZHANG	People's Republ	CN	CHN		
14	Shooting	tokyo-2020	25m Rapid Fire Pistc	Men	GOLD	Athlete		https://olympics.co	Jean QUIQUAMPOIX	France	FR	FRA		
15	Shooting	tokyo-2020	25m Rapid Fire Pistc	Men	SILVER	Athlete		https://olympics.co	Leuris PUPO	Cuba	CU	CUB		
16	Shooting	tokyo-2020	25m Rapid Fire Pistc	Men	BRONZE	Athlete		https://olympics.co	Yuehong LI	People's Republ	CN	CHN		
17	Shooting	tokyo-2020	10m Air Rifle Men	Men	SILVER	Athlete		https://olympics.co	Lihao SHENG	People's Republ	CN	CHN		
18	Shooting	tokyo-2020	10m Air Rifle Men	Men	GOLD	Athlete		https://olympics.co	William SHANER	United States of	US	USA		
19	Shooting	tokyo-2020	10m Air Rifle Men	Men	BRONZE	Athlete		https://olympics.co	Haoran YANG	People's Republ	CN	CHN		
20	Shooting	tokyo-2020	10m Air Rifle wome	Women	SILVER	Athlete		https://olympics.co	Anastasiia GALASHINA	ROC	ROC	ROC		
21	Shooting	tokyo-2020	10m Air Rifle wome	Women	BRONZE	Athlete		https://olympics.co	Nina CHRISTEN	Switzerland	CH	SUI		
22	Shooting	tokyo-2020	10m Air Rifle wome	Women	GOLD	Athlete		https://olympics.co	Qian YANG	People's Republ	CN	CHN		
23	Shooting	tokyo-2020	Trap Men	Men	GOLD	Athlete		https://olympics.co	Jiri LIPTAK	Czech Republic	CZ	CZE		
24	Shooting	tokyo-2020	Trap Men	Men	SILVER	Athlete		https://olympics.co	David KOSTELECKY	Czech Republic	CZ	CZE		
25	Shooting	tokyo-2020	Trap Men	Men	BRONZE	Athlete		https://olympics.co	Matthew John COWA	Great Britain	GB	GBR		
26	Shooting	tokyo-2020	Skeet women	Women	SILVER	Athlete		https://olympics.co	Diana BACOSI	Italy	IT	ITA		
27	Shooting	tokyo-2020	Skeet women	Women	GOLD	Athlete		https://olympics.co	Amber ENGLISH	United States of	US	USA		
28	Shooting	tokyo-2020	Skeet women	Women	BRONZE	Athlete		https://olympics.co	Meng WEI	People's Republ	CN	CHN		

olympic_medals

ภาพที่ 3.4 ชุดข้อมูล olympic_medals

ขั้นตอนการแปลงข้อมูลที่ได้รวบรวมมาและเลือกไว้ ให้อยู่ในรูปแบบที่พร้อมสำหรับนำไปวิเคราะห์ในขั้นตอนต่อไปได้ โดยการทำให้เป็นข้อมูลที่ถูกต้อง (Data cleaning) มักใช้เวลาค่านข้างมาก ระบบการรับข้อมูลป้อนเข้าสู่ระบบที่ทันสมัยในปัจจุบันจะลดการคีย์ข้อมูลจากคนให้น้อยที่สุด แต่จะใช้วิธีการสแกน การติ๊กเลือก เพื่อลดความผิดพลาดให้น้อยที่สุด เพราะขั้นตอนใช้เวลามาก การลดข้อผิดพลาดของข้อมูลได้มากเท่าใด ก็จะยิ่งมีประสิทธิภาพมากขึ้นเท่านั้น

3.1.3 การเตรียมข้อมูล (Data Preparation)

ในขั้นตอนนี้ผู้วิเคราะห์ข้อมูลได้ทำการตรวจข้อมูลและเช็คชุดข้อมูลที่ได้เลือกมา แล้วทำการตัดสินใจลบข้อมูลบางส่วนที่ไม่ได้ใช้ออก เพื่อให้ชุดข้อมูลสามารถนำไปใช้ได้มีประสิทธิภาพมากขึ้น โดยได้ทำการลบข้อมูลในส่วนของ participant_title และ athlete_url ออกไป

ในขั้นตอนต่อมาผู้วิเคราะห์ข้อมูลได้ทำการคัดเลือกแอตทริบิวต์ที่เหมาะสมออกมาเป็นชุดข้อมูลที่นำไปใช้สำหรับขั้นตอนการทำเหมืองข้อมูลต่อไป ดังนั้นจึงได้ชุดข้อมูลเหมาะสมดังนี้

จากชุดข้อมูลที่คัดเลือกมาจาก olympic_medals

A	B	C	D	E	F
Sports	Game	Gender	Team_type	Medal	Country_code
Canoe Sprint	tokyo-2020	Women	Athlete	GOLD	USA
Canoe Sprint	tokyo-2020	Women	Athlete	SILVER	CAN
Canoe Sprint	tokyo-2020	Women	Athlete	BRONZE	UKR
Canoe Sprint	tokyo-2020	Men	GameTeam	GOLD	CUB
Canoe Sprint	tokyo-2020	Men	GameTeam	GOLD	CUB
Canoe Sprint	tokyo-2020	Men	GameTeam	SILVER	CHN
Canoe Sprint	tokyo-2020	Men	GameTeam	SILVER	CHN
Canoe Sprint	tokyo-2020	Men	GameTeam	BRONZE	GER
Canoe Sprint	tokyo-2020	Men	GameTeam	BRONZE	GER
Canoe Sprint	tokyo-2020	Men	GameTeam	GOLD	GER
Canoe Sprint	tokyo-2020	Men	GameTeam	SILVER	ESP
Canoe Sprint	tokyo-2020	Men	GameTeam	BRONZE	SVK
Canoe Sprint	tokyo-2020	Men	GameTeam	GOLD	AUS
Canoe Sprint	tokyo-2020	Men	GameTeam	GOLD	AUS
Canoe Sprint	tokyo-2020	Men	GameTeam	SILVER	GER
Canoe Sprint	tokyo-2020	Men	GameTeam	SILVER	GER
Canoe Sprint	tokyo-2020	Men	GameTeam	BRONZE	CZE
Canoe Sprint	tokyo-2020	Men	GameTeam	BRONZE	CZE
Canoe Sprint	tokyo-2020	Women	Athlete	BRONZE	DEN
Canoe Sprint	tokyo-2020	Women	Athlete	GOLD	NZL
Canoe Sprint	tokyo-2020	Women	Athlete	SILVER	ESP
Canoe Sprint	tokyo-2020	Women	GameTeam	GOLD	CHN
Canoe Sprint	tokyo-2020	Women	GameTeam	GOLD	CHN
Canoe Sprint	tokyo-2020	Women	GameTeam	SILVER	UKR
Canoe Sprint	tokyo-2020	Women	GameTeam	SILVER	UKR
Canoe Sprint	tokyo-2020	Women	GameTeam	BRONZE	CAN
Canoe Sprint	tokyo-2020	Women	GameTeam	BRONZE	CAN

ภาพที่ 3.5 ชุดข้อมูลที่คัดเลือกมาจาก olympic_medal

จากรูปภาพที่ชุดข้อมูลที่คัดเลือกมาประกอบด้วย 5 แอตทริบิวต์ ได้แก่

- Sports ประกอบด้วย Canoe Sprint, Diving, Fencing, Cycling Track
- Game ประกอบด้วย tokyo-2020, rio-2016
- Gender ประกอบด้วย Women, Men
- Team_type ประกอบด้วย GameTeam, Athlete
- Medal ประกอบด้วย GOLD, SILVER, BRONZE
- County_Code ประกอบด้วย CHN, FRA, GBR, GER, HUN, ITA, RUS, UKR, USA

จากชุดข้อมูลที่ได้คัดเลือกมานั้นจะมีแอตทริบิวต์บางส่วนที่ต้องปรับแก้ไขคือ County_Code (ROC) เนื่องจาก ROC คือประเทศ รัสเซีย (RUS) เนื่องจากการแข่งขันในปี 2021 ประเทศ รัสเซีย (RUS) ได้ทำผิดกฎของ WADA (World anti-doping Agency) องค์กรอิสระเพื่อต่อต้านการใช้สารต้องห้ามทางกีฬา จึงได้ทำการตั้งคณะกรรมการโอลิมปิกรัสเซีย Russian Olympic Committee (ROC) ขึ้นมาเพื่อใช้ในการแข่งขันโอลิมปิกเกมส์เนื่องไม่สามารถใช้ชื่อประเทศ รัสเซีย (RUS) เข้าแข่งขันได้ ดังนั้นทางผู้เคราะห์จึงแก้ไขแอตทริบิวต์ County_Code (ROC) ให้เป็นชื่อประเทศ รัสเซีย (RUS) เพื่อให้ข้อมูลการวิเคราะห์และข้อมูลการพยากรณ์ออกมาคลาดเคลื่อนน้อยที่สุดและมีความแม่นยำ

3.1.4 การสร้างโมเดล (Modeling)

ผู้วิเคราะห์ข้อมูลเริ่มกระบวนการทำเหมืองข้อมูลในแบบ Data Classification เพื่อใช้ทำนายแนวโน้มการเกิดขึ้นของปัจจัยต่าง ๆ จากเทคนิคการจำแนกประเภทข้อมูลด้วยการสร้างโมเดล Decision Tree และการจัดหมวดหมู่โดยใช้หลักความน่าจะเป็น Naive Bayes ในการเปรียบแทบความแม่นยำ เพื่อจัดกลุ่มข้อมูลที่มีความสัมพันธ์กัน โดยใช้ชุดข้อมูลที่ได้คัดเลือกมาข้างต้น

ชุดข้อมูลที่คัดเลือกมาใช้ ประกอบด้วย 5 แอตทริบิวต์ ได้แก่

- Sports ประกอบด้วย Canoe Sprint, Diving, Fencing, Cycling Track
- Game ประกอบด้วย tokyo-2020, rio-2016
- Gender ประกอบด้วย Women, Men
- Team_type ประกอบด้วย GameTeam, Athlete
- Medal ประกอบด้วย GOLD, SILVER, BRONZE
- County_Code ประกอบด้วย CHN, FRA, GBR, GER, HUN, ITA, RUS, UKR, USA

การสร้างโมเดล Decision Tree จะทำการคัดเลือกแอตทริบิวต์ที่มีความสัมพันธ์กับคลาสมากที่สุดขึ้นมาเป็นโหนดของ Tree การหาความสัมพันธ์ของแอตทริบิวต์นี้จะใช้ตัววัด ที่เรียกว่า Information gain (IG) มีการคำนวณสมการดังนี้

IG (parent, child) = entropy(parent) – [p(c1) × entropy(c1) + p(c2) × entropy(c2) + ...]

โดยที่ entropy(c1) = -p(c1) log2 p(c1) และ p(c1) คือ ค่าความน่าจะเป็นของ c1

1) ผลสรุปที่ได้ จากการหาค่าของ IG ที่มากที่สุดเพื่อจะนำมาเป็น Root ที่ 1 จากผลคำนวณทั้งหมดของค่าแอตทริบิวต์ จึงได้ผลดังนี้

การคำนวณค่าแอตทริบิวต์ที่มีค่า IG มากที่สุดมาเป็น Root ของ Decision tree กับจำนวนข้อมูลทั้งหมดโดยใช้ผลลัพธ์เป็น County_code ดังนี้

1.1) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= Canoe Sprint) x Entropy(ผล= Canoe Sprint) + p(ผล= Diving) x Entropy(ผล= Diving) + p(ผล= Fencing) x Entropy(ผล= Fencing) + p(ผล= Cycling Track) x Entropy(ผล= Cycling Track)] ผลลัพธ์การคำนวณค่า IG = 0.63

1.2) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Game จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= tokyo–2020) x Entropy(ผล= tokyo–2020) + p(ผล= rio–2016) x Entropy(ผล= rio–2016)] ผลลัพธ์การคำนวณค่า IG = 0.16

1.3) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Gender จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= women) x Entropy(ผล= women) + p(ผล= men) x Entropy(ผล= men)] ผลลัพธ์การคำนวณค่า IG = 0.06

1.4) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Medal จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gold) x Entropy(ผล= gold) + p(ผล=silver) x Entropy(ผล= silver) + p(ผล= bronze) x Entropy(ผล= bronze)] ผลลัพธ์การคำนวณค่า IG = 0.14

1.5) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Team_Type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= Athlete) x Entropy(ผล= Athlete)] ผลลัพธ์การคำนวณค่า IG = 0.04

จากผลสรุปการคำนวนของค่า IG ของแต่ละแอตทริบิวต์ในการค่ามากที่สุดในการนำมาเป็น Root ที่ 1 โดยค่า IG ที่มากที่สุดคือค่าของแอตทริบิวต์ Sport โดยมีค่า IG อยู่ที่ 0.63 จึงนำมาเป็น Root ที่ 1 จากนั้นจึงหาค่าแอตทริบิวต์ของ Root ที่ 2 ต่อไป

2) ผลสรุปที่ได้ จากการหาค่าของ IG ที่มากที่สุดเพื่อจะนำมาเป็น Root ที่ 2 จากผลคำนวณทั้งหมดของค่าแอตทริบิวต์จึงได้ผลดังนี้

การคำนวณค่าแอตทริบิวต์ที่มีค่า IG มากที่สุดมาเป็น Root ของ Decision tree กับจำนวนข้อมูลทั้งหมดโดยใช้ผลลัพธ์เป็น County_code ดังนี้

2.1) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Canoe Sprint) & Medal จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gold) x Entropy(ผล= gold) + p(ผล= silver) x Entropy(ผล= silver) + p(ผล= bronze) x Entropy(ผล= bronze)] ผลลัพธ์การคำนวณค่า IG = 0.32

2.2) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Diving) & Medal จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gold) x Entropy(ผล= gold) + p(ผล= silver) x Entropy(ผล= silver) + p(ผล= bronze) x Entropy(ผล= bronze)] ผลลัพธ์การคำนวณค่า IG = 0.61

2.3) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Fencing) & Gender จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= women) x Entropy(ผล= women) + p(ผล= men) x Entropy(ผล= men)] ผลลัพธ์การคำนวณค่า IG = 0.19

2.4) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Cycling Track) & Gender จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= women) x Entropy(ผล= women) + p(ผล= men) x Entropy(ผล= men)] ผลลัพธ์การคำนวณค่า IG = 0.48

จากผลสรุปการคำนวนของค่า IG ของแต่ละแอตทริบิวต์ในการหาค่ามากที่สุดในการนำมาเป็น Root ที่ 2 โดยที่จะแบ่งค่าแอตทริบิวต์ Sports ออกเป็นแต่ละประเภทของกีฬาจากนั้นจึงนำค่า IG ที่มากที่สุดของแต่ละแอตทริบิวต์ไปเป็น Root ที่ 2 โดยที่จะแบ่งตามประเภทของกีฬา โดยผลของค่า IG ของแอตทริบิวต์นำแยกตามประเภทกีฬาได้ดังนี้ กีฬาประเภท Canoe Sprint ค่า IG ของแอตทริบิวต์ที่มากที่สุด Medal มีค่า IG = 0.32 กีฬาประเภท Diving ค่า IG ของแอตทริบิวต์ที่มากที่สุด Medal มีค่า IG = 0.61 กีฬาประเภท Fencing ค่า IG ของแอตทริบิวต์ที่มากที่สุด Gender มีค่า IG = 0.19 กีฬาประเภท Cycling Track ค่า IG ของแอตทริบิวต์ที่มากที่สุด Gender มีค่า IG = 0.48 โดยค่าที่ได้จึงนำมาเป็น Root ที่ 2 จากนั้นจึงหาค่าแอตทริบิวต์ของ Root ที่ 3 ต่อไป

3) ผลสรุปที่ได้ จากการหาค่าของ IG ที่มากที่สุดเพื่อจะนำมาเป็น Root ที่ 3 จากผลคำนวณทั้งหมดของค่าแอตทริบิวต์จึงได้ผลดังนี้

การคำนวณค่าแอตทริบิวต์ที่มีค่า IG มากที่สุดมาเป็น Root ของ Decision tree กับจำนวนข้อมูลทั้งหมดโดยใช้ผลลัพธ์เป็น County_code ดังนี้

3.1) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Canoe Sprint) & Medal(GOLD) & Gender จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= women) x Entropy(ผล= women) + p(ผล= men) x Entropy(ผล= men)] ผลลัพธ์การคำนวณค่า IG = 0.66

3.2) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Canoe Sprint) & Medal(SILVER) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.72

3.3) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Canoe Sprint) & Medal(BRONZE) & Gender จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= women) x Entropy(ผล= women) + p(ผล= men) x Entropy(ผล= men)] ผลลัพธ์การคำนวณค่า IG = 0.59

3.4) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Diving) & Medal(SILVER) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.61

3.5) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Diving) & Medal(BRONZE) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.41

3.6) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Fencing) & Gender(Women) & Medal จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gold) x Entropy(ผล= gold) + p(ผล= silver) x Entropy(ผล= silver) + p(ผล= bronze) x Entropy(ผล= bronze)] ผลลัพธ์การคำนวณค่า IG = 0.38

3.7) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Fencing) & Gender(Men) & Medal จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gold) x Entropy(ผล= gold) + p(ผล= silver) x Entropy(ผล= silver) + p(ผล= bronze) x Entropy(ผล= bronze)] ผลลัพธ์การคำนวณค่า IG = 0.35

3.8) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Cycling Track) & Gender(Women) & Medal จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gold) x Entropy(ผล= gold) + p(ผล= silver) x Entropy(ผล= silver) + p(ผล= bronze) x Entropy(ผล= bronze)] ผลลัพธ์การคำนวณค่า IG = 0.40

3.8) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Cycling Track) & Gender(Men) & Medal จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gold) x Entropy(ผล= gold) + p(ผล= silver) x Entropy(ผล= silver) + p(ผล= bronze) x Entropy(ผล= bronze)] ผลลัพธ์การคำนวณค่า IG = 0.57

จากผลสรุปการคำนวนของค่า IG ของแต่ละแอตทริบิวต์ในการหาค่ามากที่สุดในการนำมาเป็น Root ที่ 3 ของ Root ที่ 2 แต่ละ Root ที่แยกออกมาจากแอตทริบิวต์ Sports อีกที่ โดยใน Root ที่ Root ยังมีบางแอตทริบิวต์ที่ยังสามารถแตกกิ่งออกเป็น Root ที่ 4 ได้ดังนั้นจึงคำนวณหาค่า IG ของ Root ที่ 4 ต่อไป

4) ผลสรุปที่ได้ จากการหาค่าของ IG ที่มากที่สุดเพื่อจะนำมาเป็น Root ที่ 4 จากผลคำนวณทั้งหมดของค่าแอตทริบิวต์จึงได้ผลดังนี้

การคำนวณค่าแอตทริบิวต์ที่มีค่า IG มากที่สุดมาเป็น Root ของ Decision tree กับจำนวนข้อมูลทั้งหมดโดยใช้ผลลัพธ์เป็น County_code ดังนี้

4.1) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Canoe Sprint) & Medal(GOLD) & Gender(Women) & Game จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= tokyo–2020) x Entropy(ผล= tokyo–2020) + p(ผล= rio–2016) x Entropy(ผล= rio–2016)] ผลลัพธ์การคำนวณค่า IG = 0.55

4.2) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Canoe Sprint) & Medal(GOLD) & Gender(Men) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.62

4.3) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Canoe Sprint) & Medal(SILVER) & Team_type(Gameteam) & Gender จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= women) x Entropy(ผล= women) + p(ผล= men) x Entropy(ผล= men)] ผลลัพธ์การคำนวณค่า IG = 0.56

4.4) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Canoe Sprint) & Medal(SILVER) & Team_type(Athlete) & Gender จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= tokyo–2020) x Entropy(ผล= tokyo–2020) + p(ผล= rio–2016) x Entropy(ผล= rio–2016)] ผลลัพธ์การคำนวณค่า IG = 0.72

4.5) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Canoe Sprint) & Medal(BRONZE) & Gender(Women) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.92

4.6) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Canoe Sprint) & Medal(BRONZE) & Gender(Men) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.66

4.7) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Diving) & Medal(SILVER) & Team_type(Gameteam) & Gender จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= women) x Entropy(ผล= women) + p(ผล= men) x Entropy(ผล= men)] ผลลัพธ์การคำนวณค่า IG = 0.38

4.8) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Diving) & Medal(BRONZE) & Team_type(Gameteam) & Game จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= tokyo–2020) x Entropy(ผล= tokyo–2020) + p(ผล= rio–2016) x Entropy(ผล= rio–2016)] ผลลัพธ์การคำนวณค่า IG = 0.97

4.9) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Diving) & Medal(BRONZE) & Team_type(Athlete) & Game จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= tokyo–2020) x Entropy(ผล= tokyo–2020) + p(ผล= rio–2016) x Entropy(ผล= rio–2016)] ผลลัพธ์การคำนวณค่า IG = 0.57

4.10) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Fencing) & Gender(Women) & Medal(GOLD) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.91

4.11) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Fencing) & Gender(Women) & Medal(SILVER) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.99

4.12) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Fencing) & Gender(Women) & Medal(BRONZE) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.75

4.13) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Fencing) & Gender(Men) & Medal(GOLD) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.88

4.14) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Fencing) & Gender(Men) & Medal(SILVER) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.63

4.15) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Fencing) & Gender(Men) & Medal(BRONZE) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.99

4.16) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Cycling Track) & Gender(Women) & Medal(GOLD) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 0.39

4.17) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Cycling Track) & Gender(Women) & Medal(SILVER) & Game จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= tokyo–2020) x Entropy(ผล= tokyo–2020) + p(ผล= rio–2016) x Entropy(ผล= rio–2016)] ผลลัพธ์การคำนวณค่า IG = 0.62

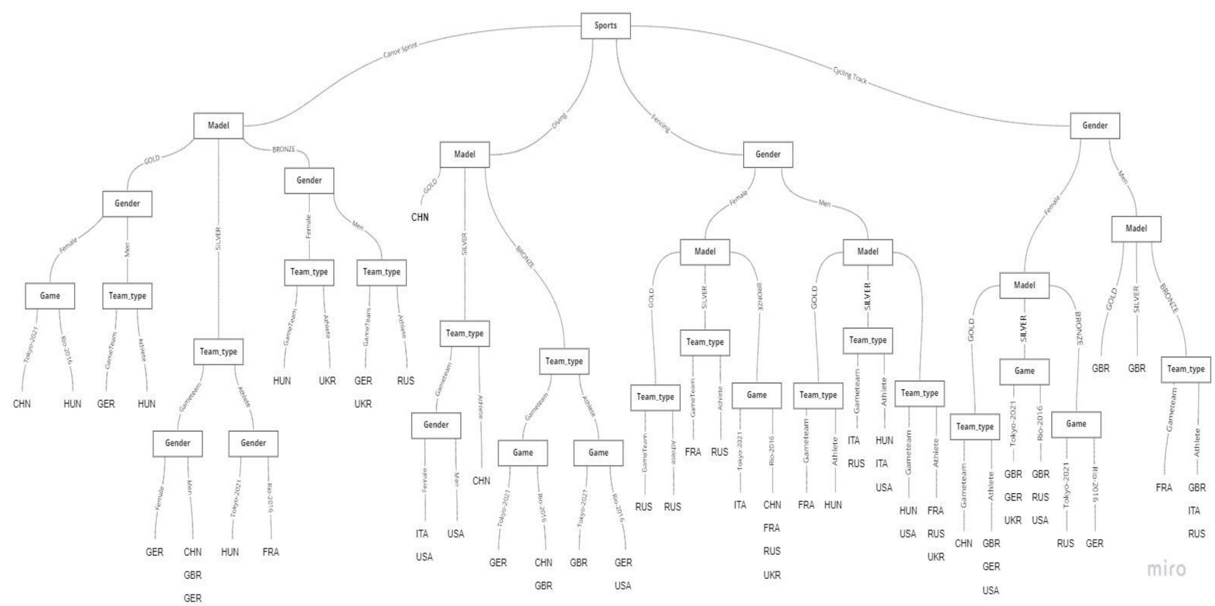
4.18) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Cycling Track) & Gender(Women) & Medal(BRONZE) & Game จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= tokyo–2020) x Entropy(ผล= tokyo–2020) + p(ผล= rio–2016) x Entropy(ผล= rio–2016)] ผลลัพธ์การคำนวณค่า IG = 1.00

4.19) สรุปการคำนวณค่า IG ของแอตทริบิวต์ Sports(Cycling Track) & Gender(Men) & Medal(BRONZE) & Team_type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

IG (parent, child) = Entropy (parent) –[p(ผล= gameteam) x Entropy(ผล= gameteam) + p(ผล= athlete) x Entropy(ผล= athlete)] ผลลัพธ์การคำนวณค่า IG = 1.00

จากการคำนวณค่า IG ของแอตทริบิวต์ Sports และ Game และ Gender และ Team_type และ Madel พบว่าข้อมูลในแต่ละโหนดมีคลาสคำตอบเดียวกันแล้ว คือ ผลลัพธ์เป็น County_code หลังจากนี้จึงนำแอตทริบิวต์ที่ได้คำนวนมานำไปทำแผนภาพ Decision Tree ต่อไป



ภาพที่ 3.6 รูปแบบโมเดล Graph Decision Tree

ในการตรวจสอบและเปรียบเทียบโมเดลของข้อมูลในการทำเหมืองข้อมูลว่ามีความแตกต่างของความแม่นยำหรือไม่ ผู้วิเคราะห์ข้อมูลจึงได้เพิ่มโมเดลอีกหนึ่งโมเดล Naive Bayes ในการใช้เพื่อเปรียบเทียบความแม่นยำของกระบวนการทำเหมืองข้อมูลในแบบ Data Classification และโมเดลที่ได้คำนวนไป โดยจะใช้ชุดข้อมูลที่คัดเลือกมาชุดข้อมูลเดียวกันในการคำนวน ดังนี้

A	B	C	D	E	F
Sports	Game	Gender	Team_type	Medal	Country_code
Canoe Sprint	tokyo-2020	Women	Athlete	GOLD	USA
Canoe Sprint	tokyo-2020	Women	Athlete	SILVER	CAN
Canoe Sprint	tokyo-2020	Women	Athlete	BRONZE	UKR
Canoe Sprint	tokyo-2020	Men	GameTeam	GOLD	CUB
Canoe Sprint	tokyo-2020	Men	GameTeam	GOLD	CUB
Canoe Sprint	tokyo-2020	Men	GameTeam	SILVER	CHN
Canoe Sprint	tokyo-2020	Men	GameTeam	SILVER	CHN
Canoe Sprint	tokyo-2020	Men	GameTeam	BRONZE	GER
Canoe Sprint	tokyo-2020	Men	GameTeam	BRONZE	GER
Canoe Sprint	tokyo-2020	Men	GameTeam	GOLD	GER
Canoe Sprint	tokyo-2020	Men	GameTeam	SILVER	ESP
Canoe Sprint	tokyo-2020	Men	GameTeam	BRONZE	SVK
Canoe Sprint	tokyo-2020	Men	GameTeam	GOLD	AUS
Canoe Sprint	tokyo-2020	Men	GameTeam	GOLD	AUS
Canoe Sprint	tokyo-2020	Men	GameTeam	SILVER	GER
Canoe Sprint	tokyo-2020	Men	GameTeam	SILVER	GER
Canoe Sprint	tokyo-2020	Men	GameTeam	BRONZE	CZE
Canoe Sprint	tokyo-2020	Men	GameTeam	BRONZE	CZE
Canoe Sprint	tokyo-2020	Women	Athlete	BRONZE	DEN
Canoe Sprint	tokyo-2020	Women	Athlete	GOLD	NZL
Canoe Sprint	tokyo-2020	Women	Athlete	SILVER	ESP
Canoe Sprint	tokyo-2020	Women	GameTeam	GOLD	CHN
Canoe Sprint	tokyo-2020	Women	GameTeam	GOLD	CHN
Canoe Sprint	tokyo-2020	Women	GameTeam	SILVER	UKR
Canoe Sprint	tokyo-2020	Women	GameTeam	SILVER	UKR
Canoe Sprint	tokyo-2020	Women	GameTeam	BRONZE	CAN
Canoe Sprint	tokyo-2020	Women	GameTeam	BRONZE	CAN

ภาพที่ 3.7 ชุดข้อมูลที่คัดเลือกมากจาก olympic_medal

ข้อมูลที่คัดเลือกมากจาก olympic_medals ชุดข้อมูลที่คัดเลือกมาคำนวน ประกอบด้วย 5 แอตทริบิวต์ ได้แก่

- Sports ประกอบด้วย Canoe Sprint, Diving, Fencing, Cycling Track
- Game ประกอบด้วย tokyo-2020, rio-2016
- Gender ประกอบด้วย Women, Men
- Team_type ประกอบด้วย GameTeam, Athlete
- County_Code ประกอบด้วย CHN, FRA, GBR, GER, HUN, ITA, RUS, UKR, USA
- Medal ประกอบด้วย GOLD, SILVER, BRONZE

โมเดลที่ผู้วิเคราะห์ได้เลือกโมเดลในลำดับที่ 2 นี่ คือโมเดล Naive Bayes โดยโมเดลนี้จะทำให้การจำแนกข้อมูลต่าง ๆ ทำได้ดีขึ้นโดยใช้พื้นฐานของ Bayes' Theorem และ สมมติฐานที่ให้การเกิดของเหตุการณ์ต่างๆ โดยมีสูตรคำนวณ ดังนี้

$$P(c \mid x) = \frac{P(x \mid c)P(c)}{P(x)}$$

$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

อธิบายสมการ แทนตัวแปร 3 ตัว คือ

c คือ Class

x คือ Attribute

P คือ Probability

P(c I x) คือ ความน่าจะเป็นที่ข้อมูลที่มีแอตทริบิวต์เป็น x จะมีคลาส C

P(x I c) คือ ความน่าจะเป็นที่ข้อมูลที่มีคลาส C และมีแอตทริบิวต์ x

P(c) คือ จำนวน Class ที่อาจจะเกิดขึ้นจำนวนClass ทั้งหมดหรือความน่าจะเป็นของ Class

P(x) คือ จำนวน Attribute ทั้งหมด

การทำนายผลของการจับคู่ของแต่ละ Attribute จะได้ผลรับออกมาเป็นปะเทศใดโดยให้ค่าผลลัพธ์จาก Class ทั้ง 10 ค่า คือ CHN, FRA, GBR, GER. HUN. ITA, ROS, UKR, USA มีคุณสมบัติ Attribute ดังนี้

กำหนดค่าตัวแปรในสมการ

c คือ Class (CHN, FRA, GBR, GER. HUN. ITA, ROS, UKR, USA)

x คือ Attribute (Sports, Game, Gender, Medal, Team_Type)

สิ่งที่ต้องการคือ อยากทำนายผลว่าหากจับคู่แต่ละ Attribute ดังกล่าวค่าสูงสุดจะแสดงผลเป็นประเทศใด

การหาค่าของ P(c) หรือ Class ค่าที่ได้ ดังนี้

P(CHN) = 38/194 = 0.196

P(FRA) = 12/194 = 0.062

P(GBR) = 31/194 = 0.160

P(GER) = 25/194 = 0.129

P(HUN) = 18/194 = 0.093

P(ITA) = 15/194 = 0.077

P(RUS) = 24/194 = 0.124

P(UKR) = 10/194 = 0.052

P(USA) = 21/194 = 0.108

การหาค่าของ P(x | c) หรือค่าของแต่ละ Attribute ค่าที่ได้ ดังนี้

ตารางที่ 3.1 ค่าของAttribute (Sports)

P(Canoe Sprint / CHN)	5 / 38		P(Canoe Sprint / FRA)	1 / 12
P(Diving / CHN)	27 / 38		P(Diving / FRA)	0 / 12
P(Fencing / CHN)	3 / 38		P(Fencing / FRA)	8 / 12
P(Cycling Track / CHN)	3 / 38		P(Cycling Track / FRA)	3 / 12
P(Canoe Sprint / GBR)	4 / 31		P(Canoe Sprint / GER)	15 / 25
P(Diving / GBR)	9 / 31		P(Diving / GER)	5 / 25
P(Fencing / GBR)	0 / 31		P(Fencing / GER)	0 / 25
P(Cycling Track / GBR)	18 / 31		P(Cycling Track / GER)	5 / 25
P(Canoe Sprint / HUN)	11 / 18		P(Canoe Sprint / ITA)	1 / 15
P(Diving / HUN)	0 / 18		P(Diving / ITA)	2 / 15
P(Fencing / HUN)	7 / 18		P(Fencing / ITA)	9 / 15
P(Cycling Track / HUN)	0 / 18		P(Cycling Track / ITA)	3 / 15

ตารางที่ 3.1 ค่าของAttribute (Sports) (ต่อ)

P(Canoe Sprint / RUS)	2 / 24		P(Canoe Sprint / UKR)	6 / 10
P(Diving / RUS)	2 / 24		P(Diving / UKR)	0 / 10
P(Fencing / RUS)	15 / 24		P(Fencing / UKR)	3 / 10
P(Cycling Track / RUS)	5 / 24		P(Cycling Track / UKR)	1 / 10
P(Canoe Sprint / USA)	1 / 21			
P(Diving / USA)	10 / 21			
P(Fencing / USA)	6 / 21			
P(Cycling Track / USA)	4 / 21			

ตารางที่ 3.2 ค่าของAttribute (Game)

P(Tokyo-2020 / CHN)	22 / 38		P(Tokyo-2020 / FRA)	7 / 12
P(Rio-2016 / CHN)	16 / 38		P(Rio-2016 / FRA)	5 / 12
P(Tokyo-2020 / GBR)	12 / 31		P(Tokyo-2020 / GER)	11 / 25
P(Rio-2016 / GBR)	19 / 31		P(Rio-2016 / GER)	14 / 25
P(Tokyo-2020 / HUN)	10 / 18		P(Tokyo-2020 / ITA)	8 / 15
P(Rio-2016 / HUN)	8 / 18		P(Rio-2016 / ITA)	7 / 15
P(Tokyo-2020 / RUS)	12 / 24		P(Tokyo-2020 / UKR)	5 / 10
P(Rio-2016 / RUS)	12 / 24		P(Rio-2016 / UKR)	5 / 10
P(Tokyo-2020 / USA)	10 / 21			
P(Rio-2016 / USA)	11 / 21			

ตารางที่ 3.3 ค่าของAttribute (Gender)

P(Men / CHN)	17 / 38	P(Men / FRA)	9 / 12
P(Female / CHN)	21 / 38	P(Female / FRA)	3 / 12
P(Men / GBR)	24 / 31	P(Men / GER)	15 / 25
P(Female / GBR)	7 / 31	P(Female / GER)	10 / 25
P(Men / HUN)	9 / 18	P(Men / ITA)	9 / 15
P(Female / HUN)	9 / 18	P(Female / ITA)	6 / 15
P(Men / RUS)	9 / 24	P(Men / UKR)	4 / 10
P(Female / RUS)	15 / 24	P(Female / UKR)	6 / 10
P(Men / USA)	11 / 21		
P(Female / USA)	10 / 21		

ตารางที่ 3.4 ค่าของAttribute (Medal)

P(GOLD / CHN)	23 / 38	P(GOLD / FRA)	3 / 12
P(SILVER / CHN)	12 / 38	P(SILVER / FRA)	4 / 12
P(BRONZE / CHN)	3 / 38	P(BRONZE / FRA)	5 / 12
P(GOLD / GBR)	14 / 31	P(GOLD / GER)	9 / 25
P(SILVER / GBR)	10 / 31	P(SILVER / GER)	6 / 25
P(BRONZE / GBR)	7 / 31	P(BRONZE / GER)	10 / 25
P(GOLD / HUN)	10 / 18	P(GOLD / ITA)	3 / 15
P(SILVER / HUN)	4 / 18	P(SILVER / ITA)	9 / 15
P(BRONZE / HUN)	4 / 18	P(BRONZE / ITA)	3 / 15

ตารางที่ 3.4 ค่าของAttribute (Medal) (ต่อ)

P(GOLD / RUS)	7 / 24		P(GOLD / UKR)	1 / 10
P(SILVER / RUS)	7 / 24		P(SILVER / UKR)	4 / 10
P(BRONZE / RUS)	10 / 24		P(BRONZE / UKR)	5 / 10
P(GOLD / USA)	3 / 21			
P(SILVER / USA)	12 / 21			
P(BRONZE / USA)	6 / 21			

ตารางที่ 3.5 ค่าของAttribute (Team_Type)

P(Athlete / CHN)	16 / 38		P(Athlete / FRA)	4 / 12
P(GameTeam / CHN)	22 / 38		P(GameTeam / FRA)	8 / 12
P(Athlete / GBR)	16 / 31		P(Athlete / GER)	4 / 25
P(GameTeam / GBR)	15 / 31		P(GameTeam / GER)	21 / 25
P(Athlete / HUN)	10 / 18		P(Athlete / ITA)	8 / 15
P(GameTeam / HUN)	8 / 18		P(GameTeam / ITA)	7 / 15
P(Athlete / RUS)	11 / 24		P(Athlete / UKR)	5 / 10
P(GameTeam / RUS)	13 / 24		P(GameTeam / UKR)	5 / 10
P(Athlete / USA)	8 / 21			
P(GameTeam / USA)	13 / 21			

จากตารางที่ได้ทำการรวบร่วมวิเคราะห์ค่าของ Attribute แต่ละ Attribute ในชุดข้อมูล ดังนั้นจึงนำข้อมูลที่ได้นำไปแปลงเป็นค่าเพื่อนำไปใช้ในการคำนวณสมการ

ตารางที่ 3.6 การแปลงค่าของAttribute (Sports)

P(Canoe Sprint / CHN)	0.132		P(Canoe Sprint / FRA)	0.084
P(Diving / CHN)	0.710		P(Diving / FRA)	0.000
P(Fencing / CHN)	0.079		P(Fencing / FRA)	0.666
P(Cycling Track / CHN)	0.079		P(Cycling Track / FRA)	0.250
P(Canoe Sprint / GBR)	0.129		P(Canoe Sprint / GER)	0.600
P(Diving / GBR)	0.290		P(Diving / GER)	0.200
P(Fencing / GBR)	0.000		P(Fencing / GER)	0.000
P(Cycling Track / GBR)	0.580		P(Cycling Track / GER)	0.200
P(Canoe Sprint / HUN)	0.611		P(Canoe Sprint / ITA)	0.067
P(Diving / HUN)	0.000		P(Diving / ITA)	0.133
P(Fencing / HUN)	0.389		P(Fencing / ITA)	0.599
P(Cycling Track / HUN)	0.000		P(Cycling Track / ITA)	0.200
P(Canoe Sprint / RUS)	0.083		P(Canoe Sprint / UKR)	0.599
P(Diving / RUS)	0.083		P(Diving / UKR)	0.001
P(Fencing / RUS)	0.625		P(Fencing / UKR)	0.300
P(Cycling Track / RUS)	0.208		P(Cycling Track / UKR)	0.100
P(Canoe Sprint / USA)	0.048			
P(Diving / USA)	0.476			
P(Fencing / USA)	0.286			
P(Cycling Track / USA)	0.190			

ตารางที่ 3.7 การแปลงค่าของAttribute (Game)

P(Tokyo-2020 / CHN)	0.579		P(Tokyo-2020 / FRA)	0.583
P(Rio-2016 / CHN)	0.421		P(Rio-2016 / FRA)	0.417
P(Tokyo-2020 / GBR)	0.387		P(Tokyo-2020 / GER)	0.440
P(Rio-2016 / GBR)	0.613		P(Rio-2016 / GER)	0.560
P(Tokyo-2020 / HUN)	0.555		P(Tokyo-2020 / ITA)	0.533
P(Rio-2016 / HUN)	0.444		P(Rio-2016 / ITA)	0.467
P(Tokyo-2020 / RUS)	0.500		P(Tokyo-2020 / UKR)	0.500
P(Rio-2016 / RUS)	0.500		P(Rio-2016 / UKR)	0.500
P(Tokyo-2020 / USA)	0.476			
P(Rio-2016 / USA)	0.524			

ตารางที่ 3.8 การแปลงค่าของAttribute (Gender)

P(Men / CHN)	0.447		P(Men / FRA)	0.749
P(Female / CHN)	0.553		P(Female / FRA)	0.250
P(Men / GBR)	0.774		P(Men / GER)	0.600
P(Female / GBR)	0.226		P(Female / GER)	0.400
P(Men / HUN)	0.500		P(Men / ITA)	0.600
P(Female / HUN)	0.500		P(Female / ITA)	0.400
P(Men / RUS)	0.375		P(Men / UKR)	0.400
P(Female / RUS)	0.625		P(Female / UKR)	0.600
P(Men / USA)	0.524			
P(Female / USA)	0.476			

ตารางที่ 3.9 การแปลงค่าของAttribute (Medal)

P(GOLD / CHN)	0.605		P(GOLD / FRA)	0.250
P(SILVER / CHN)	0.316		P(SILVER / FRA)	0.333
P(BRONZE / CHN)	0.079		P(BRONZE / FRA)	0.416
P(GOLD / GBR)	0.451		P(GOLD / GER)	0.360
P(SILVER / GBR)	0.323		P(SILVER / GER)	0.240
P(BRONZE / GBR)	0.226		P(BRONZE / GER)	0.400
P(GOLD / HUN)	0.555		P(GOLD / ITA)	0.200
P(SILVER / HUN)	0.222		P(SILVER / ITA)	0.600
P(BRONZE / HUN)	0.222		P(BRONZE / ITA)	0.200
P(GOLD / RUS)	0.292		P(GOLD / UKR)	0.100
P(SILVER / RUS)	0.292		P(SILVER / UKR)	0.400
P(BRONZE / RUS)	0.417		P(BRONZE / UKR)	0.499
P(GOLD / USA)	0.143			
P(SILVER / USA)	0.571			
P(BRONZE / USA)	0.286			

ตารางที่ 3.10 การแปลงค่าของAttribute (Team_Type)

P(Athlete / CHN)	0.421		P(Athlete / FRA)	0.333
P(GameTeam / CHN)	0.579		P(GameTeam / FRA)	0.666
P(Athlete / GBR)	0.516		P(Athlete / GER)	0.160
P(GameTeam / GBR)	0.484		P(GameTeam / GER)	0.840
P(Athlete / HUN)	0.555		P(Athlete / ITA)	0.533
P(GameTeam / HUN)	0.444		P(GameTeam / ITA)	0.467

ตารางที่ 3.10 การแปลงค่าของAttribute (Team_Type) (ต่อ)

P(Athlete / RUS)	0.458		P(Athlete / UKR)	0.500
P(GameTeam / RUS)	0.542		P(GameTeam / UKR)	0.500
P(Athlete / USA)	0.381			
P(GameTeam / USA)	0.619			

แทนค่าที่ได้ลงในสมการ

$$P(c \mid \mathbf{X}) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

โดยผลของการจับคู่แต่ละ Attribute ดังกล่าวจึงได้ค่าสูงสุด ดังนี้

1) ถ้าจับคู่ Attribute (Canoe Sprint x Rio-2016 x Men x BRONZE x GameTeam) ดังกล่าวจึงได้ค่าสูงสุดแสดงออกมาเป็นประเทศ Germany มีค่าอยู่ที่ 0.008738

2) ถ้าจับคู่ Attribute (Diving x Tokyo-2020 x Female x GOLD x GameTeam) ดังกล่าวจึงได้ค่าสูงสุดแสดงออกมาเป็นประเทศ China มีค่าอยู่ที่ 0.015608

3) ถ้าจับคู่ Attribute (Fencing x Tokyo-2020 x Men x SILVER x Athlete) ดังกล่าวจึงได้ค่าสูงสุดแสดงออกมาเป็นประเทศ Italy มีค่าอยู่ที่ 0.008850

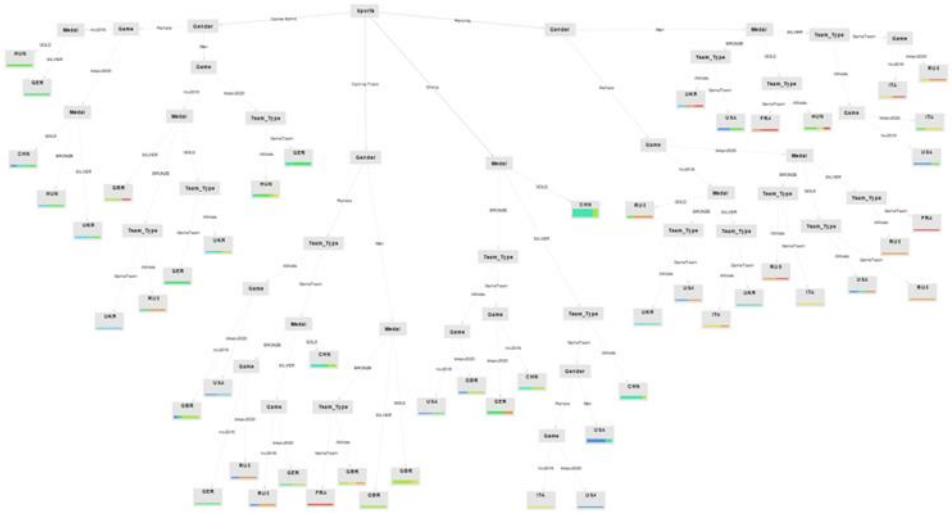
4) ถ้าจับคู่ Attribute (Cycling Track x Tokyo-2020 x Men x BRONZE x Athlete) ดังกล่าวจึงได้ค่าสูงสุดแสดงออกมาเป็นประเทศ United Kingdom of Great Britain and Northern Ireland มีค่าอยู่ที่ 0.020259

จากข้อมูลที่ได้ผ่านการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining จากการสร้างโมเดลทั้ง 2 แบบ Decision Tree และ Naive Bayes ขั้นตอนก่อนการนำผลลัพธ์ที่ได้ไปใช้งาน ด้วยการวัดประสิทธิผลของผลลัพธ์ที่ได้ตั้งไว้ในขั้นตอนแรก ว่าโมเดลทั้ง 2 แบบ Decision Tree และ Naive Bayes ว่ามีความน่าเชื่อถือมากน้อยเพียงใด ด้วยการประเมินผลจากโปรแกรม RapidMiner Studio ว่าถูกต้องหรือไม่

3.1.5 การวัดประสิทธิภาพของโมเดล (Evaluation)

โดยจะทำการวัดประสิทธิผลของผลลัพธ์ที่ได้ของโมเดลทั้ง 2 แบบ Decision Tree และ Naive Bayes โดยจะเริ่มจากโมเดล Decision Tree และตามด้วย Naive Bayes เพื่อวัดประสิทธิผลของผลลัพธ์และทำการเปรียบเทียบค่าความแม่นยำจากนั้นจึงเลือกโมเดลที่มีความแม่นยำสูงสุด

1) ผู้วิเคราะข้อมูลได้นำข้อมูลมาทำการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining โดยการใช้โมเดล Decision Tree จากการใช้โปรแกรม RapidMiner Studio ได้ข้อมูลดังนี้



ภาพที่ 3.8 รูปแบบโมเดล Graph Decision Tree ใน RapidMiner

จากการจัดทำ รูปแบบโมเดล Decision Tree ทั้งในรูปแบบเขียนมือและการใช้โปรแกรม Rapid Miner เมื่อนำมาเปรียบเทียบกันพบว่าทั้ง 2 โมเดลนั้นพบว่าได้ผลลัพธ์ที่ใกล้เคียงกัน โดยผลลัพธ์การสร้างโมเดลด้วยเทคนิค Decision Tree ในโปรแกรม RapidMiner สามารถทำนายได้จำนวน 51 กฎ ดังนี้

```
Sports = Canoe Sprint
|   Gender = Female
|   |   Game = rio-2016
|   |   |   Medal = GOLD: HUN {USA=0, UKR=0, CHN=0, GER=0, HUN=4, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   |   Medal = SILVER: GER {USA=0, UKR=0, CHN=0, GER=3, HUN=0, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   Game = tokyo-2020
|   |   |   Medal = BRONZE: HUN {USA=0, UKR=1, CHN=0, GER=0, HUN=2, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   |   Medal = GOLD: CHN {USA=1, UKR=0, CHN=2, GER=0, HUN=1, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   |   Medal = SILVER: UKR {USA=0, UKR=2, CHN=0, GER=0, HUN=1, GBR=0, ITA=0, RUS=0, FRA=0}
|   Gender = Men
|   |   Game = rio-2016
|   |   |   Medal = BRONZE
|   |   |   |   Team_Type = Athlete: RUS {USA=0, UKR=0, CHN=0, GER=1, HUN=0, GBR=0, ITA=0, RUS=2, FRA=0}
|   |   |   |   Team_Type = GameTeam: UKR {USA=0, UKR=2, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   |   Medal = GOLD
|   |   |   |   Team_Type = Athlete: UKR {USA=0, UKR=1, CHN=0, GER=1, HUN=0, GBR=1, ITA=0, RUS=0, FRA=0}
|   |   |   |   Team_Type = GameTeam: GER {USA=0, UKR=0, CHN=0, GER=5, HUN=0, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   |   Medal = SILVER: GBR {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=2, ITA=0, RUS=0, FRA=1}
|   |   Game = tokyo-2020
|   |   |   Team_Type = Athlete: HUN {USA=0, UKR=0, CHN=1, GER=0, HUN=3, GBR=1, ITA=1, RUS=0, FRA=0}
|   |   |   Team_Type = GameTeam: GER {USA=0, UKR=0, CHN=2, GER=5, HUN=0, GBR=0, ITA=0, RUS=0, FRA=0}

Sports = Diving
|   Medal = BRONZE
|   |   Team_Type = Athlete
|   |   |   Game = rio-2016: USA {USA=1, UKR=0, CHN=0, GER=1, HUN=0, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   |   Game = tokyo-2020: GBR {USA=1, UKR=0, CHN=0, GER=0, HUN=0, GBR=2, ITA=0, RUS=0, FRA=0}
|   |   Team_Type = GameTeam
|   |   |   Game = rio-2016: CHN {USA=0, UKR=0, CHN=2, GER=0, HUN=0, GBR=2, ITA=0, RUS=0, FRA=0}
|   |   |   Game = tokyo-2020: GER {USA=0, UKR=0, CHN=0, GER=4, HUN=0, GBR=0, ITA=0, RUS=2, FRA=0}
|   Medal = GOLD: CHN {USA=0, UKR=0, CHN=17, GER=0, HUN=0, GBR=4, ITA=0, RUS=0, FRA=0}
|   Medal = SILVER
|   |   Team_Type = Athlete: CHN {USA=0, UKR=0, CHN=6, GER=0, HUN=0, GBR=1, ITA=0, RUS=0, FRA=0}
|   |   Team_Type = GameTeam
|   |   |   Gender = Female
|   |   |   |   Game = rio-2016: ITA {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=2, RUS=0, FRA=0}
|   |   |   |   Game = tokyo-2020: USA {USA=2, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   |   Gender = Men: USA {USA=6, UKR=0, CHN=2, GER=0, HUN=0, GBR=0, ITA=0, RUS=0, FRA=0}
```

ภาพที่ 3.9 ภาพที่ 2 คำบรรยายลักษณะงาน Decision Tree ใน RapidMiner

```
Sports = Fencing
|   Gender = Female
|   |   Game = rio-2016
|   |   |   Medal = BRONZE
|   |   |   |   Team_Type = Athlete: UKR {USA=0, UKR=1, CHN=1, GER=0, HUN=0, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   |   |   Team_Type = GameTeam: USA {USA=1, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=1, FRA=0}
|   |   |   Medal = GOLD: RUS {USA=0, UKR=0, CHN=0, GER=0, HUN=1, GBR=0, ITA=0, RUS=3, FRA=0}
|   |   |   Medal = SILVER
|   |   |   |   Team_Type = Athlete: ITA {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=2, RUS=1, FRA=0}
|   |   |   |   Team_Type = GameTeam: UKR {USA=0, UKR=1, CHN=1, GER=0, HUN=0, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   Game = tokyo-2020
|   |   |   Medal = BRONZE
|   |   |   |   Team_Type = Athlete: RUS {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=1, FRA=1}
|   |   |   |   Team_Type = GameTeam: ITA {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=2, RUS=0, FRA=0}
|   |   |   Medal = GOLD
|   |   |   |   Team_Type = Athlete: USA {USA=1, UKR=0, CHN=1, GER=0, HUN=0, GBR=0, ITA=0, RUS=1, FRA=0}
|   |   |   |   Team_Type = GameTeam: RUS {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=2, FRA=0}
|   |   |   Medal = SILVER
|   |   |   |   Team_Type = Athlete: RUS {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=2, FRA=0}
|   |   |   |   Team_Type = GameTeam: FRA {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=0, FRA=2}
|   Gender = Men
|   |   Medal = BRONZE
|   |   |   Team_Type = Athlete: UKR {USA=0, UKR=1, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=1, FRA=1}
|   |   |   Team_Type = GameTeam: USA {USA=2, UKR=0, CHN=0, GER=0, HUN=2, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   Medal = GOLD
|   |   |   Team_Type = Athlete: HUN {USA=0, UKR=0, CHN=0, GER=0, HUN=2, GBR=0, ITA=1, RUS=0, FRA=1}
|   |   |   Team_Type = GameTeam: FRA {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=1, FRA=2}
|   |   Medal = SILVER
|   |   |   Team_Type = Athlete
|   |   |   |   Game = rio-2016: USA {USA=2, UKR=0, CHN=0, GER=0, HUN=1, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   |   |   Game = tokyo-2020: ITA {USA=0, UKR=0, CHN=0, GER=0, HUN=1, GBR=0, ITA=2, RUS=0, FRA=0}
|   |   |   Team_Type = GameTeam
|   |   |   |   Game = rio-2016: ITA {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=1, RUS=0, FRA=1}
|   |   |   |   Game = tokyo-2020: RUS {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=1, RUS=2, FRA=0}

Sports = Cycling Track
|   Gender = Female
|   |   Team_Type = Athlete
|   |   |   Game = rio-2016: GBR {USA=1, UKR=0, CHN=0, GER=1, HUN=0, GBR=4, ITA=0, RUS=0, FRA=0}
|   |   |   Game = tokyo-2020: USA {USA=1, UKR=1, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   Team_Type = GameTeam
|   |   |   Medal = BRONZE
|   |   |   |   Game = rio-2016: GER {USA=0, UKR=0, CHN=0, GER=2, HUN=0, GBR=0, ITA=0, RUS=0, FRA=0}
|   |   |   |   Game = tokyo-2020: RUS {USA=1, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=2, FRA=0}
|   |   |   Medal = GOLD: CHN {USA=0, UKR=0, CHN=3, GER=1, HUN=0, GBR=2, ITA=0, RUS=0, FRA=0}
|   |   |   Medal = SILVER
|   |   |   |   Game = rio-2016: RUS {USA=1, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=2, FRA=0}
|   |   |   |   Game = tokyo-2020: GER {USA=0, UKR=0, CHN=0, GER=1, HUN=0, GBR=1, ITA=0, RUS=0, FRA=0}
|   Gender = Men
|   |   Medal = BRONZE
|   |   |   Team_Type = Athlete: GBR {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=1, ITA=1, RUS=1, FRA=0}
|   |   |   Team_Type = GameTeam: FRA {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=0, ITA=0, RUS=0, FRA=3}
|   |   Medal = GOLD: GBR {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=6, ITA=2, RUS=0, FRA=0}
|   |   Medal = SILVER: GBR {USA=0, UKR=0, CHN=0, GER=0, HUN=0, GBR=4, ITA=0, RUS=0, FRA=0}
```

ภาพที่ 3.9 ภาพที่ 2 คำบรรยายลักษณะงาน Decision Tree ใน RapidMiner

จากข้อมูลโมเดล Decision Tree ของ RapidMiner Studio จากขั้นตอนผ่านกระบวนการเทคนิค Data Mining จึงได้ความแม่นยำที่ 26.79%

accuracy: 26.79%

	true USA	true UKR	true CHN	true GER	true HUN	true GBR	true ITA	true RUS	true FRA	class precision
pred. USA	2	1	3	0	1	0	1	0	0	25.00%
pred. UKR	0	1	0	1	1	0	0	0	0	33.33%
pred. CHN	2	0	5	3	1	3	0	2	0	31.25%
pred. GER	0	1	1	1	2	2	0	2	0	11.11%
pred. HUN	1	0	1	0	0	0	0	1	3	0.00%
pred. GBR	0	0	1	1	0	4	0	1	0	57.14%
pred. ITA	1	0	0	0	0	0	2	1	0	50.00%
pred. RUS	0	0	0	1	0	0	1	0	1	0.00%
pred. FRA	0	0	0	0	0	0	0	0	0	0.00%
class recall	33.33%	33.33%	45.45%	14.29%	0.00%	44.44%	50.00%	0.00%	0.00%	

ภาพที่ 3.10 การคำนวณความแม่นยำของโมเดล Decision Tree ใน RapidMiner

2) ผู้วิเคราะข้อมูลได้นำข้อมูลมาทำการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining โดยการใช้โมเดล Naive Bayes จากการใช้โปรแกรม RapidMiner Studio ได้ข้อมูลดังนี้

2.1) การหาค่าของ Class

SimpleDistribution

```
Distribution model for label attribute County_Code


Class USA (0.108)
5 distributions

Class UKR (0.052)
5 distributions

Class CHN (0.196)
5 distributions

Class GER (0.129)
5 distributions

Class HUN (0.093)
5 distributions

Class GBR (0.160)
5 distributions

Class ITA (0.077)
5 distributions

Class RUS (0.124)
5 distributions

Class FRA (0.062)
5 distributions
```

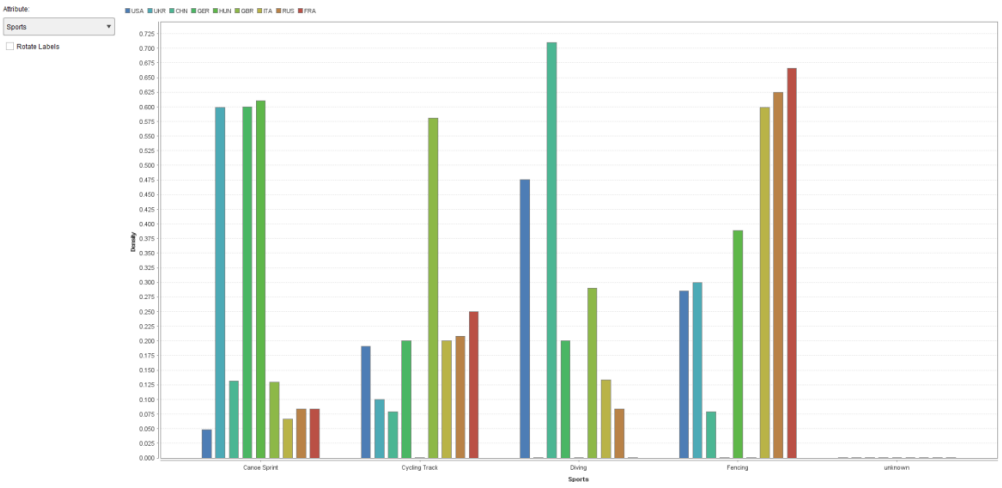
ภาพที่ 3.11 การหาค่าของ Class โดยโปรแกรม Rapid Miner

Attribute	Parameter	USA	UKR	CHN	GER	HUN	GBR	ITA	RUS	FRA
Sports	value=Canoe Sprint	0.048	0.599	0.132	0.600	0.611	0.129	0.067	0.083	0.084
Sports	value=Diving	0.476	0.001	0.710	0.200	0.000	0.290	0.133	0.083	0.000
Sports	value=Fencing	0.286	0.300	0.079	0.000	0.389	0.000	0.599	0.625	0.666
Sports	value=Cycling Track	0.190	0.100	0.079	0.200	0.000	0.580	0.200	0.208	0.250
Sports	value=unknown	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Game	value=tokyo-2020	0.476	0.500	0.579	0.440	0.555	0.387	0.533	0.500	0.583
Game	value=rio-2016	0.524	0.500	0.421	0.560	0.444	0.613	0.467	0.500	0.417
Game	value=unknown	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Gender	value=Female	0.476	0.600	0.553	0.400	0.500	0.226	0.400	0.625	0.250
Gender	value=Men	0.524	0.400	0.447	0.600	0.500	0.774	0.600	0.375	0.749
Gender	value=unknown	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Medal	value=GOLD	0.143	0.100	0.605	0.360	0.555	0.451	0.200	0.292	0.250
Medal	value=BRONZE	0.286	0.499	0.079	0.400	0.222	0.226	0.200	0.417	0.416
Medal	value=SILVER	0.571	0.400	0.316	0.240	0.222	0.323	0.600	0.292	0.333
Medal	value=unknown	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Team_Type	value=Athlete	0.381	0.500	0.421	0.160	0.555	0.516	0.533	0.458	0.333
Team_Type	value=GameTeam	0.619	0.500	0.579	0.840	0.444	0.484	0.467	0.542	0.666
Team_Type	value=unknown	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000

ภาพที่ 3.12 การแปลงค่าของ Attribute ในแต่ละ Class โดยโปรแกรม Rapid Miner

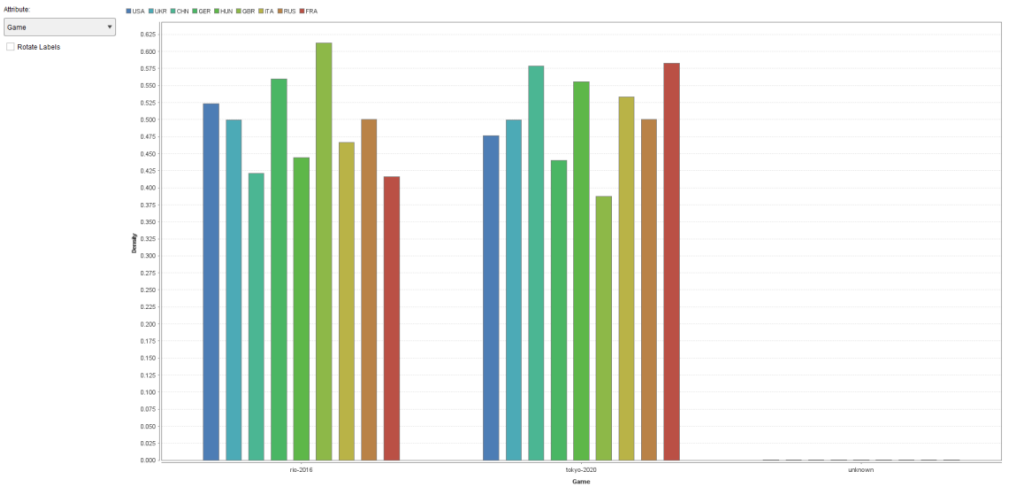
2.2) กราฟแผนภูมิแท่งของ Attribute ในแต่ละ Class ที่ได้แปลงค่า

1) กราฟแผนภูมิแท่งของ Attribute (Sports)



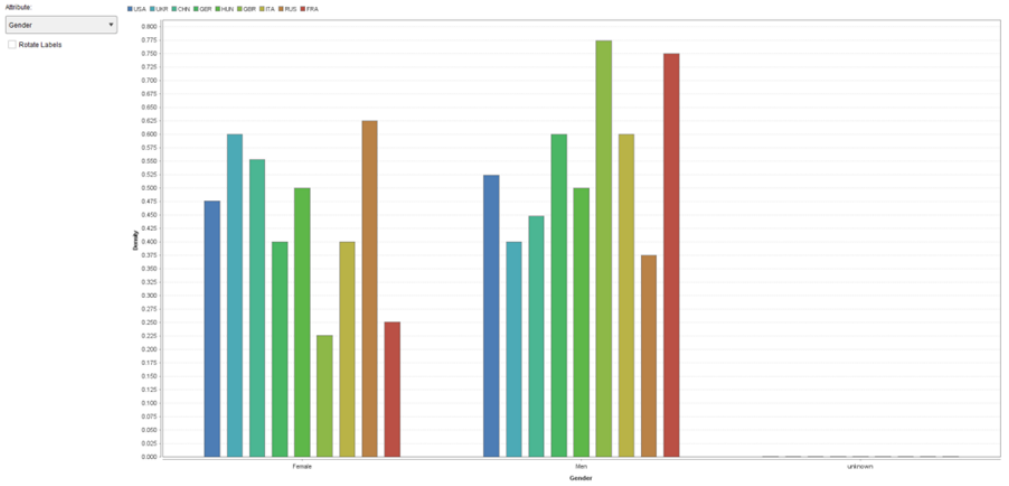
ภาพที่ 3.13 กราฟแผนภูมิแท่งของ Attribute (Sports) โดยโปรแกรม Rapid Miner

2) กราฟแผนภูมิแท่งของ Attribute (Game)



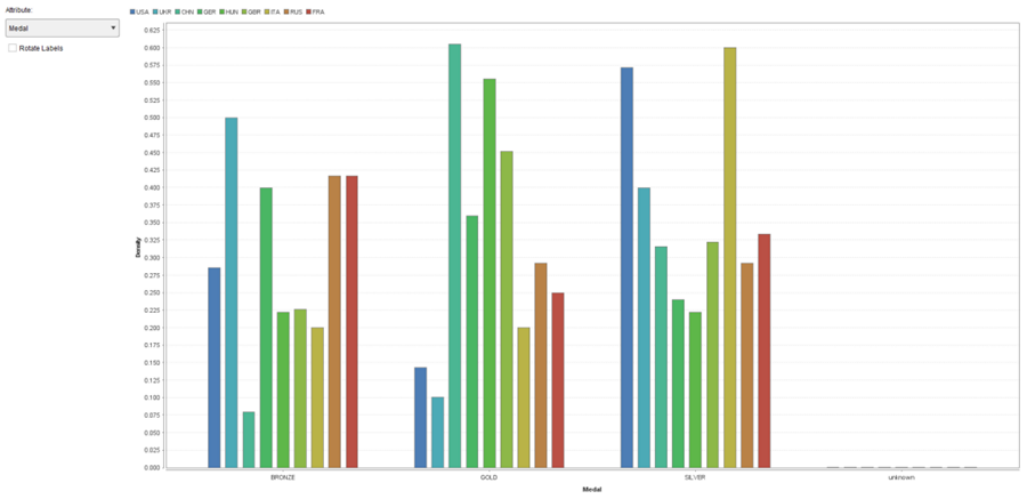
ภาพที่ 3.14 กราฟแผนภูมิแท่งของ Attribute (Game) โดยโปรแกรม Rapid Miner

3) กราฟแผนภูมิแท่งของ Attribute (Gender)



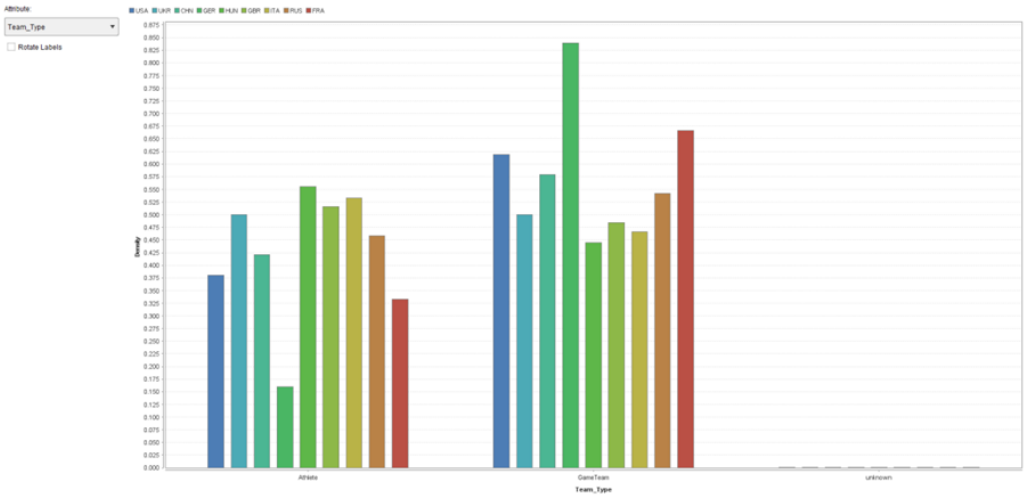
ภาพที่ 3.15 กราฟแผนภูมิแท่งของ Attribute (Gender) โดยโปรแกรม Rapid Miner

4) กราฟแผนภูมิแท่งของ Attribute (Medal)



ภาพที่ 3.16 กราฟแผนภูมิแท่งของ Attribute (Medal) โดยโปรแกรม Rapid Miner

5) กราฟแผนภูมิแท่งของ Attribute (Team_Type)



ภาพที่ 3.17 กราฟแผนภูมิแท่งของ Attribute (Team_Type) โดยโปรแกรม Rapid Miner

จากข้อมูลโมเดล Naive Bayes ของ Rapid Miner จากขั้นตอนผ่านกระบวนการเทคนิค Data Mining จึงได้ความแม่นยำที่ 48.21%

accuracy: 48.21%

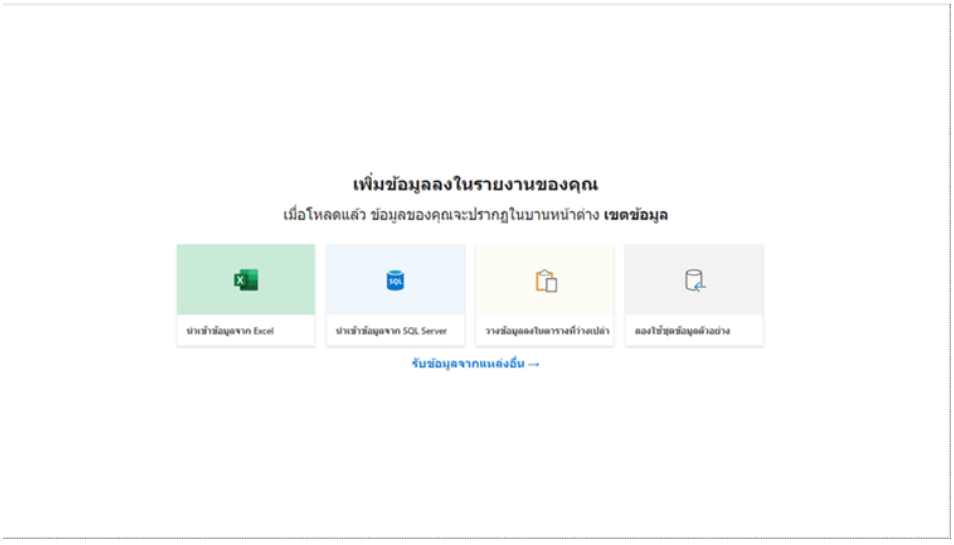
	true USA	true UKR	true CHN	true GER	true HUN	true GBR	true ITA	true RUS	true FRA	class precision
pred. USA	1	0	0	1	0	0	0	0	0	50.00%
pred. UKR	0	1	0	0	0	0	0	0	0	100.00%
pred. CHN	1	0	8	0	0	1	0	0	0	80.00%
pred. GER	0	0	1	4	1	1	0	0	0	57.14%
pred. HUN	1	1	1	0	2	0	0	1	1	28.57%
pred. GBR	1	0	0	2	0	7	0	1	1	58.33%
pred. ITA	1	0	0	0	1	0	1	0	0	33.33%
pred. RUS	0	1	1	0	1	0	2	3	2	30.00%
pred. FRA	1	0	0	0	0	0	1	2	0	0.00%
class recall	16.67%	33.33%	72.73%	57.14%	40.00%	77.78%	25.00%	42.86%	0.00%	

ภาพที่ 3.18 การคำนวณความแม่นยำของข้อมูล Naive Bayes ของ Rapid Miner

3.1.6) การนำโมเดลไปใช้งานจริง (Deployment)

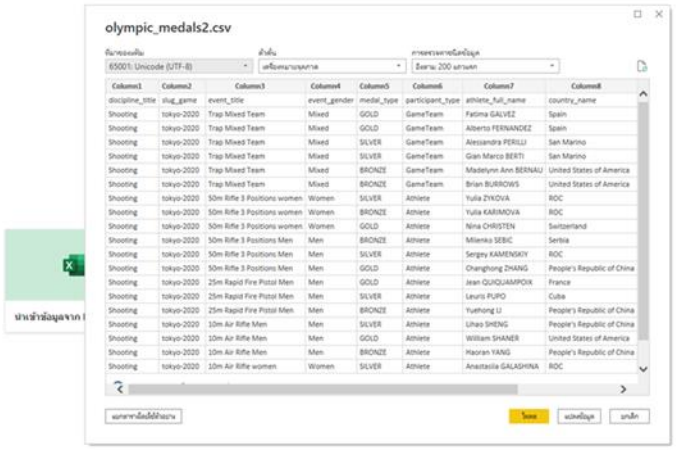
จากการวัดประสิทธิภาพของโมเดลทำให้เห็นว่ารูปแบบโมเดล Naive Bayes มีความแม่นยำที่สูงผู้วิเคราะห์ข้อมูลจึงได้เลือกรูปแบบโมเดล Naive Bayes เพื่อนำไปใช้ในการพยากรณ์ต่อไป นอกจากนี้ผลองค์ความรู้ที่ได้เหล่านี้ผู้วิเคราะห์ข้อมูลนำไปใช้ในการนำเสนอข้อมูลในรูปแบบ Vsiualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้โปรแกรม PowerBI โดยดำเนินการตามขั้นตอนดังนี้

1. Import ข้อมูลจาก Excel ลงไปใน Power BI



ภาพที่ 3.19 การ Import ข้อมูลจาก Excel

2. โหลดข้อมูลจากไฟล์ที่นำมาวิเคราะห์



ภาพที่ 3.20 การโหลดข้อมูลจาก Excel ที่นำมาวิเคราะห์

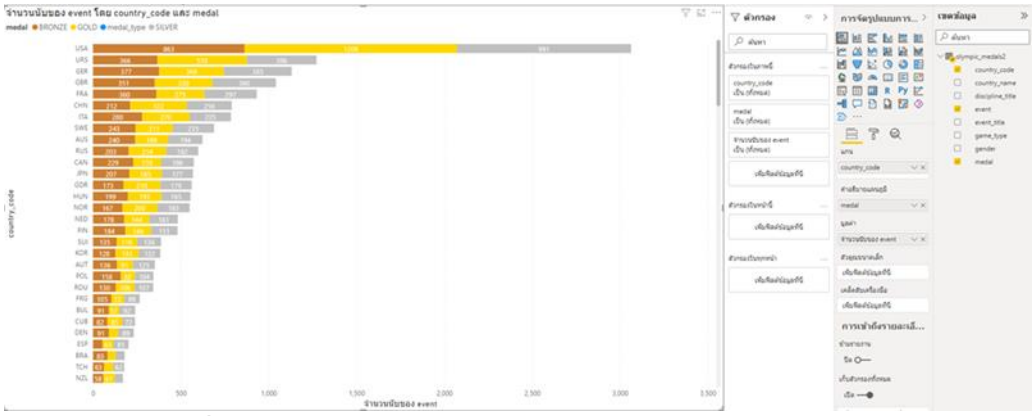
3. ทำการตรวจสอบข้อมูลอีกครั้งก่อนจะนำไปทำการวิเคราะห์

ภาพที่ 3.21 การตรวจสอบข้อมูลอีกครั้งก่อนจะนำไปทำการวิเคราะห์

4. หลังจากการตรวจสอบข้อมูลเสร็จแล้วก็จะนำข้อมูลมาจัดทำ Visualization โดยทำการเลือกเป็น Chart แผนภูมิแท่ง

ภาพที่ 3.22 การนำข้อมูลมาจัดทำเป็น Visualization

5.เลือกข้อมูล ให้ country code เป็นแกน ให้ medal เป็น คำอธิบายแผนภูมิ และ event เป็นตัวนับค่าข้อมูล



ภาพที่ 3.23 Chart แผนภูมิแท่ง country code, medal และ event

3.2 กระบวนการการพยากรณ์

ในขั้นตอนกระบวนการการพยากรณ์นี้ โดยจะทำการพยากรณ์ด้วยกัน 3 หัวข้อดังนี้

1) การพยากรณ์หาเปอร์เซ็นท์ความน่าจะเป็นของแต่ละประเทศที่จะเป็นเจ้าเหรียญทอง

2) การพยากรณ์ว่าในการแข่งขันครั้งต่อไปในอีก 4 ปี ในแต่ละประเทศจะได้รับเหรียญรางวัลรวมเท่าไร เหรียญทองเท่าไหร่ เหรียญเงินเท่าไหร่ เหรียญทองแดงเท่าไหร่

3) การพยากรณ์หาเปอร์เซ็นท์ความน่าจะเป็นในการชนะเลิศของกีฬาแต่ละประเภทของแต่ละประเทศ

โดยที่การพยากรณ์จะใช้การพยากรณ์ในรูปแบบ Time Series Analysis ด้วยวิธีการ Moving Average ในการพยากรณ์และตรวจสอบความคลาดเคลื่อนของข้อมูล โดยจะใช้ชุดข้อมูล olympic_medals ที่ได้ผ่านกระบวนการทำเนื่องข้อมูลและการคัดกรองข้อมูลมาแล้ว โดยจะเลือกแอตทริบิวต์ขึ้นมาใช้ในการพยากรณ์โดยเลือกแอตทริบิวต์ที่ตรงกับวัตถุประสงค์ของขอบเขตที่ได้ตั้งไว้โดยทำการเลือกแอตทริบิวต์มา 4 แอตทริบิวต์ ได้แก่ country_3_letter_code , medal_type, discipline_title, slug_game โดยจะนำทั้ง 4 แอตทริบิวต์นี้เข้าสู่กระบวนการพยากรณ์ต่อไป

3.2.1 การพยากรณ์หาเปอร์เซ็นท์ความน่าจะเป็นของแต่ละประเทศที่จะเป็นเจ้าเหรียญทอง

ในขั้นต่อการพยากรณ์นี้จะใช้การพยากรณ์รูปแบบของ Time Series Analysis ด้วยวิธีการ Moving Average ในการพยากรณ์และตรวจสอบความคลาดเคลื่อนของข้อมูล โดยที่จะใช้แอตทริบิวต์ที่ประกอบไปด้วย 3 แอตทริบิวต์ ได้แก่ country_3_letter_code, medal_type, slug_game โดยจะทำการพยากรณ์ด้วยกันอยู่ 2 รูปแบบ ได้แก่ การพยากรณ์ในการแข่งขัน 2 ครั้ง และ การพยากรณ์ในการแข่งขัน 4 ครั้ง เพื่อที่จะได้เปรียบเทียบความแตกต่างของเปอร์เซ็นว่ามีความแตกต่างและความคลาดเคลื่อนของข้อมูลมากน้อยเพียงใดโดยในรูปแบบการพยากรณ์หาเปอร์เซ็นท์ความน่าจะเป็นของแต่ละประเทศที่จะเป็นเจ้าเหรียญทองโดยจากการคำนวนค่าออกมานั้นจะได้ผลลัพธ์ดังนี้

ตารางที่ 3.11 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ประเทศในการแข่งขันย้อนหลัง 2 ครั้ง ได้ผลลัพธ์ ดังนี้

Country	Predict		Country	Predict
MAR	0.16%		EGY	0.16%
AFG	0.00%		ERI	0.00%
AHO	0.00%		ESP	1.55%
ALG	0.00%		EST	0.16%
ANZ	0.00%		ETH	0.31%
ARG	0.47%		EUN	0.00%
ARM	0.16%		FIJ	0.31%
AUS	3.88%		FIN	0.00%
AUT	0.16%		FRA	3.10%
AZE	0.16%		FRG	0.00%
BAH	0.31%		GAB	0.00%
BAR	0.00%		GBR	7.60%
BDI	0.00%		GDR	0.00%
BEL	0.78%		GEO	0.62%
BER	0.16%		GER	4.19%
BLR	0.16%		GHA	0.00%
BOH	0.00%		GRE	0.78%
BOT	0.00%		GRN	0.00%
BRA	2.17%		GUA	0.00%
BRN	0.16%		GUY	0.00%
BUL	0.47%		HAI	0.00%
BUR	0.00%		HKG	0.16%
CAN	1.71%		HUN	2.17%
CHI	0.00%		INA	0.31%

ตารางที่ 3.11 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ประเทศในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

Country	Predict		Country	Predict
CHN	9.92%		IND	0.16%
CIV	0.16%		IRI	0.93%
CMR	0.00%		IRL	0.31%
COL	0.47%		IRQ	0.00%
CRC	0.00%		ISL	0.00%
CRO	1.24%		ISR	0.31%
CUB	1.86%		ISV	0.00%
CYP	0.00%		ITA	2.79%
CZE	0.78%		JAM	1.55%
DEN	0.78%		JOR	0.16%
DJI	0.00%		JPN	6.05%
DOM	0.00%		KAZ	0.47%
ECU	0.31%		KEN	1.55%

ตารางที่ 3.11 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ประเทศในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

Country	Predict		Country	Predict
KGZ	0.00%		SEN	0.00%
KOR	2.33%		SGP	0.16%
KOS	0.47%		SLO	0.62%
KSA	0.00%		SMR	0.00%
KUW	0.16%		SRB	0.78%
LAT	0.16%		SRI	0.00%
LBN	0.00%		SUD	0.00%
LTU	0.00%		SUI	0.93%
LUX	0.00%		SUR	0.00%
MAS	0.00%		SVK	0.47%
MDA	0.00%		SWE	0.78%
MEX	0.00%		SYR	0.00%
MGL	0.00%		TAN	0.00%
MIX	0.00%		TCH	0.00%
MKD	0.00%		TGA	0.00%
MNE	0.00%		THA	0.47%
MOZ	0.00%		TJK	0.16%
MRI	0.00%		TKM	0.00%
NAM	0.00%		TOG	0.00%
NED	2.79%		TPE	0.47%
NGR	0.00%		TTO	0.00%
NIG	0.00%		TUN	0.16%
NOR	0.62%		TUR	0.47%
NZL	1.71%		UAE	0.00%
PAK	0.00%		UAR	0.00%

ตารางที่ 3.11 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ประเทศในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

Country	Predict		Country	Predict
PAN	0.00%		UGA	0.31%
PAR	0.00%		UKR	0.47%
PER	0.00%		URS	0.00%
PHI	0.16%		URU	0.00%
POL	0.93%		USA	13.18%
POR	0.16%		UZB	1.09%
PRK	0.31%		VEN	0.16%
PUR	0.31%		VIE	0.16%
QAT	0.31%		WIF	0.00%
ROU	0.31%		YUG	0.00%
RSA	0.47%		ZAM	0.00%
RUS	6.05%		ZIM	0.00%
SAM	0.00%			
SCG	0.00%			

จากผลลัพธ์ที่ได้มานั้นจะเห็นได้ว่าการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง ประเทศที่มีความน่าจะเป็นเจ้าเหรียญทองมากที่สุด 5 อันดับ ดังนี้

1) USA (United States of America) มีความน่าจะเป็นอยู่ที่ 13.18%

2) CHN (China) มีความน่าจะเป็นอยู่ที่ 9.92%

3) GBR (United Kingdom of Great Britain and Northern Ireland) มีความน่าจะเป็นอยู่ที่ 7.6%

4) RUS (Russia) มีความน่าจะเป็นอยู่ที่ 6.05%

5) JPN (Japanese) มีความน่าจะเป็นอยู่ที่ 6.05%

ตารางที่ 3.12 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ประเทศในการแข่งขันย้อนหลัง 4 ครั้ง

Country	Predict		Country	Predict
MAR	0.08%		CHI	0.00%
AFG	0.00%		CHN	12.03%
AHO	0.00%		CIV	0.08%
ALG	0.08%		CMR	0.08%
ANZ	0.00%		COL	0.32%
ARG	0.48%		CRC	0.00%
ARM	0.08%		CRO	0.88%
AUS	3.77%		CUB	1.60%
AUT	0.08%		CYP	0.00%
AZE	0.32%		CZE	0.96%
BAH	0.24%		DEN	0.72%
BAR	0.00%		DJI	0.00%
BDI	0.00%		DOM	0.16%
BEL	0.56%		ECU	0.16%
BER	0.08%		EGY	0.08%
BLR	0.48%		ERI	0.00%
BOH	0.00%		ESP	1.52%
BOT	0.00%		EST	0.16%
BRA	1.60%		ETH	0.72%
BRN	0.16%		EUN	0.00%
BUL	0.32%		FIJ	0.16%
BUR	0.00%		FIN	0.08%
CAN	1.28%		FRA	2.97%
FRG	0.00%		MDA	0.00%

ตารางที่ 3.12 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ประเทศในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

Country	Predict		Country	Predict
GAB	0.00%		MEX	0.24%
GBR	7.78%		MGL	0.16%
GDR	0.00%		MIX	0.00%
GEO	0.64%		MKD	0.00%
GER	4.33%		MNE	0.00%
GHA	0.00%		MOZ	0.00%
GRE	0.40%		MRI	0.00%
GRN	0.08%		NAM	0.00%
GUA	0.00%		NED	2.49%
GUY	0.00%		NGR	0.00%
HAI	0.00%		NIG	0.00%
HKG	0.08%		NOR	0.72%
HUN	2.00%		NZL	1.60%
INA	0.24%		PAK	0.00%
IND	0.16%		PAN	0.08%
IRI	1.04%		PAR	0.00%
IRL	0.24%		PER	0.00%
IRQ	0.00%		PHI	0.08%
ISL	0.00%		POL	1.04%
ISR	0.16%		POR	0.16%
ISV	0.00%		PRK	0.64%
ITA	2.73%		PUR	0.16%
JAM	1.52%		QAT	0.16%
JOR	0.08%		ROU	0.64%
JPN	4.41%		RSA	0.56%

ตารางที่ 3.12 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ประเทศในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

Country	Predict		Country	Predict
KAZ	0.64%		RUS	6.66%
KEN	1.44%		SAM	0.00%
KGZ	0.00%		SCG	0.00%
KOR	3.29%		SEN	0.00%
KOS	0.24%		SGP	0.08%
KSA	0.00%		SLO	0.48%
KUW	0.08%		SMR	0.00%
LAT	0.24%		SRB	0.48%
LBN	0.00%		SRI	0.00%
LTU	0.16%		SUD	0.00%
LUX	0.00%		SUI	0.80%
MAS	0.00%		SUR	0.00%

ตารางที่ 3.12 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ประเทศในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

Country	Predict		Country	Predict
SVK	0.48%		UAR	0.00%
SWE	0.48%		UGA	0.24%
SYR	0.00%		UKR	1.20%
TAN	0.00%		URS	0.00%
TCH	0.00%		URU	0.00%
TGA	0.00%		USA	13.39%
THA	0.40%		UZB	0.56%
TJK	0.08%		VEN	0.16%
TKM	0.00%		VIE	0.08%
TOG	0.00%		WIF	0.00%
TPE	0.40%		YUG	0.00%
TTO	0.16%		ZAM	0.00%
TUN	0.32%		ZIM	0.08%
TUR	0.40%			
UAE	0.00%			

จากผลลัพธ์ที่ได้มานั้นจะเห็นได้ว่าการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง ประเทศที่มีความน่าจะเป็นเจ้าเหรียญทองมากที่สุด 5 อันดับ ดังนี้

1) USA (United States of America) มีความน่าจะเป็นอยู่ที่ 13.39%

2) CHN (China) มีความน่าจะเป็นอยู่ที่ 12.03%

3) GBR (United Kingdom of Great Britain and Northern Ireland) มีความน่าจะเป็นอยู่ที่ 7.78%

4) RUS (Russia) มีความน่าจะเป็นอยู่ที่ 6.66%

5) JPN (Japan) มีความน่าจะเป็นอยู่ที่ 4.41%

จากผลการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง กับการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง มีผลลัพธ์ความแตกต่างและค่าความคลาดเคลื่อน ดังนี้

1) USA (United States of America) มีผลลัพธ์ความแตกต่างในการแข่งขัน 2 และ 4 ครั้งซึ่งในการแข่งขัน 4 ครั้งเพิ่มขึ้นจาก 2 ครั้งอยู่ที่ 0.21%

2) CHN (China) มีผลลัพธ์ความแตกต่างในการแข่งขัน 2 และ 4 ครั้งซึ่งในการแข่งขัน 4 ครั้งเพิ่มขึ้นจาก 2 ครั้งอยู่ที่ 2.11% ละในการแข่งขัน 4 ครั้ง ยังอยู่อันดับร่วมกับ USA (United States of America)

3) GBR (United Kingdom of Great Britain and Northern Ireland) มีผลลัพธ์ความแตกต่างในการแข่งขัน 2 และ 4 ครั้งซึ่งในการแข่งขัน 4 ครั้งลดลงจาก 2 ครั้งอยู่ที่ 0.18%

4) RUS (Russia) ผลลัพธ์ความแตกต่างในการแข่งขัน 2 และ 4 ครั้งซึ่งในการแข่งขัน 4 ครั้งเพิ่มขึ้นจาก 2 ครั้งอยู่ที่ 0.61%

5) JPN (Japan) ผลลัพธ์ความแตกต่างในการแข่งขัน 2 และ 4 ครั้งซึ่งในการแข่งขัน 4 ครั้งเพิ่มขึ้นจาก 2 ครั้งอยู่ที่ 1.64%

จากผลการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง กับการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง พบว่า อันดับ 1 ถึง 5 ไม่มีการเปลี่ยนแปลงใด ๆ โดยเมื่อได้ผลลัพธ์และความแตกต่างดังนั้นจึงนำผลลัพธ์ที่ได้ ไปทำการหาค่าความคลาดเคลื่อนเพื่อที่จะได้เลือกว่าจะใช้ข้อมูลการแข่งขันย้อนหลัง 2 ครั้ง หรือการแข่งขันย้อนหลัง 4 ครั้ง โดยได้ผลลัพธ์ ดังนี้

ในการแข่งขัน 2 ครั้ง

1) USA (United States of America) มีค่าความคลาดเคลื่อนอยู่ที่ 14.41%

2) CHN (China) มีค่าความคลาดเคลื่อนอยู่ที่ 40.59%

3) GBR (United Kingdom of Great Britain and Northern Ireland) มีค่าความคลาดเคลื่อนอยู่ที่ 19.19%

4) RUS (Russia) มีค่าความคลาดเคลื่อนอยู่ที่ 7.89%

5) JPN (Japan) มีค่าความคลาดเคลื่อนอยู่ที่ 48.15%

โดยการแข่งขัน 2 ครั้งมีผลลัพธ์ค่าความคลาดเคลื่อนรวม MAPE อยู่ที่ 65.12%

ในการแข่งขัน 4 ครั้ง

1) USA (United States of America) มีค่าความคลาดเคลื่อนอยู่ที่ 10.17%

2) CHN (China) มีค่าความคลาดเคลื่อนอยู่ที่ 23.79%

3) GBR (United Kingdom of Great Britain and Northern Ireland) มีมีค่าความคลาดเคลื่อนอยู่ที่ 20.79%

4) RUS (Russia) มีค่าความคลาดเคลื่อนอยู่ที่ 25.92%

5) JPN (Japan) มีค่าความคลาดเคลื่อนอยู่ที่ 42.13%

โดยการแข่งขันย้อนหลัง 4 ครั้งมีผลลัพธ์ค่าความคลาดเคลื่อน MAPE รวมอยู่ที่ 30.20%

จากผลที่ได้คำนวณและได้ทำการพยกรณ์ออกมาทั้งหมด 2 รูปแบบ การแข่งขันย้อนหลัง 2 ครั้ง และการแข่งขันย้อนหลัง 4 ครั้งโดยใช้รูปแบบ Time Series Analysis ด้วยวิธีการ Moving Average จึงได้ข้อสรุปในการใช้ข้อมูลในการแข่งขันย้อนหลัง 4 ครั้ง ซึ่งมีค่าความคลาดเคลื่อน MAPE รวมอยู่ที่ 30.7

3.2.2 การพยากรณ์ว่าในการแข่งขันครั้งต่อไปในอีก 4 ปี ในแต่ละประเทศจะได้รับเหรียญรางวัลรวมเท่าไร เหรียญทองเท่าไหร่ เหรียญเงินเท่าไหร่ เหรียญทองแดงเท่าไหร่

ในขั้นต่อการพยากรณ์นี้จะใช้การพยากรณ์รูปแบบของ Time Series Analysis ด้วยวิธีการ Moving Average ในการพยากรณ์และตรวจสอบความคลาดเคลื่อนของข้อมูล โดยที่จะใช้แอตทริบิวต์ที่ประกอบไปด้วย 3 แอตทริบิวต์ ได้แก่ country_3_letter_code, medal_type, slug_game โดยจะทำการพยากรณ์ด้วยกันอยู่ 2 รูปแบบ ได้แก่ การพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง และ การพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง เพื่อที่จะได้เปรียบเทียบความแตกต่างของเปอร์เซ็นว่ามีความแตกต่างและความคลาดเคลื่อนของข้อมูลมากน้อยเพียงใดโดยในรูปแบบการพยากรณ์ว่าในการแข่งขันครั้งต่อไปในอีก 4 ปี ในแต่ละประเทศจะได้รับเหรียญรางวัลรวมเท่าไร เหรียญทองเท่าไหร่ เหรียญเงินเท่าไหร่ เหรียญทองแดงเท่าไหร่โดยจากการคำนวณค่าออกมานั้นจะได้ผลลัพธ์ ดังนี้

ตารางที่ 3.13 ในรูปแบบที่ 1 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 2 ครั้ง

country	gold	silver	bronze	all
MAR	1	0	1	1
AFG	0	0	0	0
AHO	0	0	0	0
ALG	0	2	0	2
ANZ	0	0	0	0
ARG	2	1	1	4
ARM	1	3	1	4
AUS	14	15	10	38
AUT	1	0	2	2
AZE	1	6	7	13
BAH	1	0	1	2
BAR	0	0	0	0
BDI	0	1	0	1
BEL	3	2	2	6
BER	1	0	0	1
BLR	1	5	4	9
BOH	0	0	0	0
BOT	0	0	0	0
BRA	9	8	6	22
BRN	1	1	1	2
BUL	2	1	2	4
BUR	0	0	0	0

ตารางที่ 3.13 ในรูปแบบที่ 1 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

country	gold	silver	bronze	all
CAN	6	4	12	21
CHI	0	0	0	0
CHN	38	23	34	95
CIV	1	0	1	1
CMR	0	0	0	0
COL	2	2	4	7
CRC	0	0	0	0
CRO	6	3	3	12
CUB	7	2	4	12
CYP	0	0	0	0

ตารางที่ 3.13 ในรูปแบบที่ 1 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

country	gold	silver	bronze	all
CZE	3	3	7	12
DEN	3	7	7	16
DJI	0	0	0	0
DOM	0	0	2	2
ECU	1	0	1	2
EGY	1	1	2	3
ERI	0	0	0	0
ESP	7	6	7	19
EST	1	0	1	1
ETH	1	2	3	6
EUN	0	0	0	0
FIJ	1	0	0	1
FIN	0	1	1	1
FRA	11	17	15	42
FRG	0	0	0	0
GAB	0	0	0	0
GBR	28	25	21	73
GDR	0	0	0	0
GEO	2	1	5	8
GER	15	13	16	44
GHA	0	0	0	0
GRE	3	1	2	5
GRN	0	1	0	1
GUA	0	0	0	0
GUY	0	0	0	0

ตารางที่ 3.13 ในรูปแบบที่ 1 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

country	gold	silver	bronze	all
HAI	0	0	0	0
HKG	1	0	1	2
HUN	8	4	5	16
INA	2	2	1	5
IND	1	1	1	3
IOA	1	0	1	1
IRI	3	2	3	8
IRL	2	3	0	4
IRQ	0	0	0	0
ISL	0	0	0	0
ISR	1	0	1	2
ISV	0	0	0	0

ตารางที่ 3.13 ในรูปแบบที่ 1 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

country	gold	silver	bronze	all
ITA	10	13	10	33
JAM	5	4	2	10
JOR	1	0	1	1
JPN	21	10	18	48
KAZ	2	4	5	11
KEN	5	6	3	13
KGZ	0	0	1	1
KOR	8	5	7	19
KOS	2	0	0	2
KSA	0	0	1	1
KUW	0	0	0	0
LAT	1	0	0	1
LBN	0	0	0	0
LIE	0	0	0	0
LTU	0	2	3	5
LUX	0	0	0	0
MAS	0	6	1	7
MDA	0	0	0	0
MEX	0	4	1	5
MGL	0	1	1	2
MIX	0	0	0	0
MKD	0	0	1	1
MNE	0	0	0	0
MOZ	0	0	0	0

ตารางที่ 3.13 ในรูปแบบที่ 1 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

country	gold	silver	bronze	all
MRI	0	0	0	0
NAM	0	0	1	1
NED	10	7	9	25
NGR	0	0	1	1
NIG	0	1	0	1
NOR	3	0	4	7
NZL	8	10	7	24
OAR	0	0	0	0
PAK	0	0	0	0
PAN	0	0	0	0
PAR	0	0	0	0
PER	0	0	0	0
PHI	1	1	1	2
POL	4	3	7	13

ตารางที่ 3.13 ในรูปแบบที่ 1 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

country	gold	silver	bronze	all
POR	1	1	1	2
PRK	1	2	1	4
PUR	1	0	0	1
QAT	1	1	0	2
ROU	2	3	3	7
RSA	2	6	2	9
RUS	22	19	26	66
SAM	0	0	0	0
SCG	0	0	0	0
SEN	0	0	0	0
SGP	1	0	0	1
SLO	2	2	1	5
SMR	0	0	1	1
SRB	3	4	2	8
SRI	0	0	0	0
SUD	0	0	0	0
SUI	3	3	4	9
SUR	0	0	0	0
SVK	2	2	1	5
SWE	3	6	5	13
SYR	0	0	0	0
TAN	0	0	0	0
TCH	0	0	0	0
TGA	0	0	0	0

ตารางที่ 3.13 ในรูปแบบที่ 1 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

country	gold	silver	bronze	all
THA	2	2	1	5
TJK	1	0	0	1
TKM	0	0	1	1
TOG	0	0	0	0
TPE	2	0	3	5
TTO	0	0	1	1
TUN	1	0	2	3
TUR	2	3	3	7
UAE	0	0	1	1
UAR	0	0	0	0
UGA	1	0	1	2
UKR	2	5	6	12
URS	0	0	0	0
URU	0	0	0	0

ตารางที่ 3.13 ในรูปแบบที่ 1 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

country	gold	silver	bronze	all
USA	44	38	42	123
UZB	4	2	4	9
VEN	1	2	2	4
VIE	1	1	0	1
WIF	0	0	0	0
YUG	0	0	0	0
ZAM	0	0	0	0
ZIM	0	0	0	0

จากผลลัพธ์ที่ได้มานั้นจะเห็นได้ว่าการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง ประเทศที่มีความน่าจะเป็นเจ้าเหรียญทองมากที่สุด 5 อันดับ ดังนี้

1) USA (United States of America) มีเหรียญรวมอยู่ที่ 123 เหรียญ แบ่งเป็น เหรียญทองอยู่ที่ 44 เหรียญ เหรียญเงินอยู่ที่ 38 เหรียญ เหรียญทองแดงอยู่ที่ 42 เหรียญ

2) CHN (China) มีเหรียญรวมอยู่ที่ 95 เหรียญ แบ่งเป็น เหรียญทองอยู่ที่ 38 เหรียญ เหรียญเงินอยู่ที่ 23 เหรียญ เหรียญทองแดงอยู่ที่ 34 เหรียญ

3) GBR (United Kingdom of Great Britain and Northern Ireland) มีเหรียญรวมอยู่ที่ 73 เหรียญ แบ่งเป็น เหรียญทองอยู่ที่ 28 เหรียญ เหรียญเงินอยู่ที่ 25 เหรียญ เหรียญทองแดงอยู่ที่ 21 เหรียญ

4) RUS (Russia) มีเหรียญรวมอยู่ที่ 66 เหรียญ แบ่งเป็น เหรียญทองอยู่ที่ 22 เหรียญ เหรียญเงินอยู่ที่ 19 เหรียญ เหรียญทองแดงอยู่ที่ 26 เหรียญ

5) JPN (Japanese) มีเหรียญรวมอยู่ที่ 48 เหรียญ แบ่งเป็น เหรียญทองอยู่ที่ 21 เหรียญ เหรียญเงินอยู่ที่ 10 เหรียญ เหรียญทองแดงอยู่ที่ 18 เหรียญ

ตารางที่ 3.14 ในรูปแบบที่ 2 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 4 ครั้ง

country	gold pr	silver pr	bronze pr	all pr
MAR	0	0	1	1
AFG	0	0	1	1
AHO	0	0	0	0
ALG	0	1	0	1
ANZ	0	0	0	0
ARG	2	1	2	5
ARM	0	2	2	4
AUS	14	16	13	42
AUT	0	1	1	2
AZE	1	3	6	10
BAH	1	0	1	2
BAR	0	0	0	0
BDI	0	0	0	0
BEL	2	1	1	4
BER	0	0	0	0
BLR	2	5	5	12
BOH	0	0	0	0
BOT	0	0	0	0
BRA	6	6	9	21
BRN	1	0	0	1
BUL	1	1	2	4
BUR	0	0	0	0
CAN	4	6	11	22

ตารางที่ 3.14 ในรูปแบบที่ 2 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

country	gold pr	silver pr	bronze pr	all pr
CHI	0	0	0	0
CHN	44	26	31	101
CIV	0	0	0	1
CMR	0	0	0	1
COL	1	2	3	6
CRC	0	0	0	0
CRO	4	2	3	9
CUB	5	5	8	18
CYP	0	0	0	0
CZE	3	3	5	11
DEN	3	5	6	13
DJI	0	0	0	0
DOM	1	1	1	2
ECU	1	0	0	1
EGY	0	1	2	3
ERI	0	0	0	0
ESP	6	10	5	22
EST	1	1	1	2
ETH	2	2	2	7
EUN	0	0	0	0
FIJ	1	0	0	1
FIN	0	1	1	3
FRA	10	15	16	41

ตารางที่ 3.14 ในรูปแบบที่ 2 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

FRG	0	0	0	0
GAB	0	0	0	0
GBR	27	21	21	69
GDR	0	0	0	0
GEO	2	2	4	7
GER	16	16	16	47
GHA	0	0	0	0
GRE	1	2	2	5
GRN	0	0	0	1
GUA	0	0	0	0
GUY	0	0	0	0
HAI	0	0	0	0
HKG	0	0	1	1
HUN	7	4	5	16
INA	2	2	2	5
IND	1	1	2	4
IOA	0	0	0	1
IRI	3	2	2	8
IRL	1	2	2	4
IRQ	0	0	0	0
ISL	0	0	0	0
ISR	1	0	1	1
ISV	0	0	0	0
ITA	9	11	10	30
JAM	5	4	2	10

ตารางที่ 3.14 ในรูปแบบที่ 2 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

country	gold pr	silver pr	bronze pr	all pr
JOR	0	0	0	1
JPN	14	10	16	40
KAZ	2	3	5	10
KEN	5	5	4	13
KGZ	0	0	1	1
KOR	11	8	8	27
KOS	1	0	0	1
KSA	0	0	1	1
KUW	0	0	0	0
LAT	1	1	1	2
LBN	0	0	0	0
LIE	0	0	0	0
LTU	1	2	3	5
LUX	0	0	0	0
MAS	0	3	1	4
MDA	0	0	0	0
MEX	1	3	2	6
MGL	1	2	1	3
MIX	0	0	0	0
MKD	0	0	0	0
MNE	0	0	0	0
MOZ	0	0	0	0
MRI	0	0	0	0
NAM	0	0	0	0
NED	8	7	8	23

ตารางที่ 3.14 ในรูปแบบที่ 2 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

country	gold pr	silver pr	bronze pr	all pr
NGR	0	1	1	2
NIG	0	0	0	0
NOR	3	2	3	7
NZL	7	5	6	19
OAR	0	0	0	0
PAK	0	0	0	0
PAN	0	0	0	0
PAR	0	0	0	0
PER	0	0	0	0
PHI	0	0	1	1
POL	4	3	6	13
POR	1	1	1	2
PRK	2	2	2	5
PUR	1	0	0	1
QAT	1	0	1	1
ROU	3	3	3	8
RSA	2	3	1	6
RUS	23	19	27	68
SAM	0	0	0	0
SCG	0	0	0	0
SEN	0	0	0	0
SGP	0	0	1	1
SLO	2	2	2	5
SMR	0	0	1	1

ตารางที่ 3.14 ในรูปแบบที่ 2 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

country	gold pr	silver pr	bronze pr	all pr
SRB	2	2	2	5
SRI	0	0	0	0
SUD	0	0	0	0
SUI	3	2	3	7
SUR	0	0	0	0
SVK	2	2	2	6
SWE	2	5	4	10
SYR	0	0	0	0
TAN	0	0	0	0
TCH	0	0	0	0
TGA	0	0	0	0
THA	1	2	2	5
TJK	0	0	1	1
TKM	0	0	0	0
TOG	0	0	0	0
TPE	2	0	2	4
TTO	1	1	1	2
TUN	1	0	1	2
TUR	1	2	3	6
UAE	0	0	0	0
UAR	0	0	0	0
UGA	1	0	0	1
UKR	4	4	9	17
URS	0	0	0	0
URU	0	0	0	0

ตารางที่ 3.14 ในรูปแบบที่ 2 จะเป็นการพยากรณ์กีฬาในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

country	gold	silver	bronze	all
USA	44	37	39	119
UZB	2	1	3	6
VEN	1	1	1	2
VIE	0	1	0	1
WIF	0	0	0	0
YUG	0	0	0	0
ZAM	0	0	0	0
ZIM	0	1	0	1

จากผลลัพธ์ที่ได้มานั้นจะเห็นได้ว่าการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง ประเทศที่มีความน่าจะเป็นเจ้าเหรียญทองมากที่สุด 5 อันดับ ดังนี้

1) USA (United States of America) มีเหรียญรวมอยู่ที่ 119 เหรียญ แบ่งเป็น เหรียญทองอยู่ที่ 44 เหรียญ เหรียญเงินอยู่ที่ 37 เหรียญ เหรียญทองแดงอยู่ที่ 39 เหรียญ

2) CHN (China) มีเหรียญรวมอยู่ที่ 101 เหรียญ แบ่งเป็น เหรียญทองอยู่ที่ 44 เหรียญ เหรียญเงินอยู่ที่ 26 เหรียญ เหรียญทองแดงอยู่ที่ 31 เหรียญ

3) GBR (United Kingdom of Great Britain and Northern Ireland) มีเหรียญรวมอยู่ที่ 69 เหรียญ แบ่งเป็น เหรียญทองอยู่ที่ 27 เหรียญ เหรียญเงินอยู่ที่ 21 เหรียญ เหรียญทองแดงอยู่ที่ 21 เหรียญ

4) RUS (Russia) มีเหรียญรวมอยู่ที่ 68 เหรียญ แบ่งเป็น เหรียญทองอยู่ที่ 23 เหรียญ เหรียญเงินอยู่ที่ 19 เหรียญ เหรียญทองแดงอยู่ที่ 27 เหรียญ

5) GER (Germany) มีเหรียญรวมอยู่ที่ 47 เหรียญ แบ่งเป็น เหรียญทองอยู่ที่ 16 เหรียญ เหรียญเงินอยู่ที่ 16 เหรียญ เหรียญทองแดงอยู่ที่ 16 เหรียญ

จากผลการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง กับการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง มีผลลัพธ์ความแตกต่างและค่าความคลาดเคลื่อน ดังนี้

1) USA (United States of America) มีผลลัพธ์ความแตกต่างในการแข่งขัน 2 และ 4 ครั้ง ดังนี้ เหรียญทอง 0 เหรียญ เหรียญเงิน 1 เหรียญทองแดง 3 เหรียญ

2) CHN (China) มีผลลัพธ์ความแตกต่างในการแข่งขัน 2 และ 4 ครั้งดังนี้ เหรียญทอง 6 เหรียญ เหรียญเงิน 3 เหรียญทองแดง 3 เหรียญ

3) GBR (United Kingdom of Great Britain and Northern Ireland) มีผลลัพธ์ความแตกต่างในการแข่งขัน 2 และ 4 ครั้งดังนี้ เหรียญทอง 1 เหรียญ เหรียญเงิน 4 เหรียญทองแดง 0 เหรียญ

4) RUS (Russia) มีผลลัพธ์ความแตกต่างในการแข่งขัน 2 และ 4 ครั้งดังนี้ เหรียญทอง 1 เหรียญ เหรียญเงิน 0 เหรียญทองแดง 1 เหรียญ

5) ในผลลัพธ์อันดับที่ 5 มีการเปลี่ยนแปลงความน่าจะเป็นของประเทศในการแข่งขัน 2 ครั้ง เป็น JPN (Japanese) และมีการเปลี่ยนแปลงในการแข่งขัน 4 ครั้ง เป็น GER (Germany) มีผลลัพธ์ความแตกต่างของเหรียญ ดังนี้ เหรียญทอง 5 เหรียญ เหรียญเงิน 6 เหรียญดองแดง 2 เหรียญ

จากผลการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง กับการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง พบว่า อันดับ 1 ถึง 4 ไม่มีการเปลี่ยนแปลงใด ๆ แต่มีอันดับที่ 5 ที่มีการเปลี่ยนแปลง โดยเมื่อได้ผลลัพธ์

และความแตกต่างดังนั้นจึงนำผลลัพธ์ที่ได้ ไปทำการหาค่าความคลาดเคลื่อนเพื่อที่จะได้เลือกว่าจะใช้ข้อมูลการแข่งขัน 2 ครั้ง หรือการแข่งขัน 4 ครั้ง โดยได้ผลลัพธ์ ดังนี้

ในการแข่งขัน 2 ครั้ง

1) USA (United States of America) มีค่าความคลาดเคลื่อนอยู่ที่ เหรียญทองอยู่ที่ 13.19% เหรียญเงิน 16.48% เหรียญทองแดง 9.41%

2) CHN (China) มีค่าความคลาดเคลื่อนอยู่ที่ เหรียญทองอยู่ที่ 38.42% เหรียญเงิน 44.33% เหรียญทองแดง 18.42%

3)GBR (United Kingdom of Great Britain and Northern Ireland) มีค่าความคลาดเคลื่อนอยู่ที่ เหรียญทองอยู่ที่ 19.00% เหรียญเงิน 20.28% เหรียญทองแดง 9.81%

4) RUS (Russia) มีค่าความคลาดเคลื่อนอยู่ที่ เหรียญทองอยู่ที่ 7.14% เหรียญเงิน 20.31% เหรียญทองแดง 29.02%

5) JPN (Japanese) มีมีค่าความคลาดเคลื่อนอยู่ที่ เหรียญทองอยู่ที่ 51.37% เหรียญเงิน 32.14% เหรียญทองแดง 26.01%

โดยการแข่งขันย้อนหลัง 2 ครั้งมีผลลัพธ์ค่าความคลาดเคลื่อนรวม MAPA อยู่ที่ 35.53%

ในการแข่งขัน 4 ครั้ง

1) USA (United States of America) มีค่าความคลาดเคลื่อนอยู่ที่ เหรียญทองอยู่ที่ 9.07% เหรียญเงิน 15.45% เหรียญทองแดง 8.97%

2) CHN (China) มีค่าความคลาดเคลื่อนอยู่ที่ เหรียญทองอยู่ที่ 23.19% เหรียญเงิน 35.02% เหรียญทองแดง 18.42%

3)GBR (United Kingdom of Great Britain and Northern Ireland) มีค่าความคลาดเคลื่อนอยู่ที่ เหรียญทองอยู่ที่ 22.33% เหรียญเงิน 26.05% เหรียญทองแดง 14.71%

4) RUS (Russia) มีค่าความคลาดเคลื่อนอยู่ที่ เหรียญทองอยู่ที่ 21.21% เหรียญเงิน 33.85% เหรียญทองแดง 40.54%

5) GER (Germany) มีมีค่าความคลาดเคลื่อนอยู่ที่ เหรียญทองอยู่ที่ 45.00% เหรียญเงิน 47.08% เหรียญทองแดง 21.29%

โดยการแข่งขันย้อนหลัง 4 ครั้งมีผลลัพธ์ค่าความคลาดเคลื่อน MAPE รวมอยู่ที่ 39.16%

จากผลที่ได้คำนวณและได้ทำการพยกรณ์ออกมาทั้งหมด 2 รูปแบบ การแข่งขันย้อนหลัง 2 ครั้ง และการแข่งขันย้อนหลัง 4 ครั้งโดยใช้รูปแบบ Time Series Analysis ด้วยวิธีการ Moving Average จึงได้ข้อสรุปในการใช้ข้อมูลในการแข่งขันย้อนหลัง 2 ครั้ง ซึ่งมีค่าความคลาดเคลื่อน MAPE รวมอยู่ที่ 35.53%

3.2.3 การพยากรณ์หาเปอร์เซ็นท์ความน่าจะเป็นในการชนะเลิศของกีฬาแต่ละประเภทของแต่ละประเทศ

ในขั้นต่อการพยากรณ์นี้จะใช้การพยากรณ์รูปแบบของ Time Series Analysis ด้วยวิธีการ Moving Average ในการพยากรณ์และตรวจสอบความคลาดเคลื่อนของข้อมูล โดยที่จะใช้แอตทริบิวต์ที่ประกอบไปด้วย 3 แอตทริบิวต์ ได้แก่ country_3_letter_code, discipline_title, medal_type โดยจะทำการพยากรณ์ด้วยกันอยู่ 2 รูปแบบ ได้แก่ การพยากรณ์ในการแข่งขัน 2 ครั้ง และ การพยากรณ์ในการแข่งขัน 4 ครั้ง เพื่อที่จะได้เปรียบเทียบความแตกต่างของเปอร์เซ็นว่ามีความแตกต่างและความคลาดเคลื่อนของข้อมูลมากน้อยเพียงใดโดยในรูปแบบการพยากรณ์หาเปอร์เซ็นท์ความน่าจะเป็นในการชนะเลิศของกีฬาแต่ละประเภทของแต่ละประเทศโดยจากการคำนวณค่าออกมานั้นจะได้ผลลัพธ์ ดังนี้

ตารางที่ 3.15 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง

Sports	County	Predict	MAPE
3x3 basketball	LAT	50.00%	50.00%
3x3 basketball	USA	50.00%	50.00%
Archery	KOR	66.67%	0.00%
Artistic Gymnastics	CHN	20.00%	50.00%
Artistic Swimming	ROC	100.00%	50.00%
Athletics	USA	15.87%	28.57%
Sports	County	Predict	MAPE
Badminton	CHN	18.18%	50.00%
Baseball			
Baseball/Softball			
Basketball	USA	100.00%	0.00%
Beach Volleyball	BRA	20.00%	0.00%
Beach Volleyball	GER	20.00%	0.00%
Beach Volleyball	NOR	20.00%	50.00%
Beach Volleyball	USA	20.00%	0.00%
Boxing	CUB	17.39%	12.50%
Canoe Slalom	AUS	12.50%	50.00%
Canoe Slalom	CZE	12.50%	50.00%
Canoe Slalom	ESP	12.50%	0.00%

ตารางที่ 3.15 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Canoe Slalom	FRA	12.50%	0.00%
Canoe Slalom	GBR	12.50%	0.00%
Canoe Slalom	GER	12.50%	50.00%
Canoe Slalom	SLO	12.50%	50.00%
Canoe Slalom	SVK	12.50%	0.00%
Canoe Sprint	GER	15.79%	100.00%
Canoe Sprint	HUN	15.79%	0.00%
Cycling BMX	COL	50.00%	0.00%
Cycling BMX	USA	50.00%	0.00%
Cycling BMX Freestyle	AUS	50.00%	50.00%
Cycling BMX Freestyle	GBR	50.00%	50.00%
Cycling BMX Racing	GEO	50.00%	0.00%
Cycling BMX Racing	NGR	50.00%	0.00%

ตารางที่ 3.15 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Cycling Mountain Bike	GBR	33.33%	50.00%
Cycling Mountain Bike	SUI	33.33%	0.00%
Cycling Mountain Bike	SWE	33.33%	0.00%
Cycling Road	AUT	12.50%	50.00%
Cycling Road	BEL	12.50%	0.00%
Cycling Road	ECU	12.50%	50.00%
Cycling Road	NED	12.50%	0.00%
Cycling Road	SLO	12.50%	50.00%
Cycling Road	USA	12.50%	0.00%
Cycling Track	GBR	33.33%	16.67%
Diving	CHN	77.78%	0.00%
Equestrian	GER	50.00%	50.00%
Equestrian Dressage	GBR	50.00%	0.00%
Equestrian Dressage	GER	50.00%	0.00%
Equestrian Eventing	FRA	50.00%	0.00%
Equestrian Eventing	GER	50.00%	0.00%
Equestrian Jumping	FRA	33.33%	0.00%
Equestrian Jumping	GBR	33.33%	0.00%
Fencing	FRA	10.00%	25.00%
Fencing	HUN	10.00%	50.00%
Fencing	ROC	10.00%	50.00%
Fencing	RUS	10.00%	0.00%
Football	BRA	20.00%	0.00%
Golf	GBR	33.33%	0.00%

ตารางที่ 3.15 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

Golf	KOR	33.33%	0.00%
Golf	USA	33.33%	50.00%
Gymnastics Artistic	USA	3.33%	0.00%
Gymnastics Rhythmic	RUS	100.00%	0.00%
Handball	DEN	33.33%	0.00%
Handball	FRA	33.33%	25.00%
Handball	RUS	33.33%	0.00%
Hockey	ARG	20.00%	0.00%
Hockey	BEL	20.00%	50.00%
Hockey	GBR	20.00%	0.00%
Hockey	NED	20.00%	0.00%
Judo	JPN	25.00%	38.89%
Karate	BUL	12.50%	50.00%
Karate	EGY	12.50%	50.00%
Karate	ESP	12.50%	50.00%
Karate	FRA	12.50%	50.00%
Karate	IRI	12.50%	50.00%
Karate	ITA	12.50%	50.00%
Karate	JPN	12.50%	50.00%
Karate	SRB	12.50%	50.00%
Marathon Swimming	BRA	33.33%	50.00%
Marathon Swimming	GER	33.33%	50.00%
Marathon Swimming	NED	33.33%	0.00%

ตารางที่ 3.15 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Cycling Mountain Bike	GBR	33.33%	50.00%
Cycling Mountain Bike	SUI	33.33%	0.00%
Cycling Mountain Bike	SWE	33.33%	0.00%
Cycling Road	AUT	12.50%	50.00%
Cycling Road	ECU	12.50%	50.00%
Cycling Road	NED	12.50%	0.00%
Cycling Road	SLO	12.50%	50.00%
Cycling Road	SUI	12.50%	0.00%
Cycling Road	USA	12.50%	0.00%
Cycling Track	GBR	33.33%	16.67%
Diving	CHN	77.78%	0.00%
Equestrian	GER	50.00%	50.00%
Equestrian Dressage	GBR	50.00%	0.00%
Equestrian Dressage	GER	50.00%	0.00%
Equestrian Eventing	FRA	50.00%	0.00%
Equestrian Eventing	GER	50.00%	0.00%
Equestrian Jumping	FRA	33.33%	0.00%
Equestrian Jumping	GBR	33.33%	0.00%
Fencing	FRA	10.00%	25.00%
Fencing	HUN	10.00%	50.00%
Fencing	ROC	10.00%	50.00%
Fencing	RUS	10.00%	0.00%

ตารางที่ 3.15 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Football	BRA	20.00%	0.00%
Modern Pentathlon	AUS	25.00%	0.00%
Modern Pentathlon	GBR	25.00%	50.00%
Modern Pentathlon	RUS	25.00%	0.00%
Rhythmic Gymnastics	BUL	50.00%	50.00%
Rhythmic Gymnastics	ISR	50.00%	50.00%
Rowing	NZL	13.04%	0.00%
Rugby	AUS	33.33%	0.00%
Rugby	FIJ	33.33%	0.00%
Rugby Sevens	FIJ	50.00%	50.00%
Rugby Sevens	NZL	50.00%	50.00%
Sailing	GBR	13.64%	16.67%
Shooting	CHN	9.09%	50.00%
Skateboarding	JPN	66.67%	50.00%
Softball			
Sport Climbing	ESP	50.00%	50.00%
Sport Climbing	SLO	50.00%	50.00%
Surfing	BRA	50.00%	50.00%
Surfing	USA	50.00%	50.00%
Swimming	USA	30.43%	22.73%
Synchronized Swimming	RUS	100.00%	0.00%
Table Tennis	CHN	80.00%	0.00%
Taekwondo	AZE	6.25%	0.00%
Taekwondo	CHN	6.25%	0.00%

ตารางที่ 3.15 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Taekwondo	CRO	6.25%	50.00%
Taekwondo	GBR	6.25%	0.00%
Taekwondo	ITA	6.25%	0.00%
Taekwondo	JOR	6.25%	0.00%
Taekwondo	KOR	6.25%	0.00%
Taekwondo	ROC	6.25%	50.00%
Taekwondo	SRB	6.25%	0.00%
Taekwondo	THA	6.25%	50.00%
Taekwondo	USA	6.25%	50.00%
Taekwondo	UZB	6.25%	50.00%
Tennis	CRO	9.09%	50.00%
Tennis	CZE	9.09%	50.00%
Tennis	ESP	9.09%	0.00%
Tennis	GBR	9.09%	0.00%
Tennis	GER	9.09%	50.00%
Tennis	PUR	9.09%	0.00%
Tennis	ROC	9.09%	50.00%
Tennis	RUS	9.09%	0.00%
Tennis	SUI	9.09%	50.00%
Tennis	USA	9.09%	0.00%
Trampoline	BLR	14.29%	0.00%
Trampoline	CAN	14.29%	0.00%
Trampoline Gymnastics	CHN	100.00%	50.00%

ตารางที่ 3.15 ในรูปแบบที่ 1 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 2 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Triathlon	GBR	20.00%	0.00%
Triathlon	NOR	20.00%	50.00%
Triathlon	USA	20.00%	0.00%
Volleyball	BRA	20.00%	0.00%
Volleyball	CHN	20.00%	0.00%
Volleyball	FRA	20.00%	50.00%
Volleyball	USA	20.00%	50.00%
Water Polo	SRB	50.00%	0.00%
Water Polo	USA	50.00%	0.00%
Weightlifting	CHN	30.00%	14.29%
Wrestling	JPN	19.23%	10.00%

จากผลลัพธ์ที่ได้มานั้นจะเห็นได้ว่าการพยากรณ์หาเปอร์เซ็นท์ความน่าจะเป็นในการชนะเลิศของกีฬาแต่ละประเภทของแต่ละประเทศ โดยจะยกเอา 3 ประเทศที่ มีความน่าจะเป็นที่สูงและมีค่าความคลาดเคลื่อนที่ต่ำ ดังนี้

1) ประเภทกีฬา Athletics ประเทศ USA (United States of America) มีความน่าจะเป็นที่จะชนะเลิศอยู่ที่ 15.87% ค่าความคลาดเคลื่อนอยู่ที่ 28.57%

2) ประเภทกีฬา Weightlifting ประเทศ CHN (China) มีความน่าจะเป็นที่จะชนะเลิศอยู่ที่ 30.00% ค่าความคลาดเคลื่อนอยู่ที่ 14.29%

3) ประเภทกีฬา Swimming ประเทศ USA (United States of America) มีความน่าจะเป็นที่จะชนะเลิศอยู่ที่ 30.43% ค่าความคลาดเคลื่อนอยู่ที่ 22.73%

ตารางที่ 3.16 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง

Sports	County	Predict	MAPE
3x3 basketball	LAT	50.00%	50.00%
3x3 basketball	USA	50.00%	50.00%
Archery	KOR	50.00%	12.50%
Artistic Gymnastics	BEL	11.11%	50.00%
Artistic Gymnastics	BRA	11.11%	50.00%
Artistic Gymnastics	CHN	11.11%	50.00%
Artistic Gymnastics	GBR	11.11%	50.00%
Artistic Gymnastics	ISR	11.11%	50.00%
Artistic Gymnastics	JPN	11.11%	50.00%
Artistic Gymnastics	KOR	11.11%	50.00%
Artistic Gymnastics	ROC	11.11%	50.00%
Artistic Gymnastics	USA	11.11%	50.00%
Artistic Swimming	ROC	100.00%	50.00%
Athletics	USA	12.16%	85.71%

ตารางที่ 3.16 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Badminton	CHN	33.33%	50.00%
Baseball	KOR	100.00%	0.00%
Baseball/Softball			
Basketball	USA	100.00%	0.00%
Beach Volleyball	BRA	25.00%	0.00%
Beach Volleyball	GER	25.00%	0.00%
Beach Volleyball	NOR	25.00%	0.00%
Beach Volleyball	USA	25.00%	0.00%
Boxing	GBR	9.09%	0.00%
Canoe Slalom	AUS	11.11%	50.00%
Canoe Slalom	CZE	11.11%	50.00%
Canoe Slalom	ESP	11.11%	0.00%
Canoe Slalom	FRA	11.11%	0.00%
Canoe Slalom	GBR	11.11%	0.00%
Canoe Slalom	GER	11.11%	0.00%
Canoe Slalom	ITA	11.11%	0.00%
Canoe Slalom	SLO	11.11%	50.00%
Canoe Slalom	SVK	11.11%	0.00%
Canoe Sprint	GER	15.79%	100.00%
Canoe Sprint	HUN	15.79%	0.00%
Cycling BMX	COL	25.00%	0.00%
Cycling BMX	FRA	25.00%	0.00%
Cycling BMX	USA	25.00%	0.00%
Cycling BMX Freestyle	AUS	50.00%	50.00%

ตารางที่ 3.16 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Cycling BMX Freestyle	GBR	50.00%	50.00%
Cycling BMX Racing	GEO	50.00%	0.00%
Cycling BMX Racing	NGR	50.00%	0.00%
Cycling Mountain Bike	CZE	16.67%	0.00%
Cycling Mountain Bike	FRA	16.67%	0.00%
Cycling Mountain Bike	GBR	16.67%	50.00%
Cycling Mountain Bike	GER	16.67%	0.00%
Cycling Mountain Bike	SUI	16.67%	0.00%
Cycling Mountain Bike	SWE	16.67%	0.00%
Cycling Road	AUT	10.00%	50.00%
Cycling Road	BEL	10.00%	0.00%
Cycling Road	ECU	10.00%	50.00%
Cycling Road	ESP	10.00%	0.00%
Cycling Road	GBR	10.00%	0.00%
Cycling Road	NED	10.00%	0.00%
Cycling Road	SLO	10.00%	50.00%
Cycling Road	SUI	10.00%	0.00%
Cycling Road	USA	10.00%	0.00%
Cycling Track	GBR	38.46%	16.67%
Diving	CHN	63.64%	0.00%
Equestrian	GBR	33.33%	50.00%
Equestrian	GER	33.33%	50.00%
Equestrian Dressage	GBR	33.33%	0.00%
Equestrian Dressage	GER	33.33%	0.00%

ตารางที่ 3.16 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Equestrian Dressage	NED	33.33%	0.00%
Equestrian Eventing	GER	66.67%	0.00%
Equestrian Jumping	CAN	20.00%	0.00%
Equestrian Jumping	FRA	20.00%	0.00%
Equestrian Jumping	GBR	20.00%	0.00%
Equestrian Jumping	SUI	20.00%	0.00%
Equestrian Jumping	USA	20.00%	0.00%
Fencing	FRA	11.11%	0.00%
Fencing	ITA	11.11%	0.00%
Fencing	RUS	11.11%	0.00%
Football	ARG	16.67%	0.00%

ตารางที่ 3.16 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Football	BRA	16.67%	0.00%
Football	CAN	16.67%	50.00%
Football	GER	16.67%	0.00%
Football	MEX	16.67%	0.00%
Football	USA	16.67%	0.00%
Golf	GBR	33.33%	0.00%
Golf	KOR	33.33%	0.00%
Golf	USA	33.33%	0.00%
Gymnastics Artistic	CHN	20.00%	0.00%
Gymnastics Rhythmic	RUS	100.00%	0.00%
Handball	DEN	25.00%	0.00%
Handball	FRA	25.00%	25.00%
Handball	NOR	25.00%	0.00%
Handball	RUS	25.00%	0.00%
Hockey	ARG	20.00%	0.00%
Hockey	BEL	20.00%	50.00%
Hockey	GBR	20.00%	0.00%
Hockey	GER	20.00%	0.00%
Hockey	NED	20.00%	0.00%
Judo	JPN	20.00%	27.78%
Karate	EGY	12.50%	50.00%
Karate	ESP	12.50%	50.00%
Karate	IRI	12.50%	50.00%
Karate	ITA	12.50%	50.00%

ตารางที่ 3.16 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Karate	JPN	12.50%	50.00%
Karate	SRB	12.50%	50.00%
Marathon Swimming	BRA	16.67%	50.00%
Marathon Swimming	GER	16.67%	50.00%
Marathon Swimming	HUN	16.67%	0.00%
Marathon Swimming	RUS	16.67%	0.00%
Marathon Swimming	TUN	16.67%	0.00%
Modern Pentathlon	AUS	16.67%	0.00%
Modern Pentathlon	GBR	16.67%	50.00%
Modern Pentathlon	GER	16.67%	0.00%
Modern Pentathlon	LTU	16.67%	0.00%
Modern Pentathlon	RUS	16.67%	0.00%
Rhythmic Gymnastics	BUL	50.00%	50.00%
Rhythmic Gymnastics	ISR	50.00%	50.00%
Rowing	GBR	11.11%	0.00%
Rowing	NZL	11.11%	16.67%
Rugby	AUS	50.00%	0.00%
Rugby	FIJ	50.00%	0.00%
Rugby Sevens	FIJ	50.00%	50.00%
Rugby Sevens	NZL	50.00%	50.00%
Sailing	GBR	17.65%	0.00%
Shooting	CHN	10.00%	12.50%
Shooting	USA	10.00%	16.67%
Skateboarding	AUS	50.00%	50.00%

ตารางที่ 3.16 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Skateboarding	JPN	50.00%	50.00%
Softball	JPN	100.00%	0.00%
Sport Climbing	ESP	50.00%	50.00%
Sport Climbing	SLO	50.00%	50.00%
Surfing	BRA	50.00%	50.00%
Surfing	USA	50.00%	50.00%
Swimming	USA	30.43%	13.64%
Synchronized Swimming	RUS	100.00%	0.00%
Table Tennis	CHN	80.00%	0.00%
Taekwondo	KOR	10.53%	0.00%
Tennis	USA	16.67%	0.00%
Trampoline	BLR	33.33%	0.00%
Trampoline	CAN	33.33%	0.00%
Trampoline	CHN	33.33%	0.00%
Trampoline Gymnastics	CHN	100.00%	0.00%
Triathlon	BER	14.29%	50.00%
Triathlon	GBR	14.29%	0.00%
Triathlon	NOR	14.29%	50.00%
Triathlon	SUI	14.29%	0.00%
Triathlon	USA	14.29%	0.00%
Volleyball	BRA	20.00%	0.00%
Volleyball	CHN	20.00%	0.00%

ตารางที่ 3.16 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Volleyball	FRA	20.00%	50.00%
Volleyball	RUS	20.00%	0.00%
Volleyball	USA	20.00%	0.00%
Water Polo	CRO	20.00%	0.00%
Water Polo	HUN	20.00%	0.00%
Water Polo	NED	20.00%	0.00%
Water Polo	SRB	20.00%	0.00%
Water Polo	USA	20.00%	0.00%

ตารางที่ 3.16 ในรูปแบบที่ 2 จะเป็นการพยากรณ์ในการแข่งขันย้อนหลัง 4 ครั้ง (ต่อ)

Sports	County	Predict	MAPE
Weightlifting	CHN	24.00%	14.29%
Wrestling	JPN	14.29%	20.00%
Wrestling	RUS	14.29%	0.00%

จากผลลัพธ์ที่ได้มานั้นจะเห็นได้ว่าการพยากรณ์หาเปอร์เซ็นท์ความน่าจะเป็นในการชนะเลิศของกีฬาแต่ละประเภทของแต่ละประเทศ โดยจะยกเอา 3 ประเทศที่ มีความน่าจะเป็นที่สูงและมีค่าความคลาดเคลื่อนที่ต่ำ ดังนี้

1) ประเภทกีฬา Athletics ประเทศ USA (United States of America) มีความน่าจะเป็นที่จะชนะเลิศอยู่ที่ 12.16% ค่าความคลาดเคลื่อนอยู่ที่ 85.71%

2) ประเภทกีฬา Weightlifting ประเทศ CHN (China) มีความน่าจะเป็นที่จะชนะเลิศอยู่ที่ 24.00% ค่าความคลาดเคลื่อนอยู่ที่ 14.29%

3) ประเภทกีฬา Swimming ประเทศ USA (United States of America) มีความน่าจะเป็นที่จะชนะเลิศอยู่ที่ 30.43% ค่าความคลาดเคลื่อนอยู่ที่ 13.64%

จากผลการพยากรณ์ในการแข่งขัน 2 ครั้ง กับการพยากรณ์ในการแข่งขัน 4 ครั้ง โดยที่ความคลาดเคลื่อนโดยรวมของทั้งในการแข่งขัน 2 ครั้ง และ ในการแข่งขัน 4 ครั้ง อยู่ที่

1) ในการแข่งขัน 2 ครั้ง คลาดเคลื่อนโดยรวมอยู่ที่ 23%

2) ในการแข่งขัน 4 ครั้ง คลาดเคลื่อนโดยรวมอยู่ที่ 18%

จึงได้ข้อสรุปในการใช้ข้อมูลในการแข่งขันย้อนหลัง 4 ครั้ง ซึ่งมีค่าความคลาดเคลื่อน MAPE รวมอยู่ที่ 18%

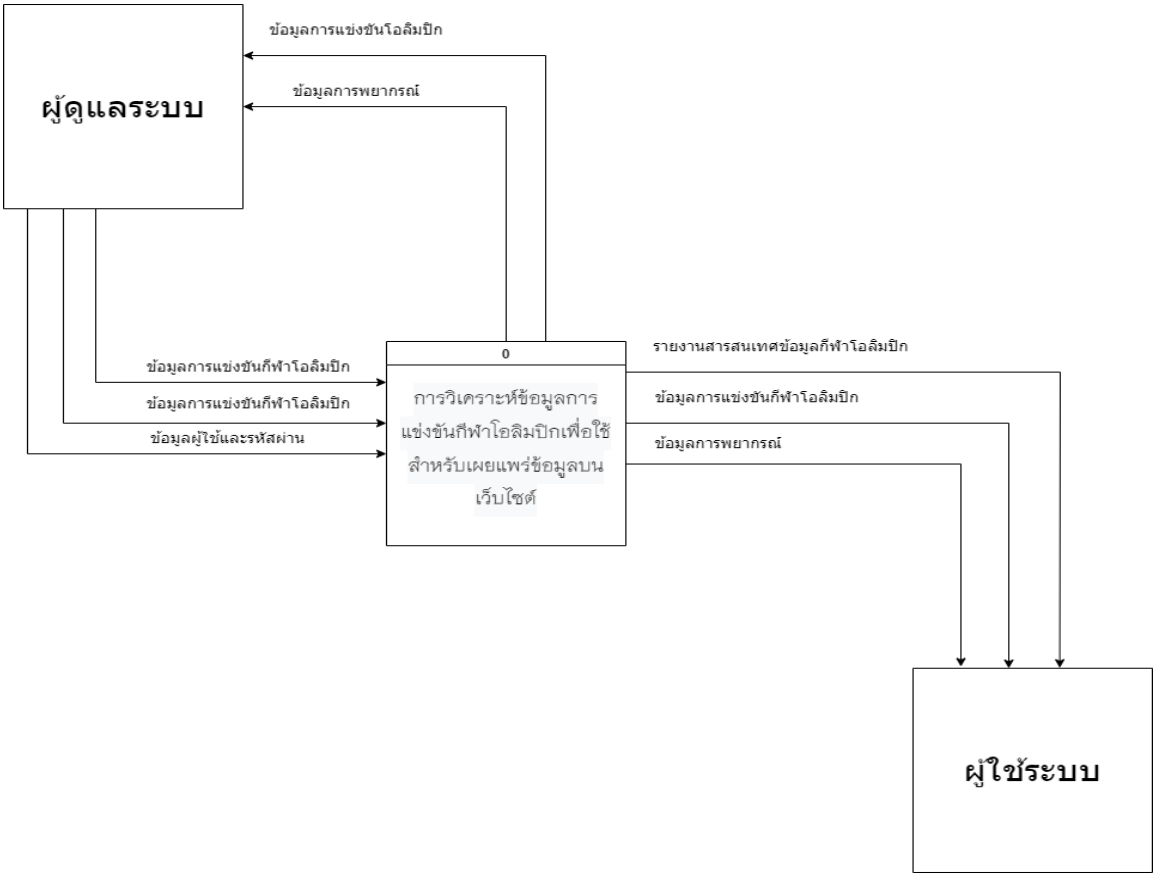
3.3 การออกแบบและพัฒนาเว็บไซต์

3.3.1 แผนภาพบริบท (Context Diagram)

แผนภาพบริบทเป็นแผนภาพที่แสดงถึงภาพรวมของระบบ และความสัมพันธ์ระบบกับสิ่งแวดล้อมที่เกี่ยวข้องกับระบบ ในเว็บไซต์เพื่อใช้สำหรับเผยแพร่ข้อมูลการแข่งขันกีฬาโอลิมปิกซึ่งประกอบด้วย ผู้ดูแลระบบ และ ผู้ใช้ระบบ ซึ่งมีรายละเอียด ดังนี้

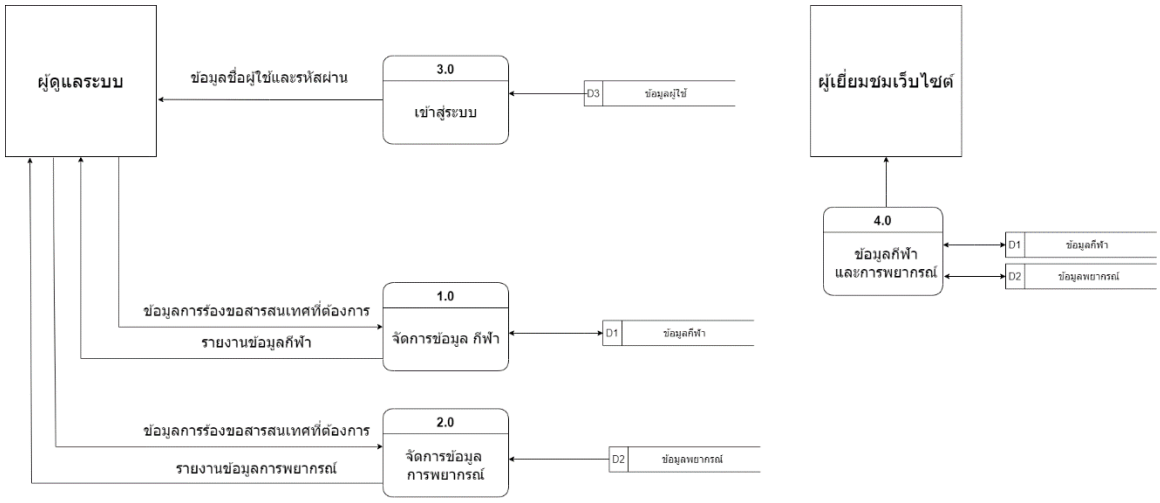
3.3.1.1 ผู้ดูแลระบบ จะทำการลงชื่อเข้าใช้ในระบบเพื่อ แก้ไข เพิ่ม ลบ ข้อมูล ต่าง ๆ และเข้าถึงข้อมูลและรายงานต่าง ๆ ที่มีในเว็บไซต์

3.3.1.2 ผู้เยี่ยมชมเว็บไซต์ จะสามารถเข้าถึงข้อมูลและรายงานต่าง ๆ ที่มีในเว็บไซต์



ภาพที่ 3.24 แผนภาพบริบท (Context Diagram) ของเว็บไซต์เพื่อใช้สำหรับเผยแพร่ข้อมูลการแข่งขันกีฬาโอลิมปิก

3.3.2 แผนภาพกระแสข้อมูลระดับที่ 0 (Data Flow Diagram Level 0)



ภาพที่ 3.25 แผนภาพกระแสข้อมูลระดับที่ 0 (Data Flow Diagram Level 0)

จากภาพที่ 3.25 แสดงแผนภาพกระแสข้อมูลระดับที่ 0 (Data Flow Diagram Level 0) ระบบจัดการข้อมูลการแข่งขันกีฬาโอลิมปิก ประกอบด้วย กระบวนการหลัก ดังนี้

ตารางที่ 3.17 แสดงคำอธิบายกระบวนการจัดการข้อมูลกีฬา

Process Description	
System	การวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิก เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์
DFD Number	1
Process Name	จัดการข้อมูลกีฬา
Input Data Flow	ข้อมูลกีฬา
Output Data Flow	ข้อมูลกีฬา รายงานข้อมูลกีฬา
Data Store Used	ข้อมูลกีฬา
Description	กระบวนการจัดการข้อมูลกีฬา

ตารางที่ 3.18 แสดงคำอธิบายกระบวนการจัดการข้อมูลการพยากรณ์

Process Description	
System	การวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิก เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์
DFD Number	2
Process Name	จัดการข้อมูลการพยากรณ์
Input Data Flow	ข้อมูลการพยากรณ์
Output Data Flow	ข้อมูลการพยากรณ์ รายงานข้อมูลการพยากรณ์
Data Store Used	ข้อมูลการพยากรณ์
Description	กระบวนการจัดการข้อมูลการพยากรณ์

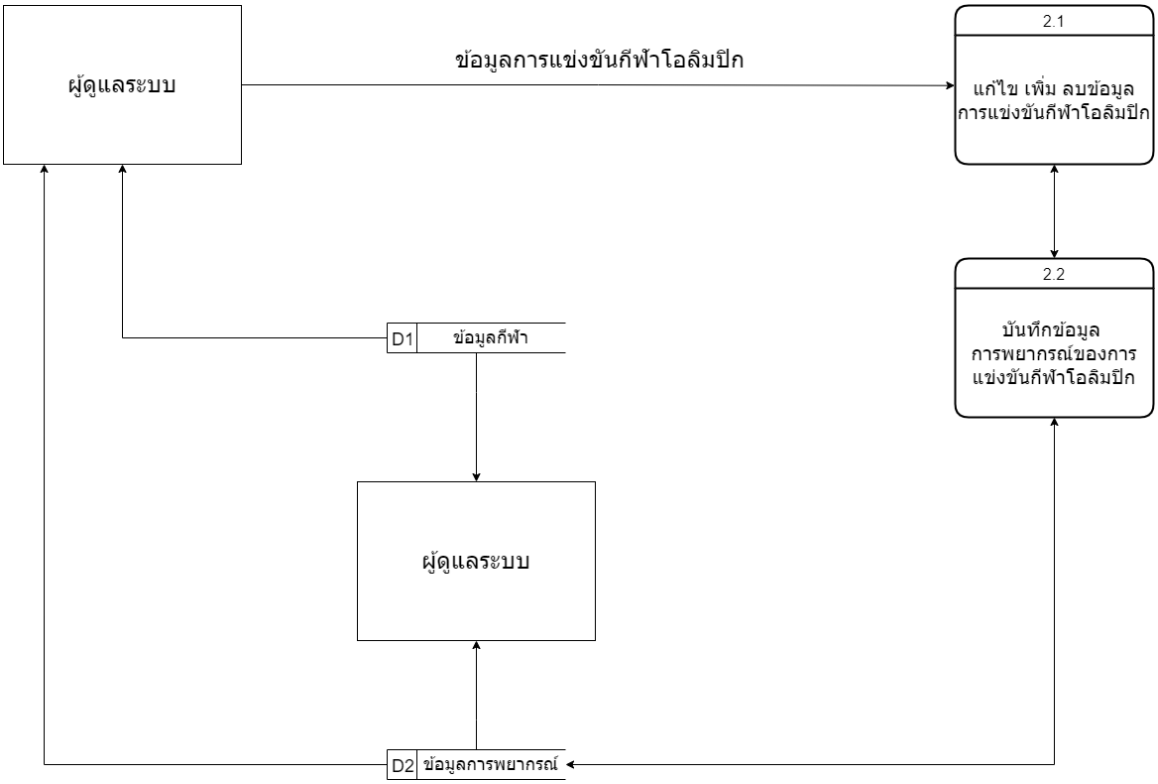
ตารางที่ 3.19 แสดงคำอธิบายกระบวนการเข้าสู่ระบบ

Process Description	
System	การวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิกเพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์
DFD Number	3
Process Name	เข้าสู่ระบบ
Input Data Flow	ข้อมูลกีฬาและการพยากรณ์
Output Data Flow	ข้อมูลชื่อและรหัสผ่าน
Data Store Used	ข้อมูลผู้ใช้
Description	กระบวนการเข้าสู่ระบบ

ตารางที่ 3.20 แสดงคำอธิบายกระบวนการแสดงข้อมูลกีฬาและการพยากรณ์

Process Description	
System	การวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิกเพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์
DFD Number	4
Process Name	ข้อมูลกีฬาและการพยากรณ์
Input Data Flow	ข้อมูลกีฬาและการพยากรณ์
Output Data Flow	ข้อมูลกีฬาและการพยากรณ์
Data Store Used	ข้อมูลกีฬาและการพยากรณ์
Description	กระบวนการแสดงข้อมูลกีฬาและการพยากรณ์

จากแผนภาพกระแสข้อมูลระดับที่ 0 กระบวนการที่ 2 สามารถแยกย่อยเป็นกระบวนการย่อยระดับที่ 1 ได้ดังนี้ แผนภาพกระแสข้อมูลระดับที่ 1 กระบวนการ จัดการข้อมูลการแข่งขันกีฬาโอลิมปิก ข้อมูลการพยากรณ์การแข่งขันกีฬาโอลิมปิก



ภาพที่ 3.26 แผนภาพกระแสข้อมูลระดับที่ 1 (Data Flow Diagram Level 1)

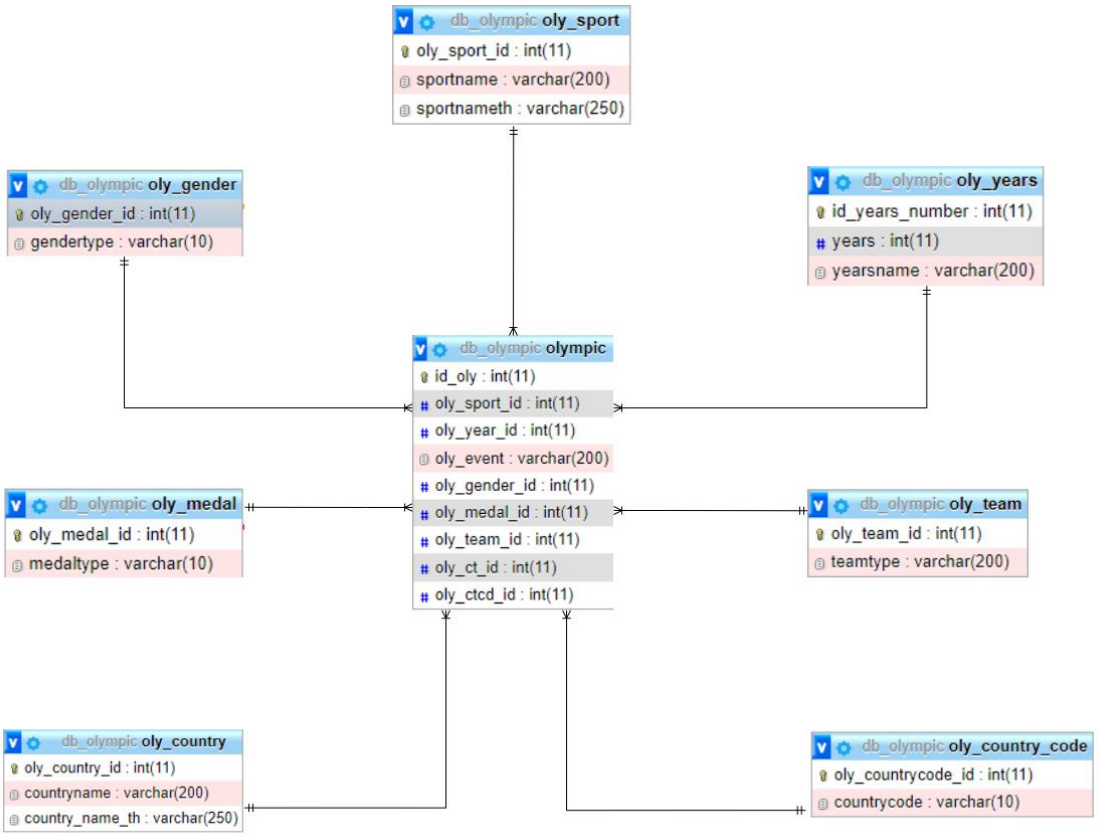
ตารางที่ 3.21 แสดงคำอธิบายกระบวนการที่ 2.1 แก้ไข เพิ่ม ลบข้อมูลการแข่งขันกีฬาโอลิมปิก

Process Description	
System	การวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิก เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์
DFD Number	2.1
Process Name	แก้ไข เพิ่ม ลบข้อมูลการแข่งขันกีฬาโอลิมปิก
Input Data Flow	ข้อมูลการแข่งขันกีฬาโอลิมปิก
Output Data Flow	ข้อมูลการแข่งขันกีฬาโอลิมปิก
Data Store Used	ข้อมูลกีฬา ข้อมูลการพยากรณ์
Description	แก้ไข เพิ่ม ลบข้อมูลการแข่งขันกีฬาโอลิมปิก ได้โดยผู้มีสิทธิ์เข้าใช้คือ ผู้ดูแลระบบ

ตารางที่ 3.22 แสดงคำอธิบายกระบวนการที่ 2.2 บันทึกข้อมูล การพยากรณ์ของการแข่งขันกีฬาโอลิมปิก

Process Description	
System	การวิเคราะห์ข้อมูลการแข่งขันกีฬาโอลิมปิก เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์
DFD Number	2.2
Process Name	บันทึกข้อมูลของการพยากรณ์ของการแข่งขันกีฬาโอลิมปิก
Input Data Flow	ข้อมูลกีฬา ข้อมูลการพยากรณ์
Output Data Flow	ข้อมูลการพยากรณ์การแข่งขันกีฬาโอลิมปิก
Data Store Used	ข้อมูลกีฬา ข้อมูลการพยากรณ์
Description	บันทึกข้อมูลของการพยากรณ์ของการแข่งขันกีฬาโอลิมปิก

3.3.3 ความสัมพันธ์ของฐานข้อมูล (ER Diagram)



ภาพที่ 3.26 ความสัมพันธ์ของฐานข้อมูล (ER Diagram)

จากภาพ 3.26 แผนภาพแสดงความสัมพันธ์ระหว่างข้อมูล (Entity Relationship Diagram) ซึ่งเป็นการออกแบบในระดับConceptual Design ที่ใช้แสดงความสัมพันธ์ระหว่าง Entity ใน Entity ที่เชื่อมความสัมพันธ์ระหว่างกัน และใช้แสดง Attributeต่างๆ ที่มีในแต่ละ Entity ซึ่งสามารถระบุได้ว่า Attribute นั้นจัดเป็น Attribute ที่เป็นคีย์หลัก (Primary Key) หรือ Attributeที่เป็นคีย์รอง (Foreign Key) หรือจัดเป็น Attribute ที่ไม่ใช่ทั้งคีย์หลักและคีย์รอง

3.3.4 พจนานุกรมข้อมูล (Data Dictionary)

จากการออกแบบระบบงานใหม่ซึ่งมีการจัดการระบบฐานข้อมูล จึงสามารถออกแบบฐานข้อมูลให้กับระบบที่ประกอบไปด้วยตารางข้อมูลต่าง ๆ ดังนี้

ตารางที่ 3.23 แสดงชื่อตารางทั้งหมดของระบบฐานข้อมูล

ลำดับ	ชื่อตาราง	ประเภท	รายละเอียด
1	olympic	Master	ข้อมูลโอลิมปิกทั้งหมด
2	oly_sport	Reference	ข้อมูลกีฬาแข่งขัน
3	oly_years	Reference	ข้อมูลปีการแข่งขัน
4	oly_team	Reference	ข้อมูลประเภทการแข่งขัน
5	oly_country_code	Reference	ข้อมูลชื่อย่อประเทศ
6	oly_country	Reference	ข้อมูลชื่อประเทศ
7	oly_medal	Reference	ข้อมูลเหรียญ
8	oly_gender	Reference	ข้อมูลเพศ

คำอธิบาย ประเภทของตาราง ได้แก่ Transaction หมายถึง ตารางที่มีการเปลี่ยนแปลงของข้อมูล Reference หมายถึง ตารางที่มีการอ้างอิงถึงข้อมูล Master ตารางข้อมูลหลัง

จากตารางระบบฐานข้อมูลสามารถแสดงรายละเอียดได้ดังนี้

ตารางที่ 3.24 แสดงรายละเอียดของตาราง olympic

ชื่อตาราง : olympic ประเภทตาราง : Master คำอธิบาย: ข้อมูลโอลิมปิกทั้งหมด คีย์หลัก : id_oly คีย์รอง : id_oly			
เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
id_oly	int (11)	รหัสข้อมูลการแข่งขันทั้งหมด	1
oly_sport_id	tinyint (3)	รหัสข้อมูลกีฬาที่มีการแข่งขัน	53 = shooting
oly_year_id	tinyint (4)	รหัสข้อมูลปีการแข่งขัน	22 = tokyo
oly_event	varchar (200)	ข้อมูลประเภทกีฬาที่แข่งขัน	Trap Mixed Team
oly_gender_id	tinyint (1)	รหัสข้อมูลเพศ	3 = Mixed
oly_medal_id	int (1)	รหัสข้อมูลเหรียญ	1 = gold
oly_team_id	int (1)	รหัสข้อมูลประเภทการแข่งขัน	2 = GameTeam
oly_ct_id	int (3)	รหัสข้อมูลประเทศ	122 = Spain
oly_ctcd_id	int (3)	รหัสข้อมูลชื่อย่อประเทศ	122 = ESP

ตารางที่ 3.25 แสดงรายละเอียดของตาราง oly_sport

ชื่อตาราง : oly_sport ประเภทตาราง : Reference คำอธิบาย: ข้อมูลกีฬาแข่งขัน คีย์หลัก : oly_sport_id คีย์รอง : oly_sport_id			
เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
oly_sport_id	tinyint (3)	รหัสข้อมูลกีฬาที่มีการแข่งขัน	53 = shooting
sportname	varchar (200)	ชื่อกีฬาภาษาอังกฤษ	Archery
sportnameth	varchar (250)	ชื่อกีฬาภาษาไทย	ยิงธนู

ตารางที่ 3.26 แสดงรายละเอียดของตาราง oly_years

ชื่อตาราง : oly_years ประเภทตาราง : Reference คำอธิบาย: ข้อมูลปีการแข่งขัน คีย์หลัก : id_years_number คีย์รอง : id_years_number			
เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
id_years_number	tinyint (2)	รหัสข้อมูลปีการแข่งขัน	1
years	tinyint (4)	ปีการแข่งขัน	1896
yearsname	varchar (200)	ชื่อปีการแข่งขัน	athens

ตารางที่ 3.27 แสดงรายละเอียดของตาราง oly_team

ชื่อตาราง : oly_team ประเภทตาราง : Reference คำอธิบาย: ข้อมูลประเภทการแข่งขัน คีย์หลัก : oly_team_id คีย์รอง : oly_team_id			
เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
oly_team_id	tinyint (1)	รหัสข้อมูลประเภทการแข่งขัน	1
teamtype	varchar (200)	ชื่อประเภทการแข่งขัน	Athlete

ตารางที่ 3.28 แสดงรายละเอียดของตาราง oly_country_code

ชื่อตาราง : oly_country_code ประเภทตาราง : Reference คำอธิบาย: ข้อมูลชื่อย่อประเทศ คีย์หลัก : oly_countrycode_id คีย์รอง : oly_countrycode_id			
เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
oly_countrycode_id	tinyint (3)	รหัสข้อมูลชื่อย่อประเทศ	17
countrycode	varchar (3)	รหัสประเทศ	BRA

ตารางที่ 3.29 แสดงรายละเอียดของตาราง oly_country

ชื่อตาราง : oly_country ประเภทตาราง : Reference คำอธิบาย: ข้อมูลชื่อประเทศ คีย์หลัก : oly_country_id คีย์รอง : oly_country_id			
เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
oly_country_id	tinyint (3)	รหัสข้อมูลชื่อประเทศ	17
countryname	varchar (200)	ชื่อประเทศภาษาอังกฤษ	Brazil
country_name_th	varchar (250)	ชื่อประเทศภาษาไทย	บราซิล

ตารางที่ 3.30 แสดงรายละเอียดของตาราง oly_medal

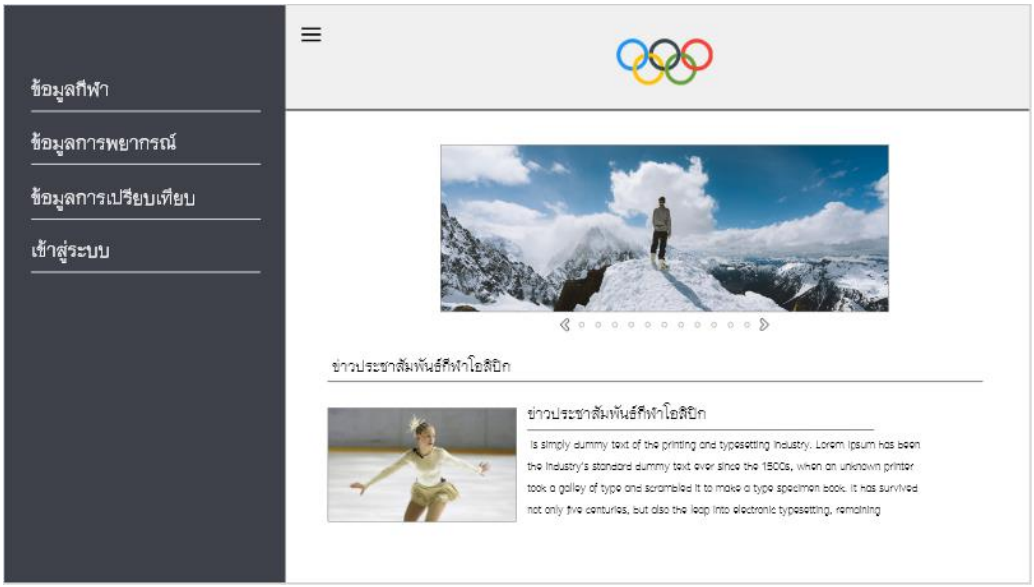
ชื่อตาราง : oly_medal ประเภทตาราง : Reference คำอธิบาย: ข้อมูลเหรียญ คีย์หลัก : oly_medal คีย์รอง : oly_medal			
เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
oly_medal_id	tinyint (1)	รหัสข้อมูลเหรียญ	1
medaltype	varchar (10)	ชื่อเหรียญ	Gold

ตารางที่ 3.31 แสดงรายละเอียดของตาราง oly_gender

ชื่อตาราง : oly_gender ประเภทตาราง : Reference คำอธิบาย: ข้อมูลเพศ คีย์หลัก : oly_gender_id คีย์รอง : oly_gender_id			
เขตข้อมูล	ชนิดและขนาด	ความหมาย	ตัวอย่าง
oly_gender_id	tinyint (1)	รหัสข้อมูลเพศ	3
gendertype	varchar (10)	ชื่อเพศ	Mixed (เพศผสม)

3.4 การออกแบบ Wireframe หน้าจอเว็บไซต์

3.4.1 หน้า Index ของเว็บไซต์ แสดงเมนูต่าง ๆ ของเว็บไซต์

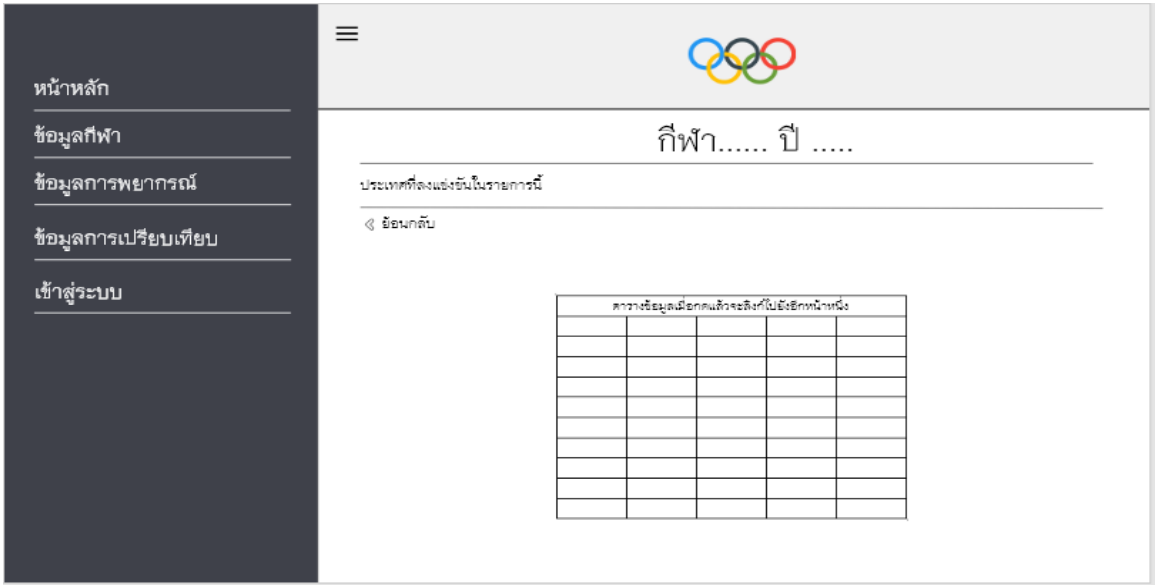


ภาพที่ 3.27 Wireframe หน้า Index ของเว็บไซต์

3.4.2 หน้า ข้อมูลกีฬา แสดงตารางกีฬาของแต่ละปี สามารถเลือกปีที่จะแสดงได้

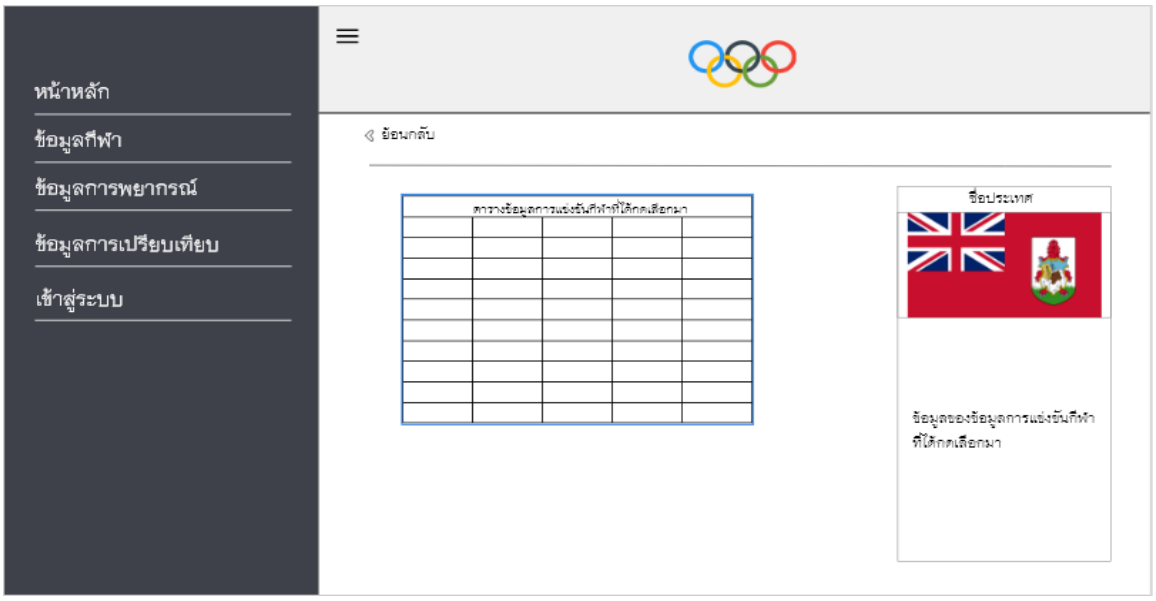
ภาพที่ 3.28 หน้า ข้อมูลกีฬา แสดงตารางกีฬาของแต่ละปี

3.4.3 หน้า ข้อมูลกีฬา แสดงตารางประเทศที่เข้าร่วมในปีนั้นและกีฬาที่ได้เลือกมา



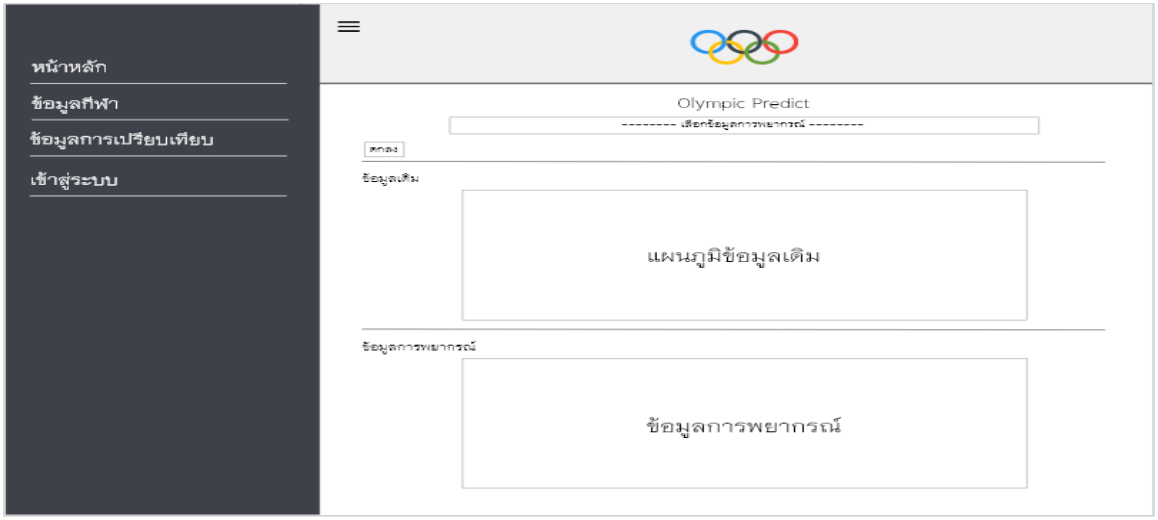
ภาพที่ 3.29 หน้า ข้อมูลกีฬา แสดงตารางประเทศที่เข้าร่วม

3.4.4 หน้า ข้อมูลกีฬา แสดงข้อมูลต่าง ๆ ที่ได้กดเลือกมาจาก 2 Wireframe ก่อนหน้า



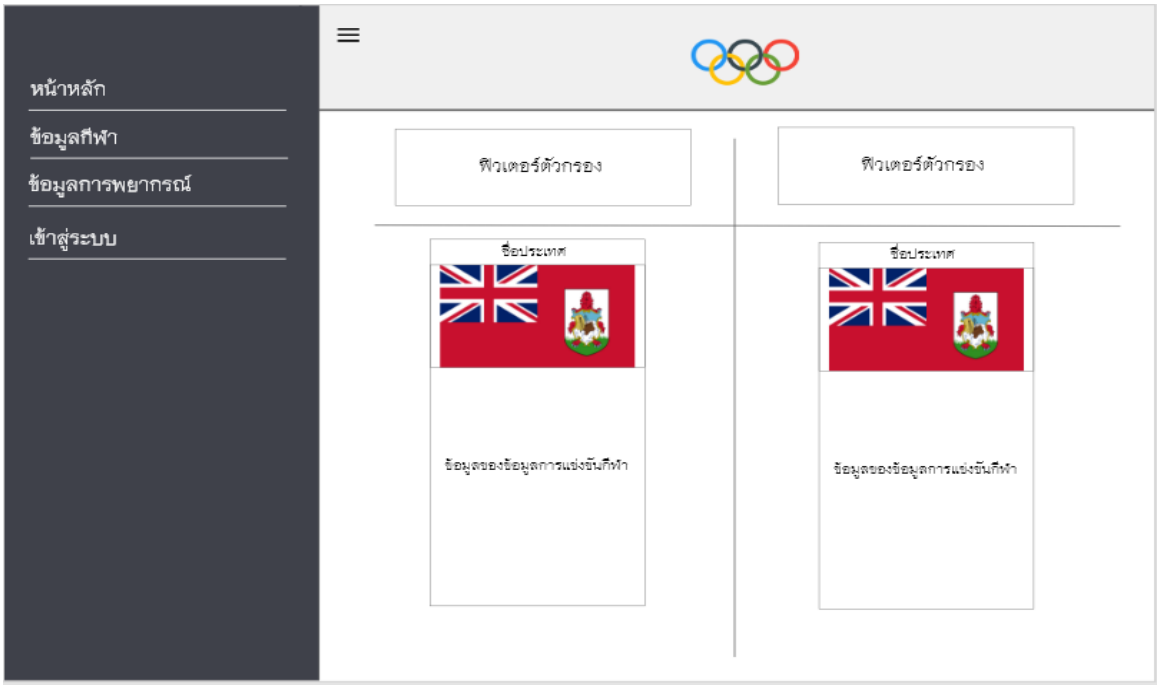
ภาพที่ 3.30 ข้อมูลกีฬา แสดงข้อมูลต่าง ๆ ที่ได้กดเลือก

3.4.5 หน้า ข้อมูลการพยากรณ์ แสดงข้อมูลแผนภูมิของข้อมูลการพยากรณ์



ภาพที่ 3.31 หน้า ข้อมูลการพยากรณ์ แสดงข้อมูลแผนภูมิของข้อมูลการพยากรณ์

3.4.6 หน้า ข้อมูลเปรียบเทียบ แสดงข้อมูลกีฬาแบบย่อเพื่อเปรียบเทียบ



ภาพที่ 3.31 หน้า ข้อมูลเปรียบเทียบ แสดงข้อมูลกีฬาแบบย่อเพื่อเปรียบเทียบ