

บทที่ 3

วิธีการดำเนินงานโครงการ

โครงการเรื่อง การวิเคราะห์ข้อมูลรายงานการฆาตกรรมปี 2011-2013 เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์ ซึ่งในบทนี้จะเป็นการวิเคราะห์ข้อมูลเทคนิค Data Mining เมื่อเสร็จสิ้นกระบวนการวิเคราะห์แล้วจะเป็นขั้นตอนการออกแบบเว็บไซต์ รูปแบบการแสดงผลและบทสรุปจากวิธีการดำเนินงาน ดังนี้

3.1 การวิเคราะห์ข้อมูลด้วย CRISP-DM

3.2 การออกแบบเว็บไซต์

3.3 บทสรุป

3.1 การวิเคราะห์ข้อมูลด้วย CRISP-DM

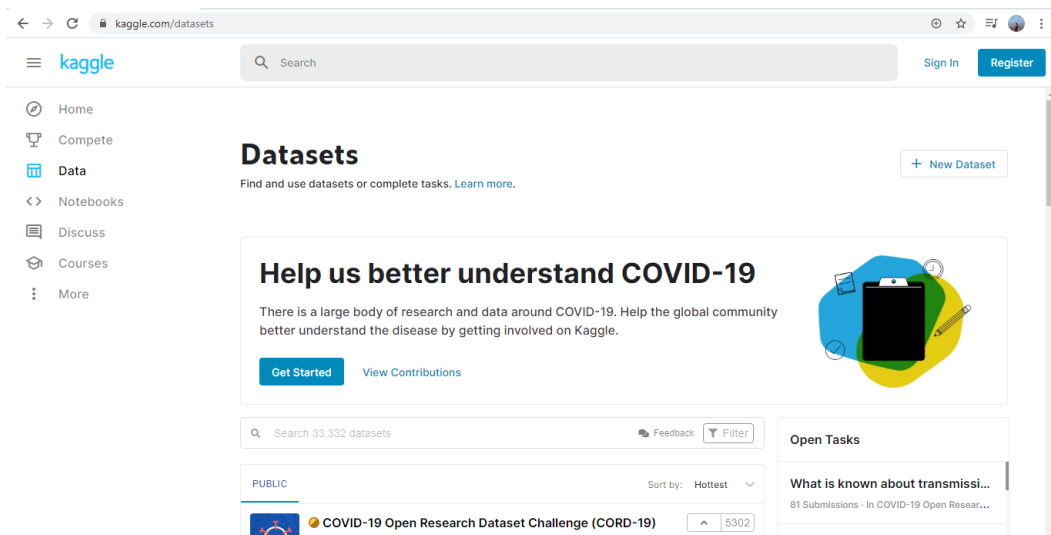
งานวิทยาการด้านข้อมูล (Data Science) ซึ่งกำลังมีบทบาทในยุคปัจจุบัน และทวีความสำคัญยิ่งขึ้นในอนาคต ได้มี CRISP-DM เป็นกระบวนการหลักในการจัดทำเหมืองข้อมูลเพื่อการวิเคราะห์และใช้ประโยชน์ในทางธุรกิจ กระบวนการวิเคราะห์ข้อมูล ด้วย CRISP-DM หรือ Cross Industry Standard Process for Data Mining พัฒนาขึ้นในปี ค.ศ. 1996 โดยความร่วมมือของ 3 บริษัทคือ Daimler Chrysler, SPSS และ NCR ประกอบด้วย 6 ขั้นตอนหลัก ได้แก่

3.1.1 รู้จักและเข้าใจในธุรกิจ (Business understanding) เป็นขั้นตอนแรกของกระบวนการที่มุ่งเน้นไปที่การทำความเข้าใจกระบวนการทางธุรกิจโดยรวม

ผู้วิเคราะห์ข้อมูลทำความเข้าใจกับปัญหาให้อยู่ในรูปของการวิเคราะห์ข้อมูลทางดาต้าไมน์นิ่งโดยการวิเคราะห์ข้อมูลรายงานการฆาตกรรมปี 2011-2013 ซึ่งมีข้อมูลทั้งหมด 638,455 รายการ ซึ่งเป็นข้อมูลที่มีจำนวนมากและทำให้ซับซ้อน

3.1.2 สร้างฐานข้อมูลให้ครบ (Data understanding) ขั้นตอนการจัดเก็บและรวบรวมข้อมูล ตลอดจนการพิจารณาตรวจสอบความถูกต้องของข้อมูลที่ได้รับ โดยเลือกที่จะใช้ข้อมูลทั้งหมดหรือบางส่วนในการวิเคราะห์ให้สอดคล้องกับวัตถุประสงค์ที่กำหนดไว้

ผู้วิเคราะห์ข้อมูลได้ทำการเก็บรวบรวมข้อมูลจากเว็บไซต์ kaggle.com จากนั้น จะทำการตรวจสอบข้อมูลที่ได้ทำการรวบรวมมาได้ เพื่อดูความถูกต้องของข้อมูล และ พิจารณาว่าข้อมูลอาชญากรรม จะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์



ภาพที่ 3.1 เว็บไซต์ kaggle.com

ซึ่งข้อมูลอาชญากรรม มีจำนวน 638,455 รายการ รายการ ประกอบด้วย 24 แดตทรี บิวท์ ข้อมูลหลัก ๆ จะประกอบด้วย ประเภทตัวแทน ปี ประเภทอาชญากรรม เพศเหยื่อ เพศ ผู้กระทำ อาวุธ เป็นต้น

Record ID	Agency	City	State	Year	Month	Incident	Crime Typ	Crime Solv	Victim Sex	Victim Age	Victim Race	Victim Ethr	Perpetrator	Perpetrator Race
1	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	January	1	Murder or Yes	Male	14	Native Am	Unknown	Male	15	Native Am Un
2	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	March	1	Murder or Yes	Male	43	White	Unknown	Male	42	White Un
3	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	March	2	Murder or No	Female	30	Native Am	Unknown	Unknown	0	Unknown Un
4	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	April	1	Murder or Yes	Male	43	White	Unknown	Unknown	42	White Un
5	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	April	2	Murder or No	Female	30	Native Am	Unknown	Unknown	0	Unknown Un
6	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	May	1	Murder or Yes	Male	30	White	Unknown	Male	26	White Un
7	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	May	2	Murder or Yes	Female	42	Native Am	Unknown	Male	27	Black Un
8	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	June	1	Murder or Yes	Female	99	White	Unknown	Male	35	White Un
9	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	June	2	Murder or No	Male	32	White	Unknown	Unknown	0	Unknown Un
10	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	June	3	Murder or Yes	Male	38	White	Unknown	Male	40	Unknown Un
11	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	July	1	Murder or No	Male	36	Native Am	Unknown	Unknown	0	Unknown Un
12	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	July	2	Murder or Yes	Male	20	White	Unknown	Male	49	White Un
13	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	July	3	Murder or Yes	Female	36	Black	Unknown	Male	39	Black Un
14	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	August	1	Murder or Yes	Male	20	Native Am	Unknown	Male	49	White Un
15	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	August	2	Murder or No	Male	48	White	Unknown	Unknown	0	Unknown Un
16	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	August	3	Murder or Yes	Male	31	Black	Unknown	Female	29	Black Un
17	AK00101	Anchorage Municipal P	Anchorage Alaska	1980	December	1	Murder or Yes	Male	16	Unknown	Unknown	Male	19	Unknown Un
18	AK00103	Juneau Municipal P	Juneau Alaska	1980	November	1	Murder or Yes	Male	33	Native Am	Unknown	Male	23	Native Am Un
19	AK00106	Nome Municipal P	Nome Alaska	1980	June	1	Murder or Yes	Male	27	Native Am	Unknown	Male	33	Native Am Un
20	AK00113	Bethel Municipal P	Bethel Alaska	1980	February	1	Murder or Yes	Male	33	Native Am	Unknown	Male	35	Native Am Un
21	AK00118	North Slope County Pol	North Slope Alaska	1980	August	1	Murder or Yes	Female	31	Unknown	Unknown	Male	29	Unknown Un
22	AK00118	North Slope County Pol	North Slope Alaska	1980	August	1	Murder or Yes	Male	21	Unknown	Unknown	Male	29	Unknown Un
23	AK00118	North Slope County Pol	North Slope Alaska	1980	June	1	Murder or Yes	Male	60	Unknown	Unknown	Male	26	Unknown Un

ภาพที่ 3.2 ข้อมูลรายงานการฆาตกรรมทั้งหมด

3.1.3 เตรียมข้อมูลให้พร้อมใช้ (Data preparation) ขั้นตอนการแปลงข้อมูลที่ได้รวบรวมมาและเลือกไว้ ให้อยู่ในรูปแบบที่พร้อมสำหรับนำไปวิเคราะห์ในขั้นตอนต่อไปได้ โดยการทำให้เป็นข้อมูลที่ถูกต้อง (Data cleaning) มักใช้เวลาค่อนข้างมาก

3.1.3.1 ทำการคัดเลือกข้อมูล (Data Selection) คือการคัดเลือกข้อมูลที่เหมาะสมเพื่อนำมาใช้ในการวิเคราะห์ข้อมูล

ผู้วิเคราะห์ข้อมูลทำการคัดเลือกข้อมูลโดยมี ส.ต.ต. วโรดม ไชยชนะ มาช่วยในการพิจารณา และทำการ Data Cleaning ข้อมูลรายงานการฆาตกรรมปี 2011-2013 โดยแยกข้อมูลออกและตัดส่วนที่ไม่จำเป็นออกให้เหลือเฉพาะข้อมูลที่จำเป็นในการวิเคราะห์ในภาพรวม จำนวน 9 แอตทริบิวท์ ได้แก่ ประเภทตัวแทน ปี ประเภทอาชญากรรม เพศเหยื่อ การแข่งขันของเหยื่อ เพศผู้กระทำผิด การแข่งขันผู้กระทำผิด อาวุธ แหล่งบันทึก จำนวน 640 รายการ ซึ่งเป็นข้อมูลที่จำเป็นในการนำไปวิเคราะห์ข้อมูล

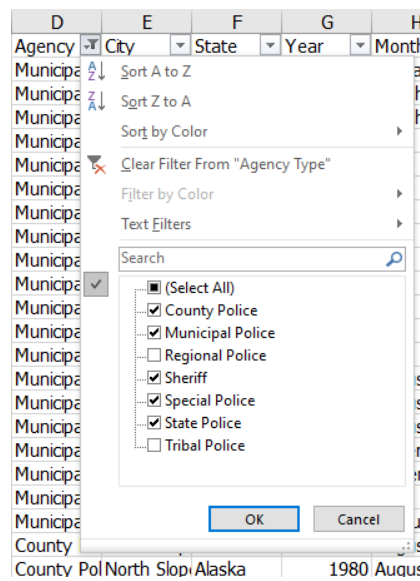
	A	B	C	D	E	F	G	H	I	J	K	L
	ประเภทตัวแทน	ปี	ประเภทอาชญากรรม	เพศเหยื่อ	การแข่งขันของเหยื่อ	เพศผู้กระทำผิด	การแข่งขันผู้กระทำผิด	อาวุธ	แหล่งบันทึก			
	Agency Type	Year	Crime Type	Victim Sex	Victim Race	Perpetrator Sex	Perpetrator Race	Weapon	Record Source			
2	Municipal Police	2011	Manslaughter by Nec	Female	White	Male	White	Drugs	FBI			
3	Municipal Police	2011	Murder or Manslaught	Male	White	Male	White	Gun	FBI			
4	Municipal Police	2011	Murder or Manslaught	Male	White	Male	White	Gun	FBI			
5	Municipal Police	2011	Murder or Manslaught	Male	White	Male	White	Gun	FBI			
6	Municipal Police	2011	Murder or Manslaught	Male	White	Male	White	Gun	FBI			
7	Municipal Police	2011	Murder or Manslaught	Male	Black	Unknown	Unknown	Drugs	FBI			
8	Municipal Police	2011	Murder or Manslaught	Female	Black	Unknown	Unknown	Gun	FBI			
9	Municipal Police	2011	Murder or Manslaught	Male	Asian/Pacific Islander	Male	Black	Drugs	FBI			
10	Municipal Police	2011	Murder or Manslaught	Female	White	Unknown	Unknown	Fire	FBI			
11	Municipal Police	2011	Murder or Manslaught	Male	Asian/Pacific Islander	Unknown	Unknown	Fire	FBI			
12	Municipal Police	2011	Murder or Manslaught	Male	White	Unknown	Unknown	Gun	FBI			
13	Municipal Police	2011	Murder or Manslaught	Female	White	Male	White	Fire	FBI			
14	Municipal Police	2011	Murder or Manslaught	Male	White	Female	White	Drugs	FBI			
15	Municipal Police	2011	Murder or Manslaught	Male	Black	Unknown	Unknown	Fire	FBI			
16	Municipal Police	2011	Murder or Manslaught	Female	Unknown	Unknown	Unknown	Fire	FBI			
17	Municipal Police	2011	Murder or Manslaught	Female	Unknown	Unknown	Unknown	Gun	FBI			
18	Municipal Police	2011	Murder or Manslaught	Male	Unknown	Unknown	Unknown	Fire	FBI			
19	Municipal Police	2011	Murder or Manslaught	Male	White	Male	White	Gun	FOIA			
20	Municipal Police	2011	Murder or Manslaught	Male	White	Male	White	Fire	FOIA			
21	Municipal Police	2011	Murder or Manslaught	Male	White	Female	White	Drugs	FOIA			
22	Municipal Police	2011	Murder or Manslaught	Female	Black	Male	Black	Fire	FOIA			
23	Municipal Police	2011	Murder or Manslaught	Male	Black	Male	Black	Fire	FBI			
24	County Police	2011	Murder or Manslaught	Female	Black	Male	Black	Fire	FBI			

ภาพที่ 3.3 ข้อมูลรายงานการฆาตกรรม

3.1.3.2 ทำการกลั่นกรองข้อมูล (Data Cleaning) คือการทำความสะอาดข้อมูล เป็นกระบวนการตรวจสอบและการแก้ไข (หรือลบ) รายการข้อมูลที่ไม่ถูกต้องออกไปจากชุดข้อมูล ตารางหรือฐานข้อมูล ซึ่งเป็นหลักสำคัญของฐานข้อมูล ทางผู้วิเคราะห์ข้อมูลได้ดำเนินการดังนี้

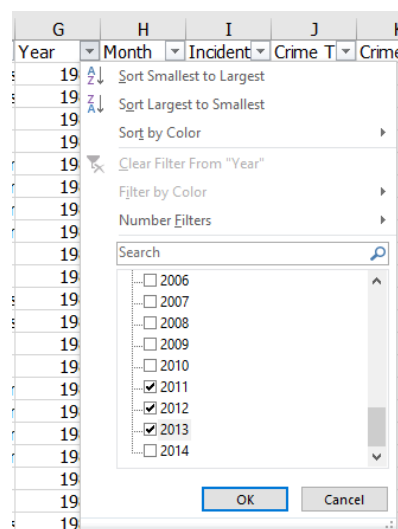
1) ข้อมูลรายงานการฆาตกรรม ผู้วิเคราะห์ข้อมูลได้ทำการแก้ไขและลบข้อมูล ซึ่งผู้วิเคราะห์ข้อมูลพบว่า ข้อมูลทั้งหมดนั้นมีจำนวนที่เยอะจึงเลือกส่วนหัวข้อย่อยในแต่ละแอดทริบิวท์มาเท่านั้น ดังนั้นผู้วิเคราะห์ข้อมูลได้ทำการลบทิ้ง ดังนี้

- Agency Type(ประเภทตัวแทน) ซึ่งมี 7 หัวข้อย่อย ผู้วิเคราะห์พบว่าควรตัด Regional Police(ตำรวจท้องถิ่น)และ Tribal Police(ตำรวจเผ่า) 2 หัวข้อย่อยเพราะ 2 หัวข้อย่อยนี้มีจำนวนที่น้อย 37 รายการ จึงทำการลบทิ้ง ดังนี้



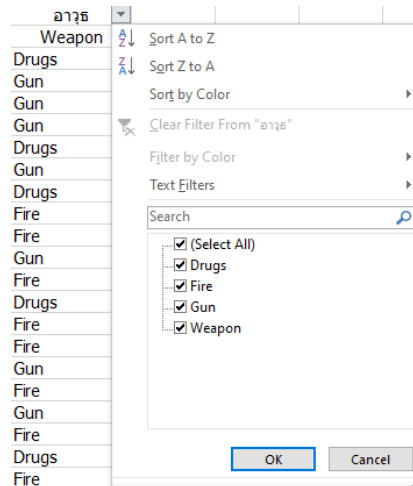
ภาพที่ 3.4 การลบรายการย่อย Agency Type(ประเภทตัวแทน)

- Year(ปี) ข้อมูลทั้งหมดมีตั้งแต่ปี 1980-2014 ผู้วิเคราะห์พบว่าคัดเลือกออกมาแสดงผล 3 ปี คือปี 2011-2013 ดังนี้



ภาพที่ 3.5 การลบรายการย่อย Year(ปี)

- Weapon(อาวุธ) มีรายการอาวุธ 14 รายการ ซึ่งมีจำนวนอาวุธ 44,199 รายการ จึงทำการเลือกอาวุธที่เป็นรายการหลัก 3 รายการ คือ Drugs(ยาเสพติด) Fire(ไฟ) Gun(ปืน) ดังนี้



ภาพที่ 3.6 การลบรายการย่อย Weapon(อาวุธ)

3.1.4 สร้างแบบจำลอง (Modeling) ขั้นตอนการสร้างตัวแบบทางคณิตศาสตร์ และสถิติ เพื่อการวิเคราะห์ข้อมูล โดยสามารถใช้เทคนิควิธีการต่าง ๆ อาทิ การจำแนก (Classification) การแบ่งกลุ่ม (Clustering) และการสร้างความสัมพันธ์ (Association rule)

ผู้วิเคราะห์ข้อมูลวิเคราะห์ข้อมูลด้วยเทคนิคทางดัด้าไมน์นิง แบบการจำแนกประเภทข้อมูล (Classification) โดยการใช้โมเดลการตัดสินใจแบบต้นไม้ (Decision Tree) ซึ่งในขั้นตอนนี้จะถูกนำมาใช้เพื่อให้ได้คำตอบที่ดีที่สุด โดยใช้โปรแกรมที่ใช้ทำเหมืองข้อมูล ด้วยชุดข้อมูลที่คัดเลือก ดังนี้

	A	B	C	D	E	F
2	Agency Type	Year	Crime Type	Victim Sex	Perpetrator Sex	Weapon
3	Municipal Police	2011	Manslaughter by Neg	Female	Male	Drugs
4	Municipal Police	2011	Murder or Manslaugh	Male	Male	Gun
5	Municipal Police	2011	Murder or Manslaugh	Male	Male	Gun
6	Municipal Police	2011	Murder or Manslaugh	Male	Male	Gun
7	Municipal Police	2011	Murder or Manslaugh	Male	Unknown	Drugs
8	Municipal Police	2011	Murder or Manslaugh	Female	Unknown	Gun
9	Municipal Police	2011	Murder or Manslaugh	Male	Male	Drugs
10	Municipal Police	2011	Murder or Manslaugh	Female	Unknown	Fire
11	Municipal Police	2011	Murder or Manslaugh	Male	Unknown	Fire
12	Municipal Police	2011	Murder or Manslaugh	Male	Unknown	Gun
13	Municipal Police	2011	Murder or Manslaugh	Female	Male	Fire
14	Municipal Police	2011	Murder or Manslaugh	Male	Female	Drugs
15	Municipal Police	2011	Murder or Manslaugh	Male	Unknown	Fire
16	Municipal Police	2011	Murder or Manslaugh	Female	Unknown	Fire
17	Municipal Police	2011	Murder or Manslaugh	Female	Unknown	Gun
18	Municipal Police	2011	Murder or Manslaugh	Male	Unknown	Fire
19	Municipal Police	2011	Murder or Manslaugh	Male	Male	Gun
20	Municipal Police	2011	Murder or Manslaugh	Male	Male	Fire
21	Municipal Police	2011	Murder or Manslaugh	Male	Female	Drugs
22	Municipal Police	2011	Murder or Manslaugh	Female	Male	Fire
23	Municipal Police	2011	Murder or Manslaugh	Male	Male	Fire
24	County Police	2011	Murder or Manslaugh	Female	Male	Fire
25	Municipal Police	2011	Murder or Manslaugh	Male	Male	Gun

ภาพที่ 3.7 ชุดข้อมูลที่คัดเลือกมาวิเคราะห์ข้อมูล

จากรูปภาพที่ ประกอบด้วย 6 แอตทริบิวต์ คือ

- Agency Type(ประเภทตัวแทน) ประกอบด้วย 4 ค่า คือ County Police, Municipal Police, Special Police, Special Police
- Year(ปี) ประกอบด้วย 3 ค่า คือ ปี2011 ปี2012 ปี2013
- Crime Type(ประเภทอาชญากรรม) ประกอบด้วย 2 ค่า คือ Murder or Manslaughter, Manslaughter by Negligence
- Victim Sex(เพศเหยื่อ) ประกอบด้วย 2 ค่า คือ Male, Female
- Perpetrator Sex(เพศผู้กระทำผิด) ประกอบด้วย 3 ค่า คือ Male, Female, Unknown
- Weapon(อาวุธ) ประกอบด้วย 3 ค่า คือ Drugs, Fire, Gun

การสร้างโมเดล Decision Tree จะทำการคัดเลือกแอตทริบิวต์ที่มีความสัมพันธ์กับคลาสมากที่สุดขึ้นมาเป็นโหนดบนสุดของ Tree (root node) หลังจากนั้นก็จะหาแอตทริบิวต์ถัดไปเรื่อยๆ ในการหาความสัมพันธ์ของแอตทริบิวต์นี้จะใช้ตัววัด ที่เรียกว่า Information Gain (IG) ค่านี้คำนวณได้จากสมการดังนี้

$$IG(\text{parent}, \text{child}) = \text{entropy}(\text{parent}) - [p(c1) \times \text{entropy}(c1) + p(c2) \times \text{entropy}(c2) + \dots]$$

โดยที่ $\text{entropy}(c1) = -p(c1) \log p(c1)$ และ $p(c2)$ คือ ค่าความน่าจะเป็นของ $c1$

การคำนวณค่าแต่ละแอตทริบิวต์เทียบกับคลาสเพื่อหาแอตทริบิวต์ที่มีค่า IG มากที่สุดมาเป็น Root ของ Decision tree กับจำนวนข้อมูลทั้งหมดโดยใช้ผลลัพธ์เป็น Perpetrator Sex(เพศผู้กระทำผิด) ชาย และหญิง ดังนี้

1) คำนวณค่า IG ของแอตทริบิวต์ Agency Type(ตัวแทนหน่วยงาน) จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned} \text{entropy}(\text{parent}) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.53 \times \log_2(0.53) + 0.19 \times \log_2(0.19) + 0.28 \times \log_2(0.28)] \\ &= -[0.53 \times -0.92 + 0.19 \times -2.40 + 0.28 \times -1.84] \\ &= -[0.49 + 0.46 + 0.52] \end{aligned}$$

$$= 1.47$$

$$\text{entropy(County Police)} = p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.76 \times \log_2(0.76) + 0.08 \times \log_2(0.08) + 0.16 \times \log_2(0.16)]$$

$$= -[0.76 \times -0.40 + 0.08 \times -0.64 + 0.16 \times -2.64]$$

$$= -[0.30 + 0.05 + 0.42]$$

$$= 0.77$$

$$\text{entropy(Municipal Police)} = p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.52 \times \log_2(0.52) + 0.19 \times \log_2(0.19) + 0.30 \times \log_2(0.30)]$$

$$= -[0.52 \times -0.94 + 0.19 \times -2.40 + 0.30 \times -1.74]$$

$$= -[0.49 + 0.46 + 0.52]$$

$$= 1.47$$

$$\text{entropy(Special Police)} = p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.25 \times \log_2(0.25) + 0.25 \times \log_2(0.25) + 0.50 \times \log_2(0.50)]$$

$$= -[0.25 \times -2 + 0.25 \times -2 + 0.50 \times -1]$$

$$= -[0.50 + 0.50 + 0.50]$$

$$= 1.50$$

$$\text{entropy(State Police)} = p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.69 \times \log_2(0.69) + 0.25 \times \log_2(0.25) + 0.06 \times \log_2(0.06)]$$

$$= -[0.69 \times -0.53 + 0.25 \times -2 + 0.06 \times -4.06]$$

$$= -[0.37 + 0.50 + 0.24]$$

$$= 1.11$$

$$\text{IG (parent, child)} = \text{entropy(parent)} - [p(\text{County Police}) \times \text{entropy}(\text{County Police}) + p(\text{Municipal Police}) \times \text{entropy}(\text{Municipal Police}) + p(\text{Special Police}) \times \text{entropy}(\text{Special Police}) + p(\text{State Police}) \times \text{entropy}(\text{State Police})]$$

$$= 1.47 - [0.04 \times 0.77 + 0.90 \times 1.47 + 0.01 \times 1.50 + 0.06 \times 1.11]$$

$$= 1.47 - [0.03 + 1.32 + 0.02 + 0.07]$$

$$= 1.47 - 1.44$$

$$= 0.03$$

2) คำนวณค่า IG ของแอตทริบิวต์ Year(ปี) จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\text{entropy(parent)} = -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.53 \times \log_2(0.53) + 0.19 \times \log_2(0.19) + 0.28 \times \log_2(0.28)]$$

$$= -[0.53 \times -0.92 + 0.19 \times -2.40 + 0.28 \times -1.84]$$

$$= -[0.49 + 0.46 + 0.52]$$

$$= 1.47$$

$$\text{entropy}(2011) = -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.55 \times \log_2(0.55) + 0.15 \times \log_2(0.15) + 0.30 \times \log_2(0.30)]$$

$$= -[0.55 \times -0.86 + 0.15 \times -2.74 + 0.30 \times -1.74]$$

$$= -[0.47 + 0.41 + 0.52]$$

$$= 1.40$$

$$\text{entropy}(2012) = -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$\begin{aligned} &= -[0.53 \times \log_2(0.53) + 0.20 \times \log_2(0.20) + 0.28 \times \\ \log_2(0.28)] \\ &= -[0.53 \times -0.92 + 0.20 \times -2.32 + 0.28 \times -1.84] \\ &= -[0.49 + 0.46 + 0.52] \\ &= 1.47 \end{aligned}$$

$$\begin{aligned} \text{entropy}(2013) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.53 \times \log_2(0.53) + 0.20 \times \log_2(0.20) + 0.27 \times \\ \log_2(0.27)] \\ &= -[0.53 \times 0.92 + 0.20 \times -2.32 + 0.28 \times -1.89] \\ &= -[0.49 + 0.46 + 0.53] \\ &= 1.48 \end{aligned}$$

$$\begin{aligned} \text{IG (parent, child)} &= \text{entropy}(\text{parent}) - [p(2011) \times \text{entropy}(2011) + p(2012) \\ \times \text{entropy}(2012) + p(2013) \times \text{entropy}(2013)] \\ &= 1.47 - [0.28 \times 1.40 + 0.34 \times 1.47 + 0.38 \times 1.48] \\ &= 1.47 - [0.39 + 0.50 + 0.56] \\ &= 1.47 - 1.45 \\ &= 0.02 \end{aligned}$$

3) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) จากข้อมูล
สามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned} \text{entropy}(\text{parent}) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.53 \times \log_2(0.53) + 0.19 \times \log_2(0.19) + 0.28 \times \\ \log_2(0.28)] \\ &= -[0.53 \times -0.92 + 0.19 \times -2.40 + 0.28 \times -1.84] \\ &= -[0.49 + 0.46 + 0.52] \\ &= 1.47 \end{aligned}$$

$$\begin{aligned}
\text{entropy(Murder or Manslaughter)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \\
&\log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.54 \times \log_2(0.54) + 0.16 \times \log_2(0.16) + 0.30 \times \\
&\log_2(0.30)] \\
&= -[0.54 \times -0.89 + 0.16 \times -2.64 + 0.30 \times -1.74] \\
&= -[0.48 + 0.42 + 0.52] \\
&= 1.42
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Manslaughter by Negligence)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \\
&\times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.45 \times \log_2(0.45) + 0.51 \times \log_2(0.51) + 0.04 \times \\
&\log_2(0.04)] \\
&= -[0.45 \times -1.15 + 0.51 \times -0.97 + 0.04 \times -4.64] \\
&= -[0.52 + 0.49 + 0.19] \\
&= 1.20
\end{aligned}$$

$$\begin{aligned}
\text{IG (parent, child)} &= \text{entropy(parent)} - [p(\text{Murder or Manslaughter}) \times \\
&\text{entropy(Murder or Manslaughter)} + p(\text{Manslaughter by Negligence}) \times \text{entropy(Manslaughter} \\
&\text{by Negligence)}] \\
&= 1.47 - [0.91 \times 1.42 + 0.09 \times 1.20] \\
&= 1.47 - [1.29 + 0.11] \\
&= 1.47 - 1.40 \\
&= 0.07
\end{aligned}$$

4) คำนวณค่า IG ของแอตทริบิวต์ Victim Sex(เพศของเหยื่อ) จากข้อมูลสามารถ
คำนวณค่า IG ได้ดังนี้

$$\begin{aligned}
\text{entropy(parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.53 \times \log_2(0.53) + 0.19 \times \log_2(0.19) + 0.28 \times \\
&\log_2(0.28)] \\
&= -[0.53 \times -0.92 + 0.19 \times -2.40 + 0.28 \times -1.84]
\end{aligned}$$

$$\begin{aligned}
&= -[0.49 + 0.46 + 0.52] \\
&= 1.47 \\
\text{entropy(Male)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.51 \times \log_2(0.51) + 0.17 \times \log_2(0.17) + 0.32 \times \\
&\log_2(0.32)] \\
&= -[0.51 \times -0.97 + 0.17 \times -2.56 + 0.32 \times -1.64] \\
&= -[0.49 + 0.44 + 0.52] \\
&= 1.45 \\
\text{entropy(Female)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.59 \times \log_2(0.59) + 0.22 \times \log_2(0.22) + 0.19 \times \\
&\log_2(0.19)] \\
&= -[0.59 \times -0.76 + 0.22 \times -2.18 + 0.19 \times -2.40] \\
&= -[0.45 + 0.48 + 0.46] \\
&= 1.39 \\
\text{IG(parent, child)} &= \text{entropy(parent)} - [p(\text{Male}) \times \text{entropy(Male)} + \\
&p(\text{Female}) \times \text{entropy(Female)}] \\
&= 1.47 - [0.71 \times 1.45 + 0.29 \times 1.39] \\
&= 1.47 - [1.03 + 0.40] \\
&= 1.47 - 1.43 \\
&= 0.04
\end{aligned}$$

5) คำนวณค่า IG ของแอตทริบิวต์ Weapon(อาวุธ) จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned}
\text{entropy(parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.53 \times \log_2(0.53) + 0.19 \times \log_2(0.19) + 0.28 \times \\
&\log_2(0.28)]
\end{aligned}$$

$$\begin{aligned}
&= -[0.53 \times -0.92 + 0.19 \times -2.40 + 0.28 \times -1.84] \\
&= -[0.49 + 0.46 + 0.52] \\
&= 1.47
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Drugs)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.52 \times \log_2(0.52) + 0.29 \times \log_2(0.29) + 0.19 \times \\
&\log_2(0.19)]
\end{aligned}$$

$$\begin{aligned}
&= -[0.52 \times -0.94 + 0.29 \times -1.79 + 0.19 \times -2.40] \\
&= -[0.49 + 0.52 + 0.46] \\
&= 1.47
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Fire)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.51 \times \log_2(0.51) + 0.20 \times \log_2(0.20) + 0.29 \times \\
&\log_2(0.29)]
\end{aligned}$$

$$\begin{aligned}
&= -[0.51 \times -0.97 + 0.20 \times -2.32 + 0.29 \times -2.06] \\
&= -[0.49 + 0.46 + 0.60] \\
&= 1.55
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Gun)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.56 \times \log_2(0.56) + 0.10 \times \log_2(0.10) + 0.34 \times \\
&\log_2(0.34)]
\end{aligned}$$

$$\begin{aligned}
&= -[0.56 \times -0.84 + 0.10 \times -3.32 + 0.34 \times -1.56] \\
&= -[0.47 + 0.33 + 0.53] \\
&= 1.33
\end{aligned}$$

$$\begin{aligned}
\text{IG(parent, child)} &= \text{entropy(parent)} - [p(\text{Drugs}) \times \text{entropy(Drugs)} + \\
&p(\text{Fire}) \times \text{entropy(Fire)} + p(\text{Gun}) \times \text{entropy(Gun)}]
\end{aligned}$$

$$\begin{aligned}
&= 1.47 - [0.29 \times 1.47 + 0.33 \times 1.55 + 0.39 \times 1.33] \\
&= 1.47 - [0.43 + 0.51 + 0.52]
\end{aligned}$$

$$= 1.47 - 1.46$$

$$= 0.01$$

จากการคำนวณค่า IG ของทุกแอตทริบิวต์พบว่าค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) มีค่ามากที่สุด (0.07) ดังนั้นจึงเลือกแอตทริบิวต์ Crime Type ขึ้นมาเป็นโหนด Root และจะต้องทำการแตกกิ่งจาก โหนด root ออกไปจนข้อมูลในแต่ละโหนด มีคลาสคำตอบเดียวกัน จึงทำการสร้างโหนดในระดับถัดไปของแอตทริบิวต์ Crime Type

การคำนวณค่าแต่ละแอตทริบิวต์ในระดับที่ 2 กับจำนวนข้อมูลทั้งหมดโดยใช้ผลลัพธ์เป็น Perpetrator Sex(เพศผู้กระทำผิด)

1) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) และ Agency Type(ตัวแทนหน่วยงาน) ของ Murder or Manslaughter จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned} \text{entropy (parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ & p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.54 \times \log_2(0.54) + 0.16 \times \log_2(0.16) + 0.30 \times \\ & \log_2(0.30)] \\ &= -[0.54 \times -0.89 + 0.16 \times -2.64 + 0.30 \times -1.74] \\ &= -[0.48 + 0.42 + 0.52] \\ &= 1.42 \end{aligned}$$

$$\begin{aligned} \text{entropy(County Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ & p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.83 \times \log_2(0.83) + 0.00 \times \log_2(0.00) + 0.17 \times \\ & \log_2(0.17)] \\ &= -[0.83 \times -0.27 + 0.00 \times -0 + 0.17 \times -2.56] \\ &= -[0.22 + 0 + 0.44] \\ &= 0.66 \end{aligned}$$

$$\begin{aligned} \text{entropy(Municipal Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ & p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.52 \times \log_2(0.52) + 0.16 \times \log_2(0.16) + \\ & 0.32 \times \log_2(0.32)] \end{aligned}$$

$$\begin{aligned}
&= -[0.52 \times -0.94 + 0.16 \times -2.64 + 0.32 \times \\
-1.64] \\
&= -[0.49 + 0.42 + 0.52] \\
&= 1.43
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Special Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.25 \times \log_2(0.25) + 0.25 \times \log_2(0.25) + 0.50 \\
\times \log_2(0.50)] \\
&= -[0.25 \times -2 + 0.25 \times -2 + 0.50 \times -1] \\
&= -[0.50 + 0.50 + 0.50] \\
&= 1.50
\end{aligned}$$

$$\begin{aligned}
\text{entropy(State Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.74 \times \log_2(0.74) + 0.21 \times \log_2(0.21) + 0.06 \times \\
\log_2(0.06)] \\
&= -[0.74 \times -0.43 + 0.21 \times -2.25 + 0.06 \times -4.06] \\
&= -[0.32 + 0.47 + 0.24] \\
&= 1.03
\end{aligned}$$

$$\begin{aligned}
\text{IG (parent, child)} &= \text{entropy(parent)} - [p(\text{County Police}) \times \text{entropy(County} \\
\text{Police)} + p(\text{Municipal Police}) \times \text{entropy(Municipal Police)} + p(\text{Special Police}) \times \text{entropy(Special} \\
\text{Police)} + p(\text{State Police}) \times \text{entropy(State Police)}] \\
&= 1.42 - [0.03 \times 0.66 + 0.90 \times 1.43 + 0.01 \times 1.50 \\
+ 0.06 \times 1.03] \\
&= 1.42 - [0.02 + 1.29 + 0.02 + 0.06] \\
&= 1.42 - 1.39 \\
&= 0.03
\end{aligned}$$

1.1) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) และ Agency Type(ตัวแทนหน่วยงาน) ของ Manslaughter by Negligence จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned}
 \text{entropy (parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
 &= -[0.45 \times \log_2(0.45) + 0.51 \times \log_2(0.51) + 0.04 \times \log_2(0.04)] \\
 &= -[0.45 \times -1.15 + 0.51 \times -0.97 + 0.04 \times -4.64] \\
 &= -[0.52 + 0.49 + 0.19] \\
 &= 1.20
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy(County Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
 &= -[0.57 \times \log_2(0.57) + 0.29 \times \log_2(0.29) + 0.14 \times \log_2(0.14)] \\
 &= -[0.57 \times -0.81 + 0.29 \times -1.79 + 0.14 \times -2.84] \\
 &= -[0.46 + 0.52 + 0.40] \\
 &= 1.38
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy(Municipal Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
 &= -[0.46 \times \log_2(0.46) + 0.52 \times \log_2(0.52) + 0.02 \times \log_2(0.02)] \\
 &= -[0.46 \times -1.12 + 0.52 \times -0.94 + 0.02 \times -5.64] \\
 &= -[0.52 + 0.49 + 0.11] \\
 &= 1.12
 \end{aligned}$$

$$\begin{aligned}
 \text{entropy(Special Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})
 \end{aligned}$$

$$\begin{aligned}
&= -[0.00 \times \log_2(0.00) + 0.00 \times \log_2(0.00) + 0.00 \\
&\times \log_2(0.00)] \\
&= -[0.00 \times -0 + 0.00 \times -0 + 0.00 \times -0] \\
&= -[0.0 + 0.0 + 0.0] \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\text{entropy(State Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.00 \times \log_2(0.00) + 1.00 \times \log_2(1.00) + 0.00 \\
&\times \log_2(0.00)] \\
&= -[0.00 \times -0.00 + 1.00 \times -0 + 0.00 \times -0.00] \\
&= -[0 + 0 + 0] \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\text{IG (parent, child)} &= \text{entropy}(\text{parent}) - [p(\text{County Police}) \times \text{entropy}(\text{County} \\
&\text{Police}) + p(\text{Municipal Police}) \times \text{entropy}(\text{Municipal Police}) + p(\text{Special Police}) \times \text{entropy}(\text{Special} \\
&\text{Police}) + p(\text{State Police}) \times \text{entropy}(\text{State Police})] \\
&= 1.20 - [0.13 \times 1.38 + 0.84 \times 1.12 + 0 \times 0 + 0 \times \\
&0] \\
&= 1.20 - [0.18 + 0.94 + 0 + 0] \\
&= 1.20 - 1.12 \\
&= 0.08
\end{aligned}$$

2) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) และ Year(ปี) ของ Murder or Manslaughter จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned}
\text{entropy (parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.54 \times \log_2(0.54) + 0.16 \times \log_2(0.16) + 0.30 \times \\
&\log_2(0.30)] \\
&= -[0.54 \times -0.89 + 0.16 \times -2.64 + 0.30 \times -1.74] \\
&= -[0.48 + 0.42 + 0.52]
\end{aligned}$$

$$\begin{aligned}
&= 1.42 \\
\text{entropy}(2011) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.54 \times \log_2(0.54) + 0.14 \times \log_2(0.14) + 0.31 \times \\
&\log_2(0.31)] \\
&= -[0.54 \times -0.89 + 0.14 \times -2.84 + 0.31 \times -1.69] \\
&= -[0.48 + 0.40 + 0.52] \\
&= 1.40 \\
\text{entropy}(2012) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.53 \times \log_2(0.53) + 0.16 \times \log_2(0.16) + 0.31 \times \\
&\log_2(0.31)] \\
&= -[0.53 \times -0.92 + 0.16 \times -2.64 + 0.31 \times -1.69] \\
&= -[0.49 + 0.42 + 0.52] \\
&= 1.43 \\
\text{entropy}(2013) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.55 \times \log_2(0.55) + 0.16 \times \log_2(0.16) + 0.29 \times \\
&\log_2(0.29)] \\
&= -[0.55 \times -0.86 + 0.16 \times -2.64 + 0.29 \times -1.79] \\
&= -[0.47 + 0.42 + 0.52] \\
&= 1.41 \\
\text{IG (parent, child)} &= \text{entropy}(\text{parent}) - [p(2011) \times \text{entropy}(2011) + p(2012) \\
&\times \text{entropy}(2012) + p(2013) \times \text{entropy}(2013)] \\
&= 1.42 - [0.29 \times 1.40 + 0.32 \times 1.43 + 0.39 \times 1.41] \\
&= 1.42 - [0.41 + 0.46 + 0.55] \\
&= 1.42 - 1.42 \\
&= 0
\end{aligned}$$

2.1) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) และ Year(ปี) ของ Manslaughter by Negligence จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned} \text{entropy (parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ &p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.45 \times \log_2(0.45) + 0.51 \times \log_2(0.51) + 0.04 \times \\ &\log_2(0.04)] \end{aligned}$$

$$= -[0.45 \times -1.15 + 0.51 \times -0.97 + 0.04 \times -4.64]$$

$$= -[0.52 + 0.49 + 0.19]$$

$$= 1.20$$

$$\begin{aligned} \text{entropy(2011)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ &p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.70 \times \log_2(0.70) + 0.20 \times \log_2(0.20) + 0.10 \times \\ &\log_2(0.10)] \end{aligned}$$

$$= -[0.70 \times -0.51 + 0.20 \times -2.32 + 0.10 \times -3.32]$$

$$= -[0.36 + 0.46 + 0.33]$$

$$= 1.15$$

$$\begin{aligned} \text{entropy(2012)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ &p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.54 \times \log_2(0.54) + 0.43 \times \log_2(0.43) + 0.04 \times \\ &\log_2(0.04)] \end{aligned}$$

$$= -[0.54 \times -0.89 + 0.43 \times -1.22 + 0.04 \times -4.64]$$

$$= -[0.48 + 0.52 + 0.19]$$

$$= 1.19$$

$$\begin{aligned} \text{entropy(2013)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ &p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.18 \times \log_2(0.18) + 0.82 \times \log_2(0.82) + 0.0 \times \\ &\log_2(0.0)] \end{aligned}$$

$$= -[0.18 \times 2.47 + 0.82 \times -0.29 + 0.0 \times -0]$$

$$= -[0.44 + 0.24 + 0.0]$$

$$= 0.68$$

$$\text{IG (parent, child)} = \text{entropy(parent)} - [p(2011) \times \text{entropy}(2011) + p(2012) \times \text{entropy}(2012) + p(2013) \times \text{entropy}(2013)]$$

$$= 1.20 - [0.18 \times 1.15 + 0.51 \times 1.19 + 0.31 \times 0.68]$$

$$= 1.20 - [0.21 + 0.61 + 0.21]$$

$$= 1.20 - 1.03$$

$$= 0.17$$

3) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) และ Victim Sex (เพศของเหยื่อ) ของ Murder or Manslaughter จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\text{entropy (parent)} = -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.54 \times \log_2(0.54) + 0.16 \times \log_2(0.16) + 0.30 \times \log_2(0.30)]$$

$$= -[0.54 \times -0.89 + 0.16 \times -2.64 + 0.30 \times -1.74]$$

$$= -[0.48 + 0.42 + 0.52]$$

$$= 1.42$$

$$\text{entropy(Male)} = -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.52 \times \log_2(0.52) + 0.14 \times \log_2(0.14) + 0.34 \times \log_2(0.34)]$$

$$= -[0.52 \times -0.94 + 0.14 \times -2.84 + 0.34 \times -1.56]$$

$$= -[0.49 + 0.40 + 0.53]$$

$$= 1.42$$

$$\text{entropy(Female)} = -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.60 \times \log_2(0.60) + 0.19 \times \log_2(0.19) + 0.21 \times \log_2(0.21)]$$

$$\log_2(0.21)]$$

$$\begin{aligned}
&= -[0.60 \times -0.74 + 0.19 \times -2.40 + 0.21 \times -2.25] \\
&= -[0.44 + 0.46 + 0.47] \\
&= 1.37
\end{aligned}$$

$$\begin{aligned}
IG(\text{parent, child}) &= \text{entropy}(\text{parent}) - [p(\text{Male}) \times \text{entropy}(\text{Male}) + \\
&p(\text{Female}) \times \text{entropy}(\text{Female})] \\
&= 1.42 - [0.72 \times 1.42 + 0.28 \times 1.37] \\
&= 1.42 - [1.02 + 0.38] \\
&= 1.42 - 1.40 \\
&= 0.02
\end{aligned}$$

3.1) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) และ Victim Sex (เพศของเหยื่อ) ของ Manslaughter by Negligence จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned}
\text{entropy}(\text{parent}) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.45 \times \log_2(0.45) + 0.51 \times \log_2(0.51) + 0.04 \times \\
&\log_2(0.04)] \\
&= -[0.45 \times -1.15 + 0.51 \times -0.97 + 0.04 \times -4.64] \\
&= -[0.52 + 0.49 + 0.19] \\
&= 1.20
\end{aligned}$$

$$\begin{aligned}
\text{entropy}(\text{Male}) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.37 \times \log_2(0.37) + 0.57 \times \log_2(0.57) + 0.07 \times \\
&\log_2(0.07)] \\
&= -[0.37 \times -1.43 + 0.57 \times -0.81 + 0.07 \times -3.84] \\
&= -[0.53 + 0.46 + 0.27] \\
&= 1.26
\end{aligned}$$

$$\begin{aligned}
\text{entropy}(\text{Female}) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})
\end{aligned}$$

$$\begin{aligned}
&= -[0.58 \times \log_2(0.58) + 0.42 \times \log_2(0.42) + 0.0 \times \\
&\log_2(0.0)] \\
&= -[0.58 \times 0.79 + 0.42 \times -1.25 + 0.0 \times -0] \\
&= -[0.46 + 0.53 + 0.0] \\
&= 0.99
\end{aligned}$$

$$\begin{aligned}
\text{IG (parent, child)} &= \text{entropy(parent)} - [p(\text{Male}) \times \text{entropy}(\text{Male}) + \\
&p(\text{Female}) \times \text{entropy}(\text{Female})] \\
&= 1.20 - [0.54 \times 1.26 + 0.46 \times 0.99] \\
&= 1.20 - [0.68 + 0.46] \\
&= 1.20 - 1.14 \\
&= 0.06
\end{aligned}$$

4) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) และ Weapon(อาวุธ) ของ Murder or Manslaughter จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned}
\text{entropy (parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.54 \times \log_2(0.54) + 0.16 \times \log_2(0.16) + 0.30 \times \\
&\log_2(0.30)] \\
&= -[0.54 \times -0.89 + 0.16 \times -2.64 + 0.30 \times -1.74] \\
&= -[0.48 + 0.42 + 0.52] \\
&= 1.42
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Drugs)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.52 \times \log_2(0.52) + 0.27 \times \log_2(0.27) + 0.21 \times \\
&\log_2(0.21)] \\
&= -[0.52 \times -0.94 + 0.27 \times -1.89 + 0.21 \times -2.25] \\
&= -[0.49 + 0.51 + 0.47] \\
&= 1.47
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Fire)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.54 \times \log_2(0.54) + 0.15 \times \log_2(0.15) + 0.30 \times \\
&\log_2(0.30)] \\
&= -[0.54 \times -0.89 + 0.15 \times -2.74 + 0.30 \times -1.74] \\
&= -[0.49 + 0.41 + 0.52] \\
&= 1.42
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Gun)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.56 \times \log_2(0.56) + 0.08 \times \log_2(0.08) + 0.36 \\
&\times \log_2(0.36)] \\
&= -[0.56 \times -0.84 + 0.08 \times -3.64 + 0.36 \times -1.47] \\
&= -[0.47 + 0.29 + 0.53] \\
&= 1.29
\end{aligned}$$

$$\begin{aligned}
\text{IG (parent, child)} &= \text{entropy(parent)} - [p(\text{Drugs}) \times \text{entropy}(\text{Drugs}) + p(\text{Fire}) \\
&\times \text{entropy}(\text{Fire}) + p(\text{Gun}) \times \text{entropy}(\text{Gun})] \\
&= 1.42 - [0.26 \times 1.47 + 0.33 \times 1.42 + 0.40 \times 1.29] \\
&= 1.42 - [0.38 + 0.47 + 0.52] \\
&= 1.42 - 1.37 \\
&= 0.05
\end{aligned}$$

4.1) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) และ Weapon(อาวุธ) ของ Manslaughter by Negligence จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned}
\text{entropy (parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.45 \times \log_2(0.45) + 0.51 \times \log_2(0.51) + 0.04 \times \\
&\log_2(0.04)] \\
&= -[0.45 \times -1.15 + 0.51 \times -0.97 + 0.04 \times -4.64] \\
&= -[0.52 + 0.49 + 0.19]
\end{aligned}$$

$$\begin{aligned}
&= 1.20 \\
\text{entropy(Drugs)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.55 \times \log_2(0.55) + 0.41 \times \log_2(0.41) + 0.03 \times \\
&\log_2(0.03)] \\
&= -[0.55 \times -0.86 + 0.41 \times -1.29 + 0.03 \times -5.06] \\
&= -[0.47 + 0.53 + 0.15] \\
&= 1.15 \\
\text{entropy(Fire)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.08 \times \log_2(0.08) + 0.85 \times \log_2(0.85) + 0.08 \\
&\times \log_2(0.08)] \\
&= -[0.08 \times -3.64 + 0.85 \times -0.23 + 0.08 \times -3.64] \\
&= -[0.29 + 0.20 + 0.29] \\
&= 0.78 \\
\text{entropy(Gun)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.62 \times \log_2(0.62) + 0.38 \times \log_2(0.38) + 0.0 \times \\
&\log_2(0.0)] \\
&= -[0.62 \times -0.69 + 0.38 \times -1.40 + 0.0 \times -0] \\
&= -[0.43 + 0.53 + 0.0] \\
&= 0.96 \\
\text{IG (parent, child)} &= \text{entropy(parent)} - [p(\text{Drugs}) \times \text{entropy(Drugs)} + p(\text{Fire}) \\
&\times \text{entropy(Fire)} + p(\text{Gun}) \times \text{entropy(Gun)}] \\
&= 1.20 - [0.53 \times 1.15 + 0.24 \times 0.78 + 0.24 \times 0.96] \\
&= 1.20 - [0.61 + 0.19 + 0.23] \\
&= 1.20 - 1.03 \\
&= 0.17
\end{aligned}$$

จากการคำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) ต่อ Agency Type(ตัวแทนหน่วยงาน) Year(ปี) Victim Sex (เพศของเหยื่อ) และWeapon(อาวุธ) พบว่าค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) ต่อ Year(ปี)และWeapon(อาวุธ) มีค่ามากที่สุด (0.17) ดังนั้นจึงเลือกแอตทริบิวต์ Year(ปี) ขึ้นมาเป็นโหนดในระดับที่ 2 ต่อจากโหนด Root และทำการแตกกิ่งจากโหนดในระดับที่ 2 ออกไปจนข้อมูลในแต่ละโหนดมีคลาสค่าตอบเดียวกัน

การคำนวณค่าแต่ละแอตทริบิวต์ในระดับที่ 3 กับจำนวนข้อมูลทั้งหมดโดยใช้ผลลัพธ์เป็น Perpetrator Sex(เพศผู้กระทำผิด)

1) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(Manslaughter by Negligence)และ Year(2012) และAgency Type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned} \text{entropy (parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ &p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.54 \times \log_2(0.54) + 0.43 \times \log_2(0.43) + 0.04 \times \\ &\log_2(0.04)] \\ &= -[0.54 \times -0.89 + 0.43 \times -1.22 + 0.04 \times -4.64] \\ &= -[0.48 + 0.52 + 0.19] \\ &= 1.19 \end{aligned}$$

$$\begin{aligned} \text{entropy(County Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ &p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[1.00 \times \log_2(1.00) + 0.00 \times \log_2(0.00) + 0.00 \times \\ &\log_2(0.00)] \\ &= -[1.00 \times -0.00 + 0.00 \times -0 + 0.00 \times -0] \\ &= -[0 + 0 + 0] \\ &= 0 \end{aligned}$$

$$\begin{aligned} \text{entropy(Municipal Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ &p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.54 \times \log_2(0.54) + 0.42 \times \log_2(0.42) + \\ &0.04 \times \log_2(0.04)] \end{aligned}$$

$$\begin{aligned}
&= -[0.54 \times -0.89 + 0.42 \times -1.25 + 0.04 \times \\
-4.64] \\
&= -[0.48 + 0.53 + 0.19] \\
&= 1.20
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Special Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.00 \times \log_2(0.00) + 0.00 \times \log_2(0.00) + 0.00 \\
\times \log_2(0.00)] \\
&= -[0.00 \times -0.00 + 0.00 \times -0.00 + 0.00 \times -0.00] \\
&= -[0.00 + 0.00 + 0.00] \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\text{entropy(State Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.00 \times \log_2(0.00) + 1.00 \times \log_2(0.00) + 0.00 \\
\times \log_2(0.00)] \\
&= -[0.00 \times -0.00 + 1.00 \times -0.00 + 0.00 \times -0.00] \\
&= -[0.00 + 0.00 + 0.00] \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\text{IG (parent, child)} &= \text{entropy(parent)} - [p(\text{County Police}) \times \text{entropy(County} \\
\text{Police)} + p(\text{Municipal Police}) \times \text{entropy(Municipal Police)} + p(\text{Special Police}) \times \text{entropy(Special} \\
\text{Police)} + p(\text{State Police}) \times \text{entropy(State Police)}] \\
&= 1.19 - [0.07 \times 0 + 0.86 \times 1.20 + 0.00 \times 0.00 + \\
0.07 \times 0] \\
&= 1.19 - [0.00 + 1.03 + 0.00 + 0.00] \\
&= 1.19 - 1.03 \\
&= 0.16
\end{aligned}$$

1.1) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(Murder or Manslaughter)และ Year(2012) และ Agency Type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned}
\text{entropy (parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.53 \times \log_2(0.53) + 0.16 \times \log_2(0.16) + 0.31 \times \\
&\log_2(0.31)] \\
&= -[0.53 \times -0.92 + 0.16 \times -2.64 + 0.31 \times -1.69] \\
&= -[0.49 + 0.42 + 0.52] \\
&= 1.43
\end{aligned}$$

$$\begin{aligned}
\text{entropy(County Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.67 \times \log_2(0.67) + 0.00 \times \log_2(0.00) + 0.33 \times \\
&\log_2(0.33)] \\
&= -[0.67 \times -0.58 + 0.00 \times -0 + 0.33 \times -1.60] \\
&= -[0.39 + 0.00 + 0.53] \\
&= 0.92
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Municipal Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.51 \times \log_2(0.51) + 0.16 \times \log_2(0.16) + \\
&0.33 \times \log_2(0.33)] \\
&= -[0.51 \times -0.97 + 0.16 \times -2.64 + 0.33 \times \\
&-1.60] \\
&= -[0.49 + 0.42 + 0.53] \\
&= 1.44
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Special Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.00 \times \log_2(0.00) + 0.50 \times \log_2(0.50) + 0.50 \\
&\times \log_2(0.50)] \\
&= -[0.00 \times -0.00 + 0.50 \times -1 + 0.50 \times -1] \\
&= -[0.00 + 0.50 + 0.50]
\end{aligned}$$

$$= 1$$

$$\text{entropy(State Police)} = p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.77 \times \log_2(0.77) + 0.15 \times \log_2(0.15) + 0.08 \times \log_2(0.08)]$$

$$= -[0.77 \times -0.38 + 0.15 \times -2.74 + 0.08 \times -3.64]$$

$$= -[0.29 + 0.41 + 0.29]$$

$$= 0.99$$

$$\text{IG (parent, child)} = \text{entropy(parent)} - [p(\text{County Police}) \times \text{entropy(County Police)} + p(\text{Municipal Police}) \times \text{entropy(Municipal Police)} + p(\text{Special Police}) \times \text{entropy(Special Police)} + p(\text{State Police}) \times \text{entropy(State Police)}]$$

$$= 1.43 - [0.02 \times 0.92 + 0.91 \times 1.44 + 0.01 \times 1 + 0.07 \times 0.99]$$

$$= 1.43 - [0.02 + 1.31 + 0.01 + 0.07]$$

$$= 1.43 - 1.41$$

$$= 0.02$$

2) Crime Type(Manslaughter by Negligence)และ Year(2012) และ Victim Sex จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\text{entropy (parent)} = -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.54 \times \log_2(0.54) + 0.43 \times \log_2(0.43) + 0.04 \times \log_2(0.04)]$$

$$= -[0.54 \times -0.89 + 0.43 \times -1.22 + 0.04 \times -4.64]$$

$$= -[0.48 + 0.52 + 0.19]$$

$$= 1.19$$

$$\text{entropy(Male)} = -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$\begin{aligned}
&= -[0.40 \times \log_2(0.40) + 0.53 \times \log_2(0.53) + 0.07 \times \\
&\log_2(0.07)] \\
&= -[0.40 \times -1.32 + 0.53 \times -0.92 + 0.07 \times -3.84] \\
&= -[0.53 + 0.49 + 0.27] \\
&= 1.29
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Female)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.69 \times \log_2(0.69) + 0.31 \times \log_2(0.31) + 0.00 \times \\
&\log_2(0.00)] \\
&= -[0.69 \times -0.54 + 0.31 \times -1.69 + 0.00 \times -0.00] \\
&= -[0.37 + 0.52 + 0.00] \\
&= 0.89
\end{aligned}$$

$$\begin{aligned}
\text{IG(parent, child)} &= \text{entropy(parent)} - [p(\text{Male}) \times \text{entropy(Male)} + \\
&p(\text{Female}) \times \text{entropy(Female)}] \\
&= 1.19 - [0.54 \times 1.29 + 0.46 \times 0.89] \\
&= 1.19 - [0.70 + 0.41] \\
&= 1.19 - 1.11 \\
&= 0.08
\end{aligned}$$

2.1) Crime Type(Murder or Manslaughter)และ Year(2012) และ Victim Sex จาก ข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned}
\text{entropy (parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.53 \times \log_2(0.53) + 0.16 \times \log_2(0.16) + 0.31 \times \\
&\log_2(0.31)] \\
&= -[0.53 \times -0.92 + 0.16 \times -2.64 + 0.31 \times -1.69] \\
&= -[0.49 + 0.42 + 0.52] \\
&= 1.43
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Male)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.51 \times \log_2(0.51) + 0.15 \times \log_2(0.15) + 0.33 \times \\
&\log_2(0.33)] \\
&= -[0.51 \times -0.97 + 0.15 \times -2.74 + 0.33 \times -1.60] \\
&= -[0.49 + 0.41 + 0.53] \\
&= 1.43
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Female)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.57 \times \log_2(0.57) + 0.20 \times \log_2(0.20) + 0.24 \times \\
&\log_2(0.24)] \\
&= -[0.57 \times -0.81 + 0.20 \times -2.32 + 0.24 \times -2.06] \\
&= -[0.46 + 0.46 + 0.49] \\
&= 1.41
\end{aligned}$$

$$\begin{aligned}
\text{IG(parent, child)} &= \text{entropy(parent)} - [p(\text{Male}) \times \text{entropy(Male)} + \\
&p(\text{Female}) \times \text{entropy(Female)}] \\
&= 1.43 - [0.76 \times 1.43 + 0.24 \times 1.41] \\
&= 1.43 - [1.09 + 0.34] \\
&= 1.43 - 1.43 \\
&= 0
\end{aligned}$$

3) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(Manslaughter by Negligence)และ Year(2012) และ Weapon จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned}
\text{entropy (parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.54 \times \log_2(0.54) + 0.43 \times \log_2(0.43) + 0.04 \times \\
&\log_2(0.04)] \\
&= -[0.54 \times -0.89 + 0.43 \times -1.22 + 0.04 \times -4.64] \\
&= -[0.48 + 0.52 + 0.19]
\end{aligned}$$

$$\begin{aligned}
&= 1.19 \\
\text{entropy(Drugs)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.63 \times \log_2(0.63) + 0.31 \times \log_2(0.31) + 0.06 \times \\
&\log_2(0.03)] \\
&= -[0.63 \times -0.67 + 0.31 \times -1.69 + 0.06 \times -4.06] \\
&= -[0.42 + 0.52 + 0.24] \\
&= 1.18 \\
\text{entropy(Fire)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.00 \times \log_2(0.00) + 1.00 \times \log_2(1.00) + 0.00 \times \\
&\log_2(0.00)] \\
&= -[0.00 \times -0.00 + 1.00 \times -0.00 + 0.00 \times -0.00] \\
&= -[0.00 + 0.00 + 0.00] \\
&= 0.00 \\
\text{entropy(Gun)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.83 \times \log_2(0.83) + 0.17 \times \log_2(0.17) + 0.0 \times \\
&\log_2(0.0)] \\
&= -[0.83 \times -0.27 + 0.17 \times -2.56 + 0.0 \times -0] \\
&= -[0.22 + 0.44 + 0.0] \\
&= 0.66 \\
\text{IG (parent, child)} &= \text{entropy(parent)} - [p(\text{Drugs}) \times \text{entropy(Drugs)} + p(\text{Fire}) \\
&\times \text{entropy(Fire)} + p(\text{Gun}) \times \text{entropy(Gun)}] \\
&= 1.19 - [0.57 \times 1.18 + 0.21 \times 0.0 + 0.21 \times 0.66] \\
&= 1.19 - [0.67 + 0.0 + 0.14] \\
&= 1.19 - 0.78 \\
&= 0.41
\end{aligned}$$

3.1) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(Murder or Manslaughter)และ Year(2012) และ Weapon จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned} \text{entropy (parent)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ &p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.53 \times \log_2(0.53) + 0.16 \times \log_2(0.16) + 0.31 \times \\ &\log_2(0.31)] \end{aligned}$$

$$= -[0.53 \times -0.92 + 0.16 \times -2.64 + 0.31 \times -1.69]$$

$$= -[0.49 + 0.42 + 0.52]$$

$$= 1.43$$

$$\begin{aligned} \text{entropy(Drugs)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ &p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.45 \times \log_2(0.45) + 0.39 \times \log_2(0.39) + 0.16 \times \\ &\log_2(0.16)] \end{aligned}$$

$$= -[0.45 \times -1.15 + 0.39 \times -1.36 + 0.16 \times -2.64]$$

$$= -[0.52 + 0.53 + 0.42]$$

$$= 1.47$$

$$\begin{aligned} \text{entropy(Fire)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ &p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.55 \times \log_2(0.55) + 0.18 \times \log_2(0.18) + 0.27 \times \\ &\log_2(0.27)] \end{aligned}$$

$$= -[0.55 \times -0.86 + 0.18 \times -2.47 + 0.27 \times -1.89]$$

$$= -[0.47 + 0.44 + 0.51]$$

$$= 1.42$$

$$\begin{aligned} \text{entropy(Gun)} &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\ &p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\ &= -[0.54 \times \log_2(0.54) + 0.09 \times \log_2(0.09) + 0.37 \times \\ &\log_2(0.37)] \end{aligned}$$

$$= -[0.54 \times -0.89 + 0.09 \times -3.47 + 0.37 \times -1.43]$$

$$= -[0.48 + 0.31 + 0.53]$$

$$= 1.32$$

$$\text{IG (parent, child)} = \text{entropy}(\text{parent}) - [p(\text{Drugs}) \times \text{entropy}(\text{Drugs}) + p(\text{Fire}) \times \text{entropy}(\text{Fire}) + p(\text{Gun}) \times \text{entropy}(\text{Gun})]$$

$$= 1.43 - [0.16 \times 1.47 + 0.27 \times 1.42 + 0.57 \times 1.32]$$

$$= 1.43 - [0.24 + 0.38 + 0.75]$$

$$= 1.43 - 1.37$$

$$= 0.06$$

จากการคำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) ต่อ Agency Type(ตัวแทนหน่วยงาน) Year(ปี) Victim Sex (เพศของเหยื่อ) และWeapon(อาวุธ) พบว่าค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) ต่อ Year(ปี)และWeapon(อาวุธ) มีค่ามากที่สุด (0.41) ดังนั้นจึงเลือกแอตทริบิวต์ Weapon(อาวุธ) ขึ้นมาเป็นโหนดในระดับที่ 3 ต่อจากโหนด Root และทำการแตกกิ่งจากโหนดในระดับที่ 3 ออกไปจนข้อมูลในแต่ละโหนดมีคลาสคำตอบเดียวกัน

การคำนวณค่าแต่ละแอตทริบิวต์ในระดับที่ 4 กับจำนวนข้อมูลทั้งหมดโดยใช้ผลลัพธ์เป็น Perpetrator Sex(เพศผู้กระทำผิด)

1) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(Manslaughter by Negligence)และ Year(2012) และWeapon(Drugs) และAgency Type จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\text{entropy (parent)} = -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$= -[0.63 \times \log_2(0.63) + 0.31 \times \log_2(0.31) + 0.06 \times \log_2(0.03)]$$

$$= -[0.63 \times -0.67 + 0.31 \times -1.69 + 0.06 \times -4.06]$$

$$= -[0.42 + 0.52 + 0.24]$$

$$= 1.18$$

$$\text{entropy(County Police)} = p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown})$$

$$\begin{aligned}
&= -[1.00 \times \log_2(1.00) + 0.00 \times \log_2(0.00) + 0.00 \times \\
&\log_2(0.00)] \\
&= -[1.00 \times -0.00 + 0.00 \times -0.00 + 0.00 \times -0.00] \\
&= -[0 + 0 + 0] \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Municipal Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.69 \times \log_2(0.69) + 0.23 \times \log_2(0.23) \\
&+ 0.08 \times \log_2(0.08)] \\
&= -[0.69 \times -0.54 + 0.23 \times -2.12 + 0.08 \times \\
&-3.64] \\
&= -[0.37 + 0.49 + 0.29] \\
&= 1.15
\end{aligned}$$

$$\begin{aligned}
\text{entropy(Special Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.00 \times \log_2(0.00) + 0.00 \times \log_2(0.00) + 0.00 \\
&\times \log_2(0.00)] \\
&= -[0.00 \times -0.00 + 0.00 \times -0.00 + 0.00 \times -0.00] \\
&= -[0 + 0 + 0] \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\text{entropy(State Police)} &= p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 \\
&p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.00 \times \log_2(0.00) + 1.00 \times \log_2(1.00) + 0.00 \times \\
&\log_2(0.00)] \\
&= -[0.00 \times -0.00 + 1.00 \times -0.00 + 0.00 \times -0.00] \\
&= -[0 + 0 + 0] \\
&= 0
\end{aligned}$$

$$\begin{aligned}
IG(\text{parent, child}) &= \text{entropy}(\text{parent}) - [p(\text{County Police}) \times \text{entropy}(\text{County Police}) + p(\text{Municipal Police}) \times \text{entropy}(\text{Municipal Police}) + p(\text{Special Police}) \times \text{entropy}(\text{Special Police}) + p(\text{State Police}) \times \text{entropy}(\text{State Police})] \\
&= 1.18 - [0.06 \times 0.00 + 0.81 \times 1.15 + 0.00 \times 0.00 + 0.13 \times 0.00] \\
&= 1.18 - [0 + 0.93 + 0 + 0] \\
&= 1.18 - 0.93 \\
&= 0.25
\end{aligned}$$

2) คำนวณค่า IG ของแอตทริบิวต์ Crime Type(Manslaughter by Negligence)และ Year(2012) และ Weapon(Drugs) และVictim Sex จากข้อมูลสามารถคำนวณค่า IG ได้ดังนี้

$$\begin{aligned}
\text{entropy}(\text{parent}) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.63 \times \log_2(0.63) + 0.31 \times \log_2(0.31) + 0.06 \times \log_2(0.03)] \\
&= -[0.63 \times -0.67 + 0.31 \times -1.69 + 0.06 \times -4.06] \\
&= -[0.42 + 0.52 + 0.24] \\
&= 1.18
\end{aligned}$$

$$\begin{aligned}
\text{entropy}(\text{Male}) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.38 \times \log_2(0.38) + 0.50 \times \log_2(0.50) + 0.13 \times \log_2(0.13)] \\
&= -[0.38 \times -1.40 + 0.50 \times -1 + 0.13 \times -2.94] \\
&= -[0.53 + 0.50 + 0.38] \\
&= 1.41
\end{aligned}$$

$$\begin{aligned}
\text{entropy}(\text{Female}) &= -p(\text{Male}) \times \log_2 p(\text{Male}) + p(\text{Female}) \times \log_2 p(\text{Female}) + p(\text{Unknown}) \times \log_2 p(\text{Unknown}) \\
&= -[0.88 \times \log_2(0.88) + 0.13 \times \log_2(0.13) + 0.00 \times \log_2(0.00)]
\end{aligned}$$

$$\begin{aligned}
&= -[0.88 \times -0.18 + 0.13 \times -2.94 + 0.00 \times -0.00] \\
&= -[0.16 + 0.38 + 0] \\
&= 0.54
\end{aligned}$$

$$\begin{aligned}
IG(\text{parent, child}) &= \text{entropy}(\text{parent}) - [p(\text{Male}) \times \text{entropy}(\text{Male}) + \\
&p(\text{Female}) \times \text{entropy}(\text{Female})] \\
&= 1.18 - [0.50 \times 1.41 + 0.50 \times 0.54] \\
&= 1.18 - [0.71 + 0.27] \\
&= 1.18 - 0.98 \\
&= 0.2
\end{aligned}$$

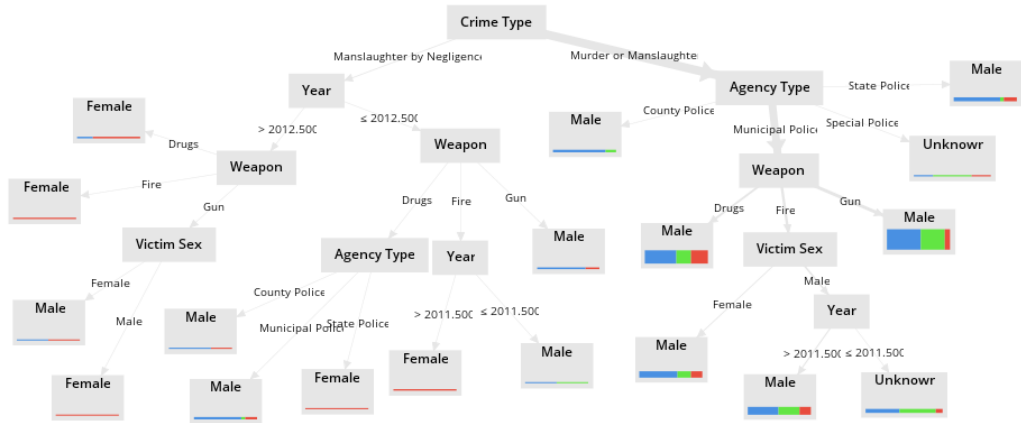
จากการคำนวณค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) ต่อ Agency Type(ตัวแทนหน่วยงาน) Year(ปี) Victim Sex (เพศของเหยื่อ) และWeapon(อาวุธ) พบว่าค่า IG ของแอตทริบิวต์ Crime Type(ประเภทอาชญากรรม) และ Year(ปี) และWeapon(อาวุธ) ต่อ Agency Type(ตัวแทนหน่วยงาน) มีค่ามากที่สุด (0.25) ดังนั้นจึงเลือกแอตทริบิวต์ Agency Type(ตัวแทนหน่วยงาน) ขึ้นมาเป็นโหนดในระดับที่ 4 ต่อจากโหนดระดับที่ 3 และโหนดสุดท้ายคือแอตทริบิวต์ Victim Sex (เพศของเหยื่อ) ทำการแตกกิ่งจากโหนดในระดับที่ 3 ออกไป

จากการคำนวณค่า IG ของแอตทริบิวต์ Crime Type และ Year และ Weapon และ Agency Type และ Victim Sex พบว่าข้อมูลในแต่ละโหนดมีคลาสค่าตอบเดียวกันแล้ว คือผลลัพธ์เป็น Perpetrator Sex

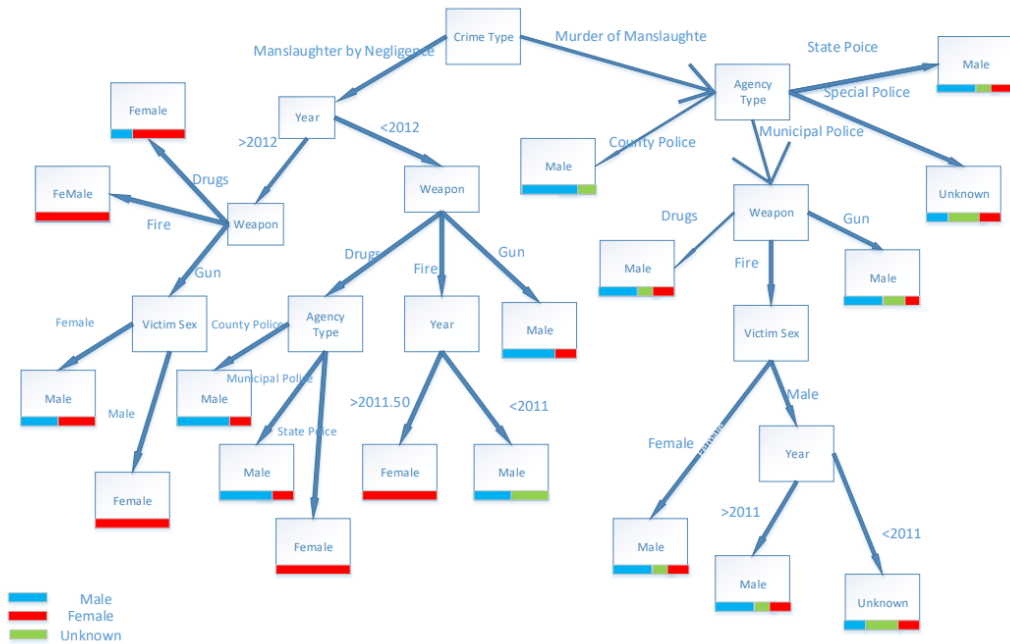
3.1.5 การประเมินผล (Evaluation) เป็นขั้นตอนก่อนนำผลลัพธ์ที่ได้จากขั้นตอนที่ 4 ไปใช้งาน ด้วยการวัดประสิทธิผลของผลลัพธ์ที่ได้กับวัตถุประสงค์ที่ตั้งไว้ในขั้นตอนแรก ว่ามีนัยสำคัญหรือความน่าเชื่อถือมากน้อยเพียงใด ด้วยการประเมินผลจากโปรแกรมว่าถูกต้องหรือไม่

ผู้วิเคราะห์ข้อมูลได้ข้อมูลที่ผ่านการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining จากการสร้างโมเดล Decision Tree ดังนี้

1.โปรแกรม RapidMiner Studio 9.6



ภาพที่ 3.8 รูปแบบโมเดล Graph Decision Tree ในโปรแกรม RapidMiner



ภาพที่ 3.9 รูปแบบโมเดล Graph Decision Tree ในการคำนวณมือ

Tree

```
Crime Type = Manslaughter by Negligence
| Year > 2012.500
| | Weapon = Drugs: Female {Male=2, Unknown=0, Female=6}
| | Weapon = Fire: Female {Male=0, Unknown=0, Female=5}
| | Weapon = Gun
| | | Victim Sex = Female: Male {Male=1, Unknown=0, Female=1}
| | | Victim Sex = Male: Female {Male=0, Unknown=0, Female=2}
| Year ≤ 2012.500
| | Weapon = Drugs
| | | Agency Type = County Police: Male {Male=2, Unknown=0, Female=1}
| | | Agency Type = Municipal Police: Male {Male=12, Unknown=1, Female=3}
| | | Agency Type = State Police: Female {Male=0, Unknown=0, Female=2}
| | Weapon = Fire
| | | Year > 2011.500: Female {Male=0, Unknown=0, Female=6}
| | | Year ≤ 2011.500: Male {Male=1, Unknown=1, Female=0}
| | Weapon = Gun: Male {Male=7, Unknown=0, Female=2}
Crime Type = Murder or Manslaughter
| Agency Type = County Police: Male {Male=15, Unknown=3, Female=0}
| Agency Type = Municipal Police
| | Weapon = Drugs: Male {Male=69, Unknown=32, Female=37}
| | Weapon = Fire
| | | Victim Sex = Female: Male {Male=37, Unknown=13, Female=11}
| | | Victim Sex = Male
| | | | Year > 2011.500: Male {Male=38, Unknown=26, Female=14}
| | | | Year ≤ 2011.500: Unknown {Male=15, Unknown=16, Female=3}
| | Weapon = Gun: Male {Male=117, Unknown=83, Female=18}
| Agency Type = Special Police: Unknown {Male=1, Unknown=2, Female=1}
| Agency Type = State Police: Male {Male=25, Unknown=2, Female=7}
```

ภาพที่ 3.10 คำบรรยายลักษณะงาน Decision Tree ของ RapidMiner

จากผลสรุปที่ได้จากการสร้างโมเดล Decision Tree จากโปรแกรม RapidMiner และจากการคำนวณมือ มีผลลัพธ์ออกมาที่สอดคล้องกัน ทำให้ข้อมูลมีความถูกต้องและน่าเชื่อถือมากยิ่งขึ้น

ดังนั้น ผู้จัดทำได้ผลลัพธ์ของกฎที่สามารถทำนายได้จำนวน 18 กฎ และสามารถนำกฎที่ได้นำไปวิเคราะห์กฎต่อไปได้ โดยสามารถจำแนกกฎได้ ดังนี้

กฎข้อที่ 1 IF Crime Type = Manslaughter by Negligence Year = >2012 Weapon = Drugs = Female หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยประมาท ในปีที่ต่ำกว่าปี 2012 โดย อาวุธ คือ การใช้ยา ส่วนมากผู้กระทำผิดจะเป็นเพศ “หญิง”

กฎข้อที่ 2 IF Crime Type = Manslaughter by Negligence Year = >2012 Weapon = Fire = Female หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยประมาท ในปีที่ต่ำกว่าปี 2012 โดย อาวุธ คือ ไฟ ส่วนมากผู้กระทำผิดจะเป็นเพศ “หญิง”

กฎข้อที่ 3 IF Crime Type = Manslaughter by Negligence Year = >2012 Weapon = Gun = Victim Sex = Female = Male หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยประมาท ในปีที่ต่ำกว่าปี 2012 โดย อาวุธ คือ ปืน เพศเหยื่อ ผู้หญิง ส่วนมากผู้กระทำผิดจะเป็นเพศ “ชาย”

กฎข้อที่ 4 IF Crime Type = Manslaughter by Negligence Year = >2012 Weapon = Gun = Victim Sex = Male = Female หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยประมาท ในปีที่ต่ำกว่าปี 2012 โดย อาวุธ คือ ปืน เพศเหยื่อ ผู้ชาย ส่วนมากผู้กระทำผิดจะเป็นเพศ “หญิง”

กฎข้อที่ 5 IF Crime Type = Manslaughter by Negligence Year = <2012 Weapon = Drugs Agency Type = County Police = Male หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยประมาท ในปีที่มากกว่าหรือในปี 2012 โดย อาวุธ คือ การใช้ยา ตัวแทนหน่วยงาน คือ ตำรวจนครบาล ส่วนมากผู้กระทำผิดจะเป็นเพศ “ชาย”

กฎข้อที่ 6 IF Crime Type = Manslaughter by Negligence Year = <2012 Weapon = Drugs Agency Type = Municipal Police = Male หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยประมาท ในปีที่มากกว่าหรือในปี 2012 โดย อาวุธ คือ การใช้ยา ตัวแทนหน่วยงาน คือ ตำรวจท้องที่ ส่วนมากผู้กระทำผิดจะเป็นเพศ “ชาย”

กฎข้อที่ 7 IF Crime Type = Manslaughter by Negligence Year = <2012 Weapon = Drugs Agency Type = State Police = Female หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยประมาท ในปีที่มากกว่าหรือในปี 2012 โดย อาวุธ คือ การใช้ยา ตัวแทนหน่วยงาน คือ ตำรวจรัฐ ส่วนมากผู้กระทำผิดจะเป็นเพศ “หญิง”

กฎข้อที่ 8 IF Crime Type = Manslaughter by Negligence Year = <2012 Weapon = Fire Year = >2012 = Female หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยประมาท ในปีที่มากกว่าหรือในปี 2012 โดย อาวุธ คือ ไฟ ในปีที่ต่ำกว่าปี 2012 ส่วนมากผู้กระทำผิดจะเป็นเพศ “หญิง”

กฎข้อที่ 9 IF Crime Type = Manslaughter by Negligence Year = <2012 Weapon = Fire Year = <2012 = Male หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยประมาท ในปีที่มากกว่าหรือในปี 2012 โดย อาวุธ คือ ไฟ ในปีที่มากกว่าหรือในปี 2012 ส่วนมากผู้กระทำผิดจะเป็นเพศ “ชาย”

กฎข้อที่ 10 IF Crime Type = Manslaughter by Negligence Year = <2012 Weapon = Gun = Male หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยประมาท ในปีที่มากกว่าหรือในปี 2012 โดย อาวุธ คือ ปืน ส่วนมากผู้กระทำผิดจะเป็นเพศ “ชาย”

กฎข้อที่ 11 IF Crime Type = Murder or Manslaughter Agency Type = County Police = Male หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยเจตนา ตัวแทนหน่วยงาน คือ ตำรวจนครบาล ส่วนมากผู้กระทำผิดจะเป็นเพศ “ชาย”

กฎข้อที่ 12 IF Crime Type = Murder or Manslaughter Agency Type = State Police = Male หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยเจตนา ตัวแทนหน่วยงาน คือ ตำรวจรัฐ ส่วนมากผู้กระทำผิดจะเป็นเพศ “ชาย”

กฎข้อที่ 13 IF Crime Type = Murder or Manslaughter Agency Type = Special Police = Unknown หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยเจตนา ตัวแทนหน่วยงาน คือ ตำรวจพิเศษ ส่วนมากผู้กระทำผิดจะเป็นเพศ “ไม่ระบุเพศ”

กฎข้อที่ 14 IF Crime Type = Murder or Manslaughter Agency Type = Municipal Police Weapon = Gun = Male หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยเจตนา ตัวแทนหน่วยงาน คือ ตำรวจท้องที่ โดย อาวุธ คือ ปืน ส่วนมากผู้กระทำผิดจะเป็นเพศ “ชาย”

กฎข้อที่ 15 IF Crime Type = Murder or Manslaughter Agency Type = Municipal Police Weapon = Drugs = Male หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยเจตนา ตัวแทนหน่วยงาน คือ ตำรวจท้องที่ โดย อาวุธ คือ การใช้ยา ส่วนมากผู้กระทำผิดจะเป็นเพศ “ชาย”

กฎข้อที่ 16 IF Crime Type = Murder or Manslaughter Agency Type = Municipal Police Weapon = Fire Victim Sex = Female = Male หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยเจตนา ตัวแทนหน่วยงาน คือ ตำรวจท้องที่ โดย อาวุธ คือ ไฟ เพศเหยื่อ ผู้หญิง ส่วนมากผู้กระทำผิดจะเป็นเพศ “ชาย”

กฎข้อที่ 17 IF Crime Type = Murder or Manslaughter Agency Type = Municipal Police Weapon = Fire Victim Sex = Male Year = >2012 = Male หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยเจตนา ตัวแทนหน่วยงาน คือ ตำรวจท้องที่ โดย อาวุธ คือ ไฟ เพศเหยื่อ ผู้ชาย ในปีที่ต่ำกว่าปี 2012 ส่วนมากผู้กระทำผิดจะเป็นเพศ “ชาย”

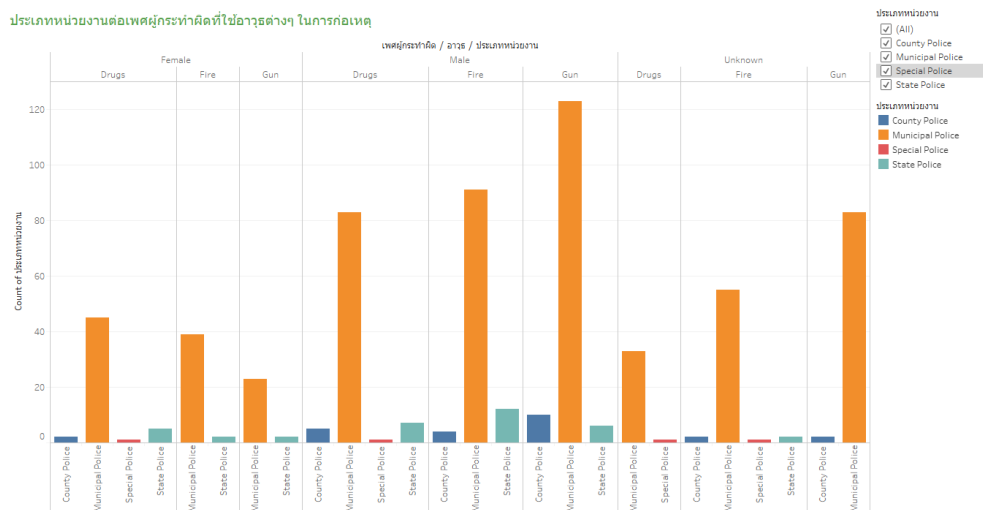
กฎข้อที่ 17 IF Crime Type = Murder or Manslaughter Agency Type = Municipal Police Weapon = Fire Victim Sex = Male Year = <2012 = Unknown หมายความว่า ถ้าประเภทอาชญากรรมเป็นการฆ่าโดยเจตนา ตัวแทนหน่วยงาน คือ ตำรวจท้องที่ โดย อาวุธ คือ

ไฟ เพศเหยื่อ ผู้ชาย ในปีที่มากกว่าหรือในปี 2012 ส่วนมากผู้กระทำผิดจะเป็นเพศ “ไม่ระบุ เพศ”

3.1.6 เผยแพร่ผลวิเคราะห์ (Deployment) ขั้นตอนการนำผลลัพธ์ที่ได้ไปใช้งานเป็นการ ทั่วไป อาจจัดทำเป็นรูปแบบของรายงาน (Report) หรือแผนภาพ (Dashboard) ที่พร้อมให้ฝ่าย ต่าง ๆ นำไปใช้ประโยชน์ในการวางแผน กำหนดกลยุทธ์ และดำเนินการต่าง ๆ ในทางธุรกิจ

ผู้วิเคราะห์ข้อมูลนำผลข้อมูลที่ทำการวิเคราะห์มาแสดงผลข้อมูลบนหน้า Web Browser โดยใช้โปรแกรม Adobe Dreamweaver ร่วมกับการนำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพโดยใช้โปรแกรม Tableau Public

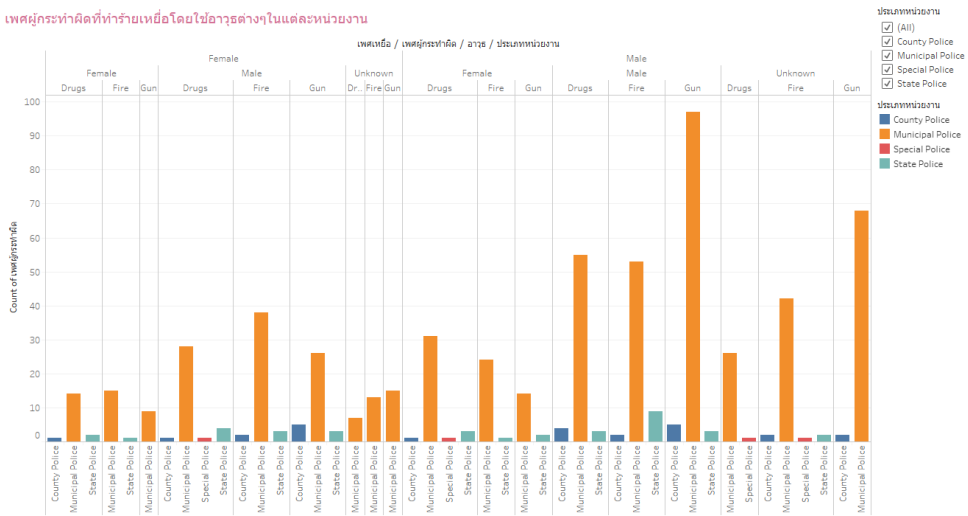
1. แสดงข้อมูลของประเภทหน่วยงาน แบบแผนภูมิแท่ง



ภาพที่ 3.11 แสดงข้อมูลกราฟของประเภทหน่วยงาน

2. แสดงข้อมูลของการฆาตกรรม แบบแผนภูมิแท่ง

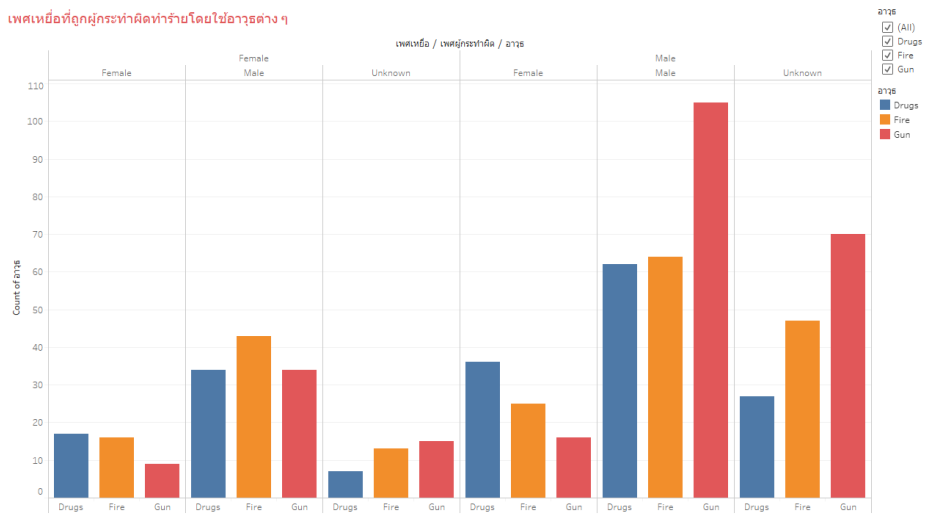
เพศผู้กระทำความผิดที่ทำร้ายเหยื่อโดยใช้อาวุธต่างๆในแต่ละหน่วยงาน



ภาพที่ 3.12 แสดงข้อมูลของการฆาตกรรม

3. แสดงข้อมูลของการใช้อาวุธ แบบแผนภูมิแท่ง

เพศเหยื่อที่ถูกผู้กระทำความผิดทำร้ายโดยใช้อาวุธต่างๆ



ภาพที่ 3.13 แสดงข้อมูลกราฟของการใช้อาวุธ

3.2 การออกแบบเว็บไซต์

3.2.1 การออกแบบ Wireframe หน้าจอเว็บไซต์

1. หน้าแรก แสดงข้อมูลเกี่ยวกับอาชญากรรม

Header	
หน้าแรก	การวิเคราะห์โมเดล ข้อมูลอาชญากรรม ▼ แบบทดสอบ แบบสอบถาม ▼ ดาวน์โหลด เข้าสู่ระบบ
หน้าแรก	
ที่อยู่	Footer ชื่อผู้จัดทำ

ภาพที่ 3.14 หน้าแรก

2. หน้าข้อมูลการวิเคราะห์โมเดล Decision Tree

Header	
หน้าแรก	การวิเคราะห์โมเดล ข้อมูลอาชญากรรม ▼ แบบทดสอบ แบบสอบถาม ▼ ดาวน์โหลด เข้าสู่ระบบ
การวิเคราะห์โมเดล	
ที่อยู่	Footer ชื่อผู้จัดทำ

ภาพที่ 3.15 หน้าข้อมูลการวิเคราะห์โมเดล Decision Tree

3. หน้าข้อมูลอาชีพการรม ของ ข้อมูลหน่วยงานตำรวจ

Header	
หน้าแรก การวิเคราะห์โมเดล ข้อมูลอาชีพการรม ▼ แบบทดสอบ แบบสอบถาม ▼ ดาวน์โหลด เข้าสู่ระบบ	
<div style="border: 1px solid black; padding: 2px;">ข้อมูลหน่วยงานตำรวจ ข้อมูลการฆาตกรรม ข้อมูลการใช้อาวุธ</div>	
หน้าข้อมูลหน่วยงานตำรวจ และแสดงกราฟข้อมูล	
ที่อยู่	Footer ชื่อผู้จัดทำ

ภาพที่ 3.16 หน้าข้อมูลอาชีพการรม ของ ข้อมูลหน่วยงานตำรวจ

4. หน้าข้อมูลอาชีพการรม ของ ข้อมูลการฆาตกรรม

Header	
หน้าแรก การวิเคราะห์โมเดล ข้อมูลอาชีพการรม ▼ แบบทดสอบ แบบสอบถาม ▼ ดาวน์โหลด เข้าสู่ระบบ	
<div style="border: 1px solid black; padding: 2px;">ข้อมูลหน่วยงานตำรวจ ข้อมูลการฆาตกรรม ข้อมูลการใช้อาวุธ</div>	
หน้าข้อมูลการฆาตกรรม และแสดงกราฟข้อมูล	
ที่อยู่	Footer ชื่อผู้จัดทำ

ภาพที่ 3.17 หน้าข้อมูลอาชีพการรม ของ ข้อมูลการฆาตกรรม

5. หน้าข้อมูลอาชีพการกรรรม ของ ข้อมูลการใช้อาวุธ

Header	
หน้าแรก การวิเคราะห์โมเดล ข้อมูลอาชีพการกรรรม ▼ แบบทดสอบ แบบสอบถาม ▼ ดาวน์โหลด เข้าสู่ระบบ	
	ข้อมูลหน่วยงานตำรวจ ข้อมูลการฆาตกรรม ข้อมูลการใช้อาวุธ
หน้าข้อมูลการใช้อาวุธ และแสดงกราฟข้อมูล	
ที่อยู่	Footer ชื่อผู้จัดทำ

ภาพที่ 3.18 หน้าข้อมูลอาชีพการกรรรม ของ ข้อมูลการใช้อาวุธ

6. หน้าแบบทดสอบ

Header	
หน้าแรก การวิเคราะห์โมเดล ข้อมูลอาชีพการกรรรม ▼ แบบทดสอบ แบบสอบถาม ▼ ดาวน์โหลด เข้าสู่ระบบ	
แบบทดสอบ Microsoft Forms	
ที่อยู่	Footer ชื่อผู้จัดทำ

ภาพที่ 3.19 หน้าแบบทดสอบ

7. หน้าแบบสอบถาม การป้องกันตนเอง

Header	
หน้าแรก การวิเคราะห์โมเดล ข้อมูลอาชญากรรม ▼ แบบทดสอบ แบบสอบถาม ▼ ดาวน์โหลด เข้าสู่ระบบ	
แบบสอบถามการป้องกัน แบบสอบถามความพึงพอใจ	
แบบสอบถามการป้องกันตนเอง Microsoft Forms	
ที่อยู่	Footer <small>ชื่อผู้จัดทำ</small>

ภาพที่ 3.20 หน้าแบบสอบถาม การป้องกันตนเอง

8. หน้าแบบสอบถาม ความพึงพอใจ

Header	
หน้าแรก การวิเคราะห์โมเดล ข้อมูลอาชญากรรม ▼ แบบทดสอบ แบบสอบถาม ▼ ดาวน์โหลด เข้าสู่ระบบ	
แบบสอบถามการป้องกัน แบบสอบถามความพึงพอใจ	
แบบสอบถามความพึงพอใจ Microsoft Forms	
ที่อยู่	Footer <small>ชื่อผู้จัดทำ</small>

ภาพที่ 3.21 หน้าแบบสอบถาม ความพึงพอใจ

9. หน้าดาวโหลดไฟล์ข้อมูล

Header	
หน้าแรก การวิเคราะห์โมเดล ข้อมูลอาชญากรรม ▼ แบบทดสอบ แบบสอบถาม ▼ ดาวโหลด เข้าสู่ระบบ	
หน้าดาวโหลดไฟล์	
ที่อยู่	Footer ชื่อผู้จัดทำ

ภาพที่ 3.22 หน้าดาวโหลดไฟล์ข้อมูล

10. หน้าเข้าสู่ระบบ

Header	
หน้าแรก การวิเคราะห์โมเดล ข้อมูลอาชญากรรม ▼ แบบทดสอบ แบบสอบถาม ▼ ดาวโหลด เข้าสู่ระบบ	
เข้าสู่ระบบ หน้า Login	
ที่อยู่	Footer ชื่อผู้จัดทำ

ภาพที่ 3.23 หน้าเข้าสู่ระบบ

11. หน้าจัดการไฟล์ เพิ่ม ลบ แก้ไข ข้อมูล

Header	
หน้าแรก	การวิเคราะห์โมเดล ข้อมูลอาชญากรรม ▼ แบบทดสอบ แบบสอบถาม ▼ อัปโหลดข้อมูล ออกจากระบบ
หน้า เพิ่ม ลบ แก้ไข ไฟล์	
ที่อยู่	Footer ชื่อผู้จัดทำ

ภาพที่ 3.24 หน้าจัดการไฟล์ เพิ่ม ลบ แก้ไข ข้อมูล

3.3 บทสรุป

จากวิธีการดำเนินงานโครงการในข้างต้น ผู้วิเคราะห์ข้อมูลได้แสดงวิธีในการจัดการกับการวิเคราะห์ข้อมูลรายงานการฆาตกรรมปี 2011-2013 ด้วยขั้นตอนกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM รวมถึงการสร้างโมเดล Decision Tree จากโปรแกรมที่ใช้ทำเหมืองข้อมูลโดยโปรแกรม RapidMiner Studio ในการสร้างโมเดล Decision Tree ซึ่งได้ผลลัพธ์ที่เหมือนกัน และนำข้อมูลสารสนเทศมาทำการแสดงผลแบบ visualization ในรูปแบบของภาพโดยใช้โปรแกรม Tableau Public และออกแบบ Wireframe ของเว็บไซต์ที่จะเผยแพร่บน web browser ด้วยโปรแกรม Adobe XD ปัญหาที่พบในบทนี้ คือการทำ Decision Tree ผ่านโปรแกรม weka ซึ่งพบปัญหาคือ ผลลัพธ์ที่ผ่านกระบวนการทำ Decision Tree นั้น รูปแบบโมเดลไม่ตรง และแสดงผลแอตทริบิวต์ออกมาไม่ครบตามที่คำนวณและ code คำบรรยายลักษณะงานก็แสดงผลออกมาไม่ครบแอตทริบิวต์ ผู้วิเคราะห์ข้อมูลจึงได้ใช้แค่ โปรแกรม RapidMiner Studio ในการทำ Decision Tree