

การตรวจจับการทุจริตของผู้ให้บริการทางการแพทย์โดยใช้เทคนิคการหาความสัมพันธ์**Fraud Detection of Medical Provider Using Association Rule Method**

จักริน วิแก้ว* และ ณิชนันท์ กิตติพัฒน์บวร

หลักสูตรการจัดการเทคโนโลยีสารสนเทศ มหาวิทยาลัยวลัยลักษณ์ อำเภอท่าศาลา จังหวัดนครศรีธรรมราช 80160

*wjakkari@gmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อหาความสัมพันธ์ของการทุจริตของผู้ให้บริการด้านสุขภาพ โดยใช้อัลกอริทึม FP-Growth ซึ่งเป็นหนึ่งในเทคนิคการทำเหมืองข้อมูลได้รับความนิยม ในการทดลองนี้ได้ใช้ข้อมูลจากชุดข้อมูลมาตรฐานจากหน่วยงาน the United States Center for Medicare and Medicaid Services หรือ CMS และ The Office of Inspector General's (OIG), Department of Health and Human Services ประเทศสหรัฐอเมริกา ตั้งแต่ปี 2012 ถึง 2016 โดยทำการสุ่มข้อมูลจำนวน 47,489 ระเบียน (จาก 9,497,892 ระเบียน) ภายหลังขั้นตอนการเตรียมข้อมูลแล้วจะมีจำนวนข้อมูล 3,778 ระเบียน เพื่อใช้ในการทดลองหาความสัมพันธ์ของข้อมูล ผลการวิจัยสามารถสรุปกฎเกณฑ์ความสัมพันธ์ของการทุจริตของผู้ให้บริการด้านการแพทย์ได้เป็น 2 กรณี ได้แก่ การกระทำความผิดที่เกี่ยวข้องกับพฤติกรรมของผู้ให้บริการด้านการแพทย์ในการใช้โปรแกรมหรือระบบคอมพิวเตอร์ และความผิดด้านการถูกเพิกถอนใบอนุญาต โดยมีค่าความสนับสนุนไม่ต่ำกว่า 0.01 และมีความเชื่อมั่นสูงกว่าหรือเท่ากับ 0.70 จากการวิจัยนี้สามารถนำกฎความสัมพันธ์ของการทุจริตไปใช้ระบุเงื่อนไขที่มีแนวโน้มเสี่ยงต่อการกระทำการทุจริต เช่น ตำแหน่งงาน ข้อมูลการรับรองของผู้ให้บริการ หรือ ความเชี่ยวชาญเฉพาะด้านของผู้ให้บริการ เป็นต้น โดยทีมวิจัยหวังว่าจะประยุกต์แนวคิดการวิจัยไปใช้ในการตรวจสอบการทุจริตของระบบตรวจสอบด้านการแพทย์ (medical audit system) ต่อไป

คำสำคัญ : การหาความสัมพันธ์ การตรวจจับการทุจริต ด้านสุขภาพ FP-Growth การทำเหมืองข้อมูล**Abstract**

The main objective of this paper is finding the association rules of fraud detection of the medical provider. This study applies the FP growth technique which the well-known method. There are two datasets from the United States Center for Medicare and Medicaid Services (CMS) and The Office of Inspector General's (OIG), Department of Health and Human Services, US (between 2012 and 2016). The sample size is 47489 (total population 9497892 records). After cleansing data, resulted in 3778 records were used to find association rule in the further state. The result of this study presented two important cases of fraud detection for the medical provider are the conviction for program-related crimes and the license revocation, suspension, or surrender with support more than 0.01 and confident more than 70%. According to the results of this study, the audition division will apply all association rules to identify the risk conditions in medical processes such as position, credentials of the provider or provider's specialty. For future work, we will adopt the concept of this research to detect fraud in medical audition system.

Keywords: Association Rules, Fraud Detection, Medical, FP-Growth, Data Mining

บทนำ

กระบวนการตรวจจับการทุจริตด้านสุขภาพ (Healthcare fraud detection) เป็นกระบวนการสำคัญในการป้องกันการสูญเสียงบประมาณด้านสุขภาพ โดยมีการคาดการณ์ว่าจะมีค่าใช้จ่ายด้านสุขภาพเพิ่มจาก 7.1 พันล้านดอลลาร์สหรัฐ ในปี 2015 เป็น 8.7 พันล้านดอลลาร์สหรัฐในปี 2020 หรือเพิ่มขึ้นร้อยละ 4.3 (Seo & Mendelevitch, 2017) แต่ในช่วงเวลาเดียวกันความสูญเสียที่เกิดจากการทุจริตภายใต้ระบบหลักประกันสุขภาพ (Healthcare group insurance) ก็มีมูลค่าความสูญเสียเพิ่มสูงขึ้นเช่นเดียวกัน เช่น ในปี 2017 ประเทศสหรัฐอเมริกามีการใช้งบประมาณด้านสุขภาพสูงถึง 3 ล้านล้านดอลลาร์สหรัฐ และมีการประมาณการความเสียหายที่เกิดจากการทุจริตสูงถึง 1 แสนล้านดอลลาร์สหรัฐ หรือร้อยละ 5 ของงบประมาณทั้งหมด (Gee et al., 2018)

ในอดีตการตรวจจับการทุจริตทำโดยการตรวจสอบจากทีมผู้เชี่ยวชาญ ซึ่งต้องอาศัยประสบการณ์ของผู้เชี่ยวชาญทำการตรวจสอบจากข้อมูลการบันทึกเวชระเบียนและข้อมูลอื่นๆที่ละรายการ โดยจะมีการสุ่มข้อมูลขึ้นมาตรวจสอบ ซึ่งจะมีข้อเสียคือต้องใช้งบประมาณสูงและเสียเวลาในการดำเนินการมาก (สำนักงานหลักประกันสุขภาพแห่งชาติ, 2562) ต่อมาผู้เชี่ยวชาญนำเสนอเทคนิคต่างๆ เช่น การใช้ PageRank algorithm ตรวจจับผู้ให้บริการทางการแพทย์ที่ทุจริตมีการเบิกจ่ายเงินจากบริษัทประกันมากกว่าผู้ให้บริการรายอื่นในการรักษาประเภทเดียวกัน โดยใช้ชุดข้อมูลทดสอบ Medicare Part B Dataset ซึ่งเก็บข้อมูลการเบิกจ่ายประกันของผู้ให้บริการในประเทศสหรัฐอเมริกา จากหน่วยงาน the United States Center for Medicare and Medicaid Services หรือ CMS ผลการวิจัยพบว่าความผิดปกติของข้อมูลจะมาจาก 3 สาเหตุคือ 1) การกรอกข้อมูลผิด (Human errors) 2) ข้อมูลจากผู้ให้บริการไม่ครบถ้วน และ 3) ความผิดปกติที่ไม่ทราบสาเหตุ (มีแนวโน้มจะเป็นการทุจริต) เช่น จำนวนความผิดปกติในการสั่งยาของอายุรแพทย์ และจักษุแพทย์ นอกจากนี้ยังพบว่าการระบุความเชี่ยวชาญ (Specialist) ของผู้ให้บริการทำให้ค่าความถูกต้อง (Accuracy) ของโมเดลสูงขึ้น (Seo & Mendelevitch, 2017)

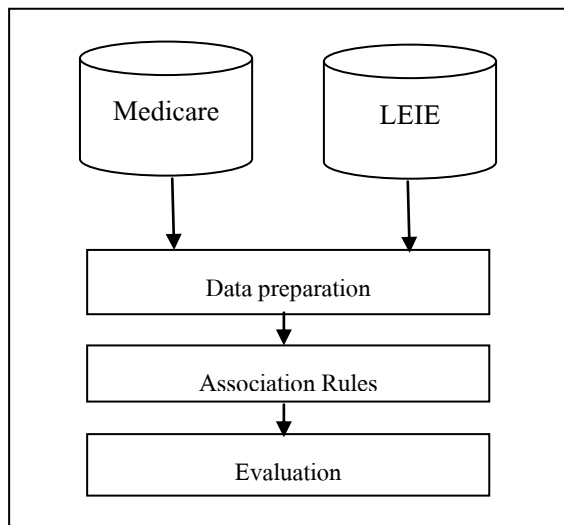
นอกจากนั้นมีการนำเสนองานวิจัยอื่นๆที่ใช้ Algorithms ต่างๆ ของการทำเหมืองข้อมูล (Data mining) ตัวอย่างเช่น การใช้เทคนิคการหาความสัมพันธ์ของข้อมูล (Association rules) เพื่อหารูปแบบการทุจริตในระบบการเรียกเก็บเงิน (Billing pattern) โดยแบ่งข้อมูลออกเป็น 2 กลุ่มคือกลุ่มที่มีรูปแบบข้อมูลที่สอดคล้องและกลุ่มที่มีข้อมูลที่ไม่สอดคล้อง เพื่อหารูปแบบความสัมพันธ์ของการทุจริต และแสดงกฎที่สามารถตรวจสอบการทุจริตจากผู้เชี่ยวชาญ โดยมีผู้เชี่ยวชาญที่ไม่กระทำผิดกฎใดๆเลย จำนวน 650 คน (จำนวนผู้เชี่ยวชาญทั้งหมด 779 คน) มีความเชื่อมั่นในกฎด้านลบ¹ (Negative rule) ร้อยละ 96 และกฎด้านบวก (Positive rule) ซึ่งโดยเฉลี่ยมีการผิดกฎเพียง 1.47 กฎเกณฑ์ต่อผู้เชี่ยวชาญ และความถูกต้องเมื่อเปรียบเทียบกับกฎที่ได้จากการใช้เทคนิคการหาความสัมพันธ์กับการตรวจสอบจากผู้เชี่ยวชาญอยู่ที่ร้อยละ 21 (Shan et al., 2008) อย่างไรก็ตามยังไม่มีงานนำเสนอความสัมพันธ์ระหว่างผู้ให้บริการและข้อมูลหน้าที่การให้บริการของผู้ให้บริการด้านการแพทย์

ดังนั้นในงานวิจัยนี้ ทีมผู้วิจัยได้ทำการศึกษารูปแบบการทุจริตของผู้ให้บริการ โดยใช้ชุดข้อมูลมาตรฐานจากหน่วยงาน CMS ประเทศสหรัฐ หรือที่รู้จักในนาม Medicare Part B ซึ่งไม่มีการเก็บฟิลด์ข้อมูลที่สามารถระบุการทุจริตเอาไว้ จึงต้องทำการเชื่อมโยงกับข้อมูลการทุจริตของ The List of Excluded Individuals/Entities (LEIE) โดยใช้เทคนิคการทำเหมืองข้อมูลในการหารูปแบบความสัมพันธ์ระหว่างข้อมูลทั้งจากข้อมูลพื้นฐานของผู้ให้บริการ เช่น ที่อยู่ ความเชี่ยวชาญ ฯลฯ และข้อมูลในกระบวนการรักษาของผู้ให้บริการทางการแพทย์ เช่น การวินิจฉัย การรักษา ฯลฯ โดยใช้กรณีศึกษาในประเทศสหรัฐอเมริกา

อุปกรณ์และวิธีการดำเนินงานวิจัย

ในหัวข้อนี้จะอธิบายเกี่ยวกับขั้นตอนการเตรียมข้อมูล และเทคนิคการหาความสัมพันธ์ของข้อมูล โดยการทดลองใช้คอมพิวเตอร์ตั้งโต๊ะที่มีหน่วยประมวลผลกลาง Intel Core i5 3.19 GHz ติดตั้งระบบปฏิบัติการ Windows 10 ใช้โปรแกรม MS Access 2013 ในการทำความสะอาดข้อมูล และใช้โปรแกรม RapidMiner 9.2 ในการเตรียมข้อมูล และหาความสัมพันธ์ ขั้นตอนดังรูปที่ 1

¹ กฎด้านลบ)Negative ruleเป็นกฎเกณฑ์ที่แสดงถึงรายการที่ไม่ถูกเรียกเก็บเงิน (



รูปที่ 1 ขั้นตอนการดำเนินงานวิจัย

การเตรียมข้อมูล

ชุดข้อมูล (Dataset)

ในการทดลองใช้ชุดข้อมูล The Medicare Provider Utilization and Payment Data: Physician and Other Supplier describes ซึ่งเก็บข้อมูลเมดิแคร์ (Medicare)² พื้นฐานของผู้ให้บริการ ข้อมูลการบริการ และสรุปข้อมูลการรักษาในประเทศสหรัฐอเมริกา ชุดข้อมูลผ่านการทำความสะอาดข้อมูล (Data cleaning) โดย The Centers for Medicare and Medicaid Services (CMS) ใช้ข้อมูลระหว่างปี 2012 ถึง 2016 จำนวน 9,497,892 ระเบียบ (Record) จำนวน 26 ฟิลด์ (Fields) และทำการสุ่มข้อมูลมาร้อยละ 5 (จำนวน 47,489 ระเบียบ) เพื่อใช้ในขั้นตอนการเตรียมข้อมูล ซึ่งประกอบด้วยการเลือกฟิลด์ข้อมูล Medicare ที่เกี่ยวข้องกับผู้ให้บริการด้านการแพทย์ (ไม่รวมข้อมูลด้านการเงิน) จำนวน 8 ฟิลด์ และข้อมูลรายชื่อสถานบริการที่ถูกยกเลิกการเข้าร่วมโครงการดูแลสุขภาพ The List of Excluded Individuals/Entities (LEIE) database โดยผ่านการทำความสะอาดข้อมูลจาก The Office of Inspector General's (OIG), Department of Health and Human Services ประเทศสหรัฐอเมริกา (Office of Inspector General, 2010) จำนวน 3 ฟิลด์ สามารถแสดงลักษณะข้อมูลได้ดังตารางที่ 1 และข้อมูลกฎการกระทำความผิดแสดงได้ดังตารางที่ 2 ก่อนทำการกรองข้อมูลที่ไม่มีความหมายออก (Field General=Individual (unaffiliate))

เนื่องจากข้อมูล Medicare จาก CMS ไม่ได้ระบุการกระทำความผิดในการทดลองจึงต้องใช้ฟิลด์ชื่อและนามสกุลของผู้ให้บริการด้านการแพทย์ในการเชื่อมโยงข้อมูลระหว่าง Medicare และ LEIE ซึ่งภายหลังจากการเชื่อมโยงจะมีจำนวนข้อมูลทั้งหมด 3,778 ระเบียบ

² เมดิแคร์ (Medicare) เป็นโครงการช่วยเหลือด้านสุขภาพสำหรับผู้สูงอายุตั้งแต่ 65 ปีขึ้นไป ผู้ทุพพลภาพและผู้ป่วยโรคไตขั้นสุดท้าย สนับสนุนโดยรัฐบาลกลางสหรัฐฯ หากระยะเวลาการทำงานในอดีตของผู้สูงอายุที่มีประวัติการทำงาน 10 ปีขึ้นไป จะได้รับการคุ้มครองทั้งหมด (เข้มทอง, 2550)

ตารางที่ 1 ลักษณะข้อมูลจาก The Medicare Provider Utilization and Payment Data: Physician and Other Supplier describes (Medicare) และ The List of Excluded Individuals/Entities (LEIE) database

Fields	Description	Dataset
NPI	Unique provider identification number	Medicare
Last_name	Provider's last name	Medicare
First_name	Provider's first name	Medicare
Hcpcs_code	The provider specialty code reported on the claim	Medicare
Provider_type	Medical provider's specialty (or practice)	Medicare
Nppes_credentials	Credentials of the Provider	Medicare
Nppes_provider_city	City of the Provider	Medicare
Nppes_provider_state	State Code of the Provider	Medicare
General	General information of provider	LEIE
Specialty	Provider's specialty	LEIE
Excltype	Exclusion rule	LEIE

ตารางที่ 2 ข้อมูลกฎการกระทำความผิดจาก The List of Excluded Individuals/Entities (LEIE)

Rule Number	Description
1128(a)(1)	Conviction of program-related crimes.
1128(a)(2)	Conviction relating to patient abuse or neglect.
1128(a)(3)	Felony conviction relating to health care fraud.
1128(b)(4)	License revocation or suspension.

การหาความสัมพันธ์ของข้อมูล

การหาความสัมพันธ์ของข้อมูลคือการหารูปแบบที่เกิดขึ้นบ่อยๆ (Frequent pattern) ในชุดข้อมูล ตัวอย่างเช่น เซตของไอเทม (Set of items) เช่น รายการของนมและขนมปังที่ปรากฏในรายการซื้อขายร่วมกันจากรายการการซื้อทั้งหมด เรียกว่า ไอเทมเซตที่ปรากฏร่วมกันบ่อย Frequent Itemset (Han et al., 2012) ในการหาความสัมพันธ์ของข้อมูล (Association rules) จะประกอบด้วยเซตไอเทม (I) ที่มีสมาชิกเป็นไอเทมทั้งหมดในฐานข้อมูล มีสมาชิกไอเทมในทรานแซคชันย่อย (T) หรือ $T \subseteq I$ และ D คืองานที่เกี่ยวข้องกับข้อมูล และ A เป็น เซตของไอเทมในทรานแซคชัน หรือ $A \subseteq T$ สามารถกำหนดกฎความสัมพันธ์ให้อยู่ในรูปแบบ $A \Rightarrow B$ เมื่อ $A \subset I$ และ $B \subset I$ และ $A \cap B = \emptyset$ โดยมีค่า Support s เมื่อ s ร้อยละข้อมูลในทรานแซคชันใน D ($A \cup B$)

ในการทดลองนี้ได้หาความสัมพันธ์โดยใช้อัลกอริทึม FP-Growth ซึ่งไม่จำเป็นต้องสร้าง Candidate itemset ใช้หลักการทำงานคือในขั้นแรกจะตัดไอเทมเซตที่มีค่าน้อยกว่าค่าสนับสนุนที่ตั้งไว้ออกไป ในขั้นถัดมาจะทำการเรียงข้อมูลไอเทมเซตตามลำดับค่าสนับสนุนจากมากไปหาน้อย โดยจัดเก็บโครงสร้างแบบ FP-Tree โดยวิธีการนี้มีข้อดีคือทำการอ่านข้อมูลทรานแซคชันเพียงสองรอบเท่านั้น

ในการเลือกกฎความสัมพันธ์ $A \Rightarrow B$ จะพิจารณาจากค่าวัดผล 3 ค่าได้แก่

ค่าสนับสนุน

$$\text{support}(A \Rightarrow B) = P(A \cup B)$$

ค่าความเชื่อมั่น

$$\text{confident}(A \Rightarrow B) = \text{support}(A \cup B) / \text{support}(A)$$

ค่าลิฟต์ ค่าบ่งบอกการเกิดกฎความสัมพันธ์

$$\text{lift}(A, B) = P(A \cup B) / P(A)P(B)$$

โดยถ้าค่าลิฟต์ มีค่าเท่ากับ 1 แสดงว่ารูปแบบกฎที่เกิดขึ้นไม่ขึ้นต่อกัน (Independent)

ผลและอภิปรายผล

จากการทดลองสามารถสร้างกฎความสัมพันธ์ของข้อมูล โดยกำหนดค่าสนับสนุนที่ 0.01 และค่าความเชื่อมั่นที่ 0.70 โดยแสดงกฎความสัมพันธ์เรียงตามค่าความเชื่อมั่นและค่าลิฟต์ ดังตารางที่ 3

ตารางที่ 3 กฎความสัมพันธ์ของข้อมูลที่เกี่ยวข้องกับการทุจริต

No.	Rules	ค่า สนับสนุน Support	ค่าความ เชื่อมั่น (Confident)	ค่า ลิฟต์ (Lift)
1	[nppes_credentials = M.D., SPECIALTY = OWNER/OPERATOR] --> [EXCLTYPE = 1128a1]	0.041	0.923	3.646
2	[SPECIALTY = OWNER/OPERATOR] --> [EXCLTYPE = 1128a1]	0.049	0.912	3.603
3	[SPECIALTY = OWNER/OPERATOR] --> [nppes_credentials = M.D., EXCLTYPE = 1128a1]	0.041	0.760	3.767
4	[GENERAL = NURSING PROFESSION] --> [EXCLTYPE = 1128b4]	0.114	0.762	1.722
5	[SPECIALTY = NURSE/NURSES AIDE, GENERAL = NURSING PROFESSION] --> [EXCLTYPE = 1128b4]	0.114	0.761	1.721
6	[GENERAL = NURSING PROFESSION] --> [EXCLTYPE = 1128b4, SPECIALTY = NURSE/NURSES AIDE]	0.114	0.760	3.149
7	[nppes_credentials = M.D., GENERAL = NURSING PROFESSION] --> [EXCLTYPE = 1128b4]	0.079	0.751	1.696
8	[nppes_credentials = M.D., SPECIALTY = NURSE/NURSES AIDE, GENERAL = NURSING PROFESSION] --> [EXCLTYPE = 1128b4]	0.079	0.751	1.696
9	[nppes_credentials = M.D., GENERAL = NURSING PROFESSION] --> [EXCLTYPE = 1128b4, SPECIALTY = NURSE/NURSES AIDE]	0.079	0.751	3.110
10	[SPECIALTY = NURSE/NURSES AIDE, GENERAL = IND- LIC HC SERV PRO] --> [EXCLTYPE = 1128b4]	0.073	0.731	1.651

จากตารางที่ 3 สามารถอธิบายกฎความสัมพันธ์ของการทุจริตได้ดังนี้ กฎข้อที่ 1 ถ้าผู้ให้บริการด้านการแพทย์มีการจذبรองว่าเป็นแพทย์และเป็นเจ้าของและ/หรือเป็นโอเปอร์เรเตอร์ด้วย แล้วจะกระทำผิดกฎข้อ 1128a1 โดยมีค่าความเชื่อมั่นร้อยละ 92.3 กฎข้อที่ 2 คือ ถ้าเป็นเจ้าของและ/หรือโอเปอร์เรเตอร์ด้วย แล้วจะกระทำผิดกฎข้อ 1128a1 โดยมีค่าความเชื่อมั่นร้อยละ 91.2 และกฎข้อที่ 3 ถ้าเป็นเจ้าของและ/หรือโอเปอร์เรเตอร์ด้วย แล้วจะมีการจذبรองว่าเป็นแพทย์และจะกระทำผิดกฎข้อ 1128a1 ในส่วนความผิดอื่น ๆ ที่น่าสนใจ เช่น การถูกเพิกถอนใบอนุญาต (1128b4) ส่วนใหญ่จะเกี่ยวข้องกับพยาบาลชำนาญการ พยาบาลทั่วไป และแพทย์ตามลำดับ

จากผลการทดลองจะเห็นว่ากฎข้อที่ 1-3 เป็นการกระทำผิดที่เกี่ยวข้องกับพฤติกรรมของผู้ให้บริการด้านการแพทย์ในการใช้โปรแกรมหรือระบบคอมพิวเตอร์ หากระบุเงื่อนไขในกรณีผู้ให้บริการด้านการแพทย์เป็นแพทย์และเป็นเจ้าของสถานบริการและ/หรือเป็นโอเปอร์เรเตอร์ของสถานบริการในการตรวจจับการทุจริตจะสามารถตรวจจับการทุจริตได้ โดยมีความเชื่อมั่นเกินร้อยละ

92 และหากต้องการตรวจสอบผู้ให้บริการที่ถูกเพิกถอนใบอนุญาต (กฎข้อ 4-10) สามารถระบุเงื่อนไขการตรวจสอบโดยมีอาชีพเป็นพยาบาลชำนาญการ พยาบาลทั่วไป หรือแพทย์แล้ว จะสามารถตรวจจับผู้กระทำผิดได้ โดยมีความเชื่อมั่นสูงกว่าร้อยละ 70

สรุปผลการวิจัย

ในงานวิจัยนี้สามารถตรวจจับผู้ที่มีแนวโน้มในการกระทำทุจริตได้แบ่งออกเป็น 2 กรณี คือ ผู้การกระทำผิดที่เกี่ยวกับกับพฤติกรรมของผู้ให้บริการด้านการแพทย์ในการใช้โปรแกรมหรือระบบคอมพิวเตอร์ และ ผู้ให้บริการที่ถูกเพิกถอนใบอนุญาต โดยการกระทำผิดในกรณีแรก กฎความสัมพันธ์ของข้อมูลทุกข้อ(ร้อยละ 100) บ่งชี้ให้เห็นว่าผู้ที่มีแนวโน้มที่จะทำการทุจริตมากที่สุดคือผู้ให้บริการด้านการแพทย์ที่เป็นเจ้าของและ/หรือโอเปอเรเตอร์ นอกจากนั้นร้อยละ 77 บ่งชี้ว่าผู้ทำการทุจริตเป็นผู้ที่ทำการจัดรับรองในตำแหน่งแพทย์อีกด้วย ส่วนผู้ที่ถูกเพิกถอนใบอนุญาตแล้วยังดำเนินการทุจริตโดยยังดำเนินการให้บริการรักษาผู้ป่วยอยู่ได้แก่ พยาบาลชำนาญการ พยาบาลทั่วไป และแพทย์ตามลำดับ

ข้อเสนอแนะ

กฎความสัมพันธ์ของข้อมูลที่น่าเสนอเป็นผลลัพธ์จากการนำเทคนิคการทำเหมืองข้อมูล ข้อดีคือสามารถนำไปใช้ในงานตรวจสอบ (Audit) ทางด้านการแพทย์และด้านอื่นๆที่เกี่ยวข้อง แต่การหาความสัมพันธ์ของข้อมูลที่น่าเสนอเป็นข้อมูล Medicare ของประเทศสหรัฐอเมริกา เพื่อเป็นการเพิ่มประสิทธิภาพของวิธีการ ทีมวิจัยมีแนวคิดในการพัฒนาวิธีการหาความสัมพันธ์ของข้อมูล โดยใช้อัลกอริทึมอื่นๆ ในการหาความสัมพันธ์ของข้อมูล เช่น Apriori, Eclat และ PrePost+ ร่วมกับเทคนิคการทำเหมืองข้อมูลด้านการจัดหมวดหมู่ (Classification) เพื่อนำมาประยุกต์ใช้ในการตรวจสอบการทุจริตในประเทศไทย

บรรณานุกรม

- เข้มทอง, ม. (2550). ระบบเฮลท์แคร์ในสหรัฐอเมริกา. Retrieved May 14, 2019, from <http://info.gotomanager.com/news/printnews.aspx?id=56049>.
- สำนักงานหลักประกันสุขภาพแห่งชาติ, ส. (2562). คู่มือแนวทางปฏิบัติในการตรวจสอบเอกสารหลักฐานการเรียกเก็บค่าใช้จ่ายเพื่อบริการสาธารณสุขของหน่วยบริการในระบบหลักประกันสุขภาพถ้วนหน้า ปี 2562.
- Bauder, R. A., & Khoshgoftaar, T. M. (2018). Medicare fraud detection using machine learning methods. *Proceedings - 16th IEEE International Conference on Machine Learning and Applications, ICMLA 2017, 2018-Janua*, 858–865. <https://doi.org/10.1109/ICMLA.2017.00-48>.
- Gee, J., Button, M., & Brooks, G. (2018). The financial cost of fraud 2018. Retrieved from <http://eprints.port.ac.uk/3988/>.
- Han, J., Kamberi, M., & Pei, J. (2012). *Data Mining Concepts and Techniques* (3rd ed). MA: Morgan Kaufmann. Retrieved from <http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>.
- Liou, F. M., Tang, Y. C., & Chen, J. Y. (2008). Detecting hospital fraud and claim abuse through diabetic outpatient services. *Health Care Management Science*, 11(4), 353–358. <https://doi.org/10.1007/s10729-008-9054-y>.
- Office of Inspector General. (2010). Background Information. https://doi.org/10.1007/978-0-387-79948-3_4303.
- Seo, J., & Mendelevitch, O. (2017). Identifying frauds and anomalies in Medicare-B dataset. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 81002, 3664–3667. <https://doi.org/10.1109/EMBC.2017.8037652>.
- Shan, Y., Jeacocke, D., Murray, D. W., & Sutinen, A. (2008). Mining medical specialist billing patterns for health service management. *Conferences in Research and Practice in Information Technology Series*, 87, 105–110.