

การประยุกต์ใช้ K-Means เพื่อจัดกลุ่มกราฟแท่งเทียน An Application of K-Means for Candle Stick Clustering

สุจิตรา ไชยกุลสินธุ์, ปริญญา มากลิ่น, และวัชระ โพธิสรณ์
มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

บทคัดย่อ

สำหรับกระบวนการในการตัดสินใจลงทุน นอกจากการพิจารณาคัดเลือกบริษัทที่จะลงทุนแล้ว การเลือกจังหวะในการลงทุนเป็นอีกปัจจัยที่จะทำให้การลงทุนประสบผลสำเร็จมากขึ้น การพิจารณากราฟแท่งเทียนเป็นส่วนหนึ่งในการพิจารณาหาจังหวะในการลงทุนที่เหมาะสมเพื่อใช้ประกอบการตัดสินใจในการลงทุน หลักการวิเคราะห์จังหวะการลงทุนของนักลงทุนแตกต่างกันไปตามความรู้และประสบการณ์ในการลงทุน ในการจัดกลุ่มกราฟแท่งเทียนเพื่อประกอบการพิจารณาจะทำให้การวิเคราะห์กราฟสะดวกยิ่งขึ้นและยังสามารถใช้ในการสร้างระบบการลงทุนในอนาคตได้ ในงานวิจัยนี้ใช้การจัดกลุ่มกราฟแท่งเทียนด้วยวิธี K-means โดยใช้ข้อมูลการซื้อขายรายวันจากหุ้นในกลุ่ม SET50 ซึ่งประกอบไปด้วย ราคาเปิด ราคาปิด ราคาสูงสุดและราคาต่ำสุดของหุ้น AOT CPALL CPN PTT และ TRUE ด้วยข้อมูลจำนวน 7500 ชุด ใช้ Weka เพื่อประมวลผลวิธีการแบบ K-means ซึ่งสามารถจัดกลุ่มกราฟแท่งเทียนออกเป็นห้ากลุ่มจากส่วนประกอบ Body Upper Wick Lower Wick ที่แตกต่างกันของกราฟแท่งเทียน ซึ่งแต่ละกลุ่มมีรูปแบบที่สะท้อนลักษณะการเคลื่อนไหวของราคาที่แตกต่างกัน

Abstract

For investment decision making process, apart from the processes of stock selection, timing for buying and selling stocks is also important for successful investment. Candle stick analysis is one of important parts for the timing. This analysis is influenced by investors' knowledge and experiences. With guideline for candle stick's analysis, the investors are able to understand the candle sticks easier. In addition, the guideline can be used for implementation of system trading. In this research, K-means clustering was applied to group candle sticks. The candle sticks were created from the opening, high low and closing prices of AOT CPALL CPN PTT and TRUE. These five stocks were listed in the SET50. The Weka was used as a tool for K-means clustering with 7500 data sets. Five clusters of candle sticks were

resulted and composed of different size of Body, Upper Wick and Lower Wick. Each cluster reflected different perceptions of price movements.

บทนำ

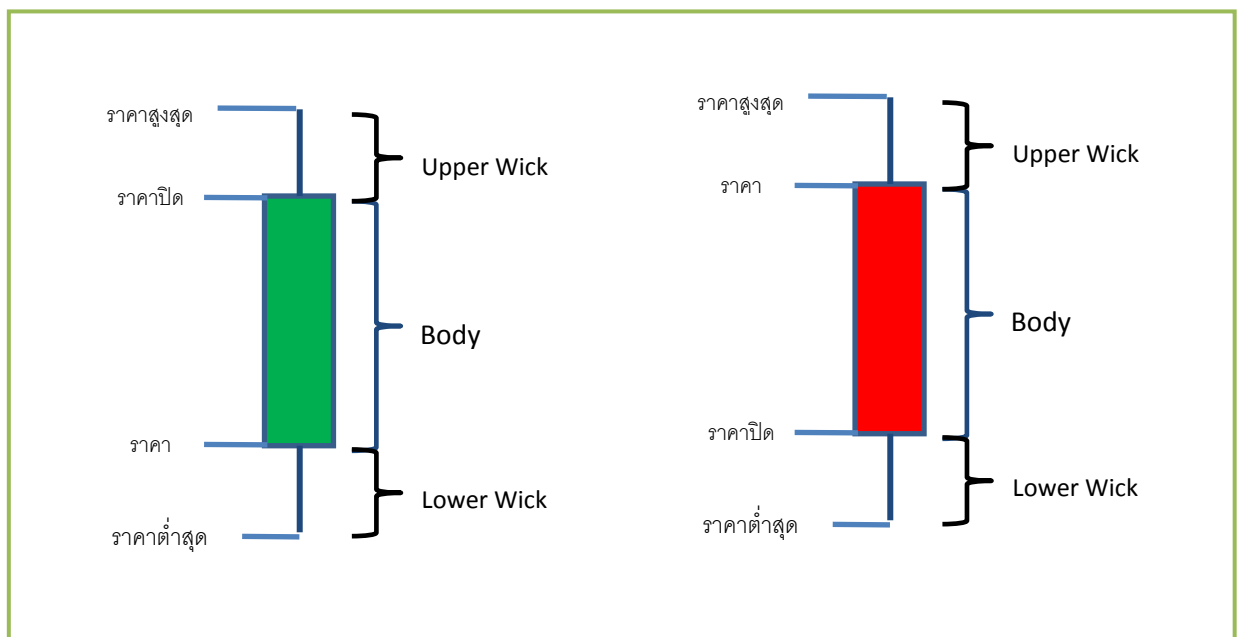
ในการลงทุนในตลาดหลักทรัพย์ของนักลงทุน นอกจากนักลงทุนจะพิจารณาเลือกหุ้นจากปัจจัยพื้นฐานทางเศรษฐกิจ การเมืองของทั้งภายในและภายนอกประเทศ ปัจจัยทางด้านคุณภาพของการดำเนินการของบริษัท และโอกาสในการเติบโตของบริษัทแล้ว การพิจารณาจังหวะในการซื้อขายก็มีส่วนสำคัญในการลงทุน ดังในแนวทางของเคล็ดลับการลงทุนสไตล์ปรมาจารย์ที่มีการกล่าวถึงการต้องรู้จักรักษาวินัยและรอคอยจังหวะและโอกาสการลงทุน (ตลาดหลักทรัพย์แห่งประเทศไทย, 2558) ซึ่งการวิเคราะห์จังหวะและโอกาสในการลงทุนที่ว่านี้ นักลงทุนแต่ละคนอาจมองเห็นจังหวะและโอกาสที่แตกต่างกันซึ่งการวิเคราะห์ดังกล่าวทำให้เกิดการซื้อขายหุ้น ดังจะเห็นว่านักลงทุนบางส่วนต้องการซื้อเนื่องมาจากการวิเคราะห์ว่าหุ้นดังกล่าวน่าจะมีโอกาสที่จะราคาสูงกว่าเดิม ในขณะที่ผู้ที่ถือครองหุ้นนั้นอยู่พิจารณาในอีกด้านและต้องการที่จะขายจึงเกิดการซื้อขายขึ้นได้ ดังนั้นจากข้อมูลที่หลากหลายประกอบกับจำนวนที่มากมายของข้อมูลที่เกิดขึ้นในแต่ละวัน และในส่วนของนักลงทุนเองซึ่งแต่ละคนมีพื้นฐานและประสบการณ์ ที่แตกต่างกัน อีกทั้งมุมมอง การรับรู้ข้อมูลและการวิเคราะห์ที่แตกต่างกันไป

ในการวิเคราะห์การลงทุนก็จะมีนักลงทุนบางส่วนที่เชื่อว่า ราคาของหุ้นนั้นๆ ได้สะท้อนข้อมูลทุกอย่างที่เกี่ยวข้องของหุ้นนั้นแล้ว ดังนั้นการพิจารณาราคาและแนวโน้มของราคา สถิติ และรูปแบบของการเคลื่อนไหวของราคาจึงเป็นส่วนหนึ่งที่ใช้ในการวิเคราะห์จังหวะการลงทุน การพิจารณาแนวโน้มของราคาซึ่งสามารถพิจารณาได้จากกราฟราคา เส้นค่าเฉลี่ยต่างๆ สร้างจากราคาหุ้น และ ทฤษฎีของ Dow ซึ่ง Charles Henry Dow ได้เผยแพร่แนวคิดและหลักการในเรื่องดังกล่าวซึ่งได้รับความนิยมมาจนถึงปัจจุบัน (พิสุทธิ์ ลิ้มวิวัฒน์กุล, 2558) นอกจากการพิจารณาแนวโน้มระยะสั้น ระยะกลางและระยะยาว ตามแนวคิดทฤษฎีของดาว (Dow's Theory) แล้ว การพิจารณาลักษณะของกราฟแท่งเทียนก็เป็นอีกวิธีที่นิยม และสามารถประยุกต์ใช้ร่วมกับการพิจารณาปัจจัยด้านอื่นๆประกอบการลงทุน

ในการพิจารณารูปกราฟแท่งเทียนซึ่งเป็นกราฟที่ให้ข้อมูลอย่างละเอียดที่สุดกราฟหนึ่งซึ่งข้อมูลของกราฟแท่งเทียนหนึ่งแท่งจะประกอบไปด้วย ราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด และสีของแท่งเทียนที่จะเป็นแตกต่างกันดังเช่น เมื่อราคาเปิดมากกว่าราคาปิด แท่งเทียนจะเป็นสีเขียวแต่ถ้าข้อมูลเป็นในทางตรงกันข้ามสีของแท่งเทียนจะเป็นสีแดง นอกจากนั้น Body ซึ่งหมายถึง ความแตกต่างระหว่างราคาเปิดกับราคาปิด Upper Wick แทนความแตกต่างระหว่างราคาปิดกับราคาที่สูงสุดในกรณีที่ราคาปิดสูงกว่าหรือความแตกต่างระหว่างราคาเปิดกับราคาที่สูงสุดในกรณีที่ราคาเปิด

สูงกว่าราคาปิด Lower Wick แทนความแตกต่างระหว่างราคาต่ำสุดกับราคาเปิดในกรณีที่ราคาเปิดต่ำกว่าราคาปิด หรือความแตกต่างระหว่างราคาต่ำสุดกับราคาปิดในกรณีที่ราคาปิดต่ำกว่าราคาเปิด นอกจากนี้รูปร่างและตำแหน่งที่ปรากฏของกราฟแท่งเทียนแต่ละรูปแบบอาจนำไปสู่การใช้ความน่าจะเป็นในการวิเคราะห์แนวโน้มของราคา

ในการพิจารณากราฟแท่งเทียนซึ่งประกอบไปด้วยข้อมูลจำนวนมากทั้งในด้านของความต้องการซื้อและความต้องการขายที่สะท้อนออกในกราฟหนึ่งแท่ง รวมทั้งพิสัยของราคาซื้อขายที่เกิดขึ้น การวิเคราะห์กราฟแท่งเทียนส่วนใหญ่ขึ้นอยู่กับความรู้และประสบการณ์ของนักลงทุน ซึ่งนักลงทุนบางส่วนใช้การพิจารณาในส่วนของแท่งเทียน (Body) เป็นหลัก แต่นักลงทุนบางส่วนอาจจะใช้ทั้งในส่วนที่เป็นเนื้อแท่งเทียนหรือแท่งเทียน (Body) และไส้เทียน (Wick) ประกอบการพิจารณาลักษณะการเคลื่อนไหวของแนวโน้มของราคา โดยทั่วไปลักษณะของแท่งเทียนเป็นดังแผนภาพ 1.1



แผนภาพ 1 กราฟแท่งเทียน

อย่างไรก็ตามในการพิจารณารูปแบบของกราฟแท่งเทียนโดยนักลงทุนที่มีความเชี่ยวชาญจะสามารถจำแนกรูปแบบลักษณะของกราฟและสามารถนำมาใช้ตัดสินใจในการเข้าซื้อหรือขายได้จากสัญญาณที่สื่อความหมายจากกราฟ แต่อย่างไรก็ตามนักลงทุนที่ยังไม่เชี่ยวชาญอาจจะยังไม่สามารถจำแนกรูปแบบลักษณะและความหมายที่สะท้อนจากกราฟได้ และนอกจากนั้นในกรณีของการพัฒนาระบบซื้อขายอัตโนมัติ นั้น การจำแนกรูปแบบลักษณะของกราฟแท่งเทียนอัตโนมัติด้วยการโปรแกรมคอมพิวเตอร์จะสามารถนำไปสู่การวิเคราะห์สัญญาณการเคลื่อนไหวอื่นๆของราคาหุ้นต่อไป ในการจัดกลุ่มข้อมูล วิธีการแบบ K-means เป็นวิธีการจัดกลุ่มที่มีประสิทธิภาพวิธีหนึ่ง ซึ่งสามารถใช้ในการวิเคราะห์ที่หลากหลายตั้งแต่การจัดกลุ่มลูกค้า ไปจนถึงการวิเคราะห์จัดกลุ่มอาชญากรรม ดังนั้น

งานวิจัยนี้จึงประยุกต์ใช้วิธีการจัดกลุ่มแบบ K-means เพื่อแยกลักษณะแห่งเทียนออกเป็นกลุ่มเพื่อนำคุณลักษณะของแต่ละเทียนแต่ละกลุ่มไปวิเคราะห์โอกาสต่อไป

วิธีการวิจัย

ลักษณะของกราฟแห่งเทียน

ในการพิจารณารูปแบบของกราฟแห่งเทียนนั้นนักวิเคราะห์แต่ละคนใช้ประสบการณ์และวิจรรย์ญาณที่แตกต่างกันออกไป โดยที่ พิสุทธิ์ ลิ้มวิวัฒน์กุล (2558) ได้ให้แนวทางการวิเคราะห์ Body ของกราฟแห่งเทียนโดยในด้านของ ระดับความผันผวนหรือความเสี่ยง อัตรากำไรในระยะสั้น ปริมาณการซื้อขายแบบฉับพลัน ระดับการตอบสนองต่อข่าว การใช้ปัจจัยทางด้านเทคนิคในการซื้อขาย การเผื่อระวังในการซื้อขายหุ้น จะสอดคล้องกันกับขนาดความยาวของ Body โดยถ้าความยาวมากก็จะส่งผลให้ด้านต่างๆที่กล่าวมาอยู่ในระดับมาก เช่น ถ้าความยาวของขนาด Body มากจะแสดงว่าขณะนั้นมีความผันผวนมาก แต่ถ้าความยาวของขนาด Body น้อยก็จะมีมีความผันผวนน้อยด้วยเช่นกัน ในขณะที่ระยะเวลาในการถือครองเพื่อทำกำไรจะแปรผกผันกับขนาดความยาวของขนาด Body โดยถ้าความยาวของ Body มาก ระยะเวลาในการถือครองเพื่อทำกำไรจะน้อย ส่วนถ้าความยาวของ Body น้อย ระยะเวลาในการถือครองเพื่อทำกำไรจะมาก โดยที่ พิสุทธิ์ ลิ้มวิวัฒน์กุล (2558) ได้แบ่งขนาดของความยาวของขนาด Body ออกเป็นห้าระดับจากราวที่สั้นที่สุดจนถึงสั้นที่สุด แต่อย่างไรก็ตามในไม่ได้มีการระบุขนาดความยาวที่แน่นอน นอกจาก พิสุทธิ์ ลิ้มวิวัฒน์กุล (2558) แล้วในการวิเคราะห์ของ Coulling (2013) ได้ให้ความสำคัญกับความยาวของ Body ด้วยเช่นกัน แต่แตกต่างกันที่ถ้าความยาวของขนาด Body มีมากแสดงว่าแนวโน้มของราคาหุ้นชัดเจนมากและเป็นไปตามทิศทางของราคา เช่นถ้าแห่งเทียนเป็นสีเขียวและขนาด Body มีความยาวมากแสดงว่าแนวโน้มของราคาหุ้นไปในทิศทางปรับตัวขึ้นชัดเจน ในทางตรงกันข้ามถ้าแห่งเทียนเป็นสีแดงและขนาด Body มีความยาวมากแสดงว่าแนวโน้มของราคาหุ้นไปในทิศทางปรับตัวลงชัดเจน ในขณะที่ขนาด Body ที่สั้นแสดงถึงการที่ทิศทางอาจจะยังไม่ชัดเจน แต่อย่างไรก็ตาม Coulling (2013) ยังให้ความสำคัญกับไส้ของแห่งเทียน (Wick) ซึ่งไส้ของแห่งเทียนจะแสดงถึงความเห็นของนักลงทุนในการเคลื่อนไหวของราคา อย่างไรก็ตามสิ่งที่ทั้งพิสุทธิ์ ลิ้มวิวัฒน์กุล (2558) และ Coulling (2013) ให้ความสำคัญตรงกันคือการที่ผู้ลงทุนจำเป็นจะต้องพิจารณาปัจจัยอื่นๆประกอบการพิจารณาในการลงทุนด้วยเช่นกัน

ในงานวิจัยนี้จึงแบ่งกราฟแห่งเทียนออกเป็นห้ากลุ่ม โดยใช้วิธีการจัดกลุ่มที่เรียกว่า K-means หรือ การวิเคราะห์กลุ่มแบบไม่เป็นขั้นตอน (Nonhierarchical Cluster Analysis) ซึ่งในที่นี้จะกำหนดให้ K เป็นห้าซึ่งหมายความว่า จะมีการแบ่งข้อมูลออกเป็นห้ากลุ่มโดยที่แต่ละกลุ่มมีความใกล้เคียงกันและไม่มีสมาชิกที่ทับซ้อนกัน

วิธีการแบ่งกลุ่มแบบ K-means

เป็นการแบ่งกลุ่มข้อมูลออกเป็นกลุ่มๆ โดยที่จำนวนกลุ่มถูกกำหนดด้วยค่า K และแต่ละกลุ่มจะมีจุดศูนย์กลาง (Centroid) และข้อมูลถูกจัดให้อยู่ในกลุ่มที่มีระยะห่างจากจุดศูนย์กลางถึงข้อมูลชุดนั้นน้อยที่สุด โดยที่ วิภาวรรณ บัวทอง (2557) ได้ให้วิธีการแบ่งกลุ่มแบบ K-means ตามขั้นตอนต่อไปนี้

1. การกำหนดจุดศูนย์กลางเริ่มต้นของแต่ละกลุ่ม โดยที่จำนวนกลุ่มตามจำนวนค่า k ที่กำหนด
2. ทหาระยะห่างระหว่างข้อมูลกับจุดศูนย์กลาง (Centroid) การจัดข้อมูลเข้ากลุ่มด้วยการเลือกกลุ่มที่ข้อมูลนั้นอยู่ใกล้จุดศูนย์กลาง (Centroid) มากที่สุด
3. นำข้อมูลในกลุ่มมาคำนวณเพื่อหาจุดศูนย์กลางใหม่ โดยใช้ค่าเฉลี่ย (mean)
4. ทำซ้ำขั้นตอนที่ 2-3 จนกระทั่งจุดศูนย์กลางไม่มีการเปลี่ยนแปลง

ในการคำนวณหาระยะห่างระหว่างข้อมูลสองชุดหรือระหว่างข้อมูลกับจุดศูนย์กลาง สามารถทำได้ด้วยการใช้วิธีการค่าด้วยสูตรของ Euclidean distance ดังต่อไปนี้

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

โดยที่ d เป็นระยะทางระหว่างข้อมูลชุด X(x₁, x₂, ..., x_n) และชุด Y(y₁, y₂, ..., y_n) โดยที่ n เป็นจำนวนคุณลักษณะของข้อมูลแต่ละชุด

ซอฟต์แวร์ที่ใช้ในการประมวลผล

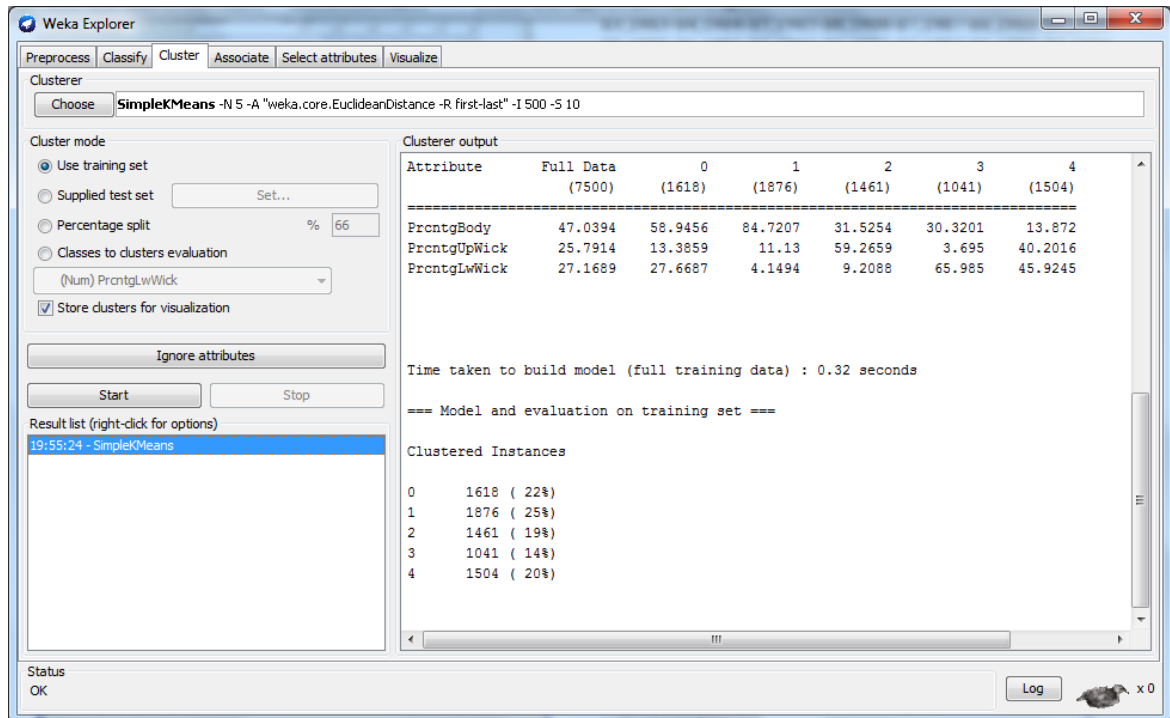
ในการแบ่งกลุ่มแบบ K-means สามารถใช้เครื่องมือหรือซอฟต์แวร์ในการประมวลผลได้หลายซอฟต์แวร์ทั้งที่เป็นแบบซอฟต์แวร์เชิงพาณิชย์ซึ่งมีค่าใช้จ่ายในการใช้ อย่างเช่น MATLAB หรือเป็นแบบไม่มีค่าใช้จ่ายอย่างเช่น Weka ซึ่งแต่ละซอฟต์แวร์ก็จะมีวิธีการใช้ และการออกแบบส่วนที่ติดต่อกับผู้ใช้แตกต่างกันออกไป ในการทำงานวิจัยนี้ผู้วิจัยได้เลือกใช้ Weka เนื่องจากไม่มีค่าใช้จ่าย และซอฟต์แวร์ดังกล่าวเป็นที่นิยมใช้ในงานวิจัยทางด้านการทำงานวิจัยทางด้านเหมืองข้อมูล เช่น Sharma, Alam and Rani (2012) ใช้ Weka ในงานวิจัยเพื่อทำ “K-Means Clustering in Spatial Data Mining” และ Patil, Deshmukh and Rajeswari (2015) ใช้ Weka ในการทำงานวิจัยเรื่อง “Analysis of Simple K-Means with Multiple Dimensions using WEKA” ด้วยเช่นกัน ในที่นี้ผู้วิจัยได้ใช้ซอฟต์แวร์ Weka เวอร์ชัน 3.6.8 เป็นเครื่องมือในการทำการแบ่งกลุ่มแบบ K-means

ข้อมูล

ข้อมูลที่ใช้เป็นข้อมูลราคารายวันของหุ้นที่ถูกสุ่มเลือกมาจำนวน 5 หุ้นที่อยู่หุ้น 50 ตัวแรกของตลาดหลักทรัพย์แห่งประเทศไทย หรือที่เรียกว่า SET50 ซึ่งเป็นข้อมูลของ SET50 ณ วันที่ 24 พฤศจิกายน 2559 โดยข้อมูลของหุ้นแต่ละตัวจะประกอบไปด้วย ราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด โดยหุ้นที่ถูกสุ่มเลือกจำนวน 5 หุ้นประกอบไปด้วย AOT CPALL CPN PTT TRUE โดยข้อมูลของหุ้นทั้งหมดถูกดึงจากตลาดหลักทรัพย์แห่งประเทศไทยโดยใช้โปรแกรม efinStockPickUp ผ่านโบรกเกอร์ MaybankKimEng ณ วันที่ 24 พฤศจิกายน 2559 โดยใช้ข้อมูลของราคา 1,500 วันของหุ้นแต่ละตัว เริ่มตั้งแต่วันที่ 18 สิงหาคม 2553 ถึงวันที่ 4 ตุลาคม 2559 โดยข้อมูลทั้งหมดจำนวน 7,500 ชุด จะถูกนำมาคำนวณเป็นแท่งเทียนดังแผนภาพที่ 1 และจะถูกจัดเป็นข้อมูลสามส่วนตามประกอบความยาวของ Body, Upper Wick และ Lower Wick และนำมาเปรียบเทียบเป็นร้อยละของความยาวของแต่ละส่วนเทียบกับพิสัยของราคาซึ่งได้จากผลต่างของราคาสูงสุดและต่ำสุด เพื่อให้ข้อมูลของหุ้นแต่ละตัวที่มีช่วงราคาที่แตกต่างกันจัดอยู่ในมาตรฐานเดียวกัน

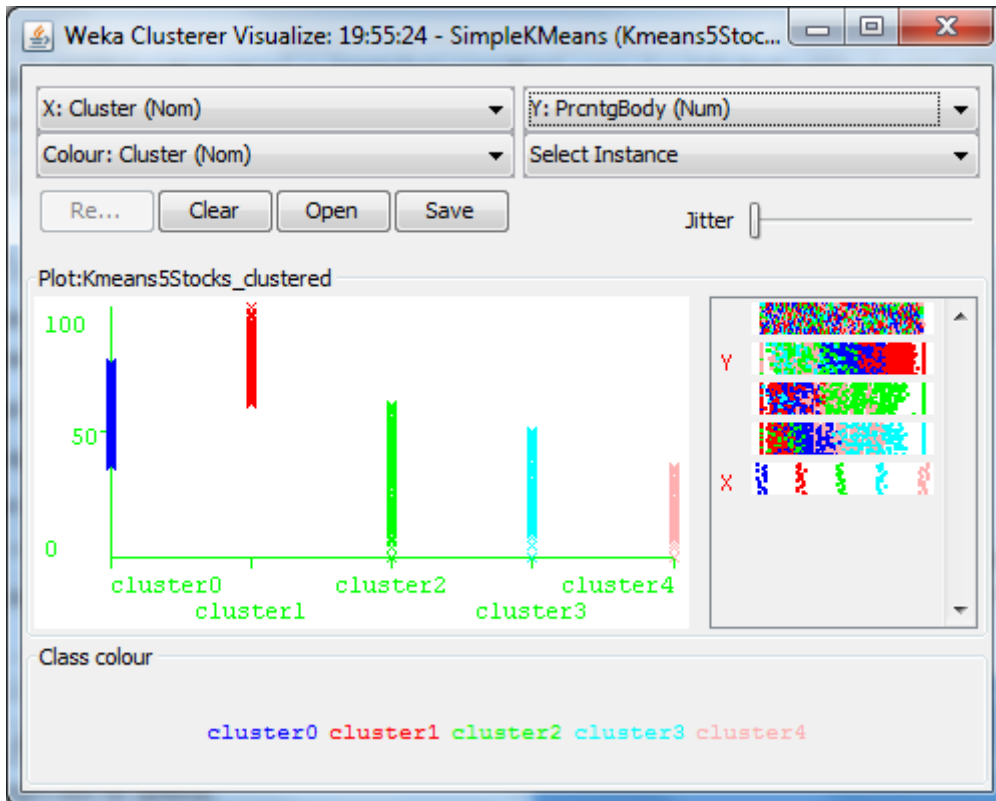
ผลการวิจัย

การนำข้อมูลการซื้อขายประจำวันของหุ้น AOT CPALL CPN PTT และ TRUE จำนวน 7,500 ชุด เข้าสู่โปรแกรม Weka ด้วยการเลือกวิธีการแบบ K-means โดยคิดระยะห่างระหว่างชุดข้อมูลด้วยวิธีการแบบ EuclideanDistance และกำหนดให้จำนวนกลุ่มหรือค่า K เป็น 5 จะได้ผลการทำงานดังในแผนภาพที่ 2 ซึ่งแสดงถึงจุดศูนย์กลางของข้อมูลแต่ละกลุ่มและจำนวนชุดข้อมูลที่อยู่ในแต่ละกลุ่ม ในที่นี้โปรแกรม Weka จะเริ่มหมายเลข Cluster จาก 0 ถึง 4 แต่ในที่นี้จะเรียกเป็นกลุ่มที่ 1-5 โดยการเทียบ cluster 0 เป็นกลุ่มที่ 1 และเทียบเคียงแบบเดียวกันกับ cluster อื่นๆ

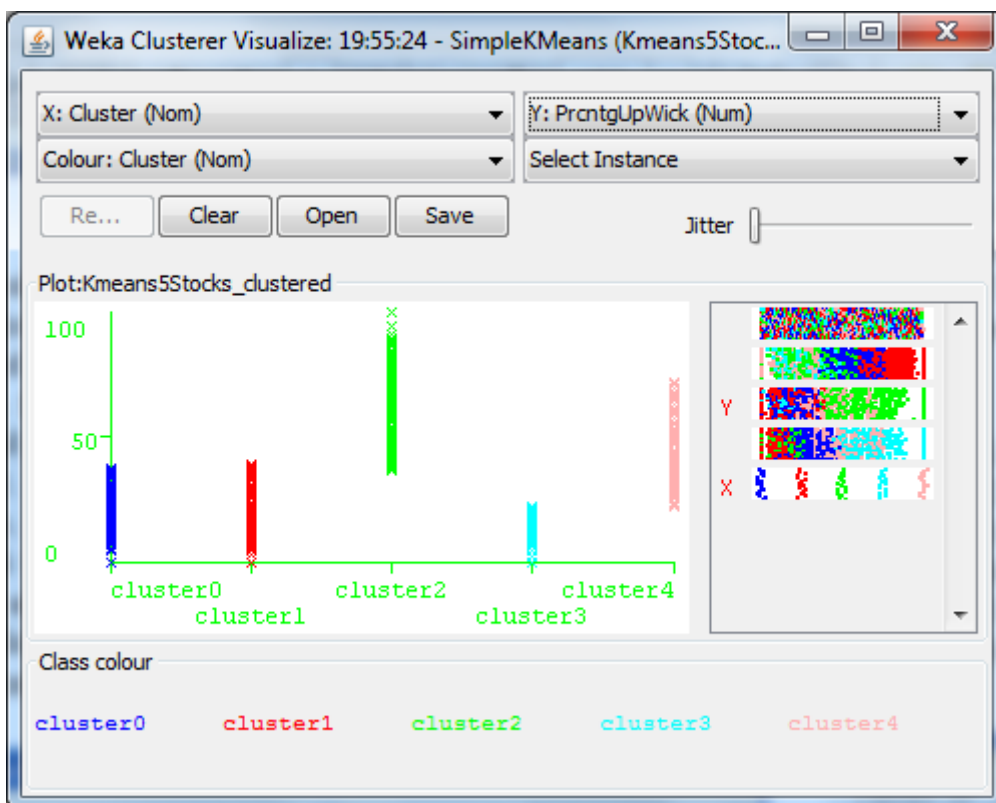


แผนภาพที่ 2 แสดงผลการทำงานของ K-means ในโปรแกรม Weka

จากแผนภาพด้านบนจะเห็นว่าข้อมูลแต่ละชุดมีจุดศูนย์กลางที่แตกต่างกันและข้อมูลที่อยู่ในกลุ่มที่ 2 หรือ Cluster 1 มีจำนวนชุดข้อมูลมากที่สุด นอกจากนั้นยังพบว่า เป็นกลุ่มที่มีส่วนของ Body ยาวที่สุดซึ่งสอดคล้องกันกับการแสดงกราฟในแผนภาพที่ 3 ซึ่งแกน X เป็น กลุ่มข้อมูล แกน Y เป็นร้อยละของความยาวของ Body ของกราฟแท่งเทียนซึ่งจะเห็นว่า cluster 1 หรือข้อมูลชุดที่สองมีความยาวของ Body มากกว่าร้อยละ 50 ไปจนถึงร้อยละ 100

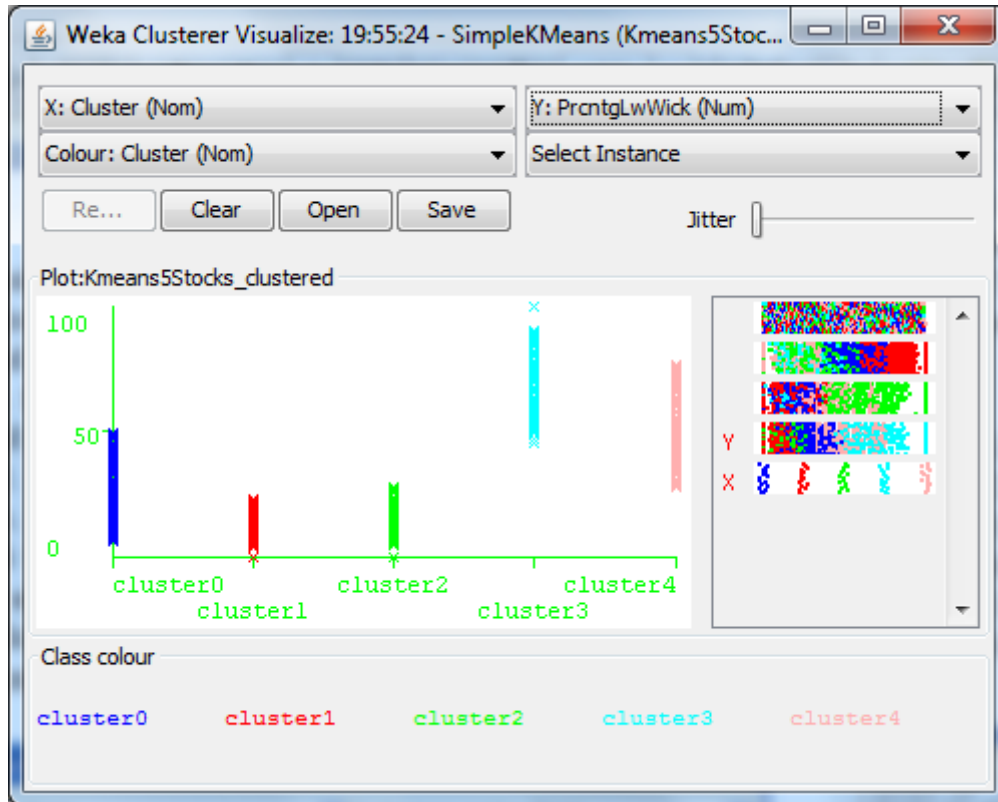


แผนภาพที่ 3 แสดงข้อมูลร้อยละของ Body ของข้อมูลในแต่ละกลุ่ม



แผนภาพที่ 4 แสดงข้อมูลร้อยละของ Upper Wick ของข้อมูลในแต่ละกลุ่ม

แผนภาพด้านบนแสดงผลการแบ่งกลุ่มด้วยวิธีการ K-means โดยแกน X เป็นกลุ่มข้อมูล แกน Y เป็นร้อยละของความยาวของ Upper Wick ของกราฟแท่งเทียน จะเห็นว่า cluster 2 หรือ กลุ่มที่ 3 มี Upper Wick สูงที่สุดและรองลงมาคือ cluster4 หรือกลุ่มที่ 5



แผนภาพที่ 5 แสดงข้อมูลร้อยละของ Lower Wick ของข้อมูลในแต่ละกลุ่ม

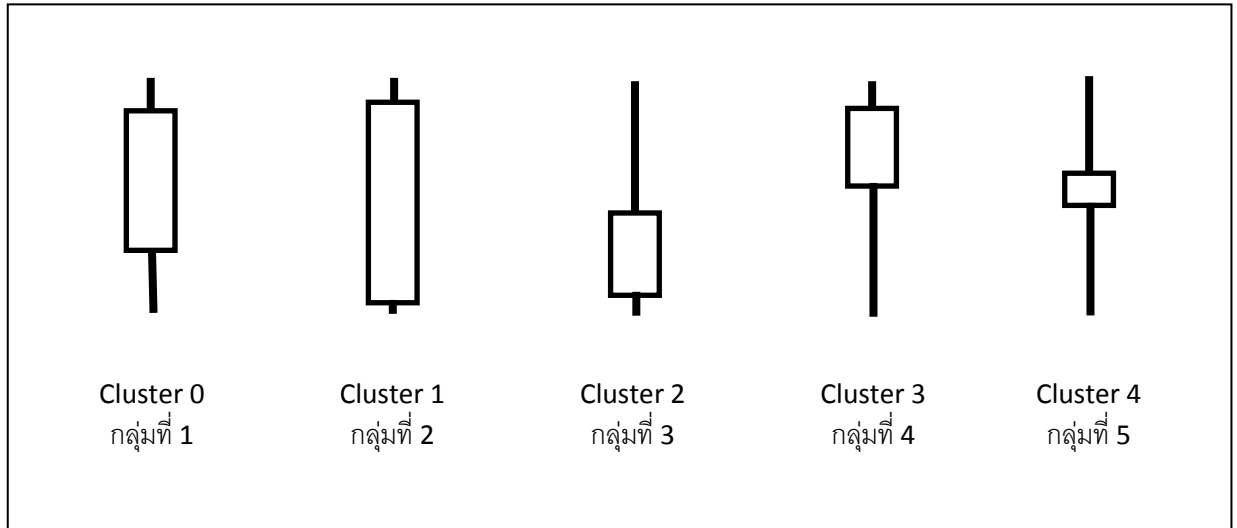
แผนภาพด้านบนแสดงผลการแบ่งกลุ่มด้วยวิธีการ K-means โดยแกน X เป็นกลุ่มข้อมูล แกน Y เป็นร้อยละของความยาวของ Lower Wick ของกราฟแท่งเทียน จะเห็นว่า cluster 3 หรือ กลุ่มที่ 4 มี Lower Wick สูงที่สุดและรองลงมาคือ cluster4 หรือกลุ่มที่ 5 ส่วน cluster1 หรือ กลุ่มที่ 2 มี Lower Wick ต่ำที่สุด

จากผลการทำงานของโปรแกรมจะสามารถนำมาเป็นข้อมูลในการแบ่งกลุ่มกราฟแท่งเทียนออกเป็น 5 กลุ่มด้วยร้อยละของความยาวของ Body, Upper Wick และ Lower Wick ดังตารางที่ 1

ตารางที่ 1 แสดงจุดศูนย์กลางของข้อมูลแต่ละกลุ่ม

	กลุ่มที่ 1	กลุ่มที่ 2	กลุ่มที่ 3	กลุ่มที่ 4	กลุ่มที่ 5
ร้อยละของ Body	59	85	32	30	14
ร้อยละของ Upper Wick	13	11	59	4	40
ร้อยละของ Lower Wick	28	4	9	66	46

เมื่อนำข้อมูลจากตารางที่ 1 ไปเขียนกราฟแท่งเทียนในสัดส่วนที่แสดงตามตารางจะได้กราฟแท่งเทียน 5 แท่งที่มีลักษณะแตกต่างกันค่อนข้างชัดเจนดังแผนภาพที่ 6



แผนภาพที่ 6 แสดงกราฟแท่งเทียนที่ได้จากการแบ่งกลุ่มด้วยวิธีการ K-means

จากแผนภาพด้านบนจะพบว่าสัดส่วนความยาวของ Body ของแต่ละกลุ่มจะมีความแตกต่างกัน และแม้ว่าขนาดของ Body ใน cluster 2 และ 3 มีความใกล้เคียงกันแต่ว่าลักษณะที่บ่งบอกในเรื่องของการเคลื่อนไหวของราคาแตกต่างกัน cluster 2 มีการแนวโน้มที่แรงขายจะมากกว่าแรงซื้อเนื่องจาก Upper Wick มีขนาดค่อนข้างยาวและราคาปิดหรือราคาเปิดค่อนข้างแตกต่างจากราคาสูงสุด ในส่วนของ cluster 3 จะมีแนวโน้มที่แรงซื้อมากกว่าแรงขายเนื่องจาก Body อยู่ในส่วนที่ด้านบนของพิสัยของราคาและ Upper Wick มีขนาดค่อนข้างสั้น แสดงว่าราคาปิดหรือราคาเปิดมีความแตกต่างจากราคาสูงสุดไม่มาก

สรุปและอภิปรายผล

การนำข้อมูลรายวันของหุ้น 5 หุ้นที่อยู่หุ้น 5 ตัวแรกของตลาดหลักทรัพย์แห่งประเทศไทย ประกอบไปด้วย AOT CPALL CPN PTT และ TRUE โดยข้อมูลรวมทั้งหมด จำนวน 7,500 ชุด ในข้อมูลแต่ละชุดประกอบด้วย ราคาเปิด ราคาปิด ราคาสูงสุด ราคาต่ำสุด เพื่อจัดในรูปแบบของกราฟแท่งเทียนที่ประกอบด้วย Body, Upper Wick และ Lower Wick

การแยกกลุ่มของกราฟแท่งเทียนโดยใช้วิธีการ K-means ด้วยการคำนวณหาระยะห่างด้วยวิธีการแบบ Euclidean Distance และกำหนดให้จำนวนกลุ่มหรือค่า K เป็น 5 โดยกราฟแท่งเทียนห้ากลุ่มมีรูปร่างที่แตกต่างกัน โดยกลุ่มที่ 2 จะมีขนาดของ Body ร้อยละ 85 เมื่อเทียบกับพิสัยของการเคลื่อนไหวของราคา กลุ่มที่ 1 ขนาดของ Body ร้อยละ 59 เมื่อเทียบกับพิสัยของการเคลื่อนไหวของราคา ซึ่งส่วนของ Upper Wick และ Lower Wick มีสัดส่วนประมาณร้อยละ 40 เทียบกับความยาวทั้งหมดของกราฟแท่งเทียน ในขณะที่กลุ่มที่ 3 และ 4 มีขนาดของ Body ใกล้เคียงกัน โดยประมาณร้อยละ 30 ซึ่งใกล้เคียงกันแต่ว่าขนาดของ Upper Wick กับ Lower Wick แตกต่างกันซึ่งจะแสดง

ถึงลักษณะของการเคลื่อนไหวของราคาที่แตกต่างกัน ดังนั้นการแยกลักษณะของกราฟแท่งเทียนจะสามารถทำได้ด้วยการใช้ K-means และการแยกลักษณะดังกล่าวสามารถใช้เป็นแนวทางในการทำงานของเครื่องมืออื่นเพื่อใช้ในการวิเคราะห์ลักษณะของกราฟแท่งเทียนต่อไปได้

ในการวิเคราะห์ลักษณะของราคาจากกราฟแท่ง นักลงทุนแต่ละคนจะใช้ความรู้และประสบการณ์ในการวิเคราะห์ซึ่งอาจจะแตกต่างกันออกไป โดยที่ พิสุทธิ ลิ้มวิวัฒน์กุล (2558) ให้แนวทางการวิเคราะห์ตามความยาวของ Body ของแท่งเทียนออกเป็นห้าระดับโดยเรียงจาก Body ยาวมากที่สุดถึงน้อยที่สุด โดยให้ปัจจัยต่อไปนี้ ระดับความผันผวนหรือความเสี่ยง ความสามารถมีอัตรากำไรในระยะสั้น และมีปริมาณการซื้อขายแบบฉับพลัน จะมีมากถ้าความยาวของ Body ยาวมาก และเมื่อความยาวของ Body น้อยจะมีระดับความเข้มข้นของปัจจัยดังกล่าวด้วยเช่นกัน (พิสุทธิ ลิ้มวิวัฒน์กุล, 2558) นอกจากนี้ระดับการตอบสนองต่อข่าว การการซื้อขายด้วยการใช้ปัจจัยทางด้านเทคนิค และความต้องการการเผื่อระวังในการซื้อขายหุ้นก็จะแปรผันตามระดับความยาวของ Body ด้วยเช่นกัน (พิสุทธิ ลิ้มวิวัฒน์กุล, 2558) ในขณะที่ความหมายของความยาวของ Body จะเป็นดังนี้ ถ้า Body ของแท่งเทียนยาวมาก จะมีระดับความผันผวนหรือความเสี่ยงมาก สามารถมีอัตรากำไรในระยะสั้นได้มาก และมีปริมาณการซื้อขายแบบฉับพลันมากด้วยเช่นกัน นอกจากนี้ระดับการตอบสนองต่อข่าว การซื้อขายด้วยการใช้ปัจจัยทางด้านเทคนิค และความต้องการการเผื่อระวังในการซื้อขายหุ้นก็จะมากด้วยเช่นกัน ในขณะที่ Coulling (2013) ให้ความสำคัญกับความยาว Body Upper Wick และ Lower Wick ด้วยเช่นกัน แต่การแปลความหมายของกราฟแท่งเทียนในแต่ละลักษณะจะแตกต่างกัน โดยที่ถ้า Body ความยาวมากจะแสดงถึงทิศทางแนวโน้มของราคาหุ้นนั้นจะชัดเจนไปตามทิศทางของราคาของกราฟแท่งนั้น ในขณะที่กราฟแท่งเทียนที่มี Body สั้นมากแสดงว่าทิศทางยังอยู่ในช่วงที่ทิศทางยังไม่ชัดเจน ซึ่งจะสอดคล้องกับรูปร่างของกราฟแท่งเทียนใน cluster 5 ดังนั้นผลการแบ่งกราฟแท่งเทียนด้วยการใช้ K-means แบบห้ากลุ่มจะเห็นว่ามียุคที่ 2 (cluster 1) กลุ่มที่ 5 (cluster 4) สอดคล้องกับแนวทางของทั้ง พิสุทธิ ลิ้มวิวัฒน์กุล (2558) และ Coulling (2013) ซึ่งอย่างไรก็ตามในการวิเคราะห์กราฟแท่งเทียนมีหลายแนวทางและการประยุกต์ใช้การวิเคราะห์ดังกล่าวขึ้นอยู่กับความรู้และประสบการณ์ในการลงทุนของนักลงทุน

ข้อจำกัดในการวิจัย

1. การนำข้อมูลของราคาเปิด ราคาปิด ราคาสูงสุด และราคาต่ำสุดของหุ้นที่แตกต่างกัน ซึ่งช่วงราคาของแต่ละหุ้นแตกต่างกันมาเทียบสัดส่วนให้อยู่ในลักษณะของร้อยละ อาจจะทำให้การจำแนกลักษณะของกราฟแท่งเทียนแตกต่างจากลักษณะเดิม เนื่องจากความยาวของ Body อาจจะเป็นร้อยละเช่นเหมือนกัน แต่เป็นร้อยละเช่นที่คำนวณมาจากตัวเลขที่แตกต่างกัน
2. การใช้ K-means นั้นการปรับค่า K เป็นผลต่อการแบ่งกลุ่ม ในการเปรียบเทียบหลายค่าเพื่อหาค่า K ที่เหมาะสมอาจจะทำให้ผลของการทำงานแตกต่างกันออกไป

ข้อเสนอแนะจากการวิจัย

การนำหุ้นที่มีช่วงราคาที่ใกล้เคียงกัน หรือการศึกษาหุ้นเพียงตัวเดียวอาจจะทำให้การทำงานของ K-means จำแนกกราฟแท่งเทียนได้เหมาะสมกับหุ้นตัวนั้น

บรรณานุกรม

- ตลาดหลักทรัพย์แห่งประเทศไทย (2558). วางแผนการลงทุน . สืบค้นจาก
https://www.set.or.th/education/th/start/start_start_3_6.pdf
- พิสุทธิ ลิ้มวิวัฒน์กุล. (2558). SYSTEM TRADING เทรดหุ้นอย่างเป็นระบบ. กรุงเทพฯ: Dream & Passion (TH)
- วิภาวรรณ บัวทอง (2557). Chapter 8 Clustering : K-means. สืบค้นจาก
<https://wipawanblog.files.wordpress.com/2014/06/chapter-8-clustering-k-means.pdf>
- Coulling, A. (2013). A Complete Guide To Volume Price Analysis (Kindle Version). CreateSpace Independent Publishing Platform
- Patil, R., Deshmukh, S.,& Rajeswari, K. (2015). Analysis of Simple K-Means with Multiple Dimensions using WEKA. International Journal of Computer Applications (0975 – 8887) Volume 110 – No. 1, January 2015
- Sharma,R. , Alam, M.A. ,& Rani, A. (2012). K-Means Clustering in Spatial Data Mining using Weka Interface. International Conference on Advances in Communication and Computing Technologies (ICACACT) 2012