

การเพิ่มประสิทธิภาพสำหรับการวิเคราะห์สถานการณ์ตลาดน้ำมันดิบโดย  
ประยุกต์ใช้เทคนิคการสกัดข้อมูลและพจนานุกรมข้อมูล  
**IMPROVEMENT THE EFFICIENCY OF CRUDE OIL MARKET TREND  
ANALYSIS BY USING INFORMATION EXTRACTION AND DATA DICTIONARY**

พีรพล สิทธิวิจารณ์<sup>1</sup>, อนิรุทธ์ ชันธสะอาด<sup>2</sup>, ยุทธนา สมสอน<sup>3</sup> และ ณัฐชัย ตั้งวิจนะโยบาย<sup>4</sup>

<sup>1</sup>อาจารย์, สาขาการจัดการโลจิสติกส์ คณะโลจิสติกส์ มหาวิทยาลัยบูรพา

169 ถนนลงหาดบางแสน ต.แสนสุข อ.เมือง จ.ชลบุรี 20131, peerapoljt@gmail.com

<sup>2</sup>อาจารย์, สาขาการค้าระหว่างประเทศฯ คณะโลจิสติกส์ มหาวิทยาลัยบูรพา

169 ถนนลงหาดบางแสน ต.แสนสุข อ.เมือง จ.ชลบุรี 20131, anirutka@gmail.com

<sup>3</sup>กัปตัน, บริษัท ซีรอน จำกัด, 292/12 หมู่บ้านวิสตาปาร์ค ถ.วิภาวดีรังสิต แขวงตลาดบางเขน

เขตหลักสี่ กทม. 10210, yuthana@cheiron.co.th

<sup>4</sup>นักวิเคราะห์และวางแผน, บริษัท ปตท. จำกัด (มหาชน), 555 ถ.วิภาวดีรังสิต แขวงจตุจักร

เขตจตุจักร กทม. 10900, nuttachai\_t@pttplc.com

Peerapol Sittivijan<sup>1</sup>, Anirut Kantasa-ard<sup>2</sup>, Yuthana Somsorn<sup>3</sup> and Nuttachai Tangwatjanayobai<sup>4</sup>

<sup>1</sup>Instructor, Logistics Management program, Faculty of Logistics Burapha University

169 Long-Haad Bangsaen Rd. Saensuk Subdistrict, Muang District, Choburi 20131,

Thailand, t\_wonginta@hotmail.com

<sup>2</sup>Instructor, International Trade and Logistics Management program, Faculty of

Logistics Burapha University, 169 Long-Haad Bangsaen Rd. Saensuk Subdistrict,

Muang District, Chonburi 20131, Thailand, anirutka@gmail.com

<sup>3</sup>Captain, Cheiron Office, 292/12 Village Vista Park Vibhavadi, Vibhavadi Rangsit Rd,

Talat Bang Khen, Lak Si, Bangkok 10210, yuthana@cheiron.co.th

<sup>4</sup>Analyst, PTT Public Company Limited, 555 Vibhavadi Rangsit Rd, Chatuchak,

Bangkok 10900, Thailand, nuttachai\_t@pttplc.com

### บทคัดย่อ

งานวิจัยนี้นำเสนอการพัฒนาระบบฐานข้อมูลสำหรับรวบรวมและจัดเก็บข้อมูลข่าวจากสื่อประเภทต่าง ๆ ร่วมกับการใช้เทคนิคการสกัดข้อมูล (Information Extraction) และการกำหนดขอบเขตกลุ่มคำด้วยพจนานุกรมข้อมูล (Data Dictionary) ซึ่งประกอบด้วยกลุ่มคำที่ระบุแนวโน้มการขึ้นลง

ของราคา (Behavior Trend) และกลุ่มคำที่ระบุปัจจัยที่เกี่ยวข้องกับความรู้สึกผู้ซื้อขาย (Sentimental Factor) เพื่อช่วยให้นักวิเคราะห์ตลาดน้ำมันดิบสามารถทำงานได้รวดเร็วขึ้น ทั้งนี้ในส่วนของการสกัดข้อมูลและการใช้พจนานุกรมข้อมูลส่งผลให้การวิเคราะห์ข่าวใช้เวลาเร็วขึ้นและผลการพยากรณ์สถานการณ์ราคาน้ำมันดิบสอดคล้องกับราคาน้ำมันดิบที่เกิดขึ้นประมาณร้อยละ 60 โดยใช้กลุ่มตัวอย่างข่าวจำนวน 30 วัน รวมถึงมีการนำเสนอสมการแบบจำลองสำหรับแนวทางการพยากรณ์แนวโน้มการขึ้นลงของราคาน้ำมันดิบและประยุกต์ใช้ผ่านระบบ Web application ที่ผู้วิจัยพัฒนาขึ้น ซึ่งจากข้อมูลดังกล่าวจะเป็นแนวทางในการพัฒนาแบบจำลองสำหรับการพยากรณ์แนวโน้มราคาน้ำมันดิบแบบอัตโนมัติในลำดับต่อไป

**คำสำคัญ:** การสกัดข้อมูล, พจนานุกรมข้อมูล, แบบจำลองการพยากรณ์, ฐานข้อมูลเชิงโครงสร้าง, ตลาดน้ำมันดิบ

### ABSTRACT

This research demonstrates the data collection system for gathering and collecting a number of news and articles in crude oil market and presents how to analyze the trend of crude oil market by using the concept of Information Extraction and Data Dictionary, which consists of Behavior Trend and Sentimental Factor contexts. The following concepts will help analyst to predict the oil market trend faster even though the prediction process still be semi-automated. The precision rate is approximately 60 percent after comparing with the actual crude oil price due to the sample news in 30 days. In addition, this paper illustrates the prediction equation model to proof how to predict the trend of oil price and implement the model via the web application. The researchers believe that both information extraction and data dictionary could be a prospect direction to improve the prediction model of crude oil price being more automation.

**KEYWORDS:** Information Extraction, Data Dictionary, Prediction Model, Textual Database Structure, Crude Oil Market

### 1. บทนำ

การคาดการณ์ราคาน้ำมันดิบในตลาดโลกเป็นสิ่งที่ได้รับความสนใจจากภาคธุรกิจที่มีการลงทุนที่เกี่ยวข้องกับธุรกิจน้ำมันดิบ เช่น โรงกลั่นน้ำมันดิบ ธุรกิจการขนส่ง เป็นต้น อย่างไรก็ตาม การคาดการณ์ราคาน้ำมันดิบในตลาดโลกค่อนข้างมีความซับซ้อน เนื่องจากมีปัจจัยที่เกี่ยวข้องหลายด้านด้วยกัน ได้แก่

- ปัจจัยพื้นฐาน (Fundamental Factors) ซึ่งเป็นปัจจัยที่กำหนดอุปสงค์ (Demand) และอุปทาน (Supply) ของน้ำมันดิบในตลาดโลก ปัจจัยที่มีผลต่อระดับอุปสงค์ของน้ำมันดิบในตลาดโลก ได้แก่ การเจริญเติบโตในภาคเศรษฐกิจ สภาพภูมิอากาศที่เปลี่ยนแปลงไป ซึ่งจะมีผลกระทบต่อความต้องการใช้น้ำมันดิบในตลาดโลกโดยตรง ในส่วนของปัจจัยที่มีผลต่อระดับอุปทานของน้ำมันดิบในตลาดโลก ได้แก่ นโยบายและกำลังการผลิตของกลุ่มผู้ผลิตน้ำมันดิบหลายใหญ่ของโลก (OPEC) ปริมาณน้ำมันดิบสำรองของประเทศผู้บริโภครายสำคัญของโลก ระดับการใช้พลังงานทดแทน เป็นต้น

- ปัจจัยทางความรู้สึกของผู้ซื้อขายในตลาดน้ำมันดิบ (Sentimental Factor) เป็นปัจจัยที่ไม่ได้มีผลกระทบโดยตรงต่อระดับอุปสงค์และอุปทานของน้ำมันดิบในตลาดโลกในปัจจุบัน แต่มีผลต่อความรู้สึกและการคาดการณ์ระดับของอุปสงค์และอุปทานของน้ำมันดิบในอนาคตของผู้เกี่ยวข้องหลัก และทำให้เกิดการเปลี่ยนแปลงของราคาน้ำมันดิบในปัจจุบัน ตัวอย่างหนึ่งของปัจจัยทางความรู้สึกของผู้ซื้อขายในตลาดน้ำมันดิบสามารถแสดงได้ตามรูปที่ 1 ซึ่งแสดงให้เห็นถึงการเพิ่มขึ้นและลดลงของราคาน้ำมันดิบในตลาดโลกจากข่าวภาวะสงครามในภูมิภาคตะวันออกกลางและเหตุการณ์อื่นๆ ในระยะเวลา 6 ปี (ปี 2010-2016) ที่ผ่านมา

Daily crude oil spot prices, 2010-16  
dollars per barrel

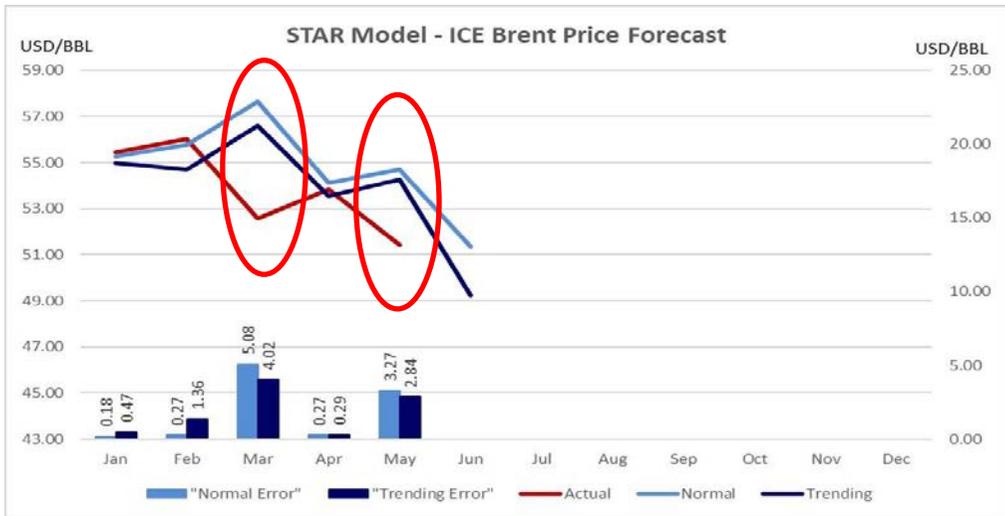


รูปที่ 1 การเพิ่มขึ้นและลดลงของราคาน้ำมันดิบในตลาดโลกอ้างอิงจากราคา BRENT และ WTI [1]

- ปัจจัยทางเทคนิค (Technical Factor) เช่น สถิติรายงานค่าเฉลี่ยย้อนหลังของราคาน้ำมันดิบและข้อมูลการซื้อขายน้ำมันดิบล่วงหน้า (future market) เป็นต้น

- ปัจจัยอื่นๆ (Miscellaneous Factor) เช่น อัตราแลกเปลี่ยนเงินตรา โดยเฉพาะค่าเงินดอลลาร์สหรัฐ ซึ่งใช้เป็นสกุลเงินมาตรฐานในการซื้อขายน้ำมันดิบในตลาดโลก

โดยส่วนใหญ่แบบจำลองที่ใช้ในการพยากรณ์ราคาน้ำมันดิบในตลาดโลกจะใช้ข้อมูลปัจจัยพื้นฐานและปัจจัยทางเทคนิคเป็นหลักในการคาดการณ์ราคาน้ำมันดิบในตลาดโลก โดยแบบจำลองจะใช้ปัจจัยเหล่านี้ในการคาดการณ์ระดับอุปสงค์และอุปทานในอนาคตของน้ำมันดิบในตลาดโลก และพยายามคำนวณจุดสมดุล (Equilibrium) ระหว่างระดับของอุปสงค์และอุปทานในอนาคต เพื่อใช้ในการคาดการณ์ราคาน้ำมันดิบในอนาคต โดยข้อมูลปัจจัยพื้นฐานและปัจจัยทางเทคนิคโดยส่วนใหญ่จะเป็นข้อมูลเชิงปริมาณ (Quantitative) และอยู่ในรูปแบบมาตรฐาน (Structured Data) ซึ่งทำให้ง่ายต่อการนำมาใช้เป็นข้อมูลนำเข้า (Input data) ของแบบจำลองสำหรับปัจจัยทางความรู้สึกของผู้ซื้อขายในตลาดน้ำมันดิบ (Sentimental factors) เป็นปัจจัยที่ยากต่อการนำมาพิจารณาไว้ในแบบจำลองเนื่องจากโดยส่วนใหญ่เป็นข้อมูลเชิงคุณภาพ (Qualitative) และไม่อยู่ในรูปแบบมาตรฐานของข้อมูล (Unstructured Data) จึงไม่ค่อยได้มีการนำเข้ามาร่วมใช้งานในการพยากรณ์ราคาน้ำมันดิบในตลาดโลกมากนัก อย่างไรก็ตามปัจจัยทางความรู้สึกของผู้ซื้อขายในตลาดน้ำมันดิบมีผลต่อระดับราคาน้ำมันดิบในตลาดโลกเป็นอย่างมาก และมักทำให้เกิดการเปลี่ยนแปลงในลักษณะที่รุนแรง (Radical Change) ต่อราคาน้ำมันดิบในตลาดโลกอยู่บ่อยครั้ง ทำให้แบบจำลองที่ใช้เฉพาะข้อมูลเชิงปริมาณมีความผิดพลาดในการพยากรณ์ในกรณีนี้ ตัวอย่างในรูปที่ 2 แสดงผลของการพยากรณ์ราคาน้ำมันดิบในตลาดโลก โดยใช้แบบจำลอง Stepwise Regression Model ที่ใช้เฉพาะข้อมูลปัจจัยพื้นฐานและปัจจัยเทคนิคซึ่งเป็นข้อมูลแบบโครงสร้างเชิงปริมาณเป็นตัวแปร ในการอธิบายราคาน้ำมันดิบ โดยจะเห็นได้ว่าผลการพยากรณ์ของบริษัทการศึกษาในเดือนมีนาคมและพฤษภาคมมีความผิดพลาดในลักษณะที่ตรงกันข้ามกับการลดลงของราคาน้ำมันดิบในตลาดโลกจริงๆ ซึ่งข้อมูลในรูปที่ 2 ได้มาจากการวิเคราะห์ร่วมกับ บริษัท ปตท. จำกัด (มหาชน) และผู้วิจัยโดยกระบวนการสัมภาษณ์และเก็บข้อมูลจากรายงานของบริษัท



รูปที่ 2 ความผิดพลาดจากการพยากรณ์ราคาน้ำมันดิบโดยใช้แบบจำลอง step-wise regression

สำหรับงานวิจัยนี้จะมุ่งเน้นการพัฒนากระบวนการพื้นฐานข้อมูลที่เหมาะสมสำหรับการจัดเก็บข้อมูล ปัจจัยทางความรู้สึกผ่านข่าวประเภทต่าง ๆ โดยจะมีการนำเทคนิคการสกัดข้อมูล (Information Extraction) และพจนานุกรมข้อมูล (Data Dictionary) มาช่วยในการสร้างแบบจำลองเพื่อใช้วิเคราะห์ข้อมูลปัจจัยทางความรู้สึกของผู้ซื้อขายในตลาดน้ำมันดิบ ซึ่งข้อมูลดังกล่าวอยู่ในรูปแบบของข้อมูลเชิงคุณภาพ ทั้งนี้ เทคนิคดังกล่าวสามารถสกัดข้อมูลปัจจัยทางความรู้สึกของผู้ซื้อขายในเบื้องต้นสำหรับช่วยวิเคราะห์สถานการณ์ตลาดน้ำมันดิบในตลาดโลกได้

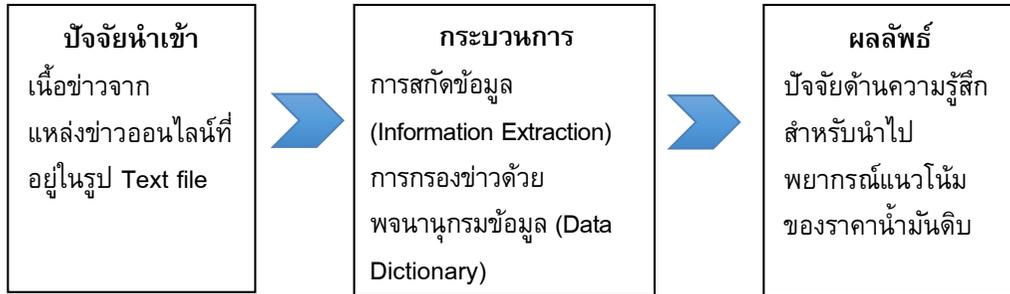
## 2. วัตถุประสงค์

1) วิเคราะห์หาความสัมพันธ์ระหว่างข้อมูลปัจจัยทางความรู้สึกของผู้ซื้อขายในตลาดน้ำมันดิบ ประเภทต่าง ๆ กับการเปลี่ยนแปลงราคาน้ำมันดิบในตลาดโลก ได้แก่ ข่าวเกี่ยวกับกลุ่มผู้ค้าน้ำมันดิบรายใหญ่ของโลกและปัจจัยที่มีผลต่อการผลิตน้ำมันดิบ กลุ่มผู้ใช้น้ำมันดิบรายใหญ่ของโลกและปัจจัยที่มีผลต่อการผลิตน้ำมันดิบ และปัจจัยอื่นๆ เป็นต้น

2) พัฒนาระบบฐานข้อมูลที่เหมาะสมและมีประสิทธิภาพเพื่อใช้ในการดึงและเก็บรวบรวมข้อมูลเกี่ยวกับปัจจัยทางความรู้สึกของผู้ซื้อขายในตลาดน้ำมันดิบจากแหล่งข่าวที่เกี่ยวข้องกับราคาน้ำมันดิบในตลาดโลก เพื่อนำมาวิเคราะห์ราคาน้ำมันดิบผ่านเทคนิคการสกัดข้อมูลและพจนานุกรมข้อมูล

### 3. กรอบแนวคิด

ในการวิจัยนี้ผู้วิจัยได้กำหนดกรอบแนวคิดในการวิจัยไว้ดังนี้



รูปที่ 3 รายละเอียดกรอบแนวคิดการวิจัย

สำหรับกรอบแนวคิดการวิจัยนั้น มีองค์ประกอบทั้งหมด 3 ส่วนคือ ปัจจัยนำเข้า กระบวนการ และผลลัพธ์ สำหรับปัจจัยนำเข้าที่มีนำมาพิจารณาในกระบวนการวิจัยครั้งนี้ประกอบด้วยเนื้อหาข่าวที่เกี่ยวข้องกับสถานการณ์การขึ้นลงของราคาน้ำมัน ซึ่งเนื้อหาข่าวดังกล่าวมาจากแหล่งข้อมูลออนไลน์และอยู่ในรูปของ Text File สำหรับกระบวนการที่นำมาใช้ในการวิเคราะห์คือ การสกัดข้อมูลข่าว (Information Extraction) และพจนานุกรมข้อมูล (Data Dictionary) ซึ่งหลักการดังกล่าวจะนำเสนอรายละเอียดในหัวข้อวิธีการวิจัย และส่วนสุดท้ายคือผลลัพธ์ที่คาดว่าจะได้รับ คือสมการและปัจจัยด้านความรู้สึก (Sentimental Factor) สำหรับนำไปใช้พยากรณ์แนวโน้มราคาน้ำมันดิบ

### 4. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

จากที่ได้มีการศึกษา พบว่าองค์ความรู้ที่เกี่ยวข้องกับแนวทางการพัฒนาระบบจัดเก็บข้อมูลข่าวที่จะนำมาใช้เป็นฐานข้อมูลสำหรับสกัดและวิเคราะห์ปัจจัยทางความรู้สึกของผู้ซื้อขายน้ำมันดิบ มีทฤษฎีและงานวิจัยที่เกี่ยวข้องแบ่งออกเป็น 3 ส่วน คือ แบบจำลองที่ใช้ในการพยากรณ์สถานการณ์น้ำมันดิบในปัจจุบัน (Existing Crude Oil Price Prediction Model) เทคนิคการสกัดข้อมูล (Information Extraction) และ พจนานุกรมข้อมูล (Data Dictionary) โดยมีรายละเอียดตามที่อธิบายด้านล่างดังนี้

#### 4.1 แบบจำลองดั้งเดิมที่ใช้ในการพยากรณ์สถานการณ์ราคาน้ำมันดิบ (Existing Crude Oil Price Prediction Model)

สถานการณ์ราคาของน้ำมันดิบปัจจุบันค่อนข้างมีความผันผวนสูงและเปลี่ยนแปลงเร็ว ทั้งในส่วนของข้อมูลราคาน้ำมันดิบและข่าวสารต่างๆ ที่เกี่ยวข้องกับตลาดน้ำมันดิบ โดยความท้าทาย

ของการพยากรณ์ราคาน้ำมันดิบคือ ความผันผวนที่สามารถเกิดขึ้นได้ตลอดเวลา ทั้งความผันผวนของราคาน้ำมันดิบระยะสั้นและระยะยาว ทั้งนี้ได้มีงานวิจัยอยู่จำนวนมากที่พัฒนาแบบจำลองสำหรับการพยากรณ์ราคาน้ำมันดิบ อาทิ

Abdullah SN ได้ทำการวิจัยในเรื่องของการพัฒนาแบบจำลองสำหรับการพยากรณ์ราคาน้ำมันดิบ โดยมีการนำแบบจำลองแนวคิดลำดับชั้น (Hierarchical Conceptual Model) มาใช้ในการพยากรณ์ข้อมูลเชิงปริมาณ เช่น กำลังการผลิตน้ำมันดิบในกลุ่มประเทศ OPEC และ Non-OPEC, อัตราการใช้น้ำมันดิบในจีนและอินเดีย, ระดับน้ำมันดิบคงคลังในสหรัฐอเมริกา, และปัจจัยทางด้านเศรษฐศาสตร์ นอกจากนี้ งานวิจัยดังกล่าวยังได้พัฒนาแบบจำลองสำหรับการพยากรณ์ข้อมูลเชิงคุณภาพด้วย อาทิ การพัฒนาแบบจำลองตามหลักการ Rule-Base Expert Model, การนำหลักการของ Fuzzy Grammar Extraction มาใช้ในการสร้างข้อกำหนดสำหรับการพยากรณ์แนวโน้มราคาซึ่งประเภทของข้อมูลที่นำมาใช้ในการพยากรณ์จะประกอบด้วย เนื้อหาข่าวที่เกี่ยวข้องกับสถานการณ์ตลาดน้ำมันดิบผ่าน Google News รวมถึงบทความวิเคราะห์ราคาน้ำมันดิบในสื่อออนไลน์ต่างๆ โดยสิ่งสำคัญที่จะทำให้การพยากรณ์เชิงคุณภาพมีความแม่นยำมากขึ้น คือการแบ่งประเภทของกลุ่มคำ ซึ่งจะใช้การนับค่าความถี่ของประเภทกลุ่มคำที่เกิดขึ้นมาเป็นปัจจัยชี้วัดว่าแนวโน้มราคาน้ำมันดิบจะเพิ่มขึ้นหรือลดลง เช่น กลุ่มคำนาม (Core Term) ที่เป็นปัจจัยทางตรงของราคาน้ำมันดิบ (Supply, Oil Price, Demand, Refinery) และกลุ่มคำกริยา (Content Category) ที่เป็นตัวกำหนดแนวโน้มการขึ้นลงของราคาน้ำมันดิบ เป็นต้น [2] นอกจากนี้งานวิจัยดังกล่าวยังได้มีการพัฒนาแบบจำลองสำหรับพยากรณ์ข้อมูลราคาน้ำมันดิบแบบผสมระหว่างข้อมูลเชิงปริมาณและข้อมูลเชิงคุณภาพ (Linguistic-Quantitative Hybrid Model) เพื่อให้เกิดความแม่นยำมากขึ้นในการพยากรณ์ราคาน้ำมันดิบ

#### 4.2 เทคนิคการสกัดข้อมูล (Information Extraction)

งานวิจัยของ Gupta ที่ได้มีการประยุกต์ใช้เทคนิคเหมืองข้อความ (Text Mining) กับเนื้อความประเภทต่างๆ ซึ่งหนึ่งในเทคนิคที่น่าสนใจและถูกนำมาประยุกต์ใช้มากที่สุดคือ เทคนิคการสกัดข้อมูล (Information Extraction) โดยมีหลักการคือ ระบุความสัมพันธ์ของตัวแปรต่างๆ อาทิ บุคคล, สถานที่, ช่วงเวลาที่มีความเกี่ยวข้องกัน ผ่านคำหรือวลีในเนื้อความนั้นๆ และนำความสัมพันธ์ดังกล่าวมาสร้างเงื่อนไขในการพยากรณ์แนวโน้ม (Pattern Matching) [3] นอกจากนี้ยังมีงานของ Mittermayer ที่ได้นำเสนอเกี่ยวกับการประยุกต์ใช้เทคนิค การเตรียมความพร้อมก่อนประมวลผลข้อความ (Text Preprocessing) ซึ่งเป็นหนึ่งในเทคนิคของ การทำเหมืองข้อความร่วมกับข้อมูลข่าวในตลาดหุ้น เพื่อทำการวิเคราะห์และพยากรณ์การขึ้นลงของราคาหุ้น โดยประเภทของข่าวจะถูกจำแนกออกเป็น 3 รูปแบบคือ กลุ่มข่าวดี (Good News) กลุ่มข่าวร้าย (Bad News) และกลุ่มข่าวที่ไม่มีผลต่อการขึ้นลงของราคาหุ้น (No Mover) [4] นอกจากนี้ยังมีงานวิจัยบางงานที่มีการพยากรณ์

ข้อมูลโดยใช้ทฤษฎีแบบผสมผสานระหว่างกระบวนการทำเหมืองข้อมูลเชิงปริมาณ(Data Mining) และการทำเหมืองข้อมูลแบบวิเคราะห์โครงสร้างเนื้อความ (Text Mining) อาทิ Kloptchenko ได้นำเสนอเกี่ยวกับการวิเคราะห์สถานการณ์การเงินโดยใช้เทคนิคเหมืองข้อมูล (Data Mining) ร่วมกันกับเทคนิคเหมืองข้อความ (Text Mining) โดยการวิเคราะห์สถานการณ์ด้านการเงินจะพิจารณาทั้งข้อมูลเชิงปริมาณและข้อมูลเชิงคุณภาพ ซึ่งเทคนิคย่อยที่นำมาใช้ในการพยากรณ์คือ Text Clustering และ Self-Organization Mapping [5] หรืองานวิจัยของ Larsen ที่มีการนำหลักการจัดกลุ่มข้อมูล (Clustering Concept) มาใช้ในการทำเหมืองข้อมูลแบบวิเคราะห์โครงสร้างเนื้อความ (Text Mining) โดยวิธีการทำ Cluster จะแบ่งเป็น 2 ส่วนคือ Feature Extraction จะทำหน้าที่ในการสกัดเนื้อหาและคำที่อยู่ในบทความ และมีการใส่ค่าน้ำหนักตามเนื้อหาที่พบ รวมถึงการเลือกวิเคราะห์เนื้อหาหรือกลุ่มคำตามค่าน้ำหนักจากมากไปน้อย, และ Clustering จะหน้าที่จับกลุ่มเนื้อหาที่มีค่าน้ำหนักใกล้เคียงกันมากที่สุดมาเป็นข้อมูลพยากรณ์แนวโน้ม [6] รวมถึงหนังสือบางเล่มที่มีการนำเสนอเทคนิคต่างๆที่ใช้ในการทำเหมืองข้อมูลแบบวิเคราะห์โครงสร้างเนื้อความ (Text Mining) โดยหนึ่งในวิธีที่นิยมนำมาประยุกต์ใช้กับงานวิจัย คือ เทคนิคการสกัดกลุ่มคำหรือวลี (Information Extraction) จากบทความ โดยสิ่งสำคัญที่จะทำให้วิธีนี้มีประสิทธิภาพยิ่งขึ้น คือ ผู้วิจัยจะต้องดำเนินการสร้างความสัมพันธ์ระหว่างกลุ่มคำให้ถูกต้อง และสามารถแสดงผลลัพธ์ได้อย่างมีประสิทธิภาพ [7]

#### 4.3 พจนานุกรมข้อมูล (Data Dictionary)

ในการพัฒนาระบบฐานข้อมูลสำหรับจัดเก็บและรวบรวมปัจจัยทางความรู้สึกของผู้ซื้อขายนั้น จะต้องมีการสร้างพจนานุกรมข้อมูล (Data Dictionary) สำหรับจัดเก็บคำสำคัญและขอบเขตของคำที่เกี่ยวข้องกับปัจจัยทางความรู้สึกของผู้ซื้อขาย เพื่อใช้เป็นข้อมูลตั้งต้นสำหรับการพัฒนาแบบจำลองพยากรณ์แนวโน้มการขึ้นลงของราคาน้ำมันดิบสำหรับข้อมูลเชิงคุณภาพในอนาคต ทั้งนี้ ได้มีงานวิจัยที่มีการนำเสนอเกี่ยวกับการประยุกต์ใช้พจนานุกรมข้อมูล อาทิ Clifton และ Thuraisingham ได้นำเสนอข้อมูลที่นำเสนอเกี่ยวกับการประยุกต์ใช้พจนานุกรมข้อมูล อาทิ Clifton และ Thuraisingham ได้นำเสนอข้อมูลที่นำเสนอเกี่ยวกับการประยุกต์ใช้พจนานุกรมข้อมูล อาทิ Clifton และ Thuraisingham ได้ นำเสนอข้อมูลที่นำเสนอเกี่ยวกับการประยุกต์ใช้พจนานุกรมข้อมูล อาทิ Clifton และ Thuraisingham ได้ นำเสนอข้อมูลที่นำเสนอเกี่ยวกับการประยุกต์ใช้พจนานุกรมข้อมูล อาทิ Clifton และ Thuraisingham ได้

มีความแม่นยำในการวิเคราะห์ผลเพิ่มมากขึ้นคือการสร้างพจนานุกรมข้อมูลสำหรับจัดเก็บปัจจัยนำเข้า ขอบเขตของตัวแปร และผลลัพธ์ที่คาดว่าจะเกิดขึ้น ซึ่งหากข้อมูลที่นำเข้ามาไม่ได้อยู่ในขอบเขตที่ถูกระบุไว้ในพจนานุกรมข้อมูล เหมืองข้อมูลชุดนั้นก็จะทำการปฏิเสธผลลัพธ์หรือแสดงผลเฉพาะข้อมูลนำเข้าที่อยู่ในขอบเขตของตัวแปรที่ระบุไว้ [8] หรืออีกงานวิจัยของ Edge ที่ได้กล่าวถึงการประยุกต์ใช้พจนานุกรมข้อมูลสำหรับจัดเก็บข้อมูลอาการของโรคมะเร็งในระยะต่างๆ เพื่อใช้ในการวินิจฉัยโรคและวางแผนในการรักษาในแต่ละระยะที่เกิดขึ้น [9] จากงานวิจัยทั้ง 2 งานเป็นตัวอย่างของการนำพจนานุกรมข้อมูลมาใช้ในการพยากรณ์ทิศทางหรือ

ผลลัพธ์ที่จะเกิดขึ้นในอนาคตซึ่งสอดคล้องกับทิศทางของงานวิจัยชั้นนี้ที่ต้องการสร้างพจนานุกรมข้อมูลสำหรับเก็บขอบเขตของปัจจัยนำเข้าที่เกี่ยวข้องกับการวิเคราะห์สถานการณ์น้ำมันดิบ

## 5. ขอบเขตการวิจัย

จากการศึกษาในงานวิจัยและบทความที่เกี่ยวข้องกับการพยากรณ์สถานการณ์ราคาน้ำมันดิบ รวมถึงเทคนิคที่ใช้ในการวิเคราะห์ข้อมูลทั้งข้อมูลเชิงปริมาณและข้อมูลเชิงคุณภาพนั้น พบว่างานวิจัยส่วนใหญ่จะทำการวิเคราะห์ข้อมูลในระยะยาว เช่น การพยากรณ์ข้อมูลล่วงหน้ารายไตรมาสหรือรายปี [10] หรือ การสร้างข้อกำหนดในการพยากรณ์ข้อมูลแบบตายตัว (Fixed Rule-Based Expert System) [11] ดังนั้น งานวิจัยนี้จะมีการนำหลักการจากงานวิจัยข้างต้นมาพัฒนาต่อยอดโดยมีรายละเอียดของขอบเขตการวิจัยดังนี้

5.1 ผู้วิจัยดำเนินการแบ่งกลุ่มของข้อมูลปัจจัยทางความรู้สึกของผู้ซื้อขายในตลาดน้ำมันดิบออกได้เป็น 3 ประเภทในเบื้องต้น ได้แก่ กลุ่มข่าวดี (Positive News) กลุ่มข่าวร้าย (Negative News) และกลุ่มที่ไม่มีเคลื่อนไหว (No Movers) โดยผู้วิจัยได้ทำการสัมภาษณ์เจ้าหน้าที่/นักวิเคราะห์ของบริษัท ปตท. จำกัด (มหาชน) และผู้เชี่ยวชาญที่มีประสบการณ์ด้านการซื้อขายน้ำมันดิบในตลาดโลก เพื่อให้สามารถสรุปปัจจัยทางความรู้สึกของผู้ซื้อขายในตลาดน้ำมันดิบและความสัมพันธ์ของปัจจัยเหล่านี้ที่มีผลต่อการเพิ่มขึ้นและลดลงของราคาน้ำมันดิบในตลาดโลก

5.2 ผู้วิจัยดำเนินการพัฒนาระบบจัดเก็บข้อมูลข่าวโดยใช้เทคนิคการสกัดข้อมูล (Information Extraction) และพจนานุกรมข้อมูล (Data Dictionary) ซึ่งพัฒนาในรูปแบบ Web application โดยใช้ภาษา PHP และ R มาใช้ในการนำเข้าและวิเคราะห์ข้อมูลข่าว และรองรับการสร้างแบบจำลองสำหรับการพยากรณ์ราคาน้ำมันดิบในระยะสั้น เช่น การพยากรณ์ราคาน้ำมันดิบรายวันและรายสัปดาห์ ซึ่งแบบจำลองที่ใช้ในการพยากรณ์จะใช้ตัวแบบการตัดสินใจแบบแผนภูมิต้นไม้ (Decision Tree Classification Model) ซึ่งเป็นหนึ่งในแบบจำลองสำหรับการวิเคราะห์ข้อมูลของ Supervised Machine Learning และข้อมูลข่าวจะอยู่ในรูปแบบของข้อความแจ้งเตือน (Pop-up Message) จากแหล่งที่มาของข่าวตั้งแต่ 2 แหล่งข่าวขึ้นไป

5.3 หลังจากนั้นผู้วิจัยดำเนินการสร้างเงื่อนไขในการกรองคำ (Text Filtering) ของคำที่ระบุในพจนานุกรมข้อมูลให้มีความยืดหยุ่นมากขึ้น เพื่อให้การพยากรณ์แนวโน้มราคาน้ำมันดิบเกิดความแม่นยำและถูกต้องมากที่สุด

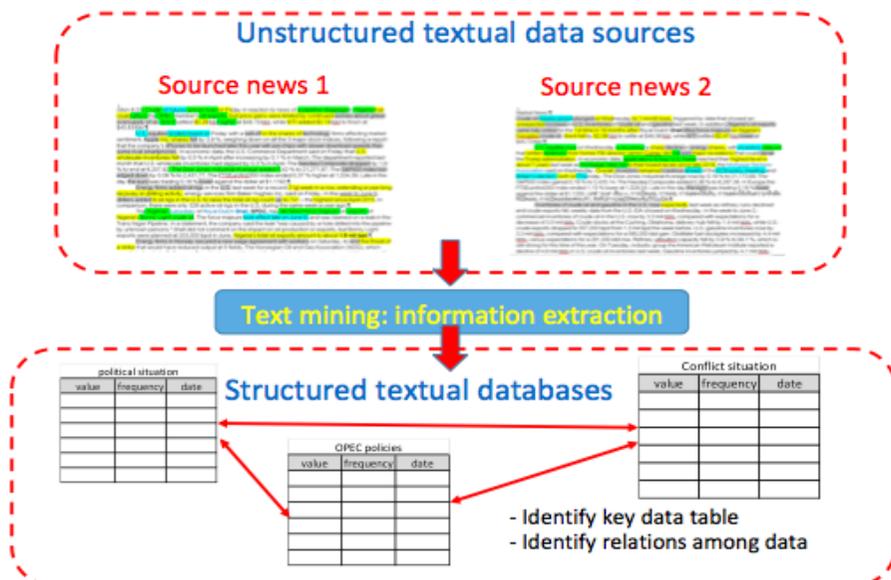
## 6. วิธีการวิจัย

ผู้วิจัยได้ร่วมกับเจ้าหน้าที่บริษัท ปตท. จำกัด (มหาชน) ตั้งแต่เดือนพฤษภาคม พ.ศ. 2560 เพื่อพัฒนาแบบจำลองที่ใช้ข้อมูลปัจจัยความรู้สึกแบบ Unstructured โดยในขอบเขตของงานวิจัยนี้

จะเน้นในเรื่องของการพัฒนาระบบฐานข้อมูลเพื่อจัดเก็บข่าว ตลอดจนการใช้เทคนิคสกัดข้อมูลและพจนานุกรมข้อมูลในการสกัดเอาปัจจัยที่เกี่ยวข้องกับความรูสึกู้ชื้อขายมาเป็นข้อมูลตั้งต้น เพื่อจะแสดงให้เห็นถึงความเป็นไปได้ในการสร้างแบบจำลองเพื่อใช้ในการคาดการณ์ราคาน้ำมันดิบในตลาดโลกโดยใช้ข้อมูลปัจจัยความรูสึกู้ชื้อขาย โดยมีรายละเอียดดังนี้

### 6.1 การสกัดข้อมูลและพัฒนาระบบสำหรับจัดเก็บข่าว (Information Extraction and Textual Structured Platform)

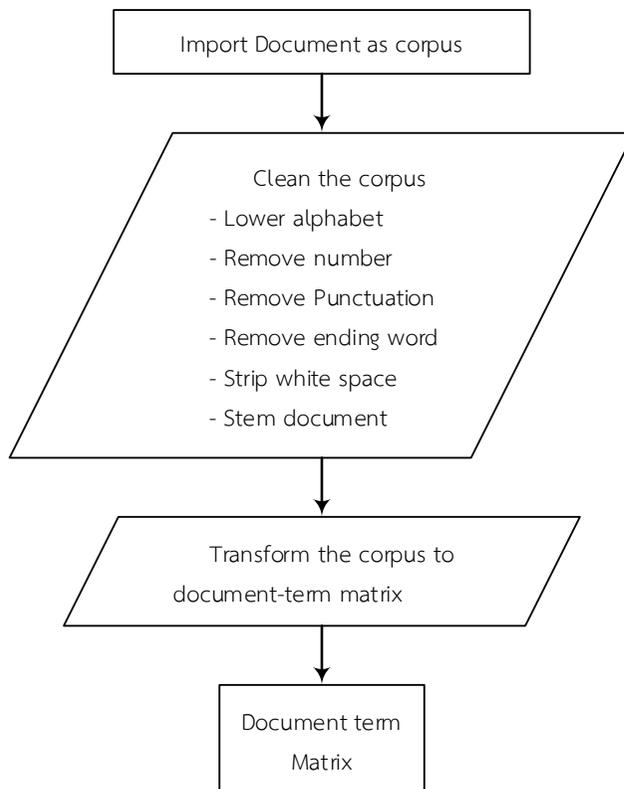
สำหรับงานในส่วนนี้ผู้วิจัยได้ทำการพัฒนาในเบื้องต้นถึงอัลกอริทึมเพื่อใช้ในการสกัด ข้อมูลจากข่าวที่เกี่ยวข้องกับราคาน้ำมันดิบในตลาดโลกที่อยู่ในรูปแบบของข้อความและประโยค (Unstructured Data) ให้มาอยู่ในรูปของข้อมูลเชิงโครงสร้าง (Textual Structured Databases) ซึ่งการสกัดข้อมูลได้มีการใช้ภาษา R ผ่านโปรแกรม RStudio ในการพัฒนาเทคนิคส่วนนี้ โดยมีรายละเอียดตามที่แสดงในด้านล่างตามรูปที่ 4 และ 5 รวมถึงแผนผังของการสกัดข้อมูล (Information Extraction Flow Chart) ดังรูปที่ 6 ซึ่งแผนผังดังกล่าวจะอธิบายแนวทางการสกัดข้อมูล โดยเริ่มจากการนำเข้ามาข้อมูลในรูปแบบ Corpus จากนั้นจะมีการตรวจสอบรูปแบบของข้อมูลที่นำเข้ามาหรือที่เรียกว่า Data Cleaning อาทิ การเปลี่ยนรูปแบบตัวอักษรให้มีขนาดเล็กทั้งหมด การลบเครื่องหมายและช่องว่างระหว่างคำออกไป รวมถึงการแปลงคำๆนั้นให้อยู่ในรูปของรากศัพท์ (Root Word) จากนั้นจะมีการแปลงข้อมูล Corpus ให้อยู่ในรูปของโครงสร้างข้อมูลเชิงเมทริกซ์ และค่อยดำเนินการวิเคราะห์ข้อมูลตามแนวคิดเหมืองข้อความ (Text Mining) ในลำดับถัดไป



รูปที่ 4 ขั้นตอนการสกัดข้อมูลและรูปแบบข้อมูลเชิงโครงสร้าง

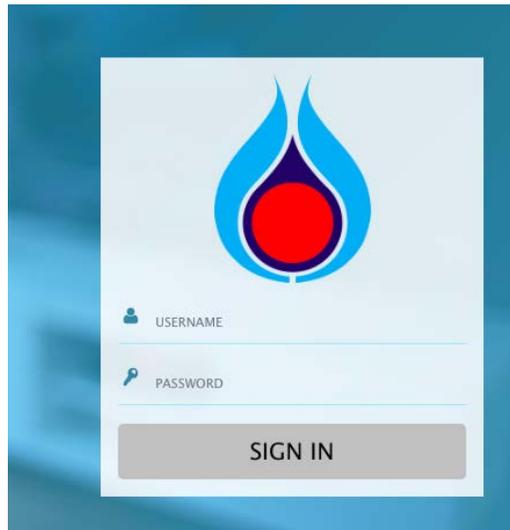
Docs	Terms	Freq
1	accord	1
4	accord	2
8	accord	1
10	accord	1
30	accord	2
31	accord	1
36	acreat	1
62	acreat	2
71	activ	1
77	activ	1
87	activ	1
106	adam	1
141	administr	5
176	advoc	1
211	aim	1
237	aim	1
246	alaska	1
281	alexandra	1
316	also	1
317	also	2
322	also	1

รูปที่ 5 ข่าวก่ถูแ่ลงให้อยู่ในรูปของข้อมูลเชิงโครงสร้าง

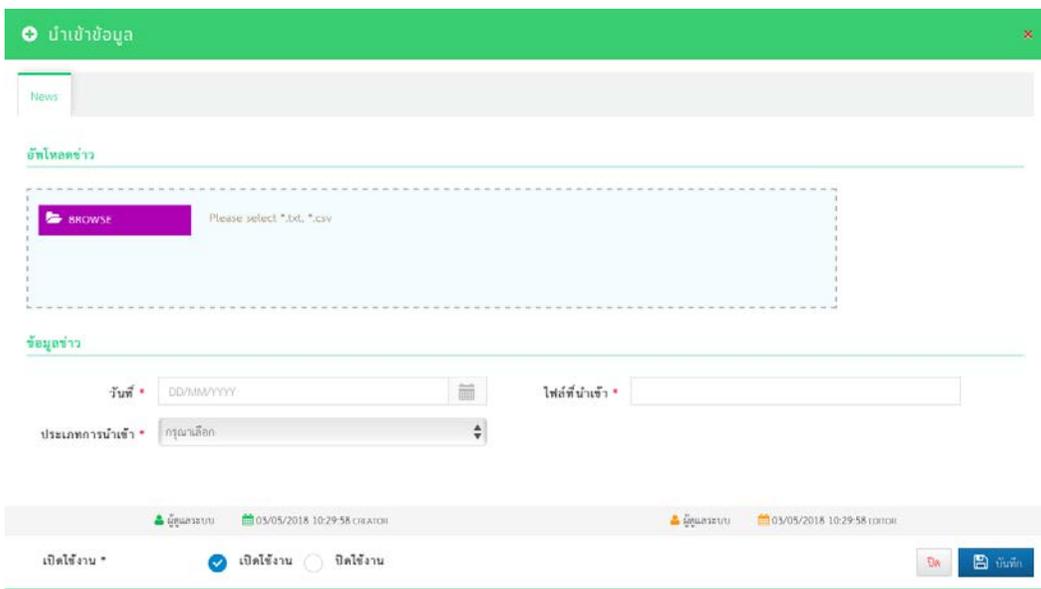


รูปที่ 6 Information Extraction Flow Chart

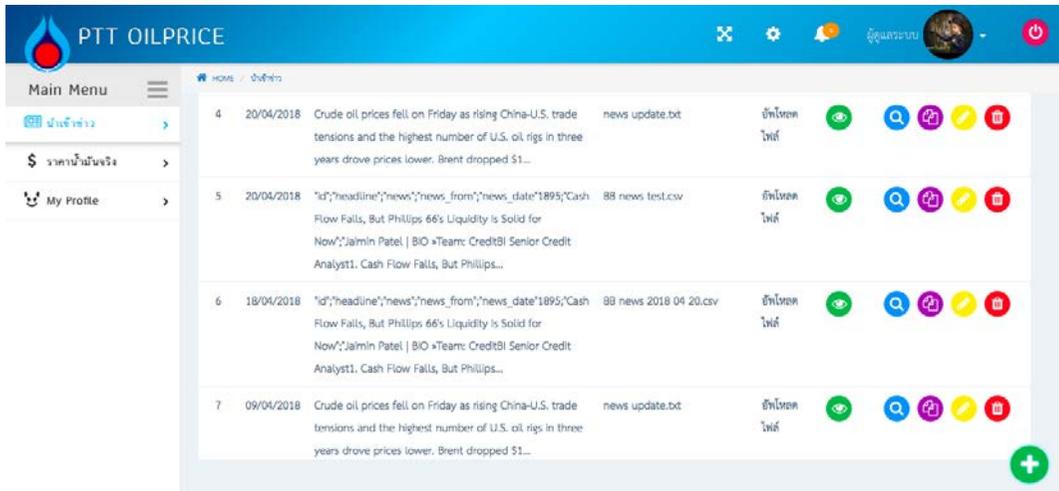
นอกจากนี้ ทางผู้วิจัยได้ดำเนินการพัฒนาระบบฐานข้อมูลสำหรับการจัดเก็บข่าว (User Interface) โดยฐานข้อมูลดังกล่าวจะถูกแบ่งการใช้งานออกเป็น 2 ส่วนคือ ส่วนที่นำเข้าข้อมูลข่าว (Back-End Interface) และส่วนที่แสดงผลลัพธ์ (Front-End Interface) ตามที่แสดงในรูปที่ 7 ถึงรูปที่ 10



รูปที่ 7 หน้าจอสำหรับ Login เข้าระบบ



รูปที่ 8 หน้าจอสำหรับนำเข้าข้อมูลข่าว



รูปที่ 9 หน้าจอแสดงผลการนำเข้าข้อมูลข่าว



รูปที่ 10 หน้าจอแสดงผลลัพท์ feed ข่าว

จากระบบดังกล่าว ปัจจุบันผู้ใช้งานสามารถนำเข้าข้อมูลข่าวได้เพียง 1 รูปแบบ คือ การนำเข้าข่าวในรูปแบบ Text File ทั้งนี้ผู้ใช้งานจะต้องดำเนินการเลือกวันที่ในการนำเข้าข่าวและประเภทการนำเข้า จากนั้นผู้ใช้งานทำการกดปุ่มบันทึกที่มุมขวาล่าง ระบบก็จะดำเนินการบันทึกข้อมูลข่าวที่เรานำเข้าเรียบร้อยแล้ว

## 6.2 การสร้างพจนานุกรมข้อมูล (Data Dictionary Platform)

นอกจากนี้ ทางผู้วิจัยได้ดำเนินการพัฒนาในส่วนของพจนานุกรมข้อมูลเพื่อจัดเก็บและระบุกลุ่มคำ รวมถึงความสัมพันธ์ระหว่างกลุ่มคำที่จะใช้ในการพยากรณ์สถานการณ์ราคาน้ำมันดิบ โดยพจนานุกรมข้อมูลจะถูกแบ่งรูปแบบการจำแนกข้อมูลออกเป็น 2 กลุ่ม ประกอบด้วย กลุ่มคำที่ระบุ

แนวโน้มการขึ้นลงของราคา (Behavior Trend) และกลุ่มคำที่ระบุปัจจัยที่เกี่ยวข้องกับความรู้สึกผู้ซื้อขาย (Sentimental Factor) สำหรับกลุ่มแรกนั้นจะแบ่งแนวโน้มการขึ้นลงของราคาน้ำมันดิบออกเป็น 3 กลุ่มย่อย อาทิ กลุ่มแนวโน้มราคาเพิ่มขึ้น (Increasing), กลุ่มแนวโน้มราคาลดลง (Decreasing) และกลุ่มที่ไม่มีการเปลี่ยนแปลงราคา (Neutral) นอกจากนี้หากต้องการที่จะเพิ่มคำนำหน้าของข้อมูลเช่นราคาเพิ่มขึ้นสูงหรือลดลงเล็กน้อย ก็จะมีการนำกลุ่มคำที่อยู่ใน Placing มาวิเคราะห์ร่วมกับแนวโน้มตามกลุ่มต่างๆ ด้วย โดยในแต่ละกลุ่มจะมีการระบุกลุ่มคำที่เกี่ยวข้องดังแสดงในรูปที่ 11

Behavior trend			
Increase	Decrease	Neutral	Placing
Rise/rose	fell	mixed	record
Increase	decrease	stable	larger-than expected
inject	drop	continue	hawkish/dovish
up	down, drawdown	gradual	examine
lift sanction	disappear	remain	force majeure
waive	sanction	unchanged	unplanned
export	lose,lost	maintain	exit strategy
expansive	cut	unchanged	balance
return	reduce, reduction		regain
open	buy		partial
resume	curb		rally
addition, add	high/higher		unscheduled
surge	shut, shut down		surprise
sell	plunged		weaken
hike	outage		scheduled
high/higher	deficit		opened fire
raise volume	dent		raise price (OSP)
gain	re-impose sanction		raise the flag
expand	slow		strong
surplus	fade		close to
restore	slide		decline to
long	short		under pressure
close short	close long		hike price/interest rate
plus			slow
			fade
			low/lower
			high/higher

รูปที่ 11 กลุ่มคำที่ระบุแนวโน้มการขึ้นลงของราคา (Behavior Trend)

สำหรับกลุ่มต่อมาคือ กลุ่มคำที่ระบุปัจจัยที่เกี่ยวข้องกับความรู้สึกผู้ซื้อขาย โดยกลุ่มดังกล่าวจะแบ่งปัจจัยออกเป็นทั้งหมด 9 กลุ่มย่อย ประกอบด้วยปัจจัยด้านการผลิต, ความต้องการผู้บริโภค, หน่วยงานควบคุมระดับน้ำมันดิบคงคลัง, ผลิตภัณฑ์มวลรวมภายในประเทศ (GDP), ดัชนีภาคอุตสาหกรรม, ดัชนีด้านฤดูกาล, กลุ่มประเทศ OPEC, กลุ่มประเทศ Non-OPEC ที่ผลิตน้ำมันดิบ, และดัชนีภูมิศาสตร์การเมือง (Geopolitics) โดยแต่ละกลุ่มจะมีการกำหนดขอบเขตของคำที่เกี่ยวข้องตามที่แสดงในรูปที่ 12 ซึ่งข้อมูลเหล่านี้ได้มาจากการสัมภาษณ์เชิงลึกกับนักวิเคราะห์การซื้อขายน้ำมันดิบ

Sentimental Factor								
Production	Geopolitic	Stock Balance	GDP	Industrial	Seasonal	Country (OPEC)	Country (non-OPEC)	Demand
Capacity	Sanction	EIA	Stock Market	PMI	Winter	Saudi Arabia	USA	Finish
Volume	Wars	API	Paper Market	Inflation	Summer	Iran	North Sea	Demand
Production	Hurricane	Custom	Asset Allocation	Consumer	Rainy	Iraq	UK, Britain	Gasoline demand
drill	Earthquake	PAJ	Volume trade	Producer	Heating	Nigeria	Norway	Stock
Export	Pipeline explosion	IEA	ICE	Price	Driving season	Kuwait	Russia	import
Produce	Protestor	ARA	WTI	Index	Traveling season	Libya	China	
Supply	Maintenance		Brent		Hydropower	Venezuela	Japan	
	Explosion		Gasoil				EU	
			NYMEX				OECD	
			RBOB				Mexico	
			Price trend				India	
			Gold				Australia	
			Fx rate					
			Labour					
			Labor					
			Jobless Claims					
			Unemployment Rate					

รูปที่ 12 ปัจจัยที่เกี่ยวข้องกับความรู้สึกผู้ซื้อขาย (Sentimental Factor)

สำหรับพจนานุกรมข้อมูลนั้น สามารถเปลี่ยนแปลงค่าที่บรรจุอยู่ในฐานข้อมูลได้ตามความเหมาะสม ขึ้นอยู่กับบริบทของการวิเคราะห์และประเภทของกลุ่มค่าที่พบเจอบ่อยครั้งจากแหล่งข่าว โดยสิ่งที่ทำให้การบรรจุค่าในพจนานุกรมมีความแม่นยำเพิ่มขึ้น จะเป็นในเรื่องของความถี่ของค่าและประเภทของเนื้อหาข่าวที่เกี่ยวข้องกับการขึ้นลงของราคาน้ำมันดิบ ทั้งนี้ในส่วนของตัวเองแบบจำลองที่ใช้ในการพยากรณ์แนวโน้มการขึ้นลงของราคาน้ำมันดิบนั้นไม่สามารถเปิดเผยได้เนื่องจากเป็นข้อมูลความลับระหว่างผู้วิจัยกับบริษัท แต่เบื้องต้นสามารถนำเสนอตัวอย่างของแนวทางการวิเคราะห์แนวโน้มการขึ้นลงของราคาน้ำมันดิบได้ โดยสมการที่ใช้ในการวิเคราะห์ แสดงดังสมการที่ (1)

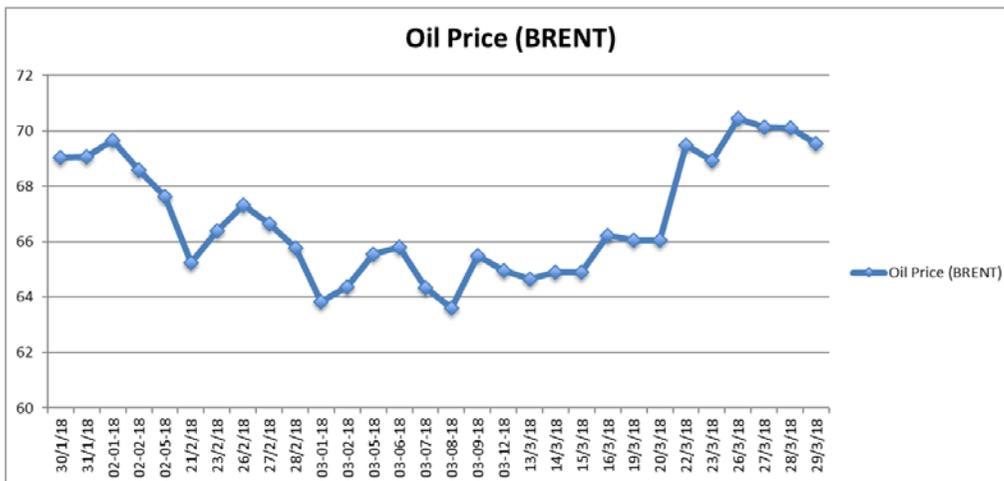
$$\text{Prediction Trend} = \text{Key of Behavior Trend} + \text{Sentimental Factor} \quad (1)$$

หมายถึง แนวโน้มการขึ้นลงของราคาน้ำมันดิบจะมาจากความสัมพันธ์ของกลุ่มค่าที่ระบุแนวโน้มการขึ้นลงของราคา (Behavior Trend) เชื่อมกับกลุ่มค่าที่ระบุปัจจัยที่เกี่ยวข้องกับความความรู้สึกผู้ซื้อขาย (Sentimental Factor) โดยการพยากรณ์จะมีการเชื่อมกลุ่มค่าระหว่าง 2 ตารางนี้ ตั้งแต่ 2 เงื่อนไขขึ้นไป เพื่อวิเคราะห์ว่าจากเงื่อนไขดังกล่าวส่งผลให้ราคาน้ำมันดิบมีแนวโน้มเพิ่มขึ้นหรือลดลง

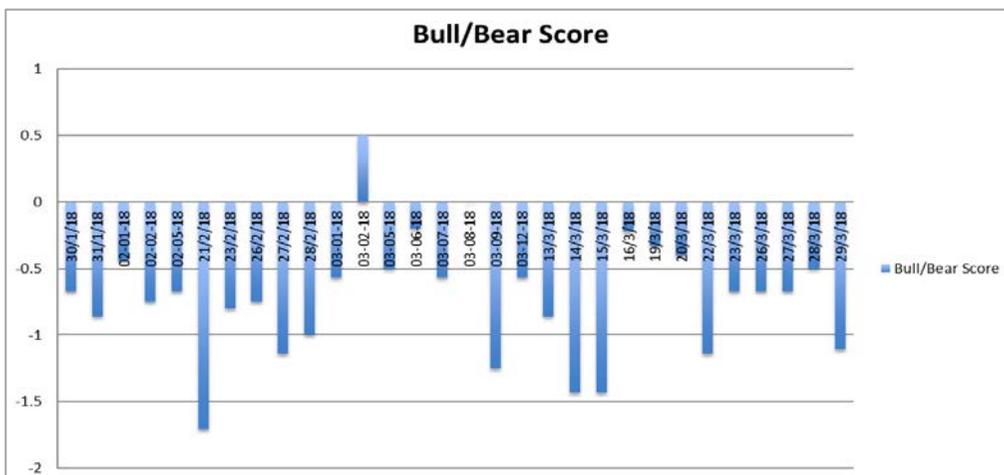
## 7. ผลการวิจัยและบทสรุป

จากที่ได้มีการพัฒนาระบบฐานข้อมูลสำหรับการจัดเก็บข่าว รวมถึงการนำเทคนิคการสกัดข้อมูล (Information Extraction) และพจนานุกรมข้อมูล (Data Dictionary) มาใช้ในการสกัดปัจจัยทางด้านความรู้สึกของผู้ซื้อขาย พบว่าผู้ใช้งานสามารถลดระยะเวลาในการวิเคราะห์ข่าวได้สั้นลงจากเดิมใช้เวลาประมาณ 3 นาทีต่อข่าว เหลือใช้เวลาเพียง 0.26 นาทีต่อข่าว (คิดจากการวิเคราะห์ข่าวจำนวน 30 ข่าวใช้เวลาเฉลี่ยประมาณ 8 นาที) หรือคิดเป็นระยะเวลาที่สั้นลงร้อยละ 90

นอกจากนี้เมื่อผู้ใช้งานได้ดำเนินการนำข่าวเข้าไปในระบบจำนวน 30 วันและทำการวิเคราะห์ผลลัพธ์ออกมา พบว่าโปรแกรมดำเนินการวิเคราะห์แนวโน้มของข่าวตรงกับแนวโน้มของราคาน้ำมันดิบจำนวน 17 วันหรือคิดเป็นความแม่นยำของการพยากรณ์ร้อยละ 60 โดยหลักในการพยากรณ์จะยึดตามขอบเขตของค่าที่พบในพจนานุกรมข้อมูลและรูปแบบค่าที่ได้จากเทคนิคการสกัดข้อมูล ซึ่งผลการเปรียบเทียบระหว่างการวิเคราะห์แนวโน้มของราคาน้ำมันดิบผ่านเนื้อหาข่าวในระบบที่ผู้วิจัยพัฒนาขึ้นกับราคาน้ำมันดิบในตลาดจริงแสดงผลตามรูปที่ 13 ถึงรูปที่ 14 โดยคำว่า Bull แทนการเพิ่มขึ้นของราคาน้ำมันดิบ (กำหนดค่า factor มากกว่า 0) และคำว่า Bear แทนการลดลงของราคาน้ำมันดิบ (กำหนดค่า factor น้อยกว่า 0)



รูปที่ 13 แนวโน้มของราคาน้ำมันดิบในตลาดโลก



รูปที่ 14 ผลการพยากรณ์แนวโน้มการขึ้นลงของราคาน้ำมันดิบที่ได้จากโปรแกรม

ทั้งนี้ การที่ผู้วิจัยได้ดำเนินการพัฒนาระบบฐานข้อมูลสำหรับรวบรวมข่าวและสกัดข้อมูลข่าว เพื่อวิเคราะห์แนวโน้มการขึ้นลงของราคาน้ำมันดิบผ่าน Web application นั้น ทำให้นักวิเคราะห์ของบริษัทใช้เวลาในการวิเคราะห์ข่าวได้เร็วขึ้น ตลอดจนนักวิเคราะห์สามารถเปรียบเทียบการวิเคราะห์ข่าวกับการขึ้นลงของราคาน้ำมันดิบได้ดีขึ้นผ่านการแสดงผลของกราฟในระบบ โดยที่ในอดีตที่ผ่านมายังไม่ได้มีการพัฒนาโปรแกรมลักษณะดังกล่าวออกมาอย่างแพร่หลายในเรื่องของการพยากรณ์แนวโน้มของราคาน้ำมันดิบในรูปแบบข้อความ Text File สำหรับการซื้อขายน้ำมันดิบของเมืองไทย อย่างไรก็ตามระบบดังกล่าวยังคงพบข้อจำกัดซึ่งจะนำเสนอในหัวข้อถัดไป

## 8. ข้อจำกัดและอุปสรรค

ทั้งนี้ในระหว่างที่มีการเก็บข้อมูลและดำเนินการวิจัยนั้น ทางผู้วิจัยพบปัญหาและอุปสรรคที่เกี่ยวข้องกับการสกัดข้อมูลและการกำหนดขอบเขตของคำในพจนานุกรมข้อมูลโดยมีรายละเอียดดังนี้

- 1) ในส่วนของการกำหนดขอบเขตของคำในพจนานุกรมข้อมูล บางกลุ่มคำยังไม่ถูกบรรจุลงในพจนานุกรมแต่พบกลุ่มคำดังกล่าวในเนื้อหาข่าว ส่งผลให้การสกัดข้อมูลเพื่อไปพยากรณ์แนวโน้มราคาน้ำมันดิบเกิดความคลาดเคลื่อน
- 2) การเพิ่มปริมาณคำศัพท์หรือกลุ่มคำเข้าไปในพจนานุกรมข้อมูลยังเป็นลักษณะการเพิ่มแบบ Manual หมายถึง ผู้พัฒนาระบบต้องเข้ามาปรับปรุงกลุ่มคำในพจนานุกรมข้อมูลด้วยตนเอง
- 3) สำหรับการพยากรณ์แนวโน้มสถานการณ์ราคาน้ำมันดิบเป็นลักษณะกึ่งอัตโนมัติ (Semi-Auto Prediction) กล่าวคือ มีการใช้นักวิเคราะห์ช่วยในการพยากรณ์ร่วมกับการสกัดข้อมูลและการเปรียบเทียบขอบเขตของคำตามที่ระบุในพจนานุกรมข้อมูลในระบบ ซึ่งทางนักวิเคราะห์ยังมีความจำเป็นต้องตรวจสอบการวิเคราะห์อีกครั้ง
- 4) ข้อมูลที่ได้ยังไม่สามารถนำไปใช้วิเคราะห์แนวโน้มได้อย่างเต็มประสิทธิภาพ

## 9. ข้อเสนอแนะ

จากที่ได้มีการนำเสนอข้อจำกัดและอุปสรรคที่พบในงานวิจัยนั้น ทางผู้วิจัยมีข้อเสนอแนะเพิ่มเติม โดยมีรายละเอียดดังนี้

- 1) กรณีแรกและกรณีที่สอง ควรมีการนำแนวคิดการเรียนรู้ของเครื่องจักร (Machine Learning) ในรูปแบบอื่นๆ มาประยุกต์ใช้ในการเก็บข้อมูลกลุ่มคำที่พบบ่อยครั้ง ทั้งกลุ่มคำที่พบใหม่และกลุ่มคำที่ค่าความถี่มีอิทธิพลต่อทิศทางการขึ้นลงของราคาน้ำมันดิบ เพื่อที่ระบบจะสามารถปรับปรุงจำนวนคำที่พบได้อย่างเหมาะสมมากขึ้น

2) กรณีที่สามและสี่ ในส่วนของการพยากรณ์แนวโน้ม ควรมีการต่อยอดเรื่องการพัฒนาแบบจำลองสำหรับพยากรณ์ราคาน้ำมันดิบแบบอัตโนมัติ (Automated Prediction Model) โดยนำหลักการของปัญญาประดิษฐ์ (Artificial Intelligence) อาทิ หลักการโครงข่ายประสาทเทียม (Artificial Neural Network) มาช่วยกระบวนการวิเคราะห์แนวโน้มแบบอัตโนมัติตามคำนำหน้าของข่าวที่นำเข้าสู่ระบบ เพื่อลดเวลาในการทำงานของนักวิเคราะห์ ตลอดจนปรับปรุงความแม่นยำในการพยากรณ์ให้เพิ่มสูงขึ้นจากเดิม

นอกจากนี้ทางผู้วิจัยมีข้อเสนอแนะเพิ่มเติมเกี่ยวกับการนำผลลัพธ์ที่ได้จากการวิเคราะห์ข้อมูลเชิงคุณภาพมาพยากรณ์แบบบูรณาการร่วมกับข้อมูลเชิงปริมาณ เช่น พยากรณ์การเปลี่ยนแปลงของราคาน้ำมันดิบในตลาดเพื่อให้ได้ผลลัพธ์ที่แม่นยำขึ้นในลำดับถัดไป

### กิตติกรรมประกาศ

สำหรับงานวิจัยฉบับนี้ได้รับการสนับสนุนทุนจากกระทรวงวิทยาศาสตร์และเทคโนโลยี ตลอดจนได้รับการสนับสนุนข้อมูลจากทางบริษัท ปตท. จำกัด (มหาชน) ทั้งนี้ผู้วิจัยขอขอบพระคุณสำหรับรายละเอียดข้อมูลต่างๆ ที่เป็นประโยชน์ต่อการวิเคราะห์และการพัฒนาต้นแบบสำหรับการพยากรณ์แนวโน้มการขึ้นลงของราคาน้ำมัน ตลอดจนข้อเสนอแนะที่เป็นประโยชน์ต่อการพัฒนาต้นแบบเพื่อการพยากรณ์ที่ดีขึ้นในอนาคต

### References

- [1] EIA.gov [Internet]. Washington: U.S. Energy Information Administration. Crude oil prices increased in 2016, still below 2015 averages; c2015-2016 [updated 2017 Jan 4; cited 2018 Jan 1]. Available from: <https://www.eia.gov/todayinenergy/detail.php?id=29412>
- [2] Abdullah SN. Machine Learning Approach For Crude Oil Price Prediction [Doctoral dissertation]. Manchester: The University of Manchester; 2013.
- [3] Gupta V, Lehal G. A survey of text mining techniques and applications. Journal of emerging technologies in web intelligence 2009;1(1):60-76.
- [4] Mittermayer MA. Forecasting intraday stock price trends with text mining techniques. In: system sciences, 2004. Proceedings of the 37th Hawaii international conference on system sciences; 2004 January 5-8; Hawaii, USA. Hawaii: USA; 2004. p. 10-9.
- [5] Kloptchenko A, Eklund T, Karlsson J, Back B, Vanharanta H, Visa A. Combining data and text mining techniques for analysing financial reports. Intelligent systems in accounting, finance and management 2004;12(1):29-41.

- [6] Larsen B, Aone C. Fast and effective text mining using linear-time document clustering. Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining; 1999 August; Virginia, USA. California: USA; 1999. p. 16-22.
- [7] Aggarwal CC, Zhai C (Eds.). Mining text data. Springer Science & Business Media; 2012.
- [8] Clifton C, Thuraisingham B. Emerging standards for data mining. Computer Standards & Interfaces 2001;23(3):187-93.
- [9] Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Annals of surgical oncology 2010;17(6):1471-4.
- [10] Kilian Lutz. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. American Economic Review 2009; 99(3): 1053-69.
- [11] Wang Shouyang, Lean Yu, Kin Keung Lai. A novel hybrid AI system framework for crude oil price forecasting. Data Mining and Knowledge Management. Springer, Berlin, Heidelberg; 2005. p. 233-42.

### ประวัติผู้เขียนบทความ



**ดร.พีรพล สิทธิวิจารณ์** อาจารย์สาขาการจัดการโลจิสติกส์ คณะโลจิสติกส์ มหาวิทยาลัยบูรพา 169 ถ.ลงหาดบางแสน ต.แสนสุข อ.เมือง จ.ชลบุรี 20131 โทรศัพท์ 087-7907920 อีเมล peerapoljt@gmail.com  
Interested Field: Operation Research and Warehouse Management



**อนิรุทธ์ ชันธสะอาด** อาจารย์สาขาการค้าระหว่างประเทศ คณะโลจิสติกส์ มหาวิทยาลัยบูรพา 169 ถ.ลงหาดบางแสน ต.แสนสุข อ.เมือง จ.ชลบุรี 20131 โทรศัพท์ 086-8587863 อีเมล anirutka@gmail.com  
Interested Field: Operation Research and Supply Chain Management



ยุทธนา สมสอน กัปตัน บริษัท ซีรอน จำกัด 292/12 หมู่บ้านวิสตาปาร์ค  
ถ.วิภาวดีรังสิต แขวงตลาดบางเขน เขตหลักสี่ กทม. 10210 อีเมลล์  
yuthana@cheiron.co.th

Interested Field: Crude Oil Maritime, Logistics Management



ณัฐชัย ตั้งวิจนะโยบาย นักวิเคราะห์และวางแผน บริษัท ปตท. จำกัด  
(มหาชน) 555 ถ.วิภาวดีรังสิต แขวงจตุจักร เขตจตุจักร กทม. 10900 อีเมลล์  
nuttachai\_t@pttplc.com

Interested Field: Crude Oil Market, Import-Export in Petroleum business

---

**Article History:**

Received: August 7, 2018

Revised: February 28, 2019

Accepted: March 4, 2019