

**GOLD PRICE VOLATILITY PREDICTION BY TEXT MINING IN  
ECONOMIC INDICATORS NEWS**



**CHANWIT ONSUMRAN**

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR  
THE DEGREE OF MASTER OF SCIENCE  
(INFORMATION TECHNOLOGY MANAGEMENT)  
FACULTY OF GRADUATE STUDIES  
MAHIDOL UNIVERSITY**

**2015**

Copyright by Mahidol University

**COPYRIGHT OF MAHIDOL UNIVERSITY**

**GOLD PRICE VOLATILITY PREDICTION BY TEXT MINING IN ECONOMIC INDICATORS NEWS**

**CHANWIT ONSUMRAN 5636623 EGIT/M**

**M.Sc. (INFORMATION TECHNOLOGY MANAGEMENT)**

**THESIS ADVISORY COMMITTEE: SOTARAT THAMMABOOSADEE, Ph.D., SUPAPORN KIATTISIN, Ph.D., ADISORN LEELASANTITHAM, Ph.D.**

**ABSTRACT**

This research focuses on the text mining approach of the gold prices volatility prediction model from the text of economic indicators news articles. The model is designed and developed to analyze how the news articles influence gold price volatility. The selected reliable source of news articles is provided by FXStreet, which offers several economic indicators. The data will be used to build text classifiers and news group affecting volatility price of gold. According to the fundamentals of the data mining process, each news article is firstly transformed in to a feature by the TF-IDF method. Then, a comparative experiment is set up to measure the accuracy of the combination of two attributes weighting approaches - which are Support Vector Machine (SVM) and Chi-Squared Statistic - and three classification algorithms - which are the k-Nearest Neighbors, SVM and Naive Bayes. The results show that the SVM classification algorithm, weighted by SVM, is the best among all tests with an accuracy of 87.52%. In the future, it can be developed to improve the classification system; it may be input to other economic news for more factors to increase the efficiency of analyzing consequences that affect the volatility of the gold price.

**KEY WORDS: TEXT MINING/ ECONOMIC INDICATORS NEWS/ GOLD PRICE/ VOLATILITY PREDICTION**

74 pages

การทำนายความผันผวนของราคาทองคำ โดยวิธีเหมืองข้อความในบทความข่าวชี้วัดทางเศรษฐกิจ  
GOLD PRICE VOLATILITY PREDICTION BY TEXT MINING IN ECONOMIC  
INDICATORS NEWS

ชาญวิทย์ อ้นสำราญ 5636623 EGIT/M

วท.ม. (การจัดการเทคโนโลยีสารสนเทศ)

คณะกรรมการที่ปรึกษาวิทยานิพนธ์ : โยทศรัตต ธรรมบุษดี, Ph.D., สุภาภรณ์ เกียรติสิน, Ph.D.,  
อดิศร ลีลาสันติธรรม, Ph.D.

บทคัดย่อ

งานวิจัยนี้มุ่งเน้นไปที่วิธีการทำเหมืองข้อความจากบทความข่าวชี้วัดทางเศรษฐกิจ เพื่อทำนายความผันผวนของราคาทองคำ โดยทำนายความผันผวนจากรูปแบบของข่าวซึ่งจะเป็นข้อความข่าวชี้วัดทางเศรษฐกิจ โดยจะออกแบบและพัฒนาขึ้นเพื่อวิเคราะห์บทความข่าวที่มีอิทธิพลต่อความผันผวนของราคาทองคำ โดยข้อมูลบทความข่าวที่นำมาใช้ในงานวิจัยนี้คือข้อมูลจากเว็บไซต์ FXStreet ซึ่งเป็นแหล่งข้อมูลที่เป็นสากลและมีความน่าเชื่อถือ โดยได้เลือกบทความข่าวทางเศรษฐกิจที่มีความเกี่ยวข้องกับงานวิจัยในครั้งนี้มาวิเคราะห์ ซึ่งข้อมูลจะถูกนำมาใช้ในการสร้างรูปแบบสำหรับการจัดหมวดหมู่ของข้อความ และกลุ่มข่าวที่มีผลต่อความผันผวนของราคาทองคำตามพื้นฐานของกระบวนการทำเหมืองข้อความ โดยใช้การให้ค่าน้ำหนักของคำในเอกสารด้วยวิธี TF/IDF จากนั้นทดลองเปรียบเทียบประเมินผลความถูกต้องด้วยการเลือกใช้ Attributes Weighting ซึ่งประกอบด้วยการของ Support Vector Machine และวิธี Chi-Squared Statistic โดยใช้เทคนิค Classification Algorithms ทั้งหมด 3 วิธีประกอบด้วย k-Nearest Neighbors, SVM และ Naive Bayes ซึ่งผลการศึกษาพบว่าวิธี SVM โดยการเลือกใช้ Attributes Weighting Support Vector Machine มีความถูกต้องมากที่สุดร้อยละ 87.52 ซึ่งการวิจัยในอนาคตสามารถที่จะพัฒนาและปรับปรุงรูปแบบการจำแนกข้อความโดยอาจจะเป็นการเพิ่มปัจจัยข้อมูลข่าวเศรษฐกิจด้านอื่น ๆ เพิ่มเติมเพื่อเพิ่มประสิทธิภาพในการวิเคราะห์ผลกระทบที่ส่งผลต่อความผันผวนของราคาทองคำต่อไป