INTERNATIONAL GOLD PRICE ANALYTICS

Adithya Srinivasan* and Nitin Agarwal

*Corresponding Author: Adithya Srinivasan et.adithya@gmail.com

This research is to predict the international gold price by analyzing consumer demand obtained from social media as well as financial newsfeed data. We started our analysis by acquiring the data that would in turn manifest into the international demand. We further cleansed the data and filtered out only the records that were relevant to our analysis. We performed text mining using a plethora of statistical tools and techniques to derive certain key sentimental insights about the data set and further proposed a regression model to ultimately predict the international gold price.

Keywords: Gold Price, Predictive Analytics, Regression Analysis, Bullion Investing, Gold Options, Gold Futures

HONESTY DECLARATION

We hereby declare that the work submitted is our own and that all passages and ideas that are not ours have been fully and properly acknowledged. Furthermore, we declare that this work has not been in parts or wholly published as a submission for another examination procedure and that all copies, both printed and electronic, are the same.

INTRODUCTION

Gold is one of the most popular investment of all the precious metals. Gold is purchased by investors to diversify their investment portfolio through the use of futures and other derivatives. The gold market is highly determined by speculation and is as volatile as any other commodity market such as oil and natural gas. The price of gold is simply driven by the international supply and demand along with speculation. Gold is sometimes used as a hedge against inflation, deflation and currency devaluation. Since the amount of new gold mined every year is insignificant when compared with the total amount of gold that exists above ground, the major demand and supply is considered with respect to its trading. Therefore, the price of gold is a direct result of its demand which is mainly affected by changes in global consumer and investor’s sentiment. This is the underlying principal of our analysis and research. Gold has also become more readily accessible due to the development of a range of products, which investors include in their own portfolios. The diversity of gold-backed and gold-related products means that gold can be used to enhance a wider variety of individual investment strategies and risk.
tolerances. Gold prices are constant and do not vary with geographies. The small difference observed in the price in different countries is only because of country specific taxes and import policies; the gold market is regulated globally.

Social media is often an important influencer on public opinion and political causes. People and investors alike on social media and national investment reports on financial news feeds give a brief and effective overview of the general trend and demand. Social media in general and twitter in particular definitely influences public opinion and decision makers.

Social media is a platform where people are able to voice their opinions and update facts liberally. In business, social media is used to market products, promote brands, and connect to current customers and foster new businesses. News reports are published throughout the day and form a key part in influencing public opinion.

The international demand therefore can be extracted by analyzing social media and financial news feed data which enables us to model the trend in prices. By finding a relationship between the international demand and international gold price, we aim to predict the future price by analysing the recent demand.

**METHODOLOGY**

To move further for our analysis, we required social media data and twitter being our primary sources. We have majorly used the Sentiment analysis technique to extract the sentiments of the tweets and News headlines, which we convert to polarities.

**TWITTER DATA**

After thorough research and analysis of user updates and tweets, we zeroed in on certain keywords which fetched most relevant data. Below are a few keywords:

- #BRIKK
- #GoldGenie
- #Miansai
- #GoldSouk
- #GoldPrice

Using above hashtags, we pulled the data for first four days of each months. This data contained many fields such as Handle/Link, Full Name, Time, Tweet/Text and @Reply 1-5 out of which only two were relevant to us that are Time, Tweet/Text. This data has so many irrelevant words or characters so required cleaning.

**FINANCIAL NEWSFEED**

For financial newsfeed data, we geographically segmented the globe to effectively understand the international demand. After identifying geographic clusters, we arrived at a single publication for each
sector, which produced significantly larger number of articles relevant for our prediction.

- *The Gulf Times* for the Middle East.
- *The Financial Times* for major parts of Europe.

The data obtained again contained many fields such as Position, URL, Title, Description, Page Authority, Total Links to Page, Total Linking Root, Domains to Page, Domain Authority and Total Links to Root but the field “Title” is relevant for our analysis.

To cleanse the data, we used the following R commands

```r
# remove retweet entities
data = gsub("(RT|via)((?:\b\W*@\w+)+)", "", data)
# remove at people
data = gsub("@\w+", "", data)
# remove punctuation
data = gsub("\[[:punct:]\]", "", data)
# remove numbers
data = gsub("\[[:digit:]\]", "", data)
# remove html links
data = gsub("http\w+", "", data)
# remove unnecessary spaces
data = gsub("\s+","",gsub("\s\s+","",gsub("[^a-zA-Z0-9]","",data)))
```

where “data” in above commands is the tweets or the news headlines.

**SENTIMENT EXTRACTION**

We consolidated the social media and financial news feed data and classified them based on their month of occurrence. Further, extracting the overall general sentiment of each using the following R functions:

```r
library(sentiment)
class_emo1 = classify_emotion(Data_Jan, algorithm="bayes", prior=1.0)
class_emo2 = classify_emotion(Data_Feb, algorithm="bayes", prior=1.0)
class_emo3 = classify_emotion(Data_Mar, algorithm="bayes", prior=1.0)
class_emo4 = classify_emotion(Data_Apr, algorithm="bayes", prior=1.0)
```

Using above techniques, we classified text into sentiments like anger, disgust, fear, joy, sadness, surprise and NA and further converted into negative, positive and neutral using below rules:

- Anger + Disgust + Fear + Sadness → Negative
- Surprise + Joy → Positive
- NA → Neutral

After analyzing the 4 days of data, we formulated a relationship between these sentiments and their corresponding prices. Then we performed regression analysis and obtained the coefficients and the regression equation:

```
Price = Negative * (-0.72087) + Positive * (0.25846) – Neutral * (0.05603) + 2417.57754
```

This enabled us to predict the international gold price on the 5th day after analyzing the previous 4 days of data. The accuracy of our predicted price when compared with the actual international market price was over 99%.
ACKNOWLEDGMENT

This work is fully supported by The International Institute of Information Technology, Bangalore. We would like to appreciate the immense help and confidence of our mentor Kalpana Subbaramappa without whose unflinching guidance, this project would not have been possible.

BUSINESS USE CASES

Investing in the bullion market: Since a century now, the most common benchmark for the international gold price has been the London gold fixing, a twice-daily telephone meeting of representatives from five bullion-trading firms of the London bullion market. Gold is traded continuously throughout the world based on the intra-day spot price, derived from over-the-counter gold-trading markets around the world (code “XAU”). The most traditional way of investing in gold is by buying bullion gold bars. Gold coins are a common way of owning gold too. Bullion coins are priced according to their fine weight. By prior knowledge of the future gold price, more informed and calculated investments can be made so as to ensure larger profits and returns.

Investing in the derivative market: Derivatives, such as gold forwards, futures and options, currently trade on various exchanges around the world. In the US, gold futures are primarily traded on the New York Commodities Exchange (COMEX). In India, gold futures are traded on the National Commodity and Derivatives Exchange (NCDEX) and Multi Commodity Exchange (MCX). The price of gold bullion is volatile, but unhedged gold shares and funds are regarded as even higher risk and even more volatile. The performance of gold bullion is often compared to stocks due to their fundamental differences. However, stock prices are more difficult to forecast since investors do not publicly announce their purchases thereby closing opportunities to track demand from social media. Gold on the other hand is a luxury commodity and people are more vocal about their acquisitions. Investing in gold futures and options in the derivative market becomes a low risk investment with prior knowledge of the future pricing trend.

REFERENCES


This article can be downloaded from http://www.ijmrbs.com/currentissue.php