THE CANARY IN THE ROOM: THE PREDICTIVE VALUE OF SOCIAL MEDIA FOR HOSPITALITY AND TOURISM STOCKS

Abstract Recently social media has been shown to provide information on tourism and politics. In recent studies that we have conducted, wine sales and tourism trends have been forecast using machine learning algorithms and hundreds of thousands of blogs, online news articles and Twitter tweets. Our previous research has shown that mining of social media can predict in real-time market penetration figures and can also forecast social and business trends in tourism. We extend the research to evaluate the predictive value of Social Media in forecasting trends for the value of publicly traded Hospitality and Tourism companies.

Keywords Social media, Hospitality stocks, Micro Blogs, Text Mining, Business Intelligence, Sentiment Mining, Business Analytics, stock valuation.

Research background In this study we investigate whether public mood as measured from large-scale collection of social media is correlated or even predictive of publicly traded Hospitality stock values. Our research to date, as well as those mentioned above suggest that models which include text mining and sentiment mining of internet based data repositories and specifically blogs, online dailies, Twitter, and Facebook can provide predictive power for financial markets. Specifically we monitor the stock measures of publicly traded Hospitality midsized and large firms. To date there has not been published work of this kind on this subgroup. This research additionally augments methodologies already established and will provide tools and models which will be transferable to other quantitative research.

As found in a number of papers, sentiment mining aims at extracting features on which users express their opinions in order to determine the user's sentiment towards the query object. There has been a flourish of articles in the academic literature on market correlations to
various sentiment measures of social media starting in about 2009. Our efforts at collection of data began in that year as well and we have published several papers in this field. However no work has been published on Hospitality and Tourism companies and the possible links that their market value may have to Social Media. Our study of the same type looked at social media sentiment about tourism in Thailand amid the unrest in that country during the early part of 2010 and further whether analysis of micro-blogs can be used to discern the effect of that unrest on Phuket's tourism environment.

There is a treasure trove of hidden information in the textual and narrative data of various social media that can be deciphered by text-mining techniques. The information provided by these methods can provide a basis for artificial intelligence and help support or improve any kind of analysis but in particular the verification of market valuations. To date we have devised a buy, hold, and sell strategy with accuracy of approximately 70 percent for seven large stocks that sells on the NYSE. In this paper we develop this methodology for medium and midsized Hospitality and Tourism companies on the NYSE by the collection of data from blogs, and other forms of social media.

**Methodology** We selected a portfolio of Hospitality and Tourism stocks whose companies hold a significant presence in English based social media and a global footprint. Over 800 million micro-blogs and 25,000 reviews were collected and analyzed to gain knowledge regarding Hospitality and Tourist sentiment on Hotels in various destinations such as New York, London and Tokyo. Sentiment was measured using a binary choice keyword algorithm and a multi-knowledge based approach was proposed using, Self-Organizing Maps and tourism domain knowledge in order to model sentiment. A visual model was developed to express this taxonomy of sentiment vocabulary and then this model was applied to maximums and minimums in the time sentiment data. The results showed actionable knowledge could be extracted.
We obtained a collection of public media (Facebook, blogs, news, Twitter) for a period of 3 years. Software was written to account for location and language and a means was designed for implementation of machine translation of text from non-English sources. Domain specific models of sentiment are developed in order to achieve robust models. The sentiment was compared with various balance sheet measures such as sales, ROI, and EBITDA. We further looked for correlations between sentiment and various financial indicators such as closing stock prices on public markets.

Testing of the models included periods of substantial volatility and periods with little volatility. Data sets of varying lengths were also tested (short -3 months, medium -8 months, and long -15 months). For each stock, a minimum of 10 data sets was tested against a minimum of 12 artificial intelligence algorithms including decision trees, Bayesian methods, Neural Networks, and Regression models. Aspects of traditional times-series modeling are also employed.

**Results and future works** The results indicate that significant correlations exist between sentiment and various business measures. The most significant of these has been between sentiment and share price. Given these correlations our next steps will involve constructing forecasting models.

Further innovations will include the incorporation of geo-location data in order to improve the precision of the predictive analytics. We have already developed software which can match a tweeter’s location based on characteristics of the tweet. With the growing prevalence of Android and IPhone smart phone technology GPS data is becoming increasingly available. Further innovation will be introduced through the inclusion of language indicators. The innovative inclusion of both of these factors may help avoid geographical and cultural sampling errors.
William Claster, PhD
Associate Professor of Information Technology Communication
Ritsumeikan Asia Pacific University, Japan
Email: wclaster@apu.ac.jp

Phillip Pardo
Professor of Accounting
Ritsumeikan Asia Pacific University, Japan
Email: pardorit@apu.ac.jp

Malcolm Cooper PhD LLM
Professor of Tourism Management and Environmental Law
Ritsumeikan Asia Pacific University, Japan
Email: cooperm@apu.ac.jp