A Project Report on

INVERSE EFFECT MODELING OF EARNING CALL TRANSCRIPTS FOR FINANCIAL ANALYSIS

Submitted in partial fulfilment for award of degree of

Masters of Business Administration (MBA)

In **Business Analytics**

Submitted by

NAGENDRA B.V

R17DM010

Under the Guidance of

Dr. J.B SIMHA

Chief Mentor-REVA

REVA Academy for Corporate Excellence

REVA University

Rukmini Knowledge Park, Kattigenahalli,

Yelahanka, Bangalore – 560064

October, 2020

REVA UNIVERSITY
Bengaluru, India

Candidate's Declaration

I, Nagendra B.V hereby declare that I have completed the project work towards the Master of Business Administration in Business Analytics at REVA University on the topic entitled "INVERSE EFFECT MODELING OF EARNING CALL TRANSCRIPTS FOR FINANCIAL ANALYSIS" under the supervision of Dr. J.B Simha, Chief Mentor, REVA Academy for Corporate Excellence (RACE), REVA UNIVERSITY. This report embodies the original work done by me in partial fulfilment of the requirements for the award of degree for the academic year 2020.

Place: Bengaluru   Name of the Student:

Date: 5th October 2020   Signature of Student

NAGENDRA B.V

REVA UNIVERSITY
Bengaluru, India

Certificate

This is to Certify that the PROJECT work entitled "INVERSE EFFECT MODELING OF EARNING CALL TRANSCRIPTS FOR FINANCIAL ANALYSIS" carried out by NAGENDRA B.V with SRN R17DM010, is a bonafide student of REVA University, is submitting the project report in fulfilment for the award of Masters of Business Administration (MBA) in Business Analytics during the academic year 2020. The Project report has been tested for plagiarism, and has passed the plagiarism test with the similarity score less than 15%. The project report has been approved as it satisfies the academic requirements in respect of PROJECT work prescribed for the said Degree.

<Signature of the Guide> <Signature of the Director>

Dr. J.B Simha <Name of the Director>

Guide Director

External Viva

Names of the Examiners

1. Ravi Shukla, Senior Advisor & Data Scientist, Virtual mode
2. Krishna Kumar Tiwari, Senior Data Scientist, CoE, AI/ML, Virtual mode

Place: Bengaluru

Date: 5th October 2020

Acknowledgement

I am thankful to many people who helped me accomplish this capstone project successfully.

First, I would like to thank the honourable Vice Chancellor, Dr Shyam Raju (REVA University) for giving me the opportunity to do my project.

I thank Dr. S.Y Kulkarni, Ex-Vice Chancellor, Dr. K. Mallikarjuna Babu, Vice-Chancellor and Dr. M. Dhanamjaya, Dr. Shinu Abhi, Director Corporate Training and Management Team, REVA University for their generous support.

I thank my academic and corporate mentor Dr. Jay Bharatheesh Simha for his magnanimous support and guidance during the course of my project. I remember him with much gratitude for his patience and motivation, but for which I could not have submitted this work.

I thank my parents for their blessings and constant support from my family members, without which this project would not have seen the light of day.

NAGENDRA B.V
(R17DM010)

Place: Bengaluru
Date: 5th October 2020

Similarity Index Report

Title of the Thesis: INVERSE EFFECT MODELING OF EARNING CALL TRANSCRIPTS FOR FINANCIAL ANALYSIS

Total No. of Pages:

Name of the Student: NAGENDRA B.V

Name of the Guide(s): Dr. J.B Simha

This is to certify that the above thesis was scanned for similarity detection. Process and outcome is given below.

Software Used: Turnitin
Date of Report Generation: 12th September 2020
Similarity Index in %: 12
Total word count: 6576

Place: Bengaluru

Name of the Student:
NAGENDRA B.V
Signature of Student

Date: 5th October 2020

Verified By:

Signature

Dr. Shinu Abhi, Director, Corporate Training

List of Abbreviations

Sl. No	Abbreviation	Long Form
1	ECT	Earning Call Transcripts
2	AGM	Annual General Meeting
3	CRISP DM	Cross Industry Standard Process for Data Mining
4	EDA	Exploratory Data Analysis
5	EPS	Earnings Per Share
6	AFFIN	Finn Arup Neilson

List of Figures

List of Tables

Abstract

This research paper intends to study the relationship between the quarterly board room discussions conducted in the form of “Earnings Call Transcripts” for the current quarter and its influence on the company’s stock performance in subsequent periods. In this study, the extraction of sentiments was undertaken from the “textual quarterly transcripts” of three major software companies for a 10 year period. The extracted sentiments were statistically analyzed and validated for patterns and sentiments. Apart from this, new features were extracted from the existing data which comprised of a response variable called the ‘Inverse Effect’, which has been modelled using various machine learning algorithms. The ‘Inverse Effect’ refers to the discordance between the present positive or negative sentiments in the boardroom discussions and performance of the stock in the stock market. For example, if the boardroom sentiments for the current quarter are positive and the stock performance in the market is also positive in the same quarter, it is considered as “concordance” and if the performance of the individual stock is opposite to the sentiments shared in the boardroom it will be called as “discordance”.

The findings emerged from the study suggest a possible causality between the sentiments and the stock market performance. It was also found that polarity, three-quarter average stock price and the previous quarter stock price are the determinants of the ‘Inverse Effect’. Based on the findings from the study, appropriate machine learning models were developed and evaluated to predict the ‘Inverse Effect’ on the performance of individual stocks in the stock market. The study also looks at to determine whether or not the inverse effect is the lead indicator of the market performance.

Key Words: Earnings Call Transcripts, Sentiments, Polarity, Earnings per Share, Market Performance, Stock Price, ‘Inverse Effect’

Contents

Chapter 1: Introduction

Market intelligence is the process of effective use of endogenous and extraneous information for making better business decisions. Organizations use market intelligence data extensively for decision making. Market intelligence enables organizations to foresee risk and accordingly initiate corrective and preventive actions. Market intelligence in the context of this study is to know in advance the likely performance of the market based on certain inputs which are generally available in the form of quarterly fundamental data and unstructured corpus like Earnings Call Transcripts (ECT) and Annual Directors Reports.

Every organization is required to measure data related to its performance. The performance can be related to employees, customers, business transactions or organization as a whole. Based on the business requirements necessary reports can be generated. In particular, Davis and Tama-Sweet found a positive relationship between an increase in positive tone within earnings press releases and the short-term stock price response to firms' earnings announcement. They also studied whether managers adjust the tone of earnings press releases to increase the value of their stock options and found that the optimal level increased before their option exercises. The results of their studies show that optimism in their tone had a positive impact on short-term stock prices and negative effect if their tone was pessimistic (Davis & Tama-Sweet, 2012). Another study by Demers and Vega found a positive association between firms' future returns and the level of tone changes in earnings press releases (Demers & Vega, 2008).

Organizations need to measure their performance for continuous improvement; else they will be at the mercy of market forces which include news media and other competitors. With the advent of technology and robust workflow systems organizations today collect voluminous data both structured and unstructured through various data sources. The process of extracting actionable insights from such voluminous data is often termed as big data analytics in the world of data science. The organization must be able to extract actionable insights from these huge data collected which in turn can be part of its metrics and measurement system. Every actionable insight must be mapped with the organization's profitability to know the effectiveness of the metric being measured. In general, big data analytics enable organizations to improve their overall profitability by extracting actionable insights (Yu & Guo, 2016).

Public invest in stock market, expecting a reasonable return on the money invested. It is important for leadership teams in organizations to protect the interest of their investors/shareholders. Many a time the public takes actions based on the announcements or statements made by the leaders concerning earnings, expenditure and their likely plans during quarterly, half-yearly and yearly held Annual General Meetings (AGMs). AGMs thus play an important role on the stock market performance. It is to be noted here that, more than the company's actual performance; stock markets are susceptible to sentiments and rely heavily on the behavioural aspects of the leadership team. This is because investors are vigilant and keep a close watch on any kind of signs from the leaders to decide their future moves (Demers & Vega, 2008).

The stock market being highly volatile, the leadership team in organizations observe market performance closely and base their strategies and tactics. The organizations need to connect strategies and performance with the actual market performances for better decision making (Li et al., 2020). Through this paper, we demonstrate whether or not what is being spoken by the leadership team in organizations manifest in-market performance and how consistent is the leadership team in terms of their behaviour during public meetings concerning the overall performance. The study also aims to explore correlations between various internal and external financial performance indicators.

Chapter 2: Literature Review

Sentiments extraction is key to text mining. To extract hidden information from the mammoth corpus, text mining algorithms are largely employed. Text mining algorithms can work on any type of input ranging from a sentence to a document or a group of documents. The text mining is a process in itself with a set of actions to be performed. It starts with a document as an input. Once the target document is inputted, some amount of pre-processing of the input takes place. The pre-processing of text is somewhat similar to fundamental data analysis wherein the numeric features are treated for outliers, missing values and more importantly the feature engineering and feature selection which is paramount to machine learning models. In the case of text, each word constitutes a feature. Similar to fundamental data, the text also needs to be treated, this exactly what is accomplished through text pre-processing. Stop words removal is an important step which facilitates removal of certain words like ”the”,” is”, “are”,” then”, “who” to name a few, which do not add any value to information extraction. Once the pre-processing is done, many mining algorithms are employed to extract information from the processed data. There are many lexicon-based techniques to extract sentiments from the text, the most commonly used are “AFINN” and “VADER (Valence Aware Dictionary and sEntiment Reasoner). Once the sentiments are extracted, they are analyzed and the insights are generated. Businesses collect a large amount of data and social media like Twitter, Facebook have customer reviews, product reviews, reviews on stock market performance reviews. There has been a surge in text data in social media and extracting insights from them is paramount to organizations in decision making. Market intelligence, visualization, document searching and analysis survey responses require text mining algorithms (Nam & Seong, 2019).

Marketing and finance domains rely heavily on social media for information extraction from text data. Various studies show the impact of sentiments analysis on various performance indicators. Apart from extracting the sentiments, text mining algorithms have wide applications in document classification, document clustering, text summarization, social network analysis and topic modelling. Even though significant progress has been made in the field of text mining, feature selection remains challenging and critical to text mining (Kumar & Ravi, 2016).

Internal controls, corporate governance and risk management have off late received a great deal of attention in organizations. A major area of focus is the robustness of internal controls which are periodically reviewed with internal audits. Adequacy of internal controls is analyzed using financial reports available annually with text mining algorithms which aim at keywords and feature selection. The results thus obtained can be useful to internal and external stakeholders in decision making (Boskou et al., 2018)

Another novel approach private sector banks in India are embarking on service quality framework with text mining techniques to assess the customer service quality and their perceptions on products offered by the banks (Chakrabarti et al., 2018).

In one of the recent studies, the Securities and Exchange Commission used rule-based text mining techniques to extract information from the plain language reported in the annual 10Ks of publicly listed firms. Based on the tone in which the business managers have spoke, future period performances are assessed (Kang et al., 2018).

Insurance companies can be prone to risks due to non-performing assets, poor governance and lack of adequate controls. Text mining algorithms can be effectively used to identify key risk factors based on the financial statements published by the banks. Periodic review of the process enables the identification of new risks not identified previously (Wang et al., 2019).

Gold price is influenced by the volatility of the News articles. Key economic indicators available in news articles are being effectively used to model volatility in gold prices with the help of text mining algorithms. Classifiers have been developed to predict as to which newsgroup is affecting the gold price (Onsumran et al., 2015).

National policies can be analyzed with the help of text mining where policy documents become the input data. Since policies are available in the written format, a large number of such documents can be analyzed to unearth patterns, extract topics and identify causal relationships (Han et al., 2019).

Another important aspect in the financial domain is when CXOs make public statements or through earning call transcripts and its impact on the stock market performance. A change in the behaviour of directors expressed through video, audio or earning call transcripts can have a considerable impact on the stock market. Such anomalies expressed by organizations can be analyzed with text mining algorithms (Nourbakhsh & Bang, 2019).

The investor community affects stock price movements; their sentiments directly impact on stock market performance. The investor sentiments can be further incorporated in improving the forecasts of a stock price. Investor behaviour is not always unbiased as opposed to the efficient market hypothesis. This is often termed as investor irrational behaviour (Shi et al., 2018; Yin et al., 2018). Their studies have demonstrated the relationship between investor sentiments and the stock market performance and used to Support Vector Machine (SVM) to model the scenarios.

Stock market movement is paramount as it has a bearing on the country's economy. Considering its impact on the economy, several studies have been conducted to identify the key drivers of volatile stock price movements. With technologies evolving and advancing novel approaches like text mining algorithms with Random Forest as machine learning models being applied to classify news articles and detect stock market directions for future decisions based on key drivers (Elagamy et al., 2018)

Several studies have been conducted to extract hidden patterns in textual information. Modeling random behaviour of the stock market performance with event-based textual data available in the form of financial reports, news articles, investor reviews, analyst reviews and annual reports have been explored by researchers. Although this study uses textual information available in the form of earnings call transcripts, we propose a novel idea of feature engineering based on a new feature called 'Inverse Effect' as a response variable. The 'Inverse Effect' refers to the discordance between the present positive or negative sentiments in the boardroom discussions and performance of the stock in the subsequent period. The study also proposes how 'Inverse Effect' can be considered as a lead indicator of the stock performance for a given stock.

The methodology followed is CRISP-DM (Cross Industry Standard Process for Data Mining) popular in the field of data science. The problem-solving methodology involves six stages, business understanding through model deployment.

Chapter 3: Problem Statement

Organizations strive to perform well in terms of improved profitability and market performance. The performance of an organization depends on various factors like sales pipeline, customer satisfaction, propensity to get new businesses and last but not least managing shareholder relationship and better public image. Apart from these factors, the market performance of an organization is often decided by public perception and company's internal activities like quarterly performance review, annual general meeting and share holders meeting. Many a times what is spoken by CXOs may have an impact on market performance which is difficult to quantify in terms of numbers. Organizations look for the early warning signs to predict their future performance based on some manifestations which have taken place. These manifestations are some events like statements made, news items published in social media and these are mostly qualitative in nature. Publicly listed companies are often watched closely for their financial performance as they are accountable to investors and public. Hence it is important that organizations are perceived well in the minds of people. Numerous analytical studies have been done in the area of text mining to extract hidden patterns available in unstructured data. Unstructured do not have a specific format and are often textual in nature. This can be minutes of financial review meetings, shareholders meeting, annual general meeting, analyst reviews, financial statements and auditor recommendations. Today, organizations have mammoth amount of textual information in the form of reports and documents and with social media at its peak; there is no dearth of information. The major challenge the organizations are often faced with is processing of such vast textual information to extract actionable insights for effective decision making and profit maximization. There is an immense need for organizations to use text mining techniques to take advantage of the data and use it for effective market intelligence. This study utilizes a novel approach of using hybrid data which enable organizations in effective decision making based on insights extracted from both structured and unstructured data.

Chapter 4: Objectives of the Study

The purpose of the study is to identify the determinants of 'Inverse Effect' and to develop suitable models to predict the 'Inverse Effect' using machine learning classification algorithms. Further the study also aims to determine if inverse effect is a lead indicator of the market performance. The following are the objectives of the study.

i. To extract the sentiments of the Earnings Call Transcripts (ECT) using Lexicon based sentiment analysis,
ii. To analyze the impact of quarterly review meetings on the subsequent period of market performance
iii. To identify the key drivers of the 'Inverse Effect',
iv. To develop various classification models to predict the 'Inverse Effect'.

Chapter 5: Project Methodology

The methodology followed is CRISP DM (Cross Industry Standard Process for Data Mining) a popular methodology in the field of data science. The problem solving methodology involves six stages, business understanding through model deployment. The pictorial representation of the framework is given below.

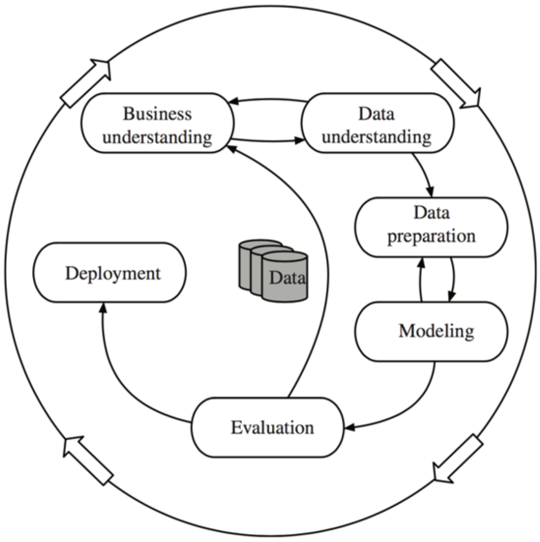

Figure 5.1: CRISP DM Methodology

The first stage in CRISP DM methodology starts with business understanding. In this phase one would focus on understanding the domain, the objectives to be accomplished from the study and converting the business problem into a potential data mining opportunity. This phase also deals with releasing a high level plan to achieve the set objectives. The second phase in the methodology is data understanding. Here one would undertake data collection and get familiarity with data identify gaps in the data and extract actionable insights from the data collected. The data preparation phase also covers making certain assumptions and testing them using hypothesis testing. Finally an analysis base table is created which goes as an input into a machine learning model. The analysis base table will be complete in aspects in terms of handling issues or gaps in the data, feature engineering, feature selection and necessary transformations. In data understanding phase, based on the data wrangling activity undertaken, one would fairly have an understanding about the kind of machine learning models likely to fit the data. In the modelling stage, several restrictive and flexible models will be explored based on the patterns observed in the data understanding phase. In the model selection phase, based on the several models explored, the best model will be selected based on parameters tuning and there is a need to revisit data understanding phase repeatedly till the best features and optimal parameters are selected. This stage warrants one to revisit data understanding stage to understand if any important business aspect was ignored during data understanding phase, which is key to model accuracy. Post model development, the model evaluation is undertaken to assess its generalizability. In model evaluation phase one would assess if the set objectives are accomplished by the model. One may revisit the previous stages if the model performance is not satisfactory. Once the adequacy of the model is ascertained, the developed model is pushed into the real time production environment. In this stage, the stakeholder participation from different functions is required for effective implementation of the model.

Chapter 6: Business Understanding

Quarterly financial performance reports help investors' judge pulse of organizations. Investors get ample insight into growth and performance of organizations by comparing reports quarterly. Earnings Call Transcripts (ECT) is one form of unstructured reports released quarterly by organizations wherein CXOs discuss past performance and growth trajectory for subsequent quarters. These are made available to investors in respective company websites in both audio and textual formats. From an investor point of view it is paramount to know whether what is spoken during the earnings call has an impact on future market performance or just a mundane activity. This empirical study aims at analyzing the impact of these reports on market performance through statistical and analytical approach. A SIPOC view of the business that is the "Earning Call Transcripts" process is given below for the benefit of the reader.

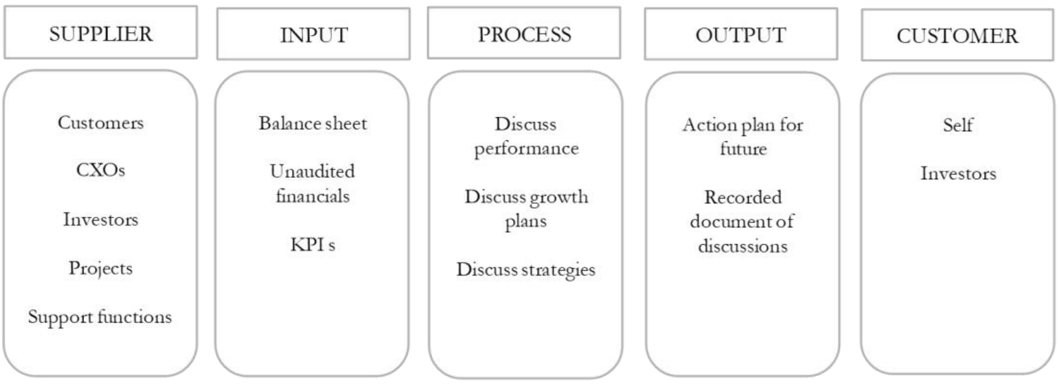

Figure 6.1: SIPOC View of ECT process

Chapter 7: Data Understanding

The data for the proposed study comprise of unstructured corpus available in the form of textual transcripts and the historical earnings per share and stock prices which are structured in nature. The data spanning last ten years i.e., for the period 2008 to till date was collected with the frequency being quarterly. A snapshot of the data dictionary is presented below.

Variable	Scale	Type	Period	Periodicity	Data Source
Percentage positive sentiments	Continuous	Endogenous	2008-2020	Quarterly	Transcripts available in ECT
Percentage negative sentiments	Continuous	Endogenous	2008-2020	Quarterly	Transcripts available in ECT
Percentage neutral sentiments	Continuous	Endogenous	2008-2020	Quarterly	Transcripts available in ECT
Earnings Per Share (EPS)	Continuous	Endogenous	2008-2020	Quarterly	ECT
Stock Price	Continuous	Exogenous	2008-2020	Quarterly	Yahoo Finance
Polarity	Continuous	Endogenous	2008-2021	Quarterly	Transcripts available in ECT

Table 7.1: Data Dictionary

Chapter 8: Data Preparation

The data preparation process involved the following steps:

1. Extraction of sentiments and polarity from quarterly transcripts
2. Extraction of Earnings Per Share (EPS) from financial statements
3. Extraction of Stock Price (SP) from Yahoo finance
4. Feature engineering (extraction new features and a response variable, which we term as **"Inverse Effect**"
5. Creation of analysis base table for analysis and model building

New features were extracted from the existing data which comprised of a response variable called the "Inverse Effect" and a set of features made of polarity and stock price. Inverse Effect is simply the discordance between the current quarter sentiment and the next period stock market performance. To illustrate further, if the current quarter sentiments are positive and if this results in positive increase in the next period stock performance, it is concordance else it is discordance. The snapshot of the analysis base table and the word cloud extracted using Vader sentiment are provided below.

Inverse_Effect	Polarity_Class	SP_3Q_ave	SP_6Q_ave	SP_9Q_ave	SP_PrevQtr
No Positive Effect	Positive	86.847	86.005	73.646	124.388
No Positive Effect	Positive	120.007	99.320	83.150	135.956
No Positive Effect	Positive	132.100	106.370	94.409	135.956
No Positive Effect	Neutral	129.627	108.237	100.546	116.969
No Positive Effect	Positive	116.959	118.483	105.199	97.950
No Positive Effect	Positive	109.044	119.336	108.506	114.591
No Positive Effect	Neutral	115.204	116.081	117.390	116.430
No Positive Effect	Neutral	115.003	112.420	118.980	113.988
No Positive Effect	Neutral	116.143	112.594	118.271	118.011
No Positive Effect	Neutral	116.670	115.937	116.278	118.011
No Positive Effect	Neutral	113.946	114.474	112.929	105.815

Table 8.1: Analysis Base Table

Figure 8.2: Transcripts Word Cloud

Chapter 9: Exploratory Data Analysis (EDA)

Post creation of the analysis base table, the data was explored using statistical techniques and various insights were generated. The individual distribution of the variables and the correlations were studied. The individual distributions plot indicates distribution of the variables being right skewed for earnings per share, stock price, positive and negative sentiments. The neutral sentiments are mostly left skewed.

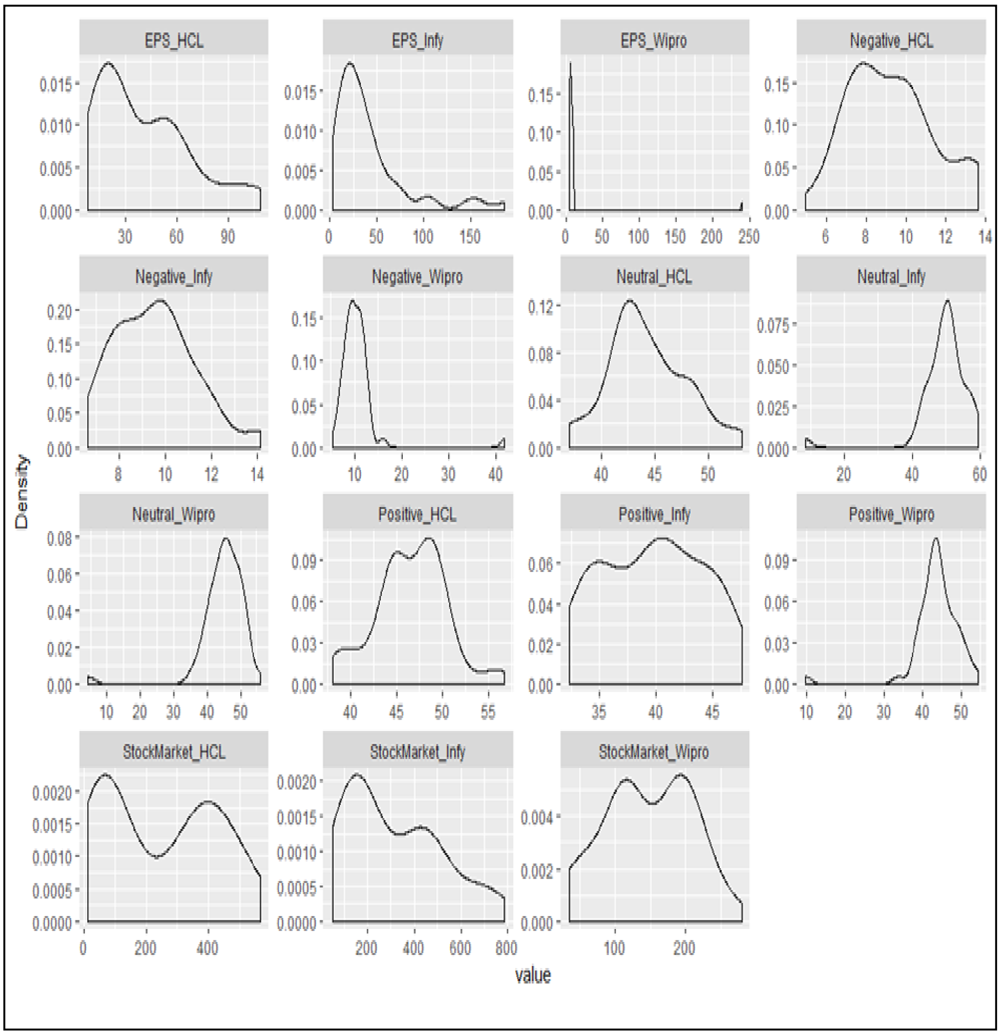

Figure 9.1: Individual Distribution Analysis of Sentiments

This study initially assumed a possible correlation between the polarity and the inverse effect. The correlations study reveals possible association between sentiments and the stock price performance. Kendall's tau statistic was employed to study correlations between the variables. Kendall's tau is a non-parametric test similar to Pearson's product moment and Spearman's rank correlation techniques. Kendall's tau and Spearman's rank correlations are more suited in this case as the underlying distributions are skewed as depicted in figure3.0. Kendall's tau takes the value between 0 and 1, where 0 denotes no relationship and the value close 1 indicates perfect relationship. Similar results can be obtained using Spearman's rank correlations as well.

Correlation-Infosys					
		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-b	.602	.084	5.000	.000
N of Valid Cases		48			

Correlation-Wipro					
		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-b	.793	.067	10.132	0.000
N of Valid Cases		56			

Correlation-HCL					
		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-b	.378	.096	2.283	.022
N of Valid Cases		37			

Table 9.2: Kendall's Tau Test Table

Kendall's tau indicates a strong correlation between polarity and the inverse effect. Polarity is a metric derived from the sentiments extracted from the quarterly earnings call transcripts. The null hypothesis of the test is two categorical variables being independent. Since the p-values in all the cases are less than the significance level, the null hypothesis is rejected and the conclusion is that there exists an association between polarity and the inverse effect. This insight was used as the basis for exploring various machine learning classification models for predicting the inverse effect. “Kendall’s tau b” is used as the contingency table formed by the two categorical variables is a square matrix.

Chapter 10: Data Modeling

The response variable "Inverse Effect" was modeled using several machine learning restrictive and flexible methods. Different models were tried and tested on three different data sets, Wipro, Infosys and HCL respectively. A general model building frame work is given below.

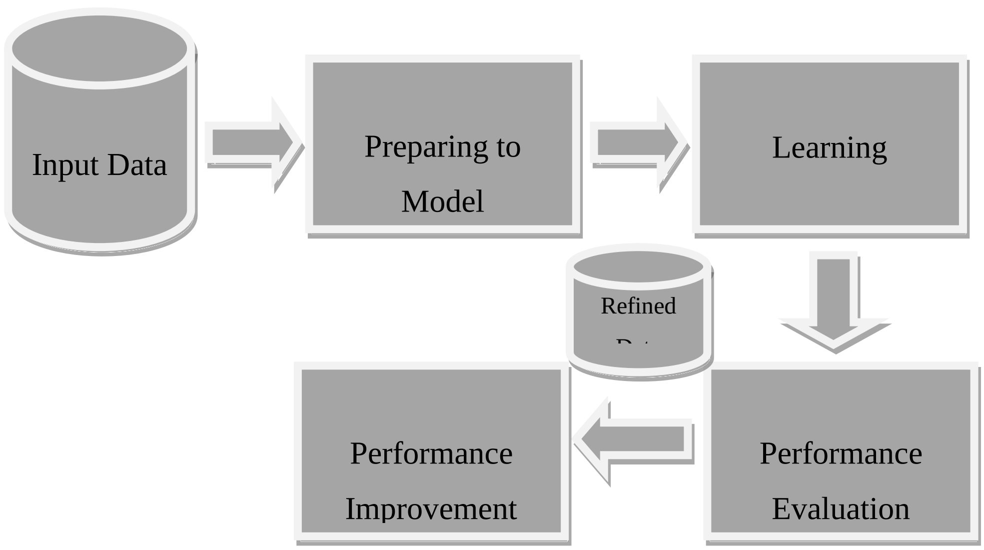

Figure 10.1: Machine Learning Process

The data modeling process starts with feeding of the input data, which is the analysis base table created as part of data preparation explained in chapter 8. The next step is to perform data partitioning. The data partitioning is to randomly divide data into train and test groups. Post data partitioning, the train group is fed into the classifier algorithm for model development. Post model development using train group, the generalizability of the model is assessed on the previously unseen test group. Finally the performance measures of several models are reported. The following models were tried and tested as part of the model development. A brief illustration of these models is given below for the benefit of reader.

Decision Tree: Decision tree learning is one of the most widely adopted algorithms for classification. It builds the model in the form of a tree structure.

A decision tree is used for multidimensional analysis with multiple classes. It is characterized by fast execution time and ease in the interpretation of rules. The goal of decision tree learning is to create a model based on the past vector that predicts the value of the response variable based on the input variables in the feature vector space.

Binary Logistics: Logistic regression is both a classification and a regression technique depending on the scenario used. Logistic regression is a type of regression used for predicting the outcome of a categorical dependent variable similar to ordinary least squares regression. In logistic regression, response variable is dichotomous.

Naïve Bayes: Naïve Bayes classifier makes the naïve assumption that the attribute values are conditionally independent given the classification of the instance. This simplifying assumption considerably reduces the calculation overhead without losing the effectiveness of the outcome.

K-Nearest Neighbor: The kNN algorithm is a simple but extremely powerful classification algorithm. In kNN algorithm, the unknown and unlabelled data which comes for a prediction problem is judged on the basis of the training dataset elements which are similar to the unknown element. So, the class label of the unknown element is assigned on the basis of the class labels of the similar training data set elements, metaphorically can be considered as neighbors of the unknown element.

Random Forest: Random forest is an ensemble classifier that is combining classifier that uses and combines many decision tree classifiers. Ensembling is usually done using the concept of bagging. Bagging stands for bootstrap and aggregation, applied on different feature sets. The reason for using large number of trees in random forest is to train the trees enough such that contribution from each feature comes in a number of models. After the random forest is generated by combining many trees, majority vote is applied to combine the output of the different trees.

Support Vector Machine (SVM): Support Vector Machines (SVM) can do both classification and regression. SVM is based on the concept of surface, called a hyperplane, which draws a boundary between data instances plotted in the multi-dimensional feature space. The output prediction of SVM algorithm is one of two conceivable classes which are already defined in the training data. In general, SVM builds an N-dimensional hyperplane model that assigns future instances into one of the two possible output classes.

Neural Network: An Artificial Neural Network (ANN) models the relationship between a set of input signals and an output signal using a model derived from our understanding of how a biological brain responds to stimuli from sensory inputs. Just as a brain uses a network of interconnected cells called neurons to create a massive parallel processor, ANN uses a network of artificial neurons or nodes to solve learning problems

The performance measures of the aforesaid classification algorithms are presented in chapter11 in the form of leaderboard.

Chapter 11: Model Evaluation

Several restrictive and flexible models explored as part of the study were validated on the test data to assess their generalizability. A leaderboard depicting different performance measures for different models for each of the data set (Wipro, Infosys and HCL) is presented below.

#	Organization	Algorithm	Sampling Method	TPR	FPR	PLR	Precision	Recall	F1-Measure	AUC	Coverage	Rank
1	Wipro	J48_Pruned Tree	Cross Validation	0.964	0.036	26.78	0.967	0.964	0.964	0.94	0.9642	1
2	Wipro	Random Forest	Cross Validation	0.964	0.036	26.78	0.967	0.964	0.964	1	0.9821	1
3	Wipro	SGD	Cross Validation	0.911	0.089	10.24	0.924	0.911	0.91	0.911	0.9107	3
4	Wipro	SVM(Polydot)	Cross Validation	0.946	0.054	17.52	0.952	0.946	0.946	0.946	0.9464	2
5	Wipro	Binary Logistics	Cross Validation	0.946	0.054	17.52	0.952	0.946	0.946	0.909	0.9464	2
6	Wipro	Nnet(MLP)	Cross Validation	0.946	0.054	17.52	0.952	0.946	0.946	0.898	0.9464	2
7	Wipro	Naïve Bayes	Cross Validation	0.821	0.179	4.59	0.868	0.821	0.816	0.923	0.9464	4
8	Wipro	KNN	Cross Validation	0.946	0.054	17.52	0.952	0.946	0.946	0.935	0.9464	2
1	Infosys	J48_Pruned Tree	Cross Validation	0.833	0.174	4.79	0.842	0.833	0.832	0.877	0.9583	4
2	Infosys	Random Forest	Cross Validation	0.896	0.11	8.15	0.902	0.896	0.895	0.935	0.9583	1
3	Infosys	SGD	Cross Validation	0.813	0.193	4.21	0.817	0.813	0.811	0.81	0.8125	5
4	Infosys	SVM(Polydot)	Cross Validation	0.75	0.251	2.99	0.75	0.75	0.75	0.75	0.75	7
5	Infosys	Binary Logistics	Cross Validation	0.854	0.148	5.77	0.855	0.854	0.854	0.864	0.9375	3
6	Infosys	Nnet(MLP)	Cross Validation	0.875	0.136	6.43	0.899	0.875	0.872	0.84	0.9583	2
7	Infosys	Naïve Bayes	Cross Validation	0.792	0.216	3.67	0.799	0.792	0.789	0.758	0.8333	6
8	Infosys	KNN	Cross Validation	0.896	0.11	8.15	0.902	0.896	0.895	0.867	0.8958	1
1	HCL	J48_Pruned Tree	Cross Validation	0.649	0.413	1.571428571	0.787	0.649	0.583	0.521	1	2
2	HCL	Random Forest	Cross Validation	0.378	0.643	0.587869362	0.366	0.378	0.37	0.337	0.9459	6
3	HCL	SGD	Cross Validation	0.649	0.413	1.571428571	0.787	0.649	0.583	0.618	0.6486	2
4	HCL	SVM(Polydot)	Cross Validation	0.649	0.413	1.571428571	0.787	0.649	0.583	0.618	0.6486	2
5	HCL	Binary Logistics	Cross Validation	0.622	0.392	1.586734694	0.62	0.622	0.619	0.653	0.9729	1
6	HCL	Nnet(MLP)	Cross Validation	0.595	0.433	1.374133949	0.592	0.595	0.582	0.576	1	3
7	HCL	Naïve Bayes	Cross Validation	0.568	0.482	1.178423237	0.569	0.568	0.521	0.356	1	4
8	HCL	KNN	Cross Validation	0.378	0.625	0.6048	0.381	0.378	0.379	0.368	0.3783	5

Table 11.1: Model Leader Board

The following tables provides a view of various classification performance measures and their relevance

#	Performance Measure	Definition
1	True Positive Rate	These are cases which the model predicted as positive when we also said positive. Normally a classification algorithm should predict more of true positive rates
2	False Positive Rate	These are the cases which are actually not true, but the model predicted as true
3	Positive Likelihood Ratio	It is the ratio of sensitivity to one minus specificity. A value greater than 1 indicates the test

		result being associated with true positives
4	Precision	Precision gives the proportion of positive predictions which are truly positive
5	Recall	Recall gives the proportion of true positive cases over all actually positive cases
6	F1-Measure	F1-measure is the harmonic mean of precision and recall with the assumption of equal weights
7	Area Under Curve	It is a measure of discriminatory power of the model in maximizing the true positive rate by minimizing the false positive rate. The value of AUC ranges from 0 to1
8	Coverage	It is the percentage of the target values which lie within 95% confidence interval

Table 11.2: Classification Performance Measures

Based on the performance measures reported in the above table, different models have been proposed to predict the inverse effect for each of the companies. Following are the model recommendations in the order of PLR values (higher is desirable):

1. Decision Tree/Random forest and SVM are the models recommended in the case of Wipro
2. Random forest/KNN and neuralnet are the models recommended in the case of Infosys
3. Binary Logistics, Decision tree/SVM are the models recommended in the case of HCL

To summarize, flexible models seems to be capturing the patterns in the data well than the restrictive models. This suggests the presence of non-linear relationships between the target variable and the predictors. The proposed models can be used in real time basis to understand if inverse effect is present or not.

Chapter 12: Deployment

In chapter 11 we have demonstrated the performance measures of various models which can predict the inverse effect. Mostly flexible models are doing phenomenally well compared to restrictive ones. Now these models can be used by leadership team in predicting the inverse effect based on the quarterly meetings and earning call transcripts. Given the past performance one can predict the inverse effect and this market intelligence is useful organizations and the investor community in decision making. A mathematical of the form given below can be deployed to get the required output.

Inverse Effect = f(Polarity, Stock Price_Prev Qtr, Stock Price_3Qtr_Avg, Stock Price_6Qtr_Avg, Stock Price_9Qtr_Avg)

Figure 12.1: Functional form of Inverse Effect

Inverse Effect- This is the target variable. Inverse effect is the discordance between the current period sentiments and the next period market performance. It is dichotomous with two levels namely “Positive Increase” and “Not Positive Increase”

The details of the predictors are as below:

Polarity –a measure deduced from the sentiments extracted from the earning call transcripts

Stock Price_Prev Qtr- a feature which measures the value of the previous quarter stock price

Stock Price_3Qtr_Avg-a feature which measures the average stock price for the previous 3 quarters

Stock Price_6Qtr_Avg- a feature which measures the average stock performance for the previous 6 quarters

Stock Price_9Qtr_Avg- a feature which measures the average stock performance for the previous 9 quarters

Given the value of the predictors and the respective beta coefficients, simple restrictive models can be deployed to predict the inverse effect. There is no need for implementation of R, Python and other open source applications, model democratization can be achieved with simple spreadsheet implementations. Depending on the magnitude of data and the features, different levels of implementations can be achieved with the help of open source applications.

Chapter 13: Analysis and Results

Though the end objective of the study was to model the inverse effect, certain hypotheses were tested and few actionable insights were drawn using exploratory data analysis and decision tree models. Initially, hypotheses were tested to study if any relationship exists between sentiments and the market performance. It was found that there exists a relationship between the sentiments and the stock market performance. In the case of Wipro, there exists a correlation 66% between the negative sentiments and the stock price. With Infosys, a 65% correlation existed between the positive sentiments and the market performance, which implies a negative impact on the stock price. In the case of HCL, the correlations are not significant, but a moderate negative correlation of 39% existed between the neutral sentiments and the market performance.

The decision tree analysis reveals the factors (key drivers) contributing to the inverse effect. Following are the findings from decision tree models.

1. In the case of HCL, the Inverse effect is influenced by the sentiments
2. In the case of Wipro, Inverse Effect is influenced by the Previous three quarters average Stock Price
3. In the case of Infosys, Inverse Effect is influenced by sentiments, previous quarter stock price and previous three quarters average stock price

Wipro	Infosys	HCL
1.Inverse Effect is solely driven by the previous 3 quarter average stock price 2. Whenever the previous 3 quarter average stock price is less than equal to 109.83INR, there is an 80% chance of observing positive increase in the market in the subsequent period 3. Whenever the previous 3 quarter average stock price is greater than 109.83INR, there is a 59% chance of not observing a positive increase in the next period 4. Good to invest when the previous 3 quarters average is less than or equal to 109.83INR	1.Polarity, previous quarter stock price and previous 3 quarters average stock price are the key drivers of inverse effect 2.There is a 78% chance of observing a positive increase in the market whenever the current quarter polarity is positive and the previous quarter end price is <=INR444.744 and the previous 3 quarters average is <=157.45 3.In all the other cases, market will not react positively 4.Good to invest whenever condition(2) is satisfied	1.Inverse Effect is solely driven by the polarity 2.There is 60% chance that the market would positively increase in the next period whenever the polarity of the current period is positive 3.Whenever the current quarter polarity is positive, there is almost certainty that the market will not increase in the positive direction 4.Good to invest when the current quarter polarity is positive

Table 13.1: Decision Tree Rules

Further analysis revealed, if inverse effect is a lead indicators of the market performance. Based on the data, in the case of Wipro, whenever the ‘Inverse Effect’ is not positive, one can expect an average change in the stock price of 8.5%, this is because negative sentiments have a bearing on the stock price. In the case of Infosys, a positive increase in the ‘Inverse Effect’ lead to an average 12.5% change in stock price and a 10.3% change for a not positive increase. Similarly for HCL, one can expect an average change of 20.8% for a positive increase in the ‘Inverse Effect’ and 9% otherwise.

Chapter 14: Conclusions and Recommendations for future work

Stock markets being volatile, robust techniques are required to capture volatility and thus make predictions for investment decision making. There are umpteen ways of modelling volatility. Models like ARIMA, regression based forecasting, ARCH-GARCH, neural net models and other machine learning models come handy in modelling time series data. However, they do not factor endogenous variables. The concept of inverse effect modelling is a novel approach in itself wherein one can observe the effect of inverse effect on the stock market performance in the subsequent quarters. From a trading point of view one can decide how much will be gain or loss in the short term. Based on the data, in the case of Wipro, whenever the inverse effect is not positive, one can expect an average change in stock price of 8.5%, this is due to the fact that negative sentiments have a bearing on stock price. In the case of Infosys, a positive increase in the inverse effect lead to an average 12.5% change in stock price and a 10.3% change for not positive increase. Similarly for HCL, one can expect an average change of 20.8% for positive increase in the inverse effect and 9% otherwise.

The limitation of this study is that it is confined only to software companies. The future scope of this study can be extended to various sectors and different events, news articles, analysts reports can be considered to model random behaviour of the stock market.

Bibliography

[1] Boskou, G., Kirkos, E., & Spathis, C. (2018). Assessing internal audit with text mining. *Journal of Information and Knowledge Management*, *17*(2), 1–22. https://doi.org/10.1142/S021964921850020X

[2] Boskou, G., Kirkos, E., & Spathis, C. (2018). Assessing internal audit with text mining. *Journal of Information and Knowledge Management*, *17*(2), 1–22. https://doi.org/10.1142/S021964921850020X

[3] Chakrabarti, S., Trehan, D., & Makhija, M. (2018). Assessment of service quality using text mining – evidence from private sector banks in India. *International Journal of Bank Marketing*, *36*(4), 594–615. https://doi.org/10.1108/IJBM-04-2017-0070

[4] Davis, A. K., & Tama-Sweet, I. (2012). Managers' Use of Language Across Alternative Disclosure Outlets: Earnings Press Releases Versus MD&A. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1866369

[5] Demers, E., & Vega, C. (2008). Soft Information in Earnings Announcements : News or Noise? *International Finance Discussion Paper*, *2008*(951), 1–56. https://doi.org/10.17016/ifdp.2008.951

[6] Elagamy, M. N., Stanier, C., & Sharp, B. (2018). Stock market random forest-text mining system mining critical indicators of stock market movements. *2nd International Conference on Natural Language and Speech Processing, ICNLSP 2018*, 1–8. https://doi.org/10.1109/ICNLSP.2018.8374370

[7] Han, H., Wang, Q., & Chen, C. (2019). Policy Text Analysis Based on Text Mining and Fuzzy Cognitive Map. *Proceedings - 2019 15th International Conference on Computational Intelligence and Security, CIS 2019*, 142–146. https://doi.org/10.1109/CIS.2019.00038

[8] Kang, T., Park, D. H., & Han, I. (2018). Beyond the numbers: The effect of 10-K tone on firms' performance predictions using text analytics. *Telematics and Informatics*, *35*(2), 370–381. https://doi.org/10.1016/j.tele.2017.12.014

[9] Kumar, B. S., & Ravi, V. (2016). A survey of the applications of text mining in financial domain. *Knowledge-Based Systems*, *114*, 128–147. https://doi.org/10.1016/j.knosys.2016.10.003

[10] Li, X., Wu, P., & Wang, W. (2020). Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong. *Information Processing and Management*, *July 2019*, 102212. https://doi.org/10.1016/j.ipm.2020.102212

[11] Nam, K. H., & Seong, N. Y. (2019). Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market. *Decision Support Systems*, *117*, 100–112. https://doi.org/10.1016/j.dss.2018.11.004

[12] Nourbakhsh, A., & Bang, G. (2019). *A framework for anomaly detection using language modeling, and its applications to finance*. http://arxiv.org/abs/1908.09156

[13] Onsumran, C., Thammaboosadee, S., & Kiattisin, S. (2015). Gold Price Volatility Prediction by Text Mining in Economic Indicators News. *Journal of Advances in Information Technology*, *6*(4), 243–247. https://doi.org/10.12720/jait.6.4.243-247

[14] Shi, Y., Tang, Y. R., Cui, L. X., & Long, W. (2018). A text mining based study of investor sentiment and its influence on stock returns. *Economic Computation and Economic Cybernetics Studies and Research*, *52*(1), 183–199. https://doi.org/10.24818/18423264/52.1.18.11

[15] Wang, Y., Li, B., Li, G., Zhu, X., & Li, J. (2019). Risk factors identification and evolution analysis from textual risk disclosures for insurance industry. *Procedia Computer Science*, *162*(Itqm 2019), 25–32. https://doi.org/10.1016/j.procs.2019.11.253

[16] Yin, L., Zhang, N., He, L., & Fang, W. (2018). A Study of Relationship between Investor Sentiment and Stock Price Based on Text Mining. *Proceedings - 2016 International Conference on Identification, Information and Knowledge in the Internet of Things, IIKI 2016*, *2018-Janua*, 536–539. https://doi.org/10.1109/IIKI.2016.49

[17] Yu, S., & Guo, S. (2016). Big data concepts, theories, and applications. In *Big Data Concepts, Theories, and Applications* (Issue December). https://doi.org/10.1007/978-3-319-27763-9

Appendix

Plagiarism Report[1]

Publications in a Journal/Conference Presented/White Paper[2]

Publication of this study has been initiated in the following journals.

1. Global business Review

Any Additional Details

The following sources were referred to collect stock performance data of the companies considered for this study.

https://www.wipro.com/quarterly-results/

http://conferencecalltranscripts.org/?co=INFY

https://www.infosys.com/investors/reports-filings/quarterly-results.html

https://www.hcltech.com/investors/results-reports

https://in.finance.yahoo.com/

The data analysis in this study is done using Weka open source tool.

https://sourceforge.net/projects/weka/

[1] Turnitn report to be attached from the University.

[2] URL of the white paper/Paper published in a Journal/Paper presented in a Conference/Certificates to be provided.

INVERSE EFFECT MODELING OF EARNING CALL TRANSCRIPTS FOR FINANCIAL ANALYSIS

by Nagendra Bv

INVERSE EFFECT MODELING OF EARNING CALL TRANSCRIPTS FOR FINANCIAL ANALYSIS

ORIGINALITY REPORT

12%

SIMILARITY INDEX

7%

INTERNET SOURCES

6%

PUBLICATIONS

6%

STUDENT PAPERS

PRIMARY SOURCES

1 Taeyoung Kang, Do-Hyung Park, Ingoo Han. "Beyond the numbers: The effect of 10-K tone on firms' performance predictions using text analytics", Telematics and Informatics, 2018

Publication

www.caedes.net

Internet Source %

3 Souvik Chowdhury, Shibakali Gupta. "A comparative study of text mining in big data analytics using deep learning and other machine learning algorithms", International Journal of Hybrid Intelligence, 2019

Publication

Submitted to Coventry University
Student Paper %

Submitted to Sogang University
Student Paper %

6 "Data Management, Analytics and Innovation", Springer Science and Business Media LLC,

7

8

9

10

11

12

13

14

15

2021
Publication

Kapitanova, Krasimira, and Sang Son. "Machine Learning Basics", Intelligent Sensor Networks The Integration of Sensor Networks Signal Processing and Machine Learning, 2012.
Publication

Submitted to University of Lancaster
Student Paper

digitalcommons.calpoly.edu
Internet Source

mafiadoc.com
Internet Source

www.edilms.eti.br
Internet Source

ir.cuea.edu
Internet Source

"Opinion Mining and Client Feedback Mining for Quality Improvement", International Journal of Engineering and Advanced Technology, 2019
Publication

archive.org
Internet Source

www.termpaperwarehouse.com

Internet Source

1%

<1%

<1%

<1%

<1%

<1%

<1%

<1%

<1%

17

18

19

20

21

22

23

24

Cornelia Caragea, Jivko Sinapov, Vasant Honavar, Drena Dobbs. "Assessing the Performance of Macromolecular Sequence Classifiers", 2007 IEEE 7th International Symposium on BioInformatics and BioEngineering, 2007
Publication

halshs.archives-ouvertes.fr
Internet Source

"The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2018)", Springer Science and Business Media LLC, 2018
Publication

www.inmybangalore.com
Internet Source

Submitted to University of Durham
Student Paper

student-friendly.blogspot.com
Internet Source

www.researchsquare.com
Internet Source

news.vidyaacademy.ac.in
Internet Source

John M. Williamson, Sara B. Crawford, Hung-

<1%

<1%
<1%

<1%
<1%
<1%

<1%
<1%

25

26

Mo Lin. "A Multiple Imputation Approach for Estimating Rank Correlation With Left-Censored Data", Statistics in Biopharmaceutical Research, 2010
Publication

www.infosys.com
Internet Source

docplayer.net
Internet Source

‘INVERSE EFFECT’ MODELING OF EARNINGS CALL TRANSCRIPTS FOR FINANCIAL ANALYSIS

Nagendra.B.V
NagendraBV.BA02@reva.edu.in

Dr. J.B Simha
Chief Technology Officer
ABIBA Systems
Bengaluru
jbsimha@gmail.com

Dr. Shinu Abhi
Director
REVA Academy for Corporate Excellence
shinuabhi@reva.edu.in

Abstract— This research paper intends to study the relationship between the quarterly board room discussions conducted in the form of "Earnings Call Transcripts" and its influence on the company's stock performance. In this study, the extraction of sentiments was undertaken from the "textual quarterly transcripts" of three major software companies for 10 years. The extracted sentiments were statistically analyzed for patterns and sentiments. Apart from this, new features were extracted from the existing data which comprised of a response variable called the 'Inverse Effect'. The 'Inverse Effect' refers to the discordance between the present positive or negative sentiments in the boardroom discussions and performance of the stock in the stock market. For example, if the boardroom sentiments for the current quarter are positive and the stock performance in the market is also positive in the same quarter, it is considered as "concordance" and if the performance of the individual stock is opposite to the sentiments shared in the boardroom it will be called as "discordance".

The findings emerged from the study suggest a possible causality between the sentiments and the stock market performance. It was also found that polarity, three-quarter average stock price and the previous quarter stock price are the determinants of the 'Inverse Effect'. Based on the findings from the study, appropriate machine learning models were developed and evaluated to predict the 'Inverse Effect' on the performance of individual stocks in the stock market.

Key Words: Earnings Call Transcripts, Sentiments, Polarity, Earnings per Share, Market Performance, Stock Price, 'Inverse Effect', Behavioural Finance

Introduction

Market intelligence is the process of effective use of endogenous and extraneous information for making better business decisions. Organizations, use market intelligence data extensively for decision making. Market intelligence enables organizations to foresee risk and accordingly initiate corrective and preventive actions. Market intelligence in the context of this study is to know in advance the likely performance of the market based on certain inputs which are generally available in the form of quarterly fundamental data and unstructured corpus like Earnings Call Transcripts (ECT) and Annual Directors Reports.

Every organization is required to measure data related to its performance. The performance can be related to employees, customers, business transactions or organization as a whole. Based on the business requirements necessary reports can be generated. In particular, Davis and Tama-Sweet found a positive relationship between an increase in positive tone within earnings press releases and the short-term stock price response to firms' earnings announcement. They also studied whether managers adjust the tone of earnings press releases to increase the value of their stock options and found that the optimal level increased before their option exercises. The results of their studies show that optimism in their tone had a positive impact on short-term stock prices and negative effect if their tone was pessimistic (Davis & Tama-Sweet, 2012). Another study by Demers and Vega found a positive association between firms' future returns and the level of tone changes in earnings press releases (Demers & Vega, 2008).

Organizations need to measure their performance for continuous improvement; else they will be at the mercy of market forces which include news media and other competitors. With the advent of technology and robust workflow systems organizations today collect voluminous data both structured and unstructured through various data sources. The process of extracting actionable insights from such voluminous data is often termed as big data analytics in the world of data science. The organization must be able to extract actionable insights from these huge data collected which in turn can be part of its metrics and measurement system. Every actionable insight must be mapped with the organization's profitability to know the effectiveness of the metric being measured. In general, big data analytics enable organizations to improve their overall profitability by extracting actionable insights (Yu & Guo, 2016).

Public invest in stock market, expecting a reasonable return on the money invested. It is important for leadership teams in organizations to protect the interest of their investors/shareholders. Many a time the public takes actions based on the announcements or statements made by the leaders concerning earnings, expenditure and their likely plans during quarterly, half-yearly and yearly held Annual General Meetings (AGMs). AGMs thus play an important role on the stock market performance. It is to be noted here that, more than the company's actual performance; stock markets are susceptible to sentiments and rely heavily on the behavioural aspects of the leadership team. This is because investors are vigilant and keep a close watch on any kind of signs from the leaders to decide their future moves (Demers & Vega, 2008).

The stock market being highly volatile, the leadership team in organizations observe market performance closely and base their strategies and tactics. The organizations need to connect strategies and performance with the actual market performances for better decision making (Li et al., 2020). Through this paper, we demonstrate whether or not what is being spoken by the leadership team in organizations manifest in-market performance, and how consistent is the leadership team in terms of their behaviour during public meetings concerning the overall performance. The study also aims to explore correlations between various internal and external financial performance indicators.

OBJECTIVES OF THE STUDY

At a high level, the purpose of the study is to identify the determinants of 'Inverse Effect' and to develop suitable models to predict the 'Inverse

Effect' using machine learning algorithms. The following are the objectives of the study.

v. To extract the sentiments of the Earnings Call Transcripts (ECT) using Lexicon based sentiment analysis,
vi. To analyze the impact of quarterly review meetings on the subsequent period of market performance,
vii. To identify the key drivers of the 'Inverse Effect',
viii. To develop various classification models to predict the 'Inverse Effect'.

Literature Review

Sentiments extraction is key to text mining. To extract hidden information from the mammoth corpus, text mining algorithms are largely employed. Text mining algorithms can work on any type of input ranging from a sentence to a document or a group of documents. The text mining is a process in itself with a set of actions to be performed. It starts with a document as an input. Once the target document is inputted, some amount of pre-processing of the input takes place. The pre-processing of text is somewhat similar to fundamental data analysis wherein the numeric features are treated for outliers, missing values and more importantly the feature engineering and feature selection which is paramount to machine learning models. In the case of text, each word constitutes a feature. Similar to fundamental data, the text also needs to be treated, this exactly what is accomplished through text pre-processing. Stop words removal is an important step which facilitates removal of certain words like "the"," is", "are"," then", "who" to name a few, which do not add any value to information extraction. Once the pre-processing is done, many mining algorithms are employed to extract information from the processed data. There are many lexicon-based techniques to extract sentiments from the text, the most commonly used are "AFINN" and "VADER (Valence Aware Dictionary and sEntiment Reasoner). Once the sentiments are extracted, they are analyzed and the insights are generated. Businesses collect a large amount of data and social media like Twitter, Facebook have customer reviews, product reviews, reviews on stock market performance reviews. There has been a surge in text data in social media and extracting insights from them is paramount to organizations in decision making. Market intelligence, visualization, document searching and analysis survey responses require text mining algorithms (Nam & Seong, 2019).

Marketing and finance domains rely heavily on social media for information extraction from text data. Various studies show the impact of sentiments analysis on various performance indicators. Apart from extracting the sentiments, text mining algorithms have wide applications in document classification, document clustering, text summarization, social network analysis and topic modelling. Even though significant progress has been made in the field of text mining, feature selection remains challenging and critical to text mining (Kumar & Ravi, 2016).

Internal controls, corporate governance and risk management have off late received a great deal of attention in organizations. A major area of focus is the robustness of internal controls which are periodically reviewed with internal audits. Adequacy of internal controls is analyzed using financial reports available annually with text mining algorithms which aim at keywords and feature selection. The results thus obtained can be useful to internal and external stakeholders in decision making (Boskou et al., 2018)

Another novel approach private sector banks in India are embarking on service quality framework with text mining techniques to assess the customer service quality and their perceptions on products offered by the banks (Chakrabarti et al., 2018).

In one of the recent studies, the Securities and Exchange Commission used rule-based text mining techniques to extract information from the plain language reported in the annual 10Ks of publicly listed firms. Based on the tone in which the business managers have spoke, future period performances are assessed (Kang et al., 2018).

Insurance companies can be prone to risks due to non-performing assets, poor governance and lack of adequate controls. Text mining algorithms can be effectively used to identify key risk factors based on the financial statements published by the banks. Periodic review of the process enables the identification of new risks not identified previously (Wang et al., 2019).

Gold price is influenced by the volatility of the News articles. Key economic indicators available in news articles are being effectively used to model volatility in gold prices with the help of text mining algorithms. Classifiers have been developed to predict as to which newsgroup is affecting the gold price (Onsumran et al., 2015).

National policies can be analyzed with the help of text mining where policy documents become the input data. Since policies are available in the written format, a large number of such documents can be analyzed to unearth patterns, extract topics and identify causal relationships (Han et al., 2019).

Another important aspect in the financial domain is when CXOs make public statements or through earning call transcripts and its impact on the stock market performance. A change in the behaviour of directors expressed through video, audio or earning call transcripts can have a considerable impact on the stock market. Such anomalies expressed by organizations can be analyzed with text mining algorithms (Nourbakhsh & Bang, 2019).

The investor community affects stock price movements; their sentiments directly impact on stock market performance. The investor sentiments can be further incorporated in improving the forecasts of a stock price. Investor behaviour is not always unbiased as opposed to the efficient market hypothesis. This is often termed as investor irrational behaviour (Shi et al., 2018; Yin et al., 2018). Their studies have demonstrated the relationship between investor sentiments and the stock market performance and used to Support Vector Machine (SVM) to model the scenarios.

Stock market movement is paramount as it has a bearing on the country's economy. Considering its impact on the economy, several studies have been conducted to identify the key drivers of volatile stock price movements. With technologies evolving and advancing novel approaches like text mining algorithms with Random Forest as machine learning models being applied to classify news articles and detect stock market directions for future decisions based on key drivers (Elagamy et al., 2018)

Several studies have been conducted to extract hidden patterns in textual information. Modeling random behaviour of the stock market performance with event-based textual data available in the form of financial reports, news articles, investor reviews, analyst reviews and annual reports have been explored by researchers. Although this study uses textual information available in the form of earnings call transcripts, we propose a novel idea of feature engineering based on a new feature called 'Inverse Effect' as a response variable. The 'Inverse Effect' refers to the discordance between the present positive or negative sentiments in the boardroom discussions and performance of the stock in the subsequent period. The study also proposes how 'Inverse Effect' can be considered as a lead indicator of the stock performance for a given stock.

The methodology followed is CRISP-DM (Cross Industry Standard Process for Data Mining) popular in the field of data science. The problem-solving methodology involves six stages, business understanding through model deployment.

Business Understanding

Quarterly financial performance reports help investors' judge the pulse of organizations. Investors get ample insight into the growth and performance of organizations by comparing reports quarterly. Earnings Call Transcripts (ECT) is one form of unstructured reports released quarterly by organizations wherein CXOs discuss past performance and growth trajectory for subsequent quarters. These are made available to investors in respective company websites in both audio and textual formats. From an investor point of view, it is paramount to know what is spoken during the earnings call and whether the talk has an impact on future market performance. This empirical study aims at analyzing the impact of these reports on market performance through a statistical and analytical approach.

Data Understanding

The data for the proposed study comprise of unstructured corpus available in the form of textual transcripts and the historical earnings per share and stock prices which are structured in nature. The data spanning the last ten years i.e., for the period

2008 to till date was collected with the frequency being quarterly. The data for the proposed study comprised of both endogenous and exogenous variables. Endogenous data namely text information available in the form of reports was collected from the company website. These transcripts were further processed to extract sentiments using text mining algorithms. This way, three features namely positive, neutral and negative sentiments were created. Earnings Per Share (EPS) was collected from quarterly audited financial reports. The quarterly stock performance data for the

A snapshot of the data dictionary is presented in Table 1.0.

Variable	Scale	Type	Period	Periodicity	Source
Percentage positive sentiments	Continuous	Endogenous	2008-2020	Quarterly	Transcripts available in ECT
Percentage negative sentiments	Continuous	Endogenous	2008-2020	Quarterly	Transcripts available in ECT
Percentage neutral sentiments	Continuous	Endogenous	2008-2020	Quarterly	Transcripts available in ECT
Earnings Per Share (EPS)	Continuous	Endogenous	2008-2020	Quarterly	ECT
Stock Price	Continuous	Extraneous	2008-2020	Quarterly	Yahoo Finance
Polarity	Continuous	Endogenous	2008-2021	Quarterly	Transcripts available in ECT

Table1.0: Data Dictionary

Data Preparation

The data preparation process involved the following steps:

i. Extraction of sentiments and polarity from quarterly transcripts
ii. Extraction of Earnings Per Share (EPS) from the quarterly unaudited financial statements
iii. Extraction of Stock Price (SP) from Yahoo finance
iv. Feature engineering (extraction new set of features and a response variable, which we term as ‘Inverse Effect’)
v. Creation of analysis base table for analysis and model building

A snapshot of the analysis base table (2.0) and a word cloud (Fig.1.0) are presented below:

Inverse_Effect	Polarity_Class	SP_3Q_ave	SP_6Q_ave	SP_9Q_ave	SP_PrevQtr
No Positive Effect	Positive	86.847	86.005	73.646	124.388
No Positive Effect	Positive	120.007	99.320	83.150	135.956
No Positive Effect	Positive	132.100	106.370	94.409	135.956
No Positive Effect	Neutral	129.627	108.237	100.546	116.969
No Positive Effect	Positive	116.959	118.483	105.199	97.950
No Positive Effect	Positive	109.044	119.336	108.506	114.591
No Positive Effect	Neutral	115.204	116.081	117.390	116.430
No Positive Effect	Neutral	115.003	112.420	118.980	113.988
No Positive Effect	Neutral	116.143	112.594	118.271	118.011
No Positive Effect	Neutral	116.670	115.937	116.278	118.011
No Positive Effect	Neutral	113.946	114.474	112.929	105.815

Table 2.0: Analysis Base Table

Figure 1.0: Transcripts Word Cloud

Exploratory Data Analysis (EDA)

Post creation of the analysis base table, the data were explored using statistical techniques and various insights were generated. The individual distribution of the variables and the correlations were studied.

The individual distributions plot indicates the distribution of the variables being right-skewed for earnings per share, stock price, positive and negative sentiments. The neutral sentiments are mostly left-skewed.

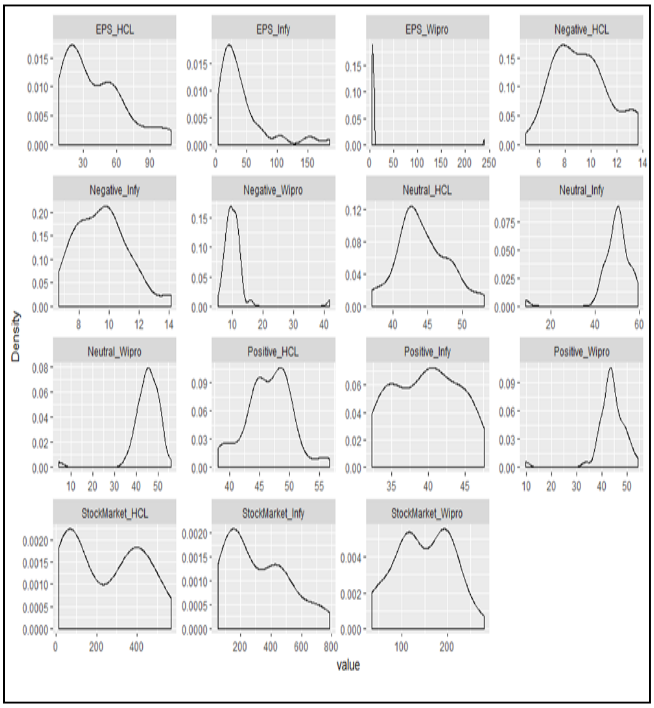

Figure 2.0: Individual Distribution Analysis

Correlation-Infosys

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-b	.602	.084	5.000	.000
N of Valid Cases		48			

Correlation-Wipro

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-b	.793	.067	10.132	0.000
N of Valid Cases		56			

Correlation-HCL

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-b	.378	.096	2.283	.022
N of Valid Cases		37			

Table 3.0: Kendall’s Tau Test Table

This study initially assumed a possible correlation between the polarity and the ‘Inverse Effect’. The correlations study reveals a possible association between sentiments and stock price performance. Kendall’s tau statistic was employed to study correlations between the variables. Kendall’s tau is a non-parametric test similar to Pearson’s product-moment and Spearman’s rank correlation techniques. Kendall’s tau and Spearman’s rank correlations are more suited in this case as the underlying distributions are skewed as depicted in figure3.0. Kendall’s tau takes the value between 0 and, where 0 denotes no relationship and the value close 1 indicates a perfect relationship. Similar results can be obtained using Spearman’s rank correlations as well.

Kendall’s tau indicates a strong correlation between polarity and the ‘Inverse Effect’. Polarity is a metric derived from the sentiments extracted from the quarterly earnings call transcripts. The null hypothesis of the test is two categorical variables being independent. Since the p-values in all the cases are less than the significance level, the null hypothesis is rejected and the conclusion is that there exists an association between polarity and the ‘Inverse Effect’. This insight was used as the basis for exploring various machine learning classification models for predicting the ‘Inverse Effect’. Kendall’s tau b is used as the contingency table formed by the two categorical variables is a square matrix.

A decision tree model with tenfold cross validation was run to understand the key drivers of ‘Inverse Effect’ organization wise. The organization wise tree diagrams are presented below.

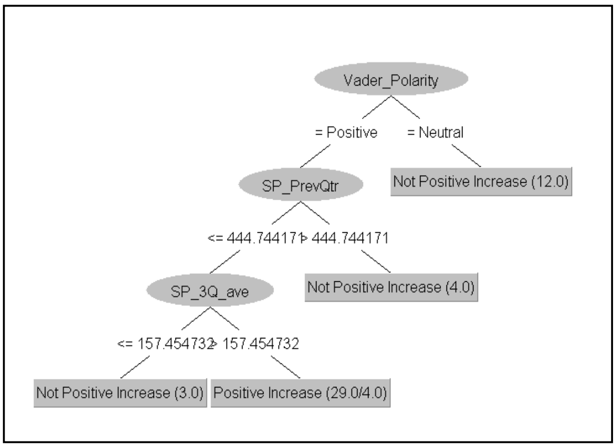

Figure 3.0: Decision Tree Diagram-Infosys

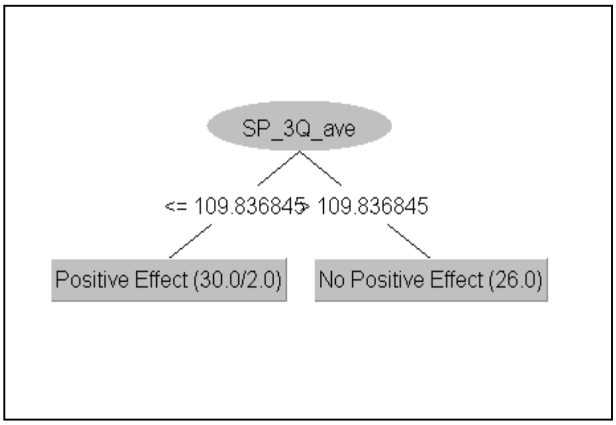

Figure 4.0: Decision Tree Diagram-Wipro

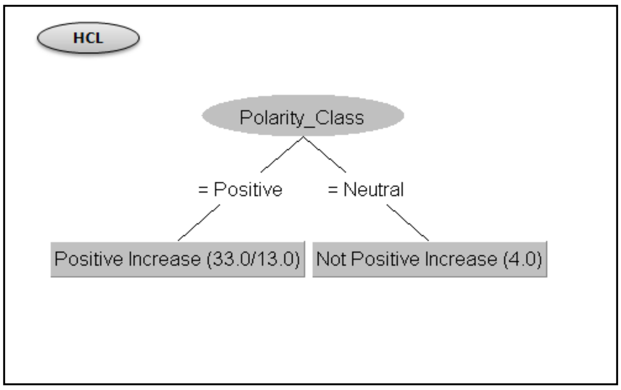

Figure 5.0: Decision Tree Diagram-Wipro

Figures 3.0, 4.0 and 5.0 represent the decision trees for Infosys, Wipro and HCL respectively. In the case of Infosys, the nodal split has happened with polarity, previous quarter stock price and the previous three-quarter average stock price respectively. Hence these variables can be considered significant. In the case of Wipro, the nodal split has happened with a variable which is the previous three-quarter average stock price; hence it is the significant factor explaining the 'Inverse Effect'. Similarly, with HCL, the split has taken place with the variable polarity. Hence the 'Inverse Effect' is largely influenced by the polarity. In general is can be concluded that sentiments have some bearing on the 'Inverse Effect'.

The decision rules generated from the decision tree diagrams 3.0, 4.0 and 5.0 are summarized in the form of a table presented in table 4.0.

Wipro	Infosys	HCL
1.Inverse Effect is solely driven by the previous 3 quarter average stock price 2. Whenever the previous 3 quarter average stock price is less than equal to 109.83INR, there is an 80% chance of observing positive increase in the market in the subsequent period 3. Whenever the previous 3 quarter average stock price is greater than 109.83INR, there is a 59% chance of not observing a positive increase in the next period 4. Good to invest when the previous 3 quarters average is less than or equal to 109.83INR	1.Polarity, previous quarter stock price and previous 3 quarters average stock price are the key drivers of inverse effect 2.There is a 78% chance of observing a positive increase in the market whenever the current quarter polarity is positive and the previous quarter end price is <=INR444.744 and the previous 3 quarters average is <=157.45 3.In all the other cases, market will not react positively 4.Good to invest whenever condition(2) is satisfied	1.Inverse Effect is solely driven by the polarity 2.There is 60% chance that the market would positively increase in the next period whenever the polarity of the current period is positive 3.Whenever the current quarter polarity is positive, there is almost certainty that the market will not increase in the positive direction 4.Good to invest when the current quarter polarity is positive

Table 4.0: The Decsion Rules for three stocks

Model Building

#	Organization	Algorithm	Sampling Method	TPR	FPR	PLR	Precision	Recall	F1-Measure	AUC	Coverage	Rank
1	Wipro	J48_Pruned Tree	Cross Validation	0.964	0.036	26.78	0.967	0.964	0.964	0.94	0.9642	1
2	Wipro	Random Forest	Cross Validation	0.964	0.036	26.78	0.967	0.964	0.964	1	0.9821	1
3	Wipro	SGD	Cross Validation	0.911	0.089	10.24	0.924	0.911	0.91	0.911	0.9107	3
4	Wipro	SVM(Polydot)	Cross Validation	0.946	0.054	17.52	0.952	0.946	0.946	0.946	0.9464	2
5	Wipro	Binary Logistics	Cross Validation	0.946	0.054	17.52	0.952	0.946	0.946	0.909	0.9464	2
6	Wipro	Nnet(MLP)	Cross Validation	0.946	0.054	17.52	0.952	0.946	0.946	0.898	0.9464	2
7	Wipro	Naïve Bayes	Cross Validation	0.821	0.179	4.59	0.868	0.821	0.816	0.923	0.9464	4
8	Wipro	KNN	Cross Validation	0.946	0.054	17.52	0.952	0.946	0.946	0.935	0.9464	2
1	Infosys	J48_Pruned Tree	Cross Validation	0.833	0.174	4.79	0.842	0.833	0.832	0.877	0.9583	4
2	Infosys	Random Forest	Cross Validation	0.896	0.11	8.15	0.902	0.896	0.895	0.935	0.9583	1
3	Infosys	SGD	Cross Validation	0.813	0.193	4.21	0.817	0.813	0.811	0.81	0.8125	5
4	Infosys	SVM(Polydot)	Cross Validation	0.75	0.251	2.99	0.75	0.75	0.75	0.75	0.75	7
5	Infosys	Binary Logistics	Cross Validation	0.854	0.148	5.77	0.855	0.854	0.854	0.864	0.9375	3
6	Infosys	Nnet(MLP)	Cross Validation	0.875	0.136	6.43	0.899	0.875	0.872	0.84	0.9583	2
7	Infosys	Naïve Bayes	Cross Validation	0.792	0.216	3.67	0.799	0.792	0.789	0.758	0.8333	6
8	Infosys	KNN	Cross Validation	0.896	0.11	8.15	0.902	0.896	0.895	0.867	0.8958	1
1	HCL	J48_Pruned Tree	Cross Validation	0.649	0.413	1.571428571	0.787	0.649	0.583	0.521	1	2
2	HCL	Random Forest	Cross Validation	0.378	0.643	0.587869362	0.366	0.378	0.37	0.337	0.9459	6
3	HCL	SGD	Cross Validation	0.649	0.413	1.571428571	0.787	0.649	0.583	0.618	0.6486	2
4	HCL	SVM(Polydot)	Cross Validation	0.649	0.413	1.571428571	0.787	0.649	0.583	0.618	0.6486	2
5	HCL	Binary Logistics	Cross Validation	0.622	0.392	1.586734694	0.62	0.622	0.619	0.653	0.9729	1
6	HCL	Nnet(MLP)	Cross Validation	0.595	0.433	1.374133949	0.592	0.595	0.582	0.576	1	3
7	HCL	Naïve Bayes	Cross Validation	0.568	0.482	1.178423237	0.569	0.568	0.521	0.356	1	4
8	HCL	KNN	Cross Validation	0.378	0.625	0.6048	0.381	0.378	0.379	0.368	0.3783	5

Table 5.0: Classification models leaderboard

The modelling process involved creating a functional form of the model as given below. The response variable is the 'Inverse Effect' and the independent variables being the features extracted from stock price abbreviated at "SP" lagged by different periods.

InverseEffect=f(Polarity,PrevQtrSP,Prev3QtrAvgSP,Prev6QtrAvgSP,Prev9QtrAvgSP)

Several models were explored with a popular tenfold cross-validation technique. Tenfold cross-validation is a generalization of a k-fold cross validation wherein the whole data is divided into a non-overlapping set called folds using stratified random sampling. Each fold contains approximately 10% of the observations. A model is trained on the (k-1) folds barring the one which will be used as a holdout fold for testing. Each time a different fold is selected randomly as a holdout. This experiment will be repeated ten times and the final output is the aggregation of all the ten folds. A leaderboard is presented inTable5.0. Several models with both restrictive and flexible were explored using weka software. Different metrics like True Positive Rate (TPR), False Positive Rate (FPR), Positive Likelihood Rate (PLR), Precision, Recall, F1-measure, Area Under the Curve (AUC) and Coverage were measured to assess the model performance done on the holdout data.

Leaderboard presented in table 5.0 depicts various models with the corresponding performance measures. Each model was replicated thrice across three companies. In total eight models were developed for each organization. Based on the positive likelihood ratio and area under the curve, the models were ranked. It can be observed that in the case of Wipro and Infosys flexible models like random forest and n-net seems to be picking the patterns in the data well were as restrictive models like binary logistics seems to be working well in the case of HCL. Further, the positive likelihood ratio, coverage and the area under the curve associated with these models are high indicating model adequacy. The positive likelihood ratio is the ratio of true positive rate to a false positive rate ranging from zero to infinity, higher values are desirable. Coverage is the percentage of true cases in the test data falling within a 95% confidence level and higher the value better the coverage is. The area under the curve (AUC) is a measure of discriminatory power of the model in separating the classes. AUC takes the value between 0 and 1. A value greater than or equal to0.80 is desirable.

The models thus developed can be used to predict the 'Inverse Effect' for the subsequent periods based on the sentiments or polarity extracted from the quarterly meeting transcripts.

Company	Inverse Effect	%Change in Next Quarter Stock Price
Wipro	Positive Increase	NA
Wipro	Not Positive Increase	0.085
Infosys	Positive Increase	0.125
Infosys	Not Positive Increase	0.103
HCL	Positive Increase	0.208
HCL	Not Positive Increase	0.09

Table 6.0: ‘Inverse Effect’ a lead indicator of market performance

Table 6.0 illustrates the hypothesis whether or not the ‘Inverse Effect’ a lead indicator of market performance. In the case of Wipro, it can be observed that whenever there is no positive increase in the ‘Inverse Effect’, there is an average positive growth of 8.5% in the stock price. This is due to the fact that stock prices are largely influenced by negative sentiments in the case of Wipro. Similarly with Infosys, ‘Inverse Effect’ being a positive increase, an average increase of 12.5% in stock prices can be observed. With HCL, there is a average increase of 20.8% in the stock price, with a positive increase in the ‘Inverse Effect’.

Conclusion

Stock markets being volatile, robust techniques are required to capture volatility and thus make predictions for investment decision making. There are umpteen ways of modelling volatility. Models like ARIMA, regression-based forecasting,

ARCH-GARCH, neural net models and other machine learning models come handy in modeling time series data. However, they do not factor endogenous variables. The concept of ‘Inverse Effect’ is a novel approach in itself wherein one can observe the effect of ‘Inverse Effect’ on the stock market performance in the subsequent quarters. From a trading point of view, one can decide how much will be gain or loss in the short term using this model. Based on the data, in the case of Wipro, whenever the ‘Inverse Effect’ is not positive, one can expect an average change in the stock price of 8.5%, this is because negative sentiments have a bearing on the stock price. In the case of Infosys, a positive increase in the ‘Inverse Effect’ lead to an average 12.5% change in stock price and a 10.3% change for a not positive increase. Similarly for HCL, one case expects an average change of 20.8% for a positive increase in the ‘Inverse Effect’ and 9% otherwise.

The limitation of this study is that it is confined only to software companies The future scope of this study can be extended to various sectors and different events, news articles, analysts reports can be considered to model random behaviour of the stock market.

References

[18] Boskou, G., Kirkos, E., & Spathis, C. (2018). Assessing internal audit with text mining. *Journal of Information and Knowledge Management*, *17*(2), 1–22. https://doi.org/10.1142/S021964921850020X

[19] Chakrabarti, S., Trehan, D., & Makhija, M. (2018). Assessment of service quality using text mining – evidence from private sector banks in India. *International Journal of Bank Marketing*, *36*(4), 594–615. https://doi.org/10.1108/IJBM-04-2017-0070

[20] Davis, A. K., & Tama-Sweet, I. (2012). Managers' Use of Language Across Alternative Disclosure Outlets: Earnings Press Releases Versus MD&A. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1866369

[21] Demers, E., & Vega, C. (2008). Soft Information in Earnings Announcements : News or Noise? *International Finance Discussion Paper*, *2008*(951), 1–56. https://doi.org/10.17016/ifdp.2008.951

[22] Elagamy, M. N., Stanier, C., & Sharp, B. (2018). Stock market random forest-text mining system mining critical indicators of stock market movements. *2nd International Conference on Natural Language and Speech Processing, ICNLSP 2018*, 1–8. https://doi.org/10.1109/ICNLSP.2018.8374370

[23] Han, H., Wang, Q., & Chen, C. (2019). Policy Text Analysis Based on Text Mining and Fuzzy Cognitive Map. *Proceedings - 2019 15th International Conference on Computational Intelligence and Security, CIS 2019*, 142–146. https://doi.org/10.1109/CIS.2019.00038

[24] Kang, T., Park, D. H., & Han, I. (2018). Beyond the numbers: The effect of 10-K tone on firms' performance predictions using text analytics. *Telematics and Informatics*, *35*(2), 370–381. https://doi.org/10.1016/j.tele.2017.12.014

[25] Kumar, B. S., & Ravi, V. (2016). A survey of the applications of text mining in financial domain. *Knowledge-Based Systems*, *114*, 128–147. https://doi.org/10.1016/j.knosys.2016.10.003

[26] Li, X., Wu, P., & Wang, W. (2020). Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong. *Information Processing and Management*, *July 2019*, 102212. https://doi.org/10.1016/j.ipm.2020.102212

[27] Nam, K. H., & Seong, N. Y. (2019). Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market. *Decision Support Systems*, *117*, 100–112. https://doi.org/10.1016/j.dss.2018.11.004

[28] Nourbakhsh, A., & Bang, G. (2019). *A framework for anomaly detection using language modeling, and its applications to finance*. http://arxiv.org/abs/1908.09156

[29] Onsumran, C., Thammaboosadee, S., & Kiattisin, S. (2015). Gold Price Volatility Prediction by Text Mining in Economic Indicators News. *Journal of Advances in Information Technology*, *6*(4), 243–247. https://doi.org/10.12720/jait.6.4.243-247

[30] Shi, Y., Tang, Y. R., Cui, L. X., & Long, W. (2018). A text mining based study of investor sentiment and its influence on stock returns. *Economic Computation and Economic Cybernetics Studies and Research*, *52*(1), 183–199. https://doi.org/10.24818/18423264/52.1.18.11

[31] Wang, Y., Li, B., Li, G., Zhu, X., & Li, J. (2019). Risk factors identification and evolution analysis from textual risk disclosures for insurance industry. *Procedia Computer Science*, *162*(Itqm 2019), 25–32. https://doi.org/10.1016/j.procs.2019.11.253

[32] Yin, L., Zhang, N., He, L., & Fang, W. (2018). A Study of Relationship between Investor Sentiment and Stock Price Based on Text Mining. *Proceedings - 2016 International Conference on Identification, Information and Knowledge in the Internet of Things, IIKI 2016*, *2018-Janua*, 536–539. https://doi.org/10.1109/IIKI.2016.49

[33] Yu, S., & Guo, S. (2016). Big data concepts, theories, and applications. In *Big Data Concepts, Theories, and Applications* (Issue December). https://doi.org/10.1007/978-3-319-27763-9

From: SAGE peer review [mailto:Noreply.Peerreview@sagepub.com]
Sent: 08 September 2020 11:53
To: Dr. Shinu Abhi <shinuabhi@reva.edu.in>
Subject: GBR-2020-1491: Your manuscript submitted to Global Business Review

Please be cautious
This email was sent outside of your organisation

08-Sep-2020

Dear Author(s),

Thank you for submitting manuscript GBR-2020-1491 entitled INVERSE EFFECT' MODELING OF EARNINGS CALL TRANSCRIPTS FOR FINANCIAL ANALYSIS for Global Business Review.

Global Business Review is a refereed journal and we publish only empirical /research based papers. The paper will be scrutinized by the Editor and if finds satisfactory, it will be sent for blind review by experts in the concerned area and the decision will then be sent to the author. This process may take 5 -6 months' time.

In case you have any questions, please feel free to contact us.

You may view the status of your manuscript anytime by accessing your author desk at the https://peerreview.sagepub.com/.

Thank you again for submitting your manuscript to Global Business Review.

Sincerely,
Global Business Review Editorial Office