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

Nagendra B.V

Under the guidance of
Dr. Jay Bharateesh Simha

REVA UNIVERSITY

WHY THE STUDY

- The stock market is always driven by new information emerging from the listed companies. Any positive or negative information emanating from the listed company will have significant impact on the performance of the stock.
- It becomes imperative for the top leadership to perform consistently and assure the stake holders of the sustained growth.
- Many a times what is spoken by the leadership team has an impact on the stock performance as market participants will be all eager to analyze every information emerging from the company.
- Hence, this project focuses on one such area called “Earning Call Transcripts” to study the effect of these quarterly meetings on the market performance.

REVA UNIVERSITY

ONTOLOGICAL PREMISE

Sl. No	Authors	Findings
1	Nam & Seong, 2019	There has been a surge in text data in social media and extracting insights from them is paramount to organizations in decision making
2	Boskou et al., 2018	financial reports available annually with text mining can be useful to internal and external stakeholders in decision making.
3	Chakrabarti et al., 2018	Private sector banks in India are embarking on service quality framework with text mining techniques to assess the customer service quality and their perceptions.
4	Kang et al., 2018	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.
5	Wang et al., 2019	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.
6	Han et al., 2019	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.

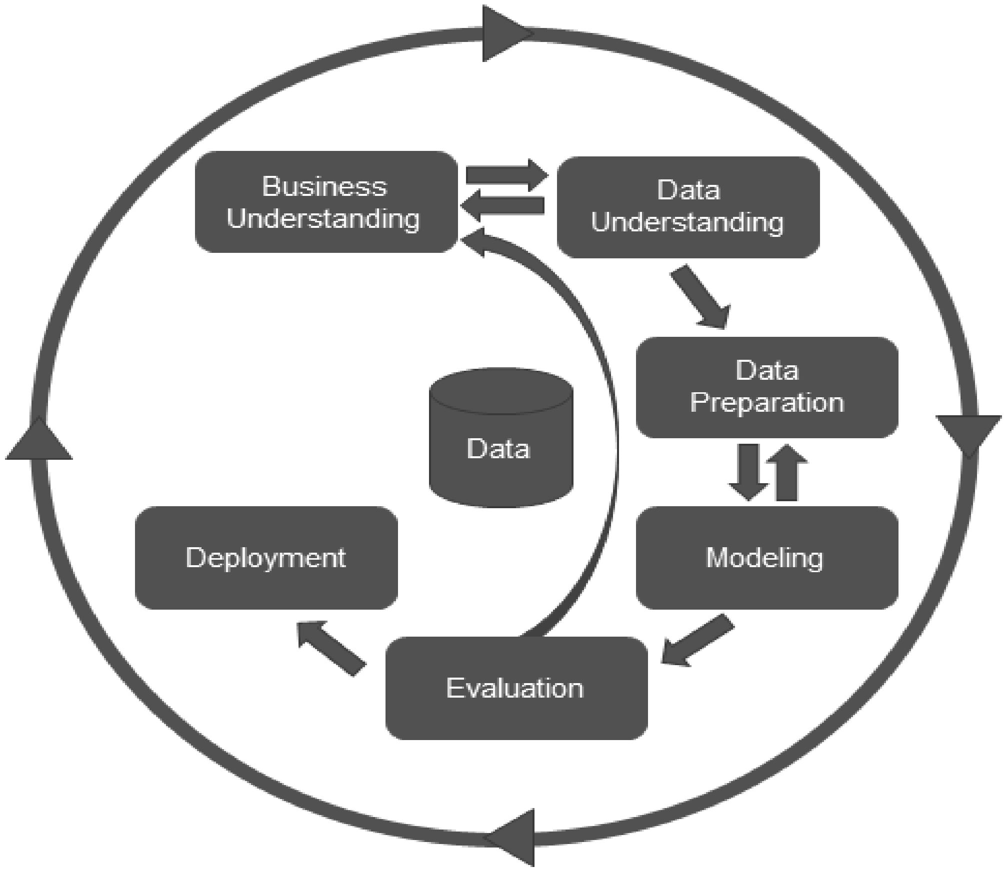
APPROACH

- This study uses both unstructured corpus and the fundamental data for the purposes of analysis and model building
- We use unstructured data available in the form of transcripts and the fundamental data like historical stock prices and the Earnings Per Share (EPS) of three leading Indian software companies for the period 2008 till date
- The unstructured corpus was converted into numeric score using lexicon based modeling and then integrated with the fundamental data. Hence the approach can be termed hybrid

BUSINESS UNDERSTANDING

- Quarterly financial performance reports help investors judge pulse of organizations
- Investors get ample insights into growth and performance of organizations by comparing reports on a quarterly basis
- 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.
- Earnings Call Transcripts are accessible to investors and general public in text and audio formats
- From an investor point of view it is paramount to know whether what is being 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 using CRISP DM methodology.

CRISP DM METHODOLOGY



BUSINESS UNDERSTANDING-A SIPOC VIEW

SUPPLIER	INPUT	PROCESS	OUTPUT	CUSTOMER
Customers CXOs Investors Projects Support functions	Balance sheet Unaudited financials KPI s	Discuss performance Discuss growth plans Discuss strategies	Action plan for future Recorded document of discussions	Self Investors General Public

SPECIFIC OBJECTIVES AND SCOPE

Objectives

1. Lexicon based sentiment analysis of the Earnings Call Transcripts (ECT)
2. To analyze the impact of quarterly review meetings on the subsequent period market performances
3. To identify the key drivers of the stock market performance
4. To develop various classification models to predict the Inverse Effect

Scope

- The scope of this project is limited to top 3 Indian service majors Wipro, Infosys and HCL only

DATA UNDERSTANDING AND DATA PREPARATION

The data preparation process involved the following :

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

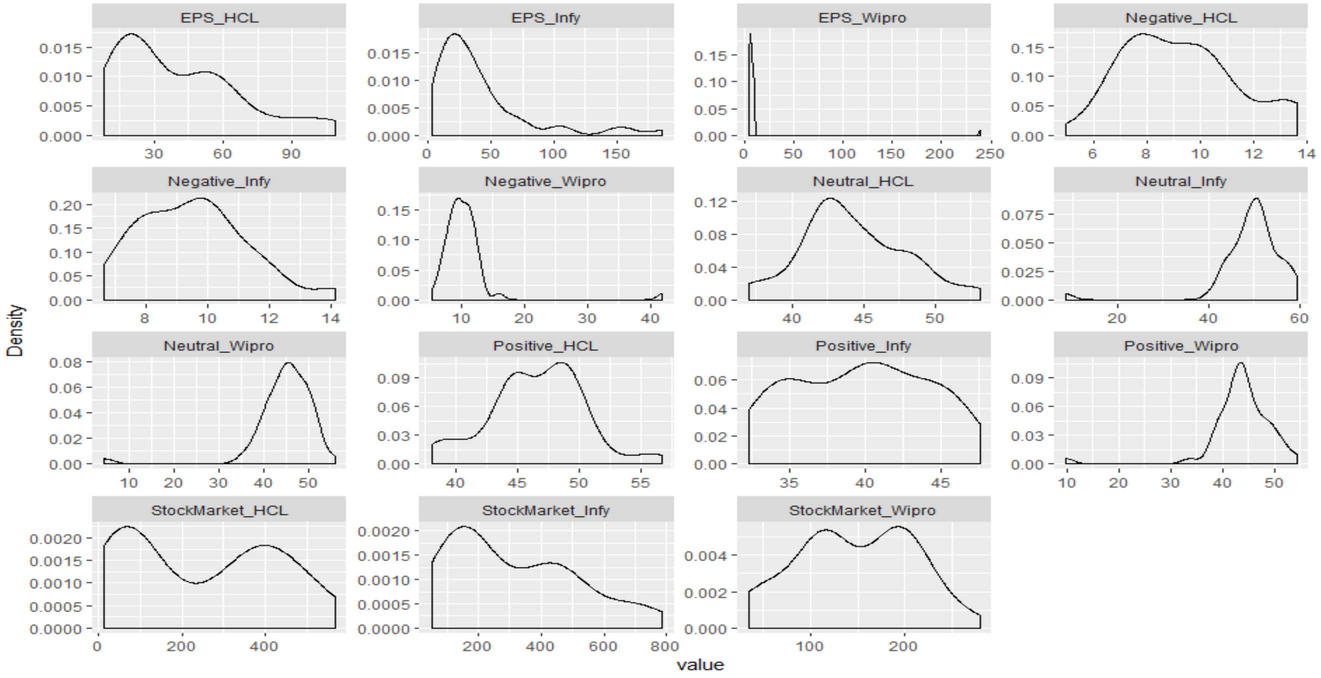
Inverse Effect Definition: New features were extracted from the existing data which comprised of a response variable called the "Inverse Effect" and a set of features made of polarity and stock price. Inverse Effect is simply the discordance between the current quarter sentiment and the next period stock 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 "Positive Increase" else it is "No Positive Increase".

DATA UNDERSTANDING AND DATA PREPARATION

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	Extraneous	2008-2020	Quarterly	Yahoo Finance
Polarity	Continuous	Endogenous	2008-2021	Quarterly	Transcripts available in ECT

Reporting Period	Inverse Effect	Polarity	SP_3Q_ave	SP_6Q_ave	SP_9Q_ave	SP_PrevQtr
ECT_Q2_10	Positive Increase	Positive	110.9897307	95.90060733	84.07563056	145.08493
ECT_Q3_10	Not Positive Increase	Positive	120.91156	113.4964968	93.59980056	158.574387
ECT_Q4_10	Positive Increase	Positive	162.4988453	130.924207	107.5098724	183.837219
ECT_Q1_11	Not Positive Increase	Positive	157.3186037	134.1541672	116.3732728	129.544205
ECT_Q2_11	Positive Increase	Positive	159.41805	140.164805	128.8036812	164.872726
ECT_Q3_11	Positive Increase	Positive	145.9701537	154.2344995	135.9395226	143.49353
ECT_Q4_11	Not Positive Increase	Positive	157.454732	157.3866678	141.9210221	163.99794
ECT_Q1_12	Positive Increase	Positive	160.190567	159.8043085	146.840059	173.080231
ECT_Q2_12	Not Positive Increase	Neutral	165.936203	155.9531783	158.1350673	160.730438
ECT_Q3_12	Positive Increase	Positive	165.3969013	161.4258167	160.0567457	162.380035
ECT_Q4_12	Positive Increase	Positive	159.8529407	160.0217538	159.8205192	156.448349
ECT_Q1_13	Positive Increase	Positive	162.8907367	164.4134698	158.2656978	169.843826
ECT_Q2_13	Positive Increase	Positive	170.647471	168.0221862	164.4997014	185.650238

EDA-INDIVIDUAL DISTRIBUTIONS



Earnings Per Share are mostly right skewed

Stock Prices exhibit bimodal effect

Sentiments are both right and left skewed

EDA- EFFECT OF SENTIMENTS ON MARKET PERFORMANCE

Correlations-Wipro						
			Positive_Wipro	Negative_Wipro	Neutral_Wipro	StockMarket_Wipro
Kendall's tau_b	StockMarket_Wipro	Correlation Coefficient	-.081	**.462****	-.127	1.000
		Sig. (2-tailed)	.419	.000	.204	
		N	48	48	48	48

Correlations-HCL						
			Positive_HCL	Negative_HCL	Neutral_HCL	StockMarket_HCL
Kendall's tau_b	StockMarket_HCL	Correlation Coefficient	.163	.173	**-.250***	1.000
		Sig. (2-tailed)	.107	.086	.013	
		N	47	47	47	47

Correlations-Infy						
			Positive_Infy	Negative_Infy	Neutral_Infy	StockMarket_Infy
	StockMarket_Infy	Correlation Coefficient	.474**	-.168	-.337**	1.000
		Sig. (2-tailed)	.000	.095	.001	
		N	47	47	47	47

Kendall's tau indicates a strong correlation between sentiments and market performance in the case of Wipro and Infosys.

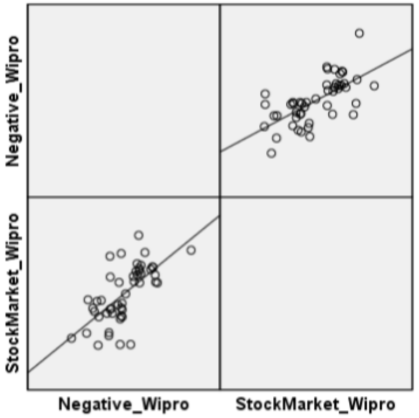
EXPLORATORY DATA ANALYSIS continued…

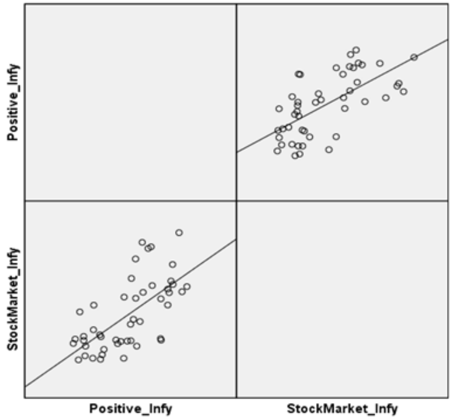
Correlation-Infosys					
		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-b	.602	.084	5.000	.000
N of Valid Cases		48			

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			

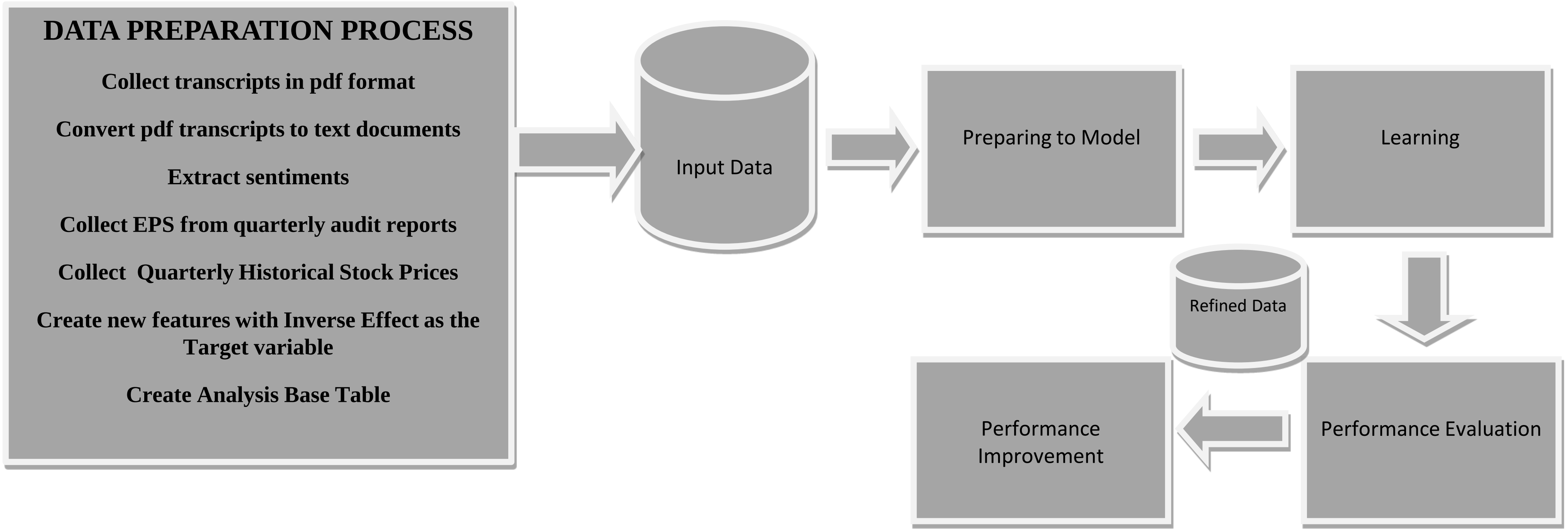
Correlation-HCL					
		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-b	.378	.096	2.283	.022
N of Valid Cases		37			

The significance of the tests indicate possible correlations between sentiments and market performance





MODELING APPROACH

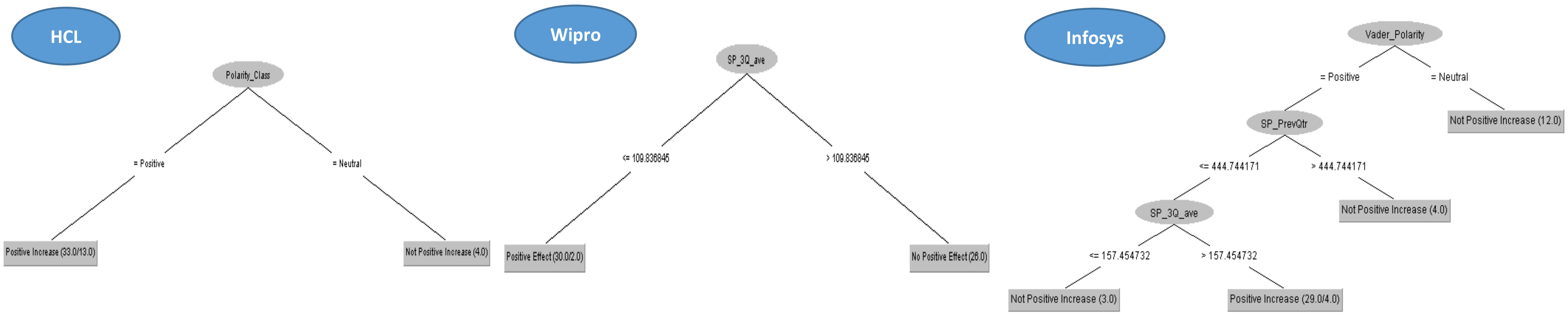


MODELING APPROACH

A novel idea of direction detection of the market performance was conceived with the help of existing features. The original idea that a positive sentiment will have a positive impact on the market performance was further explored and a new response variable was derived , which is termed as “**Inverse Effect**”. Inverse Effect comprise of two levels namely “Positive Increase” and “Not Positive Increase” in the market performance. The final mathematical model looks as below:

$$\text{Inverse Effect} = f(\text{Polarity}, \text{Stock Price\_Prev Qtr}, \text{Stock Price\_3Qtr\_Avg}, \text{Stock Price\_6Qtr\_Avg}, \text{Stock Price\_9Qtr\_Avg})$$

KEY DRIVERS OF INVERSE EFFECT BASED ON DECISION TREE MODEL



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

MODELS EVALUATION

#	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

MODEL RECOMMENDATIONS

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

ANALYSIS INSIGHTS

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 or 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

IS INVERSE EFFECT A LEAD INDICATOR OF MARKET PERFORMANCE?

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

DATA SOURCES FOR THE 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/

LITERATURE REFERENCES

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

PUBLICATION POTENTIAL AND FUTURE SCOPE

Future scope of this study include:

1. Sector wise analysis
2. Nifty 50 companies
3. Augmentation of analyst reviews and social media reviews

This paper has been submitted to “Global Business Review” A SCOPUS Indexed journal for publication

REVA UNIVERSITY

SAGE track

Select journal: Global Business Review

Global Business Review (GBR)

Home Author desk Help

Return to home

Complete submission draft (1)

Manuscripts under review (1)

Manuscripts under Review

Show 10 entries

Filter:

Manuscript ID	Manuscript Title	Date Created	Date Submitted	Status
GBR-2020-1491	INVERSE EFFECT' MODELING OF EARNINGS CALL TRANSCRIPTS FOR FINANCIAL ANALYSIS	08-Sep-2020	08-Sep-2020	(ADM) Awaiting administrative check-in ; Leena Prakasan

Showing 1 to 1 of 1 entries

Previous 1 Next

THANK YOU