For the exclusive use of V. Rudrappa, 2020.

भारतीय प्रबंध संस्थान बेंगलूर
INDIAN INSTITUTE OF MANAGEMENT BANGALORE

IMB 577

PREDICTING EARNINGS MANIPULATION BY INDIAN FIRMS USING MACHINE LEARNING ALGORITHMS

U DINESH KUMAR, TOUSIF AHMED INAYATH SYED AND SURESH GANESHAN

U Dinesh Kumar, Professor of DSIS, Tousif Ahmed Inayath Syed and Suresh Ganeshan, prepared this case for class discussion. This case is not intended to serve as an endorsement, source of primary data, or to show effective or inefficient handling of decision or business processes.

Copyright © 2016 by the Indian Institute of Management Bangalore. No part of the publication may be reproduced or transmitted in any form or by any means – electronic, mechanical, photocopying, recording, or otherwise (including internet) – without the permission of Indian Institute of Management Bangalore.

This document is authorized for use only by Vindhya Rudrappa in IDS 572 Spring 2020 taught by NEGAR SOHEILI AZAD, University of Illinois at Chicago from Jan 2020 to Jul 2020.

For the exclusive use of V. Rudrappa, 2020.

Predicting Earnings Manipulation by Indian Firms using Machine Learning Algorithms

I worked in accounting for two and half years, realized that was not what I wanted to do with rest of my life and decided I was just going to give comedy a try.

- Bob Newhart

August 8, 2016, Sachin Kumar, Director at MCA Technology Solutions was preparing for the meeting that he had to attend the following week with one of his clients, a private bank that provided commercial loans to small and medium enterprises (SMEs) in India. The value of the loan provided by the bank ranged from INR 10 million to 500 million ($1 = INR 66.82, August 16, 2016). MCA Technology Solutions was assisting the bank with auditing services. In addition to auditing services, the bank had enquired if Sachin could help in checking whether any of the bank's customers were involved in earning manipulation. The bank collected financial statements from its customers to predict credit worthiness of the customers. The bank also wanted to ensure that the figures in the financial statements were genuine and not manipulated. Being an analytics company that specialized in auditing services, MCA Technology Solutions received several such requests from many of its customers in the banking sector to assist them with predicting earnings manipulations.

In India, earning manipulations by corporates has been a major concern for banks. In 2009, the chairman of Satyam Computers, Ramalinga Raju, confessed that the company's accounts had been manipulated by USD 1.47 billion.[1] A study by Securities and Exchange Board of India (SEBI) in 2013 revealed that small firms in India involved relatively more in earning manipulations than medium and large firms.[2] According to the report, the average earning manipulation in India was 2.9% of the total assets, whereas the earning manipulation for small firms was 10.6%. During the same period, Indian media also started to report high percentages of bad loans or non-performing assets (NPAs) among the public sector banks. As on March 31, 2015, the total NPAs of public sector banks was approximately USD 60 billion dollars.[3] Increasing NPAs also forced banks to check for any accounting manipulations by the firms.

Sachin Kumar and his data science team explored different approaches for checking accounting manipulations such as Benford's law and Jones's model.[4] However, Sachin Kumar thought that Beneish model may be more suitable for predicting earnings manipulations. The original model by Beneish was developed in 1999 based on US data, and there is no guarantee that the same model would apply for Indian companies. Sachin said:

[1] Source: Madan Bhasin, Corporate accounting scandal at Satyam: A case study of India's Enron, *European Journal of Business and Social Sciences*, 2013, **1(**12**)**, 25-47.
[2] D Ajit, Sarat Malik and V K Verma, Earnings management in India, SEBI DRG Study, 2013. Accessed on August 1, 2016 at http://www.sebi.gov.in/cms/sebi_data/DRG_Study/EMiM.pdf
[3] Source: The ticking NPA clock: How many crores in bad loans are weighing on Indian banks' balance sheets? Accessed on August 1, 2016 at http://www.firstpost.com/business/the-ticking-npa-clock-what-is-the-cost-of-the-escalating-bad-loans-pile-on-indian-banks-balance-sheets-2669746.html
[4] J J Jones, Earnings management during import relief investigations, *Journal of Accounting Research*, 1991, **20(**2**)**, 193-228.

This document is authorized for use only by Vindhya Rudrappa in IDS 572 Spring 2020 taught by NEGAR SOHEILI AZAD, University of Illinois at Chicago from Jan 2020 to Jul 2020.

For the exclusive use of V. Rudrappa, 2020.

> The numbers of earnings manipulators are usually very low compared to non-manipulators and thus the data is likely to be an unbalanced data. Unbalanced data sets are difficult to analyze using traditional statistical techniques; we may have to use machine learning algorithms to accurately classify manipulators and non-manipulators.

Sachin Kumar and his analytics team at MCA Technology Solutions started to explore a reliable way of identifying earning manipulators and/or accounting fraud by Indian companies using publically available data such as Prowess database maintained by the Centre for Monitoring Indian Economy (CMIE). Prowess provided details of financial performance of Indian companies since 1998-1999.

MCA TECHNOLOGY SOLUTIONS

MCA Technology Solutions Private Limited was established in 2015 in Bangalore with an objective to integrate analytics and technology with business. MCA Technology Solutions helped its clients in areas such as customer intelligence, forecasting, optimization, risk assessment, web analytics and text mining, and cloud solutions. Risk assessment vertical at MCA technology solutions focused on problems such as fraud detection and credit scoring. Within a short period of its existence, MCA Technology Solutions had built a good clientele; the fraud analytics division of the company had seen considerable growth over the years. The division dealt with detection of internal and external money laundering, loan frauds, customer transaction frauds, online banking frauds, etc.

Predicting fraud is highly data intensive and for this reason, MCA Technology Solutions employed a dedicated team of data scientists to collect relevant data and develop analytics models that could be used for predicting various types of fraud.

EARNINGS MANAGEMENT, EARNINGS MANIPULATION, AND FINANCIAL REPORTING FRAUD

Earnings management, earnings manipulation, and fraudulent reporting have been frequently used in accounting literature and the differences between these terms are very slim (see **Exhibit 1**). Accounting literature differentiates earnings management from financial reporting fraud. According to Healy and Wahlen (1999),[5] earnings management is defined as follows:

> *Earnings management occurs when managers use their judgment in financial reporting and in structuring transactions to alter financial reports to either mislead stakeholders about the underlying economic performance of the company or otherwise to influence contractual outcomes that depend on reported accounting numbers.*

[5] Source: P M Healy and J M Wahlen, A review of the earnings management literature and its implications for standard setting, *Accounting Horizon*, 1999, **13(4),** 365-383.

This document is authorized for use only by Vindhya Rudrappa in IDS 572 Spring 2020 taught by NEGAR SOHEILI AZAD, University of Illinois at Chicago from Jan 2020 to Jul 2020.

For the exclusive use of V. Rudrappa, 2020.

Earnings manipulations imply that the company undertakes deliberate steps to bring reported earnings to a desired level.[6] Rocco (1998) citing US National Commission on fraudulent financial reporting (1987) defines fraudulent financial reporting as an intentional or reckless conduct, whether by act or omission that leads to materially misleading financial statements[7]. However, fraud could differ from country to country and depends on the regulatory/auditing framework of the country. According to Wells (2009),[8] the following four features are common to any fraud.

1. A material false statement
2. Intent to deceive
3. Reliance on the false statement by the victim
4. Subsequent damages

Many authors have tried to differentiate earnings management/manipulation from fraud (Marai and Pavlovic, 2013).[9] According to Stlowy and Breton (2004),[10] the earnings manipulation that falls outside the law and standards constitutes fraud.

Dechow *et al.* (2010)[11] claimed that most firms manipulated more than one income statement line item based on a study of 680 firms with alleged manipulations in their financial statements. The top three line items used by manipulating companies were revenue, inventory, and cost of goods sold. Dechow *et al.* (2010) also claimed that manipulating firms have significantly high accruals (earnings component that does not generate cash flows), deteriorating cash margins, and deteriorating earnings growth.

EARNINGS MANIPULATIONS AND FRAUD BY INDIAN COMPANIES

In 2015, India had 5,835 listed companies, and studies on earnings management by Indian companies were very limited. Public sector banks performed very poorly during 2012-2015 and the level of NPAs reported by the banks were far from reality and made possible owing to reporting flexibilities. During the same period, NPAs worth INR 1,140 billion were written off by the government.[12] Sachin Kumar called his chief data scientist, Saurabh Rishi, who had joined MCA Technology Solutions recently from one of the e-commerce companies in India. Saurabh was an alumnus of the Indian Institute of Management and had worked on fraud-related projects in the past. Sachin said:

[6] D Ajit, Sarat Malik and V K Verma, Earnings management in India, SEBI DRG Study, 2013. Accessed on August 1, 2016 at http://www.sebi.gov.in/cms/sebi_data/DRG_Study/EMiM.pdf
[7] Source: R V Rocco, Fraud Auditing, *Managerial Auditing Journal*, 1998, **13(1),** 4-71.
[8] Source: J T Wells, Fraud – The occupational hazards, *Accountancy Age*, 2009, available at https://www.accountancyage.com/aa/feature/1748964/fraud-occupational-hazards

[9] Source: A Marai and V Pavlovic, Earnings management vs financial reporting fraud – key features for distinguishing, *Economics and Organization*, 2013, **10(1),** 39-47.
[10] H Stowley and G Breton, Accounts manipulation: a literature review and proposed conceptual framework, *Review of Accounting and Finance*, 2004, **3(1),** 5-92.
[11] Source: P M Dechow, W Ge, C R Larson, and R G Sloan, Predicting material accounting material manipulations, 2010, working paper available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=997483
[12] Source: *Indian Express*, February 9, 2016.

This document is authorized for use only by Vindhya Rudrappa in IDS 572 Spring 2020 taught by NEGAR SOHEILI AZAD, University of Illinois at Chicago from Jan 2020 to Jul 2020.

For the exclusive use of V. Rudrappa, 2020.

> Saurabh, can we predict the earnings manipulations or accounting fraud accurately and if yes, what data do we need to collect? One of our important clients has asked us whether we can assist them to predict earnings manipulations, and we have to respond to them quickly.

Saurabh, who was sipping his South Indian filter coffee said:

> I have heard about Beneish model that was designed to predict accounting fraud. But it was developed in 1999 using US data. I think that could be our starting point, we need to collect financial statement of the companies. I guess Prowess database will have the data that we are looking for. I am sure that we will end up with an unbalanced data set since the number of manipulators will be very small. So, we may have to use machine learning algorithms to improve the classification accuracy.

Sachin responded:

> Great Saurabh, can you give me an update in a couple of days? I have a meeting with the client next week and I would like to confirm with them whether we can undertake the project.

Both Sachin and Saurabh realized that classification accuracy is important since classifying a manipulator as non-manipulator may encourage the companies to repeat the practice, whereas misclassifying a non-manipulator as manipulator may involve wasting the resources of regulators.

DATA COLLECTION

Accounting literature has sufficient points about how companies manipulate the earnings. Saurabh examined the Beneish model[13] carefully. Beneish had used eight financial indices (**Exhibit 2**) to predict earnings manipulations by companies by treating it as a classification problem. Also, Beneish used Probit model to predict earnings manipulation, which is a supervised learning algorithm. It means that one needs to know the manipulators and non-manipulators in the training dataset. Saurabh's first task was to identify the earnings manipulators among Indian companies.

Saurabh started collecting information related to earnings manipulators from SEBI reports during 2005 to 2015 and the Lexis Nexis database. In 2016, 5,622 companies were listed in Bombay Stock Exchange (BSE) and National Stock Exchange (NSE) of India. A total of 39 earnings manipulators were identified through SEBI reports and Lexis Nexis database. As expected, the number of manipulators was relatively a smaller proportion of listed companies. Saurabh decided to collect the data related to these 39 companies

[13] M D Beneish, The detection of earnings manipulations, *Financial Analyst Journal*, 1999, **55(5),** 24-36.

This document is authorized for use only by Vindhya Rudrappa in IDS 572 Spring 2020 taught by NEGAR SOHEILI AZAD, University of Illinois at Chicago from Jan 2020 to Jul 2020.

For the exclusive use of V. Rudrappa, 2020.

along with 1,200 non-manipulators based on sales volume from the Prowess database. The number of manipulators was about 3.14% of the data collected, resulting in a highly unbalanced data set.

MODEL DEVELOPMENT

Armed with data, Saurabh developed several models using techniques such as logistic regression, classification trees and machine learning algorithms. Saurabh started sharing his results with Sachin Kumar, who was happy with the job performed by Saurabh. They also noticed that the accuracy of machine learning algorithms was much higher than Beneish model. However, they realized the challenge in deploying machine learning algorithms for predicting earnings manipulations. They wanted to develop an easily deployable solution for predicting manipulators.

This document is authorized for use only by Vindhya Rudrappa in IDS 572 Spring 2020 taught by NEGAR SOHEILI AZAD, University of Illinois at Chicago from Jan 2020 to Jul 2020.

For the exclusive use of V. Rudrappa, 2020.

Exhibit 1

Earning Management, Earning Manipulation and Fraud

Basic differences between earnings management, earnings manipulations and fraud are as follows:

Earnings Management	Earnings Manipulations	Earnings Fraud
Accounting techniques (known as creative accounting) used to make a company's financial reports look better constitute earnings management.	When a management takes deliberate steps to bring reported earnings to a desired level, it is called earnings manipulations.[14]	Earnings fraud refers to the earnings manipulation by violating accounting standards and corporate laws, and/or structuring activities in such a way that reduces expected firm value.[15]

1) Earnings Management:

Earnings management is one of the accounting techniques used to make a company's financial reports look better. Earnings management refers to the earnings manipulation through exercising the discretion accorded by accounting standards and corporate laws, and/or structuring activities in such a way that expected firm value is not affected negatively.

Some might interpret earnings management as a nice way of saying "earnings manipulation." It usually happens when management feels the pressure to conceal earnings fluctuations to keep stock prices up. Investors may not fully fathom the impact of certain accounting tactics. Every company wants its financial statements to look as good as they can. The financial statements are what potential investors and creditors look at when they make the decision whether or not to lend the company money or to become an investor. This is where the concept of earnings management comes into play. Earnings management is a creative use of different accounting techniques to make financial statements look better.

Earnings management is the acceleration or deferral of expenses or revenue through operating or accounting practices with the objective to produce consistent growth in earnings. These earnings may not reflect the underlying economics of the enterprise for the time-period.

Basic FAQs on Earnings Management:

- **Who is involved in earnings management?**
  Earnings management is done by management; it is an internal management technique.

[14] D Ajit, Sarat Malik and V K Verma, Earnings management in India, SEBI DRG Study, 2013. Accessed on August 1, 2016 at http://www.sebi.gov.in/cms/sebi_data/DRG_Study/EMiM.pdf
[15] Ibid

This document is authorized for use only by Vindhya Rudrappa in IDS 572 Spring 2020 taught by NEGAR SOHEILI AZAD, University of Illinois at Chicago from Jan 2020 to Jul 2020.

For the exclusive use of V. Rudrappa, 2020.

- **Is there any intervention of employee's in earnings management?**
  As employees cannot take managerial decision; therefore, there is no employee's intervention in earnings management.

- **What is the main purpose of earnings management?**
  Earnings management is basically done for better financial indicators and to depict better financial position.

- **What are the financial and legal implications of earnings management?**
  Financial Implications: Earnings of the company may be inflated without actual revenue being earned and thereby increasing their income.
  Legal Implications: If the company has violated any of the legal requirements as required by any of the statutes, it may have adverse effect on its reporting and company goodwill value may be affected in the market. It will lead to stringent actions by various statutory authorities.
  Note: If earnings management is legally valid or if the management has not violated any of the laws, there will not be any legal liability as such.
  Public companies and **listed companies** are more exposed to legal liability for any stringent violations of the statute.

2) Earnings Manipulation:

Management may manipulate real activities of the business to meet desired targets such as providing discounts to temporarily increase sales, overproducing to report lower cost of goods sold, and reducing discretionary expenses such as R&D, advertising expenditures, sales of profitable assets, etc.[16] Earnings manipulation is linked to earnings management.

Basic FAQs on Earnings Manipulation:

- **Who is involved in earnings manipulations?**
  As discussed earlier, earnings manipulation is also done by management; it is an internal management technique.

- **What is the main purpose of earnings manipulation?**
  Earnings manipulation is done to bring financial position to a predetermined financial figure.

- **Is there any intervention of employees in earnings manipulation?**
  As employees cannot take managerial decision, there is no employee intervention in earnings manipulations.

- **What are the financial and legal implications of earnings manipulation?**

[16] Ibid

This document is authorized for use only by Vindhya Rudrappa in IDS 572 Spring 2020 taught by NEGAR SOHEILI AZAD, University of Illinois at Chicago from Jan 2020 to Jul 2020.

For the exclusive use of V. Rudrappa, 2020.

Implications are the same as discussed above in earnings management.

3) Earnings Fraud:

Earnings management becomes fraud when companies intentionally provide materially misstated information. Earnings fraud is more concerned with violations of rules and regulations which may lead to stringent actions by regulatory authorities.

Basic FAQs on Earnings Fraud:

- **Who is involved in earnings fraud?**
  Earnings fraud is done by the management.

- **What is the main purpose of getting involved in earnings fraud?**
  Earnings fraud may be done for various purposes depending on the company's specific requirement. Earnings fraud may be done for inflating profits or for decreasing the profits; it may also be done for manipulating actual financial figures with fictitious figures.

- **What are the legal implications for companies involved in earnings fraud?**
  Earnings fraud may lead to stringent punishment by various statutes such as Securities and Exchange Board of India (SEBI), Companies Act 2013, FEMA Act 1999, Competition Act 2002, Prevention of Money Laundering Act 2002, Information Technology Act 2000, and many more.

Source: MCA Technology Solutions Pvt. Ltd.

This document is authorized for use only by Vindhya Rudrappa in IDS 572 Spring 2020 taught by NEGAR SOHEILI AZAD, University of Illinois at Chicago from Jan 2020 to Jul 2020.

For the exclusive use of V. Rudrappa, 2020.

Exhibit 2

Financial indices used by Beneish

Sl. No.	Financial Index	Computation Equation
1	Days Sales to Receivables Index (DSRI)	$\text{DSRI} = \frac{\text{Receivables}_t/\text{Sales}_t}{\text{Receivables}_{t-1}/\text{Sales}_{t-1}}$ DSRI greater than 1 may indicate revenue inflation
2	Gross Margin Index (GMI)	$\text{GMI} = \frac{(\text{Sales}_{t-1} - \text{Cost of Good Sold}_{t-1})/\text{Sales}_{t-1}}{(\text{Sales}_t - \text{Cost of Good Sold}_t)/\text{Sales}_t}$ GMI greater than 1 indicates gross margin is deteriorating
3	Asset Quality Index (AQI)	$\text{AQI} = \frac{1-((\text{Current Assets}_t + \text{PP\&E}_t)/\text{Total Assets}_t)}{1-((\text{Current Assets}_t + \text{PP\&E}_t)/\text{Total Assets}_t)}$ AQI greater than 1 may indicate the tendencies of capitalizing and deferring costs that should have been expensed
4	Sales Growth Index (SGI)	$\text{SGI} = \frac{\text{Sales}_t}{\text{Sales}_{t-1}}$ SGI greater than or less than 1 may indicate that the firm is under possible pressure to manipulate earnings to keep up appearances
5	Depreciation Index (DEPI)	$\text{DEPI} = \frac{\text{Depreciation}_{t-1}/(\text{Depreciation}_{t-1} + \text{PP\&}E_{t-1})}{\text{Depreciation}_t/(\text{Depreciation}_t + \text{PP\&}E_t)}$ DEPI greater than 1 may indicate tendencies of the assets being depreciated at a slower rate to boost earnings
6	Sales General and Administrative Index (SGAI)	$\text{SGAI} = \frac{\text{Sales, General and Administrative Expenses}_t/\text{Sales}_t}{\text{Sales, General and Administrative Expenses}_{t-1}/\text{Sales}_t}$ SGAI less than 1 may indicate that the company may manipulate earnings to defer costs
7	Accruals to Total Assets (ACCR)	$\text{ACCR} = \frac{\text{Profit after tax}_t - \text{Cash from operations}_t}{\text{Total Assets}_t}$ ACCR greater than 1 may indicate that the accruals can possibly be used to manipulate earnings
8	Leverage Index (LEVI)	$\text{LVGI} = \frac{(\text{LTD}_t + \text{Current Liabilities}_t)/\text{Total Assets}_t}{(\text{LTD}_{t-1} + \text{Current Liabilities}_{t-1})/\text{Total Assets}_{t-1}}$ LTD = Long-term debt

Source: M D Beneish, The detection of earnings manipulations, *Financial Analyst Journal*, 1999, **55**(5), 24-36.

This document is authorized for use only by Vindhya Rudrappa in IDS 572 Spring 2020 taught by NEGAR SOHEILI AZAD, University of Illinois at Chicago from Jan 2020 to Jul 2020.