

Using machine learning to detect misstatements

Jeremy Bertomeu, Edwige Cheynel, Eric Floyd and Wenqiang Pan

Review of Accounting Studies, 2021.6

Lv Manni

2021. 9. 12

Contents

- Introduction
 - Background & Motivation
 - Literature Review
 - Research Problem
 - Contribution
- Data and Model Design
- Empirical Results
- Conclusion

Backgrounds & Motivation

- Machine learning algorithms are a key to unlocking the large - and growing - financial data sources to make better predictions and smarter decisions.
- We want to offer preliminary steps to applying this technology in accounting by answering a practical question: How do we detect ongoing accounting misstatements?
- What misstatements/restatement items?
- Items 4.02(a) “Non-Reliance on Previously Issued Financial Statements or a Related Audit Report or Completed Interim Review”
- Accounting Auditing and Enforcement releases (AAERs)

Literature Review

- Dechow et al. (2011) is an archetype of research in this area predict misstatements with ML tools. It develops a prediction model that outputs a scaled logistic probability of accounting irregularities for each firm-year using financial statement variables.
- Perols (2011) and Perols et al. (2016), the first studies use machine learning to predict Accounting Auditing and Enforcement releases.
- Bao et al. (2020) further extends this methodology by using a wider set of ratios and variables and comparing various ML methods.
- Other studies have examined measures that predict misstatements or irregularities, such as deferred tax liabilities (Ettredge et al.2008), audit effort (Lobo and Zhao2013) and accounting quality (Hribar et al.2014).
- The study closest to ours is by Dutta et al. (2017).

• Research Problem

- Can ML methods help detect and interpret patterns present in ongoing accounting misstatements?
- What kinds of variables help? What variables help?
- Can misstatement models predict AAERs? Any differences between misstatements and irregularities?
- How is the model predictive ability over longer horizons?
- How to interpret?

Contribution

- We focus on material non-reliance restatements Item 4.02(a), which are important events both to an audit committee and to investors, not restricted to frauds, and enables us to exploit the richness in restatement data that would be difficult to replicate with SEC enforcement actions.
- The scope of our research question is broader. And we analyze differences between the usage of misstatements and irregularities.

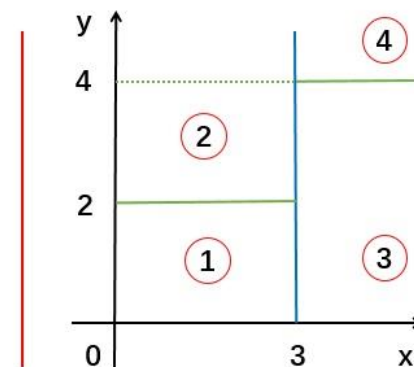
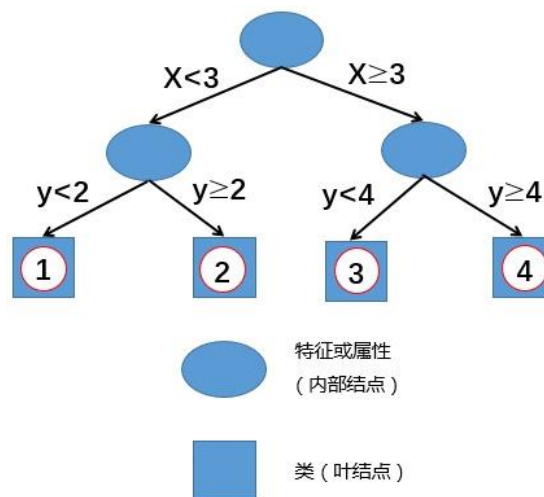
Data and Model Design

- Baseline Method: gradient boosted regression tree (GBRT)

- Friedman (2001)

- 梯度提升决策树

- regression tree



- gradient boosting

- 4 parameters: tree depth, bagging, number of trees, and shrinkage

Data and Model Design

- Data source:

Audit Analytics **Non-Reliance Restatement** database (restated in an 8-K item 4.02 filing , starts in 2001)

- Sample (2001 - 2014):

Sample selection	Compustat firm-years	Restatement firm-years
Original restatement filings sample		
Merged to Compustat with nonmissing CIK	133,047	
Main US exchange after December 2007	70,901	
Remove missing start or end year	70,901	23,772
<u>Remove SAB 108 and FIN 48 records</u>	70,901	20,239
Remove nonmaterial restatements	70,901	9,278
Unique firm-years	70,901	8,147
Firm-years between fiscal year 2001-2014	70,901	5,679
<u>Firm-years with two-year asset value</u>	65,380	5,217
Firm-years with one-year return history	55,145	4,390
Remove firm-years including a misstatement and an announcement of restatement during the same year	54,354	3,599

Data and Model Design

- Key summary statistics:

Panel A: Firm-years characteristics

Variable	Misstatement firm-years	No-Misstatement Firm-years	Compustat Firm-years
Number	3,599	50,755	54,354
Total assets (in \$ millions)	7,135	9,627	9,462
Market value (in \$ millions)	3,117	4,126	4,059
Book value (in \$ millions)	1,189	1,713	1,679

Panel B: Restatement income effects

Restatement income effect	Freq	Percent	Average (in million)	Average (scaled by average asset)
Negative	2,201	61.2	-15.2	-2.23%
Zero	649	18.0	0	0
Positive	749	20.8	9.0	1.93%
Total	3,599	100	-7.4	-0.96%

Panel A: Frequency of firm-year restatements by industry

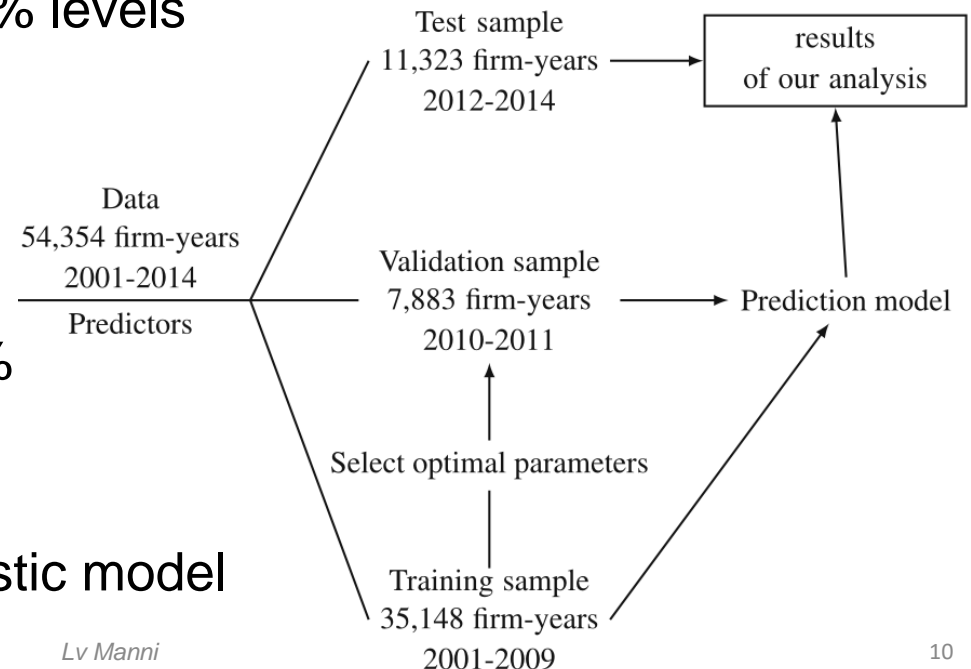
Industry	Restatement firms	Compustat population
Agriculture	0.2	0.2
Mining & Construction	3.6	3.1
Food & Tobacco	2.2	2.1
Textiles and Apparel	1.0	1.0
Lumber, Furniture, & Printing	1.1	2.4
Chemicals	1.4	2.3
Refining & Extractive	4.6	4.1
Durable Manufacturers	17.3	18.3
Computers	20.2	14.0
Transportation	5.1	5.3
Utilities	1.9	3.2
Retail	12.6	7.9
Services	12.3	8.5
Banks & Insurance	9.1	20.3
Pharmaceuticals	7.5	7.4
Total	100	100

Panel B: Frequency of firm-year restatements by size deciles

Decile rank of market value of Compustat population	Frequency	Percentage
1	313	8.7
2	304	8.5
3	343	9.5
4	370	10.3
5	385	10.7
6	437	12.1
7	433	12.0
8	425	11.8
9	313	8.7
10	276	7.7
Total	3,599	100.0

Data and Model Design

- Over 100 potential predictor variables:
financial variables, audit variables, credit rating variables, opinion divergence variables, and corporate governance variables
auditor opinion, indicator variable for the existence of a management forecast, analyst consensus forecast, short interest, and indicator variables for foreign firms and current or past restatement announcements
- winsorized at the 1% and 99% levels
- Schematic for data:
tree depth: 9
number of trees: 6,376
Shrinkage: 1%, baggage of 70%
- Other models: RUSBoost, Random Forest, backward logistic model

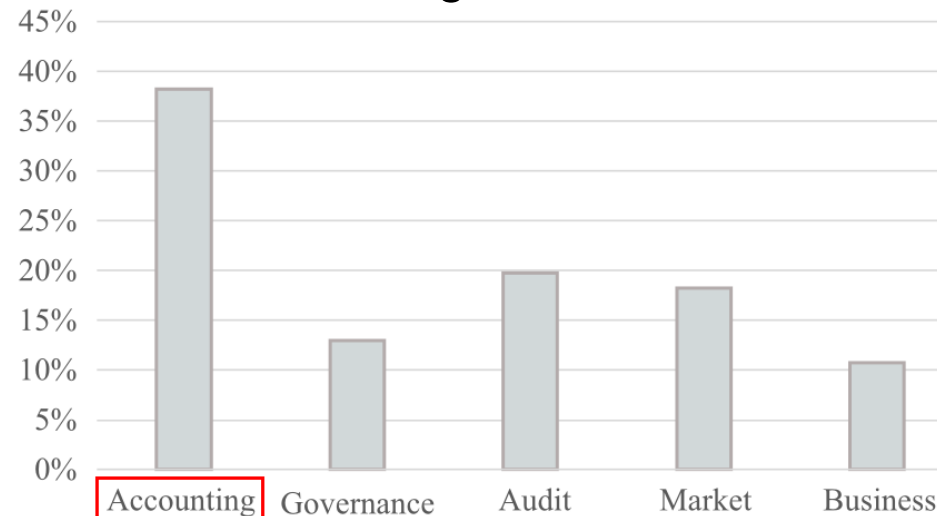


Empirical Results - GBRT Detection model

- predictive power of variables

Predictor	Importance	Cumulative
% Soft assets	3.25	3.25
Bid ask spread	2.91	6.17
Non-audit fee / total fee	2.76	8.93
Qualified opinion (internal control)	2.72	11.65
Change in operating lease activity	2.69	14.34
Short interest	2.57	16.91
Stock return volatility	2.40	19.31
Log of non-audit fee	2.35	21.66
Percentile rank of audit fee by auditor	2.30	23.97
Leverage	2.30	26.27
Level of finance raised	2.24	28.50
Abnormal change in employees	2.18	30.69
WC accruals	2.17	32.86

- Do accounting variables matter?



- By its relation with other vars?

Model	R^2	AUC	Catch rate of restatement	Catch Rate of AAER	Importance of accounting variables
Business only	0.8%	55.5%	37.1%	56.8%	0 %
Governance only	6.0%	57.2%	42.5%	35.1%	0 %
Market only	5.0%	60.6%	46.5%	45.9%	0 %
Audit only	8.5%	61.7%	48.4%	48.6%	0 %
Accounting only	3.5%	58.3%	41.6%	59.5%	100 %
Accounting+ Business	4.7%	61.6%	45.1%	59.5%	82.8%
Accounting+ Governance	7.2%	62.0%	46.7%	45.9%	74.4%
Accounting + Market	5.3%	62.0%	46.7%	51.4%	55.2%
Accounting + Audit	10.1%	66.5%	53.7%	62.2%	68.6%
Full model	14.1%	72.8%	64.3%	78.4%	36.2%

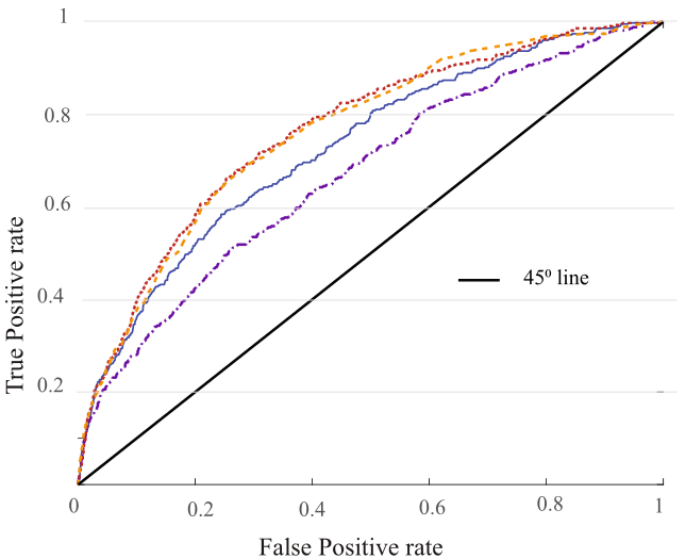
Empirical Results - GBRT Detection model

- predictive power of single variables

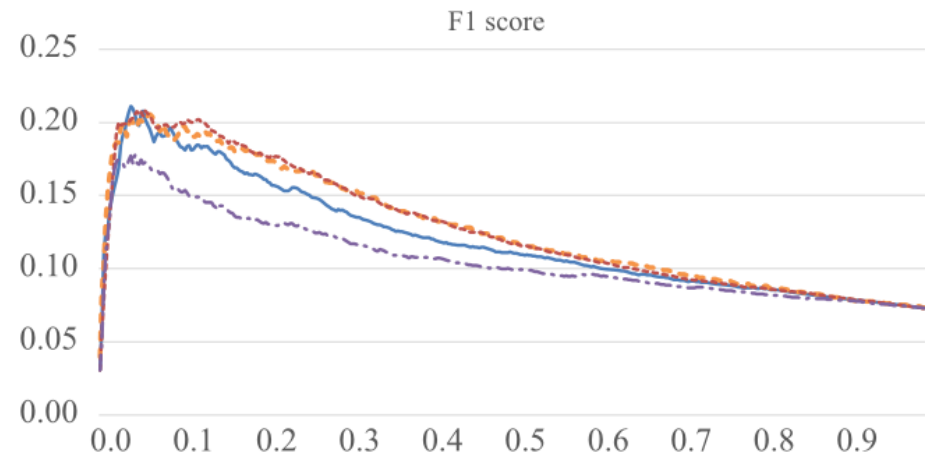
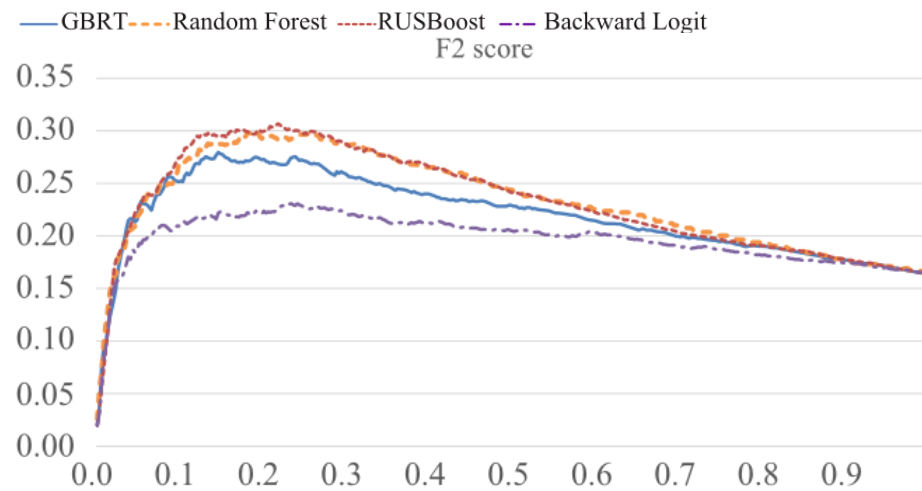
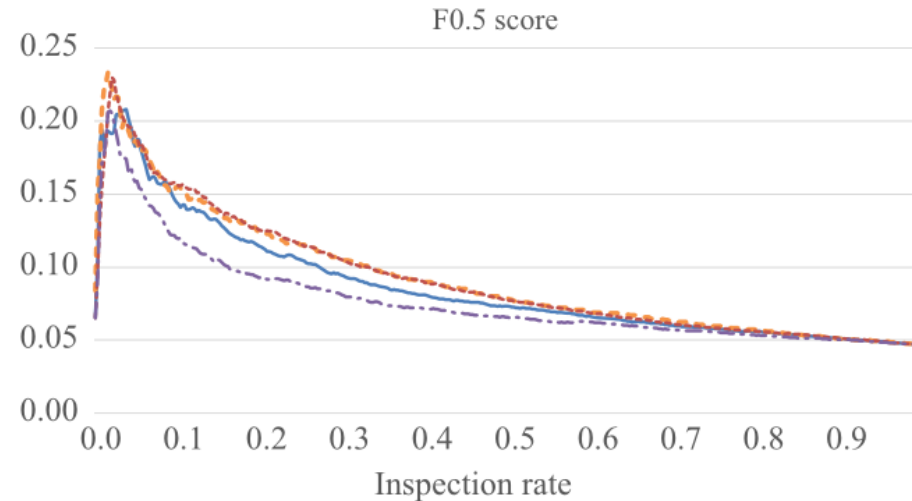
Predictors	Res Mean	Non Res Mean	Res - Non Res Diff in mean	Two-tailed p-value
% Soft assets	55.2%	57.7%	-2.5%	0.0001
Bid ask spread	1.2%	1.1%	0.0%	0.1904
Non-audit fee / total fee	25.3%	19.6%	5.7%	0.0001
Qualified opinion (internal control)	11.3%	2.4%	8.9%	0.0001
Change in operating lease activity	0.6%	0.3%	0.3%	0.0001
Short interest	3.7%	3.5%	0.2%	0.0280
Stock return volatility	3.8%	3.3%	0.5%	0.0001
Log of non audit fee	11.1	10.6	0.5	0.0001
Percentile rank of audit fee by auditor	69.8%	66.7%	3.1%	0.0001
Leverage	16.9%	15.5%	1.3%	0.0001

Empirical Results - Model comparisons

- ROC curve:



- $$F_{\beta} = (1 + \beta^2) \frac{\text{precision} \times \text{recall}}{\beta^2 \text{precision} + \text{recall}}$$



Empirical Results - Model comparisons

- Catch rate when using small bandwidths:

Inspection rate	GBRT	Random forest	RUSBoost	Backward logistic
1%	7.9%	8.4%	6.3%	7.0%
2%	11.9%	14.0%	13.6%	12.9%
3%	17.1%	17.5%	17.8%	15.4%
4%	21.7%	20.6%	20.8%	18.2%
5%	22.9%	23.1%	23.8%	19.9%
6%	25.7%	25.9%	26.6%	22.0%
7%	27.1%	28.3%	28.0%	23.8%
8%	29.7%	30.1%	30.8%	25.7%
9%	32.7%	31.5%	33.2%	26.2%
10%	33.4%	35.3%	36.4%	27.8%

Empirical Results - Model comparisons

- Performance for the top 1/3 predicted probabilities:

Model	GBRT	Random forest	RUSBoost	Backward Logistic
-------	------	---------------	----------	-------------------

Panel A: Detection rate for the top 1/3 predicted probabilities of firm—years

Catch rate	64.3%	70.8%	72.2%	55.4%
Catch number	275	303	309	237

Panel C: Restatement income effect for the top 1/3 predicted probabilities of firm—years

Number of unique firm—years	98	126	132	60
Average absolute income effect (million)	8.5	12.3	8.9	9.8
Average scaled absolute income effect	1.22%	1.24%	1.19%	1.99%
Number of negative firm—years	58	79	76	32
Average income effect (million)	−10.7	−15.5	−10.8	−13.8
Average scaled income effect	−1.45%	−0.92%	−0.89%	−2.50%
Number of positive firm—years	19	27	29	12
Average income effect (million)	11.0	12.2	12.1	12.6
Average scaled income effect	1.85%	3.08%	3.11%	3.28%

Panel D: Average detection times relative to start year and filing year for the top 1/3 predicted probabilities of firm—years

	Nb.	Mean	Nb.	Mean	Nb.	Mean	Nb.	Mean
Relative to filing year	224	1.88	239	1.92	241	1.97	197	1.77
Relative to starting year	224	1.74	239	1.67	241	1.63	197	1.57

Empirical Results - Detecting AAERs

- AAERs sample selection and description

Panel A: AAERs sample selection

Number of AAERs Firm-years	Number
All AAERs firm-years from 2001-2014	865
Less: restatement sample selection filters	(373)
Less: are not in restatement files	(107)
Total	<u>385</u>

Panel B: Income effects of AAERs

Income effect	Freq	Percent	Average (in million)	Average (scaled by average asset)
Negative	302	78.4	-42.2	-2.48%
Zero	36	9.4	0	0
Positive	47	12.2	22.4	1.49%
Total	385	100	<u>-30.4</u>	<u>-1.76%</u>

- AAERs catch rates on test dataset

Model	Catch	Percentage	Total AAER Firm-years
GBRT	29	78.4%	37
Random forest	35	94.6%	37
RUSBoost	33	89.2%	37
Backward logistic	23	62.2%	37

Empirical Results - Detecting AAERs

- Top 10 explanatory variables:

% soft Assets
bid ask spreads
non-audit fee / total fee
short interest
stock returns
percent of audit fee

GBRT	Random forest	RUSBoost
Panel A: Restatement models		
<i>% Soft assets</i>	<i>Bid-ask spread</i>	<i>% Soft assets</i>
<i>Bid-ask spread</i>	<i>Chg. in operating leases</i>	Return
<i>Non-audit fee / total fee</i>	<i>Non-audit fee / total fee</i>	Lag one year return
Qualified opinion (controls)	<i>% Soft assets</i>	<i>Bid-ask spread</i>
<i>Chg. in operating leases</i>	<i>Level of finance raised</i>	Auditor tenure
<i>Short interest</i>	Lag one year return	Book-to-market
<i>Stock return volatility</i>	<i>Stock return volatility</i>	Firm age
Log of non-audit fee	<i>Perc. rank of audit fee by auditor</i>	<i>Level of finance raised</i>
<i>Perc. rank of audit fee by auditor</i>	Earnings-to-price	<i>Short interest</i>
Leverage	<i>Perc. rank of total fee by auditor</i>	Change in receivables
Panel B: AAER models		
<i>% Soft assets</i>	<i>Non-audit fee / total fee</i>	Auditor tenure
<i>Non-audit fee / total fee</i>	<i>% Soft assets</i>	Return
<i>Return</i>	<i>Perc. rank of audit fee by auditor</i>	<i>Lag one year return</i>
<i>Log of non-audit fee</i>	<i>Log of non audit fee</i>	<i>% Soft assets</i>
<i>Perc. rank of audit fee by auditor</i>	<i>Bid-ask spread</i>	Change in receivables
<i>Short interest</i>	Perc. rank of total fee by auditor	WC accruals
Deferred tax expense	Log of total fee	<i>Log of non-audit fee</i>
<i>Bid-ask spread</i>	<i>Short interest</i>	Chg. in cash sales
<i>Lag one year return</i>	<i>Perc. rank of non-audit fee by auditor</i>	Chg. in cash margin
Firm age	Log of audit fee	Chg. in operating leases

Empirical Results - Further analyses

- Predictive ability over longer horizons:

Model	R^2	AUC	Catch rate of restatement	Catch rate of AAER
Current year	14.8%	72.3%	63.5%	82.8%
One-year-ahead	11.7%	68.0%	56.9%	79.3%
Two-year-ahead	7.5%	59.9%	42.2%	48.3%

- Importance of predictors:

Predictor	Importance	Cumulative			
Panel A: One-year ahead			Panel B: Two-year ahead		
Short interest	3.26	3.26	Short interest	10.34	10.34
% Soft assets	3.25	6.51	Stock return volatility	10.00	20.34
Stock return volatility	3.24	9.74	Non-audit fee / total fee	5.13	25.47
Non-audit fee / total fee	3.09	12.84	Industry: Banks & Insurance	4.16	29.62
Bid-ask spread	3.01	15.85	Percentile rank of total fee by auditor	3.72	33.35
Change in operating leases activity	2.62	18.48	Bid-ask spread	3.54	36.89
Lag one year return	2.56	21.04	Industry: Computers	3.21	40.09
Percentile rank of audit fee by auditor	2.36	23.41	Log of total fee	2.95	43.04
Book-to-market	2.30	25.70	Log of non-audit fee	2.69	45.74
Log of non-audit fee	2.28	27.98	% Outsiders appointed	2.43	48.17
Return	2.23	30.21	Change in operating leases activity	2.23	50.40
WC accruals	2.20	32.41	Missing or unaudited internal control	2.23	52.63
% Outsiders own	2.19	34.60	Percentile rank of audit fee by auditor	2.09	54.73
Abnormal change in employees	2.17	36.76	Industry: Retail	1.88	56.61
Change in inventory	2.14	38.91	% Board Inside	1.78	58.39

Empirical Results - Interpretation

- Rules obtained by applying InTrees(Deng (2018)) to GBRT:

Max var Frequency of rule Variables			Predicted Error rate	
Top 10 variables				
3	0.005	% soft assets≤0.513 & Qualified opinion (controls)= 1 & Leverage≤0.017	40.89%	24.17%
5	0.006	0.011<% soft assets≤0.602 & Bid-ask spread>0.005 & Qualified opinion (controls)=1 & Pct. rank of audit fee≤0.987	37.50%	23.44%
3	0.051	% soft assets>0.088 & Non-audit fee / total fee>0.323 & Chg. in operating lease>0.008	16.30%	13.64%
5	0.047	Non-audit fee / total fee>0.307 & Qualified opinion (controls)=0 & Chg. in operating lease>-0.018 & Short interest rate≤0 & Non audit fee>11.833	16.86%	14.02%
3	0.104	Non-audit fee / total fee>0.401 & Qualified opinion (controls)=0 & Stock return volatility>0.024	14.54%	12.43%
5	0.104	% soft assets≤0.886 & Non-audit fee / total fee>0.399 & Qualified opinion (controls)=0 & Chg. in operating lease≤0.033 & Stock return volatility≤0.053	11.11%	9.87%
3	0.199	Non-audit fee / total fee>0.249 & Chg. in operating lease≤0.008 & Stock return volatility>0.02	10.20%	9.16%
5	0.196	Non-audit fee / total fee>0.247 & Qualified opinion (controls)=0 & Chg. in operating lease≤0.032 & Stock return volatility>0.02 & Pct. rank of audit fee>0.401	10.93%	9.74%

Conclusion

- With a wide set of variables from accounting, capital markets, governance, and auditing dataset, we show that these methods help detect and interpret patterns present in ongoing accounting misstatements.
- We find that accounting variables, while they do not detect misstatements well on their own, become important with suitable interactions with audit and market variables.
- We also analyze differences between misstatements and irregularities, examine one-year- and two-year-ahead predictions and interpret groups at greater risk of misstatements.