AI-Assisted Finance and Management Decisions

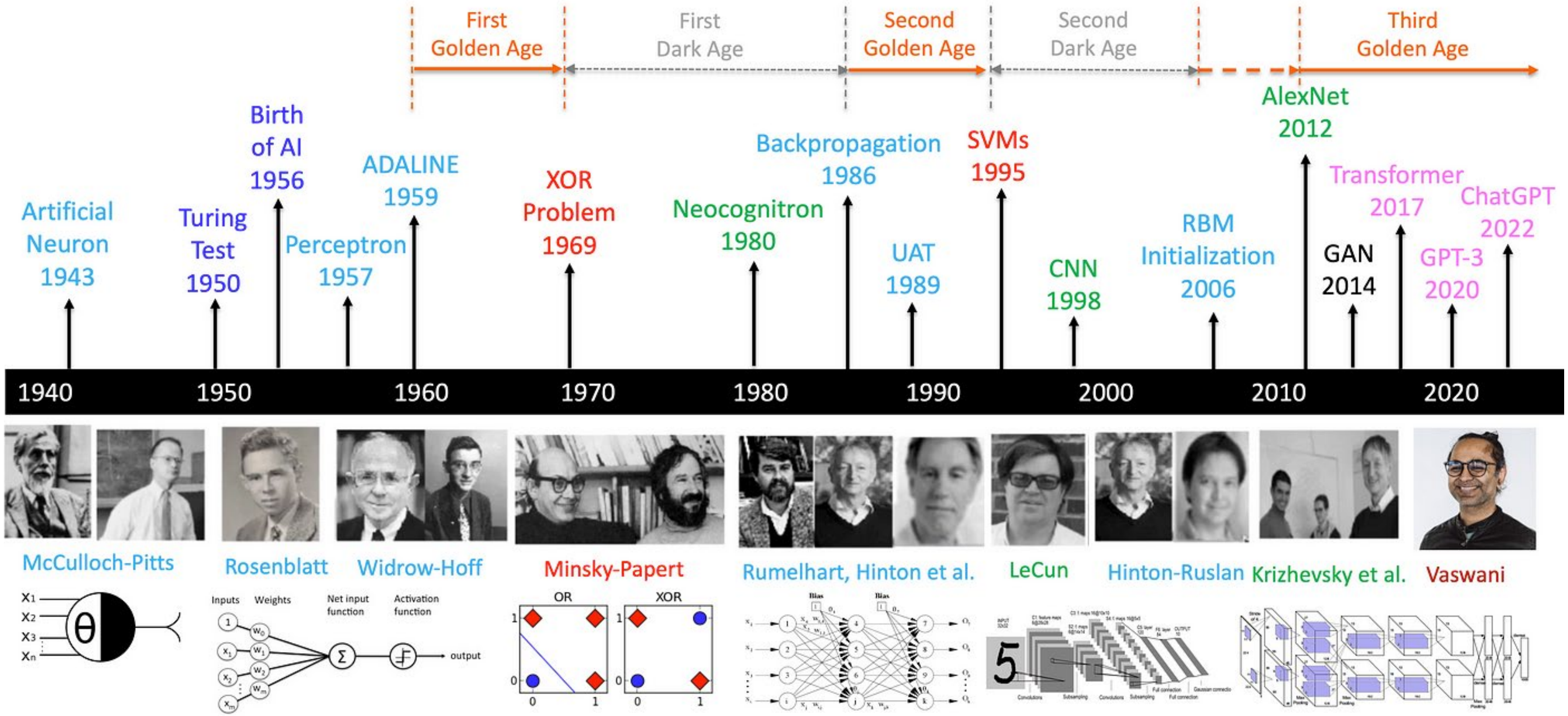
2025 New Zealand Capital Market Symposium

17 June 2025

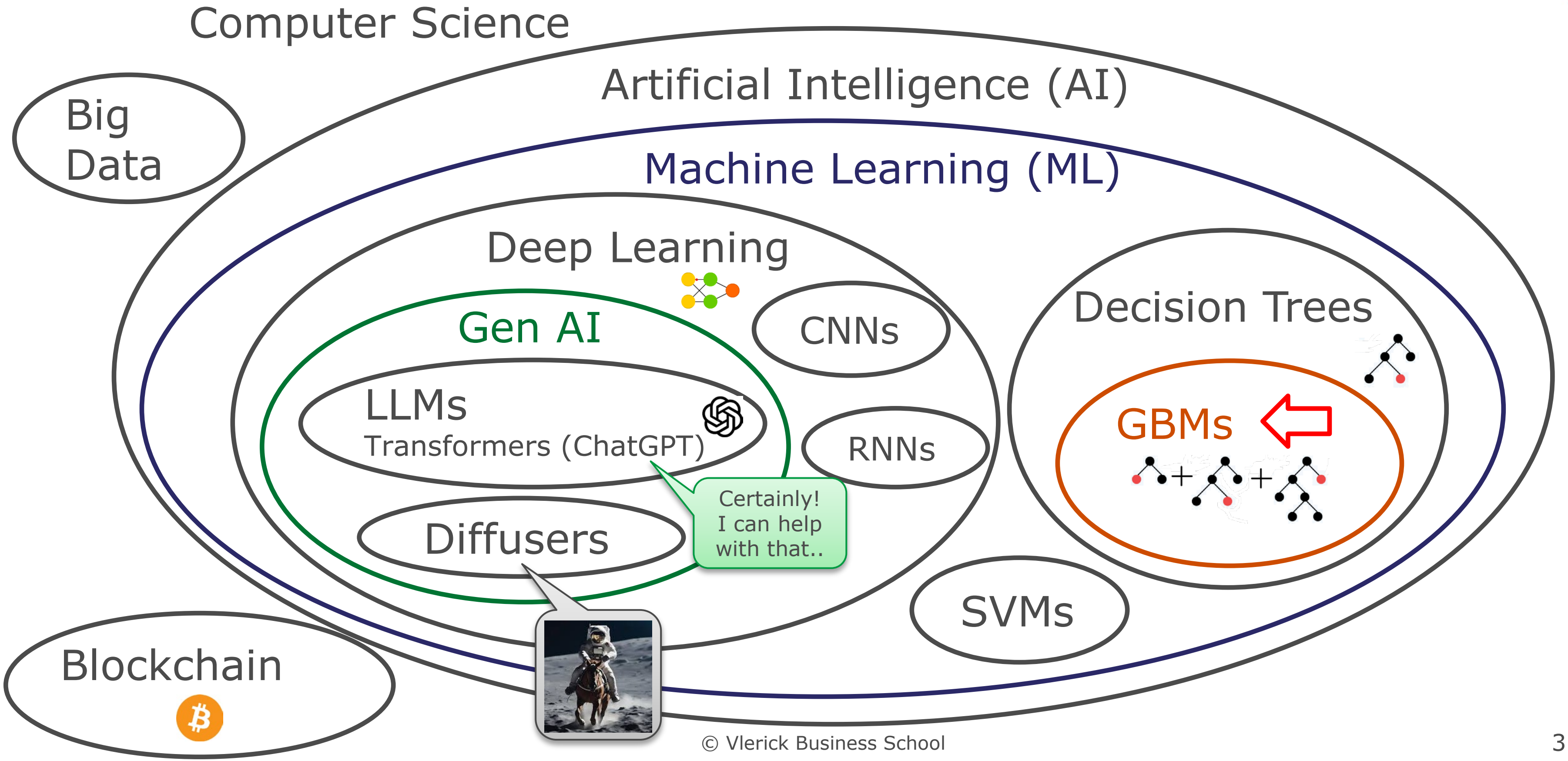
Prof Dr Helen Lu

VLERICK BUSINESS SCHOOL

A brief history of AI



What is Machine Learning and AI?



Classic AI vs Generative AI

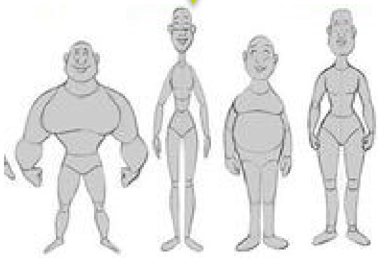
- **Supervised classic AI**
  - Data = (Features, Target)
  - Data + Training → Model
  - Model → Predictions
  - → *Use accurate prediction to do something better*

	A	B	C
6	CUSTOMER	REGION	ORDER DATE
7	Acme, inc.	NORTH	13-04-15
8	Widget Corp	NORTH	21-12-15
9	123 Warehousing	EAST	15-02-15
10	Demo Company	CENTRAL	28-06-15
11	Smith and Co.	SOUTH	22-08-15
12	Foo Bars	WEST	31-12-15
13	ABC Telecom	NORTH	13-04-15

- **Generative AI**
  - Data = text, images, video, signals
  - Model → More of same data
  - → *Produce more of something*

UNSUPERVISED CLASSIC AI: CLUSTERING

ID	Gender	Weight	Height	Age	Shoulder	Chest	Waist	Hip	…
1									
2									
.									
.									



AI + HUMANS > HUMANS

The Wall Street Journal

Journal of Accounting Research

DOI: 10.1111/1475-679X.12464
Journal of Accounting Research
Vol. No. October 2022
Printed in U.S.A.

CHICAGO BOOTH

Relative Valuation with Machine Learning

PAUL GEERTSEMA* AND HELEN LU*

Received 30 November 2020; accepted 8 September 2022

ABSTRACT

We use machine learning for relative valuation and peer firm selection. In out-of-sample tests, our machine learning models substantially outperform traditional models in valuation accuracy. This outperformance persists over time and holds across different types of firms. The valuations produced by machine learning models behave like fundamental values. Overvalued stocks decrease in price and undervalued stocks increase in price in the following month. Determinants of valuation multiples identified by machine learning models are consistent with theoretical predictions derived from a discounted cash flow approach. Profitability ratios, growth measures, and efficiency ratios are the most important value drivers throughout our sample period. We derive a novel method to express valuation multiples predicted by our machine learning models as weighted averages of peer firm multiples. These weights

HARVARD BUSINESS REVIEW

AI Can Make the Relative-Valuation Process Less Subjective

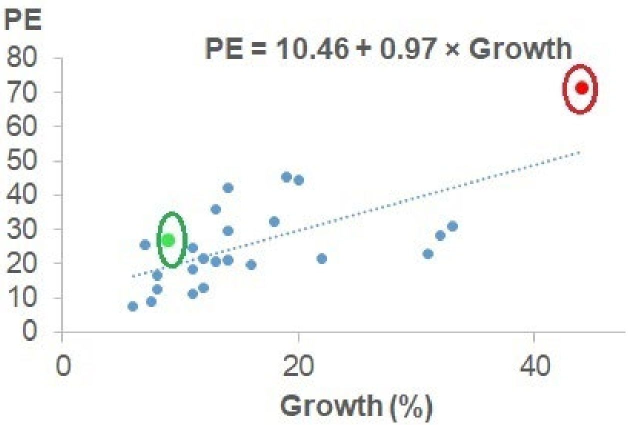
by Paul Geertsema, Helen Lu and Kristof Stouthuysen

April 29, 2025

VLERICK BUSINESS SCHOOL

AI-ASSISTED VALUATION ("MULTIPLES" "COMPS")
THE MAGIC: LEARNING FROM THE DATA

Specify a model

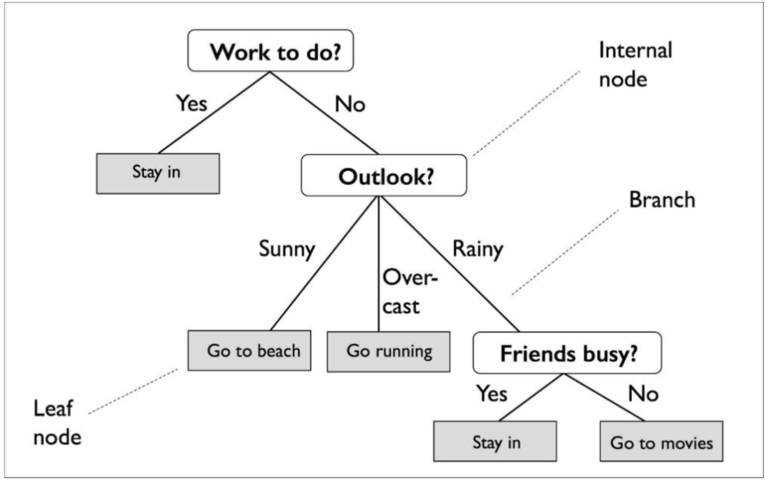


Let the machine learn from the data

- Sales growth > 20%
  - Y: Sales growth> 30%
    - Y: Price/sales = 18
    - N: Price/sales = 15
  - N: Gross margin> 60%
    - Y: Price/sales = 20
    - N: Price/sales = 18

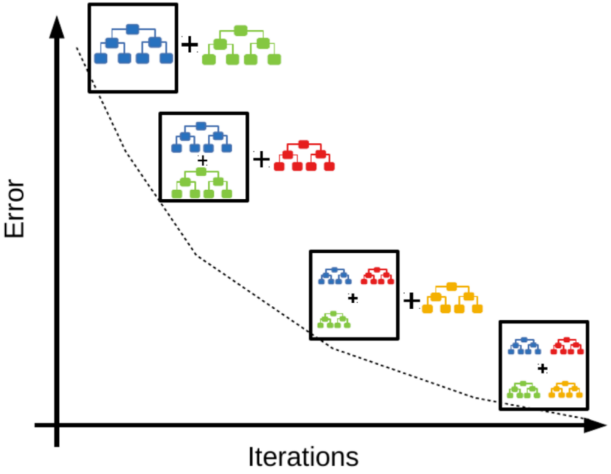
AI-ASSISTED VALUATION (“MULTIPLES” “COMPS”)
USING GRADIENT BOOSTING MACHINES

Tree predictions



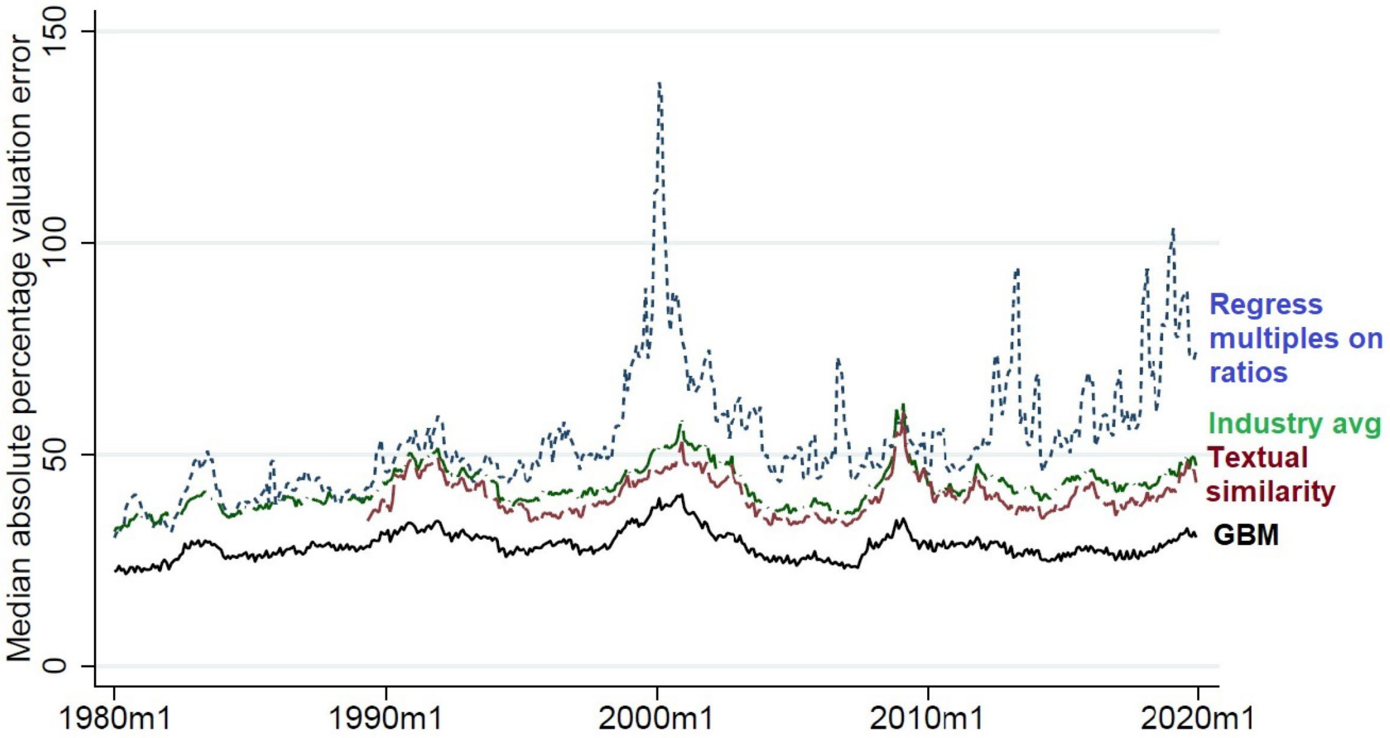
medium.com

Gradient boosting machines



Neptune.ai

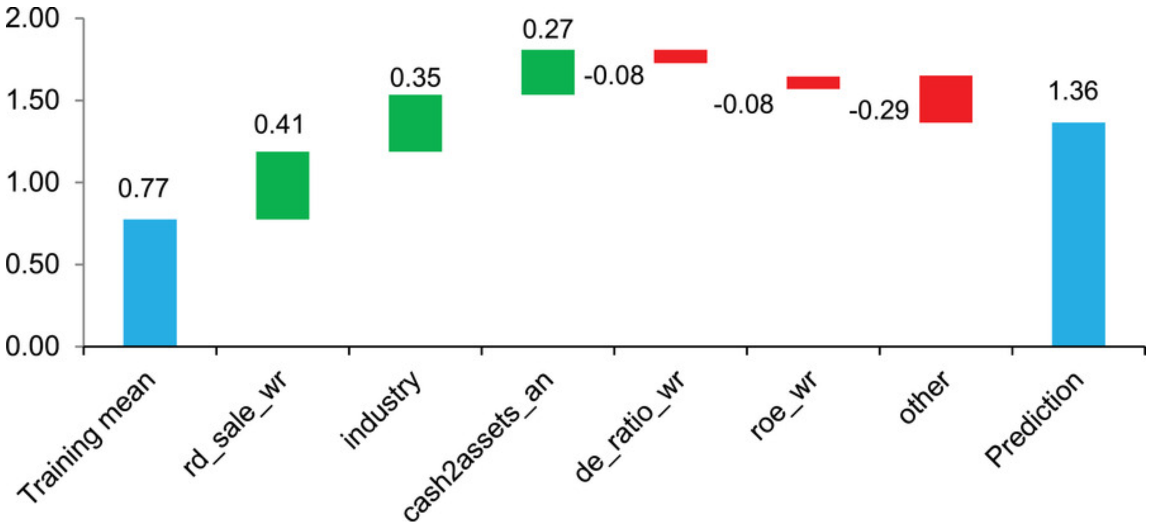
AI-ASSISTED VALUATION
HIGHER ACCURACY IN "SMALL DATA"



Geertsema and Lu (2023)

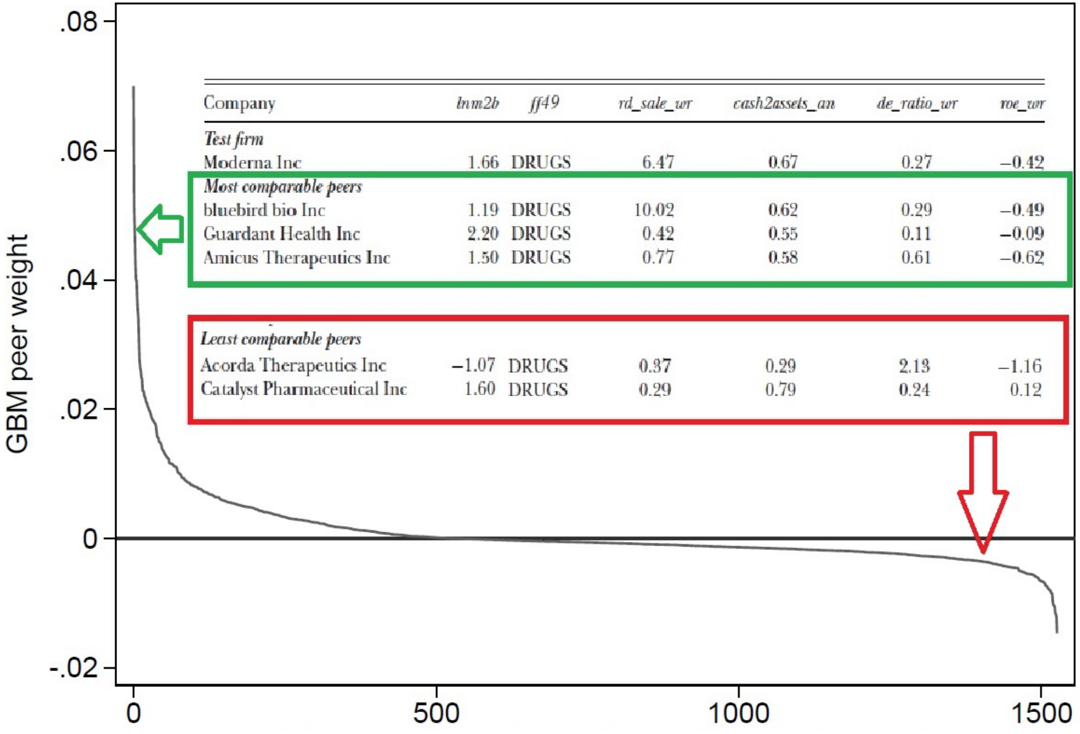
AI-ASSISTED VALUATION
BUILD TRUST THROUGH UNDERSTANDING: SHAP

- Explainability: SHAP values (SHapley Additive exPlanations)
- Individual SHAP values for Moderna (December 2019).
- Target: the log of the market-to-book multiple (lnm2b)



AI-ASSISTED VALUATION
BUILD TRUST THROUGH UNDERSTANDING: PEER WEIGHTS

"Peer" is not a "Yes/No" choice anymore

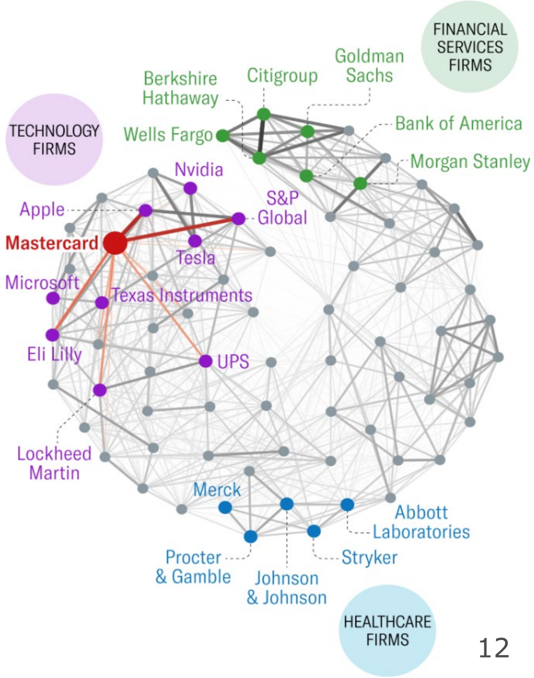


Company	lnm2b	ff49	rd_sale_wr	cash2assets_an	de_ratio_wr	roe_wr
Test firm						
Moderna Inc	1.66	DRUGS	6.47	0.67	0.27	−0.42
Most comparable peers						
bluebird bio Inc	1.19	DRUGS	10.02	0.62	0.29	−0.49
Guardant Health Inc	2.20	DRUGS	0.42	0.55	0.11	−0.09
Amicus Therapeutics Inc	1.50	DRUGS	0.77	0.58	0.61	−0.62
Least comparable peers						
Acorda Therapeutics Inc	−1.07	DRUGS	0.37	0.29	2.13	−1.16
Catalyst Pharmaceutical Inc	1.60	DRUGS	0.29	0.79	0.24	0.12

HARVARD BUSINESS REVIEW

AI Can Make the Relative-Valuation Process Less Subjective

by Paul Geertsema, Helen Lu and Kristof Stouthuysen



AI-ASSISTED FINANCIAL FORECASTING
THE WORKING PAPER

Projecting Financial Statements with Artificial Intelligence

Paul Geertsema,[*] Helen Lu,[†] Guang Ma[‡]

First version: 25th October 2023

This version: 8th June 2025

Abstract

We introduce a novel artificial intelligence framework for projecting financial statements. Our approach integrates multi-target learning and chained learning to predict interdependent financial statement items, capturing the intricate relationships across income statement and balance sheet components. Leveraging gradient boosting machines (GBMs) as the base learner, the framework employs a four-step process to optimise chaining sequences and feature sets, in order to effectively model inter-item correlations. Empirical validation using out-of-sample predictions for a large sample of U.S. public firms demonstrates the model's ability to produce accurate and internally consistent financial statement projections. Line-item analyses reveal that the incremental signal gained at each stage of the chain consistently outweighs the small losses from error propagation, so predictive information builds rather than erodes as the forecast moves through the statement. This pattern proves robust across items, periods, and alternative specifications. Furthermore, we establish the utility of these projections for detecting financial irregularities out-of-sample. Our methodology can be adapted to other tasks involving the prediction of complex and interdependent outputs. [168 words]

Keywords: artificial intelligence; machine learning; multi-target forecasting; chained learning; financial statement projection; restatements

JEL codes: C45; C53; G17; M41

AI-ASSISTED FINANCIAL FORECASTING

FINANCIALS HAVE INTERNAL STRUCTURE

Balance Sheet	**2023**	**2024**
Current assets	*214.93*	*233.17*
Cash	189.22	198.00
Debtors	15.71	15.17
Stock	10.00	20.00
	12.60	-
Long-term assets	*179.40*	*179.40*
Fixed assets	167.40	167.40
Investments	12.00	12.00
TOTAL ASSETS	*394.33*	*412.57*
Current liabilities	44.67	44.67
Creditors	41.67	41.67
Short term debt	3.00	3.00
Long-term liabilities	27.00	27.00
Debt	27.00	27.00
TOTAL LIABILITIES	71.67	71.67
Equity	322.66	340.90
Capital	300.00	300.00
Retained Earnings	22.66	40.90
TOTAL EQUITY + LIABILITIES	*394.33*	*412.57*

Profit and Loss	**2023**	**2024**
Revenue	348.00	336.00
Less: Cost of goods sold	240.00	240.00
Gross profit	108.00	96.00
Less: Admin expenses	7.70	7.70
Less: Salaries	25.00	25.00
EBITDA	75.30	63.30
Less: Interest expense	1.20	1.20
Less: Depreciation	12.60	12.60
Net income before tax	61.50	49.50
Less: Tax	27.68	22.28
Net profit after tax	33.83	27.23
Less: Dividends	11.16	8.98
Retained earnings	22.66	18.24

Cash flow statement	**2023**	**2024**
Operating activities	62.38	30.37
Cash received from sales	332.29	320.83
Last year debtors received	-	15.71
Cash paid for stock	- 208.33	- 208.33
Last year creditors paid	-	- 41.67
Admin expenses	- 7.70	- 7.70
Salaries	- 25.00	- 25.00
Interest paid	- 1.20	- 1.20
Tax	- 27.68	- 22.28
Investment activities	- 192.00	- 12.60
Net investments made	- 12.00	-
Net investments in fixed assets	- 180.00	- 12.60
Financing activities	318.84	- 8.98
Capital contributions	300.00	-
Net debt issued	30.00	-
Dividends paid	- 11.16	- 8.98
Net cash flow	**189.22**	**8.78**

AI-ASSISTED FINANCIAL FORECASTING

FINANCIALS HAVE INTERNAL STRUCTURE

Balance Sheet	2023	2024
Current assets	*214.93*	*233.17*
Cash	189.22	198.00
Debtors	15.71	15.17
Stock	10.00	20.00
	12.60	-
Long-term assets	*179.40*	*179.40*
Fixed assets	167.40	167.40
Investments	12.00	12.00
TOTAL ASSETS	*394.33*	*412.57*
Current liabilities	44.67	44.67
Creditors	41.67	41.67
Short term debt	3.00	3.00
Long-term liabilities	27.00	27.00
Debt	27.00	27.00
TOTAL LIABILITIES	71.67	71.67
Equity	322.66	340.90
Capital	300.00	300.00
Retained Earnings	22.66	40.90
TOTAL EQUITY + LIABILITIES	*394.33*	*412.57*

Profit and Loss	2023	2024
Revenue	348.00	336.00
Less: Cost of goods sold	240.00	240.00
Gross profit	108.00	96.00
Less: Admin expenses	7.70	7.70
Less: Salaries	25.00	25.00
EBITDA	75.30	63.30
Less: Interest expense	1.20	1.20
Less: Depreciation	12.60	12.60
Net income before tax	61.50	49.50
Less: Tax	27.68	22.28
Net profit after tax	33.83	27.23
Less: Dividends	11.16	8.98
Retained earnings	22.66	18.24

Cash flow statement	2023	2024
Operating activities	62.38	30.37
Cash received from sales	332.29	320.83
Last year debtors received	-	15.71
Cash paid for stock	208.33	208.33
Last year creditors paid	-	41.67
Admin expenses	- 7.70	- 7.70
Salaries	- 25.00	- 25.00
Interest paid	- 1.20	- 1.20
Tax	- 27.68	- 22.28
Investment activities	- 192.00	- 12.60
Net investments made	- 12.00	-
Net investments in fixed assets	- 180.00	- 12.60
Financing activities	318.84	- 8.98
Capital contributions	300.00	-
Net debt issued	30.00	-
Dividends paid	- 11.16	- 8.98
Net cash flow	**189.22**	**8.78**

AI-ASSISTED FINANCIAL FORECASTING
CORE ITEMS ARE ALL YOU NEED

Income Statement

Line Item	Variable	Calculation	Sign
Sales	**sale**		+
Less: COGS	**cogs**		+
Gross Profit	gp	sale-cogs	+/-
Less: Other expenses	**xsga**		+
EBITDA	ebitda	gp-xsga	+/-
Less: Depreciation and amortisation	**dp**		+
Earnings before interest and tax	ebit	ebitda-dp	+/-
Less: Interest expense	**xint**		+
Earnings before tax	ebt	ebit-xint	+/-
Add: Non-operating income	**nopi**		+/-
Add: Special items	**spi**		+/-
Income before tax	pi	ebt+nopi+spi	+/-
Less: Tax	**txt**		+/-
Net income before minority interest	ibmii	pi-txt	+/-
Less: Minority interest	**mii**		+/-
Income	ib	ibmii-mii	+/-
Less: Dividends (preference/preferred)	**dvp**		+
Add: Extraordinary, discontinued, other	**xido**		+/-
Net income to shareholders	niadj	ib-dvp+xido	+/-
Less: Dividends to shareholders	**dvc**		+
Retained income	retain	niadj-dvc	+/-

Core items in **bold**

Balance Sheet

Line Item	Variable	Calculation	Sign
Current assets	act	che+rect+invt+aco	+
Cash	che	l_che + retain + dp - (ppent - l_ppent + dp) - (rect - l_rect) - (invt - l_invt) - (ao - l_ao) + (dlc - l_dlc) + (ap - l_ap) + (txp - l_txp) + (lco - l_lco) - (iva - l_iva) - (intan - l_intan) - (lo - l_lo) + (dltt - l_dltt) + (txditc - l_txditc) + (lo - l_lo) + nei	+
Receivables	**rect**		+
Inventory	**invt**		+
Other current assets	**aco**		+
Long-term assets	alt	ppent+iva+intan+ao	+
Plant, property and equipment	**ppent**		+
Investments	**iva**		+
Intangible Assets	**intan**		+
Other long-term assets	**ao**		+
Total assets	at	act+alt	+
Current liabilities	lct	dlc+ap+txp+lco	+
Current debt	**dlc**		+
Payables	**ap**		+
Income tax payable	**txp**		+
Other current liabilities	**lco**		+
Long-term liabilities	llt	dltt+txditc+lo	+
Long-term debt	**dltt**		+
Deferred taxes and ITC	**txditc**		+
Other long-term liabilities	**lo**		+
Total liabilities	lt	lct+llt	+
Equity	eq	at-lt	+/-
Shareholders equity (parent)	seq	eq-mi	+/-
Non-controlling interest	**mi**		+/-
Net Equity Issuance	**nei**		+/-
Other Comprehensive Income	oci	eq-l_eq-retain-nei	+/-

AI-ASSISTED FINANCIAL FORECASTING

YOU CAN CALCULATE THE REST

Income Statement

Line Item	Variable	Calculation	Sign
Sales	**sale**		+
Less: COGS	**cogs**		+
Gross Profit	gp	sale-cogs	+/-
Less: Other expenses	**xsga**		+
EBITDA	ebitda	gp-xsga	+/-
Less: Depreciation and amortisation	**dp**		+
Earnings before interest and tax	ebit	ebitda-dp	+/-
Less: Interest expense	**xint**		+
Earnings before tax	ebt	ebit-xint	+/-
Add: Non-operating income	**nopi**		+/-
Add: Special items	**spi**		+/-
Income before tax	pi	ebt+nopi+spi	+/-
Less: Tax	**txt**		+/-
Net income before minority interest	ibmii	pi-txt	+/-
Less: Minority interest	**mii**		+/-
Income	ib	ibmii-mii	+/-
Less: Dividends (preference/preferred)	**dvp**		+
Add: Extraordinary, discontinued, other	**xido**		+/-
Net income to shareholders	niadj	ib-dvp+xido	+/-
Less: Dividends to shareholders	**dvc**		+
Retained income	retain	niadj-dvc	+/-

Balance Sheet

Line Item	Variable	Calculation	Sign
Current assets	act	che+rect+invt+aco	+
Cash	che	l_che + retain + dp - (ppent - l_ppent + dp) - (rect - l_rect) - (invt - l_invt) - (ao - l_ao) + (dlc - l_dlc) + (ap - l_ap) + (txp - l_txp) + (lco - l_lco) - (iva - l_iva) - (intan - l_intan) - (lo - l_lo) + (dltt - l_dltt) + (txditc - l_txditc) + (lo - l_lo) + nei	+
Receivables	**rect**		+
Inventory	**invt**		+
Other current assets	**aco**		+
Long-term assets	alt	ppent+iva+intan+ac	+
Plant, property and equipment	**ppent**		+
Investments	**iva**		+
Intangible Assets	**intan**		+
Other long-term assets	**ao**		+
Total assets	at	act+alt	+
Current liabilities	lct	dlc+ap+txp+lco	+
Current debt	**dlc**		+
Payables	**ap**		+
Income tax payable	**txp**		+
Other current liabilities	**lco**		+
Long-term liabilities	llt	dltt+txditc+lo	+
Long-term debt	**dltt**		+
Deferred taxes and ITC	**txditc**		+
Other long-term liabilities	**lo**		+
Total liabilities	lt	lct+llt	+
Equity	eq	at-lt	+/-
Shareholders equity (parent)	seq	eq-mi	+/-
Non-controlling interest	**mi**		+/-
Net Equity Issuance	**nei**		+/-
Other Comprehensive Income	oci	eq-l_eq-retain-nei	+/-

Calculated items

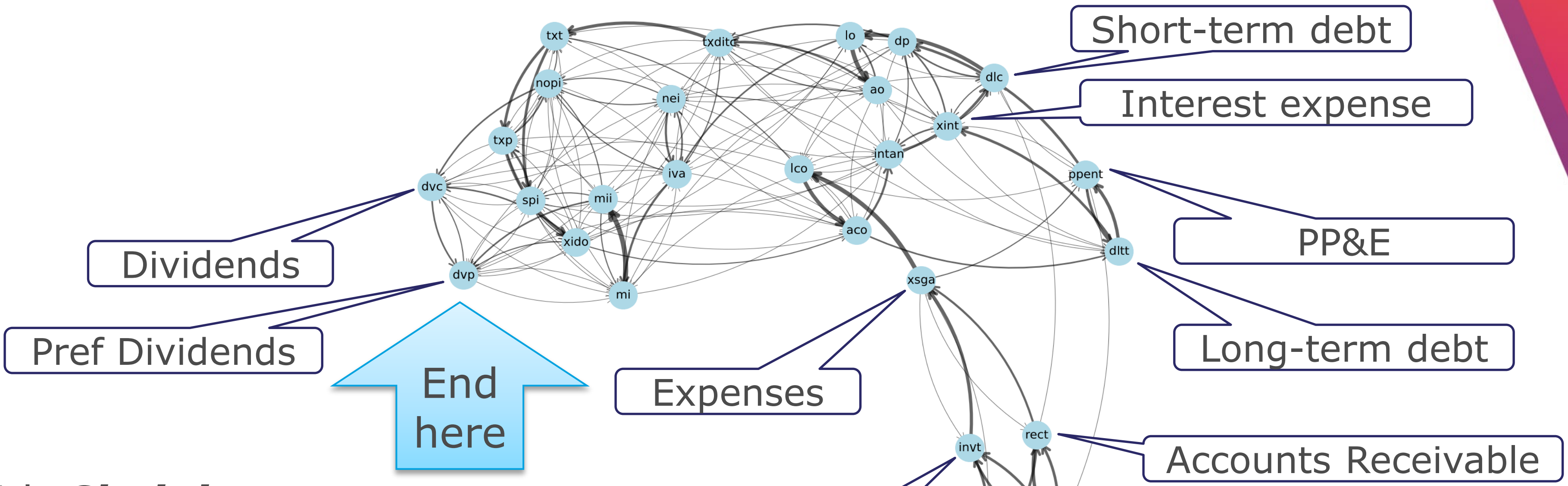
HIGH INTERDEPENDENCIES AMONG LINE ITEMS

	sale	cogs	xsga	dp	xint	nopi	spi	txt	mii	dvp	xido	dvc	rect	invt	aco	ppent	iva	intan	ao	dlc	ap	txp	lco	dltt	txditc	lo	mi	nei
sale	100	97	75	61	56	33	-25	58	23	7	-1	51	72	80	54	58	36	49	45	37	85	39	70	59	36	56	25	-45
cogs	97	100	59	48	46	27	-20	45	20	6	-0	37	65	76	44	49	28	37	37	31	85	30	58	47	30	48	22	-34
xsga	75	59	100	61	52	31	-28	59	14	7	-0	59	57	68	59	47	36	54	44	38	57	42	72	59	25	51	17	-48
dp	61	48	61	100	74	44	-30	54	24	12	-2	58	59	46	59	81	40	59	48	43	47	48	68	75	57	62	30	-44
xint	56	46	52	74	100	33	-31	44	28	18	-2	53	57	46	51	67	38	62	51	47	43	39	63	90	55	60	34	-38
nopi	33	27	31	44	33	100	-10	38	18	5	1	37	36	29	36	40	46	28	32	22	28	38	39	35	27	34	22	-24
spi	-25	-20	-28	-30	-31	-10	100	-7	-6	-4	2	-27	-26	-21	-30	-17	-16	-35	-22	-29	-22	-16	-29	-34	-15	-32	-9	13
txt	58	45	59	54	44	38	-7	100	25	3	-0	61	53	50	50	52	39	42	40	29	43	49	58	49	41	45	21	-55
mii	23	20	14	24	28	18	-6	25	100	2	-2	23	27	19	20	22	17	28	23	15	20	21	27	29	18	22	57	-19
dvp	7	6	7	12	18	5	-4	3	2	100	-1	9	8	6	6	11	6	9	10	7	7	8	7	13	5	8	8	-2
xido	-1	-0	-0	-2	-2	1	2	-0	-2	-1	100	-0	-1	-0	-1	-1	-1	1	-1	-4	-1	0	-0	-0	-0	-3	-1	-3
dvc	51	37	59	58	53	37	-27	61	23	9	-0	100	50	45	51	51	43	53	43	40	39	45	64	66	36	55	22	-46
rect	72	65	57	59	57	36	-26	53	27	8	-1	50	100	59	60	46	44	56	54	48	71	44	73	59	31	64	26	-39
invt	80	76	68	46	46	29	-21	50	19	6	-0	45	59	100	45	42	29	39	37	34	72	39	58	48	22	44	20	-37
aco	54	44	59	59	51	36	-30	50	20	6	-1	51	60	45	100	42	41	56	51	41	47	41	74	56	30	56	23	-39
ppent	58	49	47	81	67	40	-17	52	22	11	-1	51	46	42	42	100	35	34	42	35	43	45	54	66	74	53	29	-34
iva	36	28	36	40	38	46	-16	39	17	6	-1	43	44	29	41	35	100	29	35	38	32	34	40	42	26	39	25	-27
intan	49	37	54	59	62	28	-35	42	28	9	1	53	56	39	56	34	29	100	45	51	40	28	69	74	43	57	31	-42
ao	45	37	44	48	51	32	-22	40	23	10	-1	43	54	37	51	42	35	45	100	36	39	36	58	53	29	64	26	-31
dlc	37	31	38	43	47	22	-29	29	15	7	-4	40	48	34	41	35	38	51	36	100	32	25	46	49	29	46	21	-24
ap	85	85	57	47	43	28	-22	43	20	7	-1	39	71	72	47	43	32	40	39	32	100	32	55	47	27	45	22	-34
txp	39	30	42	48	39	38	-16	49	21	8	0	45	44	39	41	45	34	28	36	25	32	100	42	36	34	37	21	-26
lco	70	58	72	68	63	39	-29	58	27	7	-0	64	73	58	74	54	40	69	58	46	55	42	100	71	32	76	27	-50
dltt	59	47	59	75	90	35	-34	49	29	13	-0	66	59	48	56	66	42	74	53	49	47	36	71	100	53	63	35	-48
txditc	36	30	25	57	55	27	-15	41	18	5	-0	36	31	22	30	74	26	43	29	29	27	34	32	53	100	31	22	-28
lo	56	48	51	62	60	34	-32	45	22	8	-3	55	64	44	56	53	39	57	64	46	45	37	76	63	31	100	25	-35
mi	25	22	17	30	34	22	-9	21	57	8	-1	22	26	20	23	29	25	31	26	21	22	21	27	35	22	25	100	-16
nei	-45	-34	-48	-44	-38	-24	13	-55	-19	-2	-3	-46	-39	-37	-39	-34	-27	-42	-31	-24	-34	-26	-50	-48	-28	-35	-16	100

LEVERAGE THE INTERDEPENDENCIES: CHAINED PREDICTIONS

- Step 1: predicting sales
- Step 2: given the predicted sales, what should be the predicted costs of goods sold?
- …(the chaining method!)

Predict each core item in turn

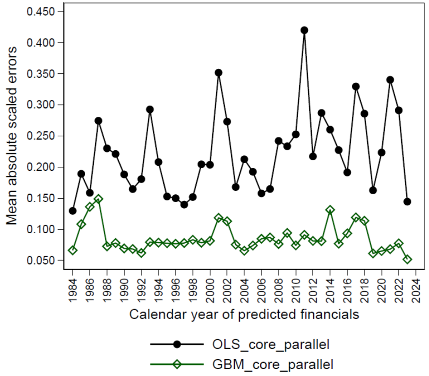


With ***Chaining***

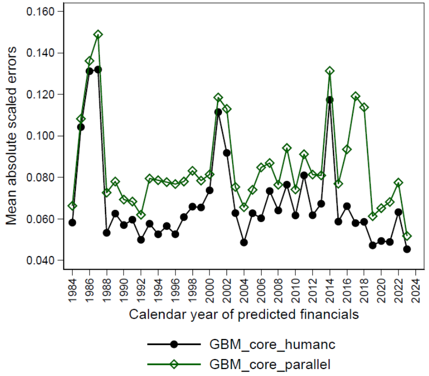
(Use previously predicted items as features for the next prediction)

AI/DATA-DRIVEN + HUMAN ACCOUNTING KNOWLEDGE

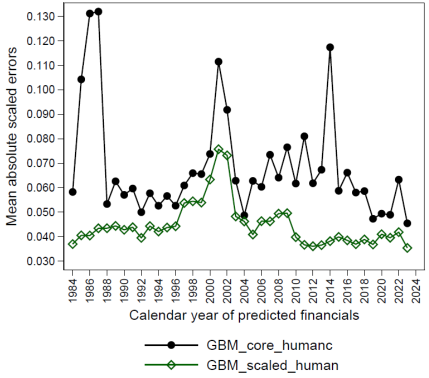
A: OLS vs GBM



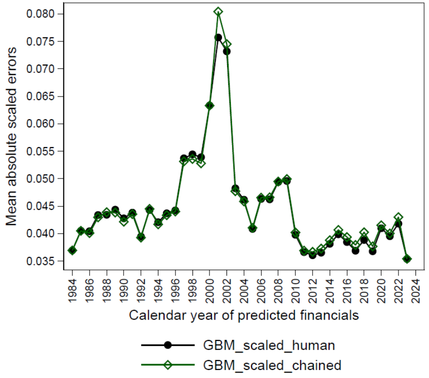
B: Human-chaining vs parallel



C: Core vs Scaled (Levels vs Ratios)



D: Human chaining vs data chaining



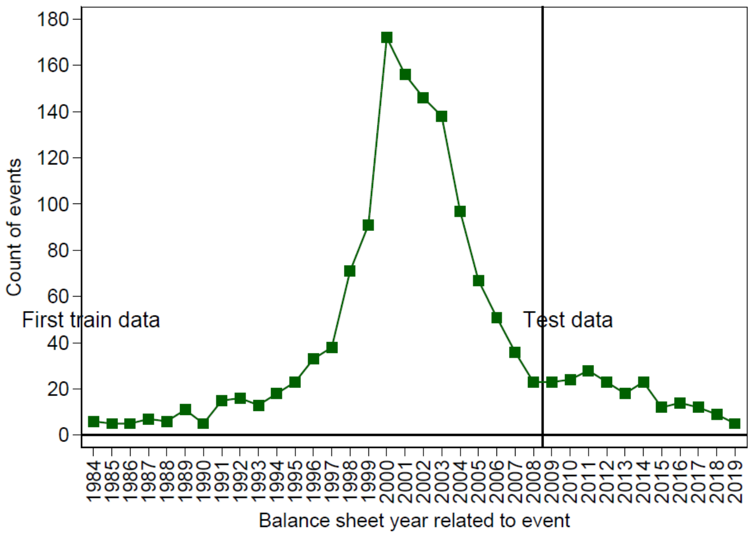
DOES "CHAINING" MAKE A DIFFERENCE? *PERFORMANCE "GAP"*

- Gap: reduction in errors moving from parallel to chained predictions
- Matters more for difficult to predict firms (small, distressed, loss-making, low-investment, young, high sales=prediction error)

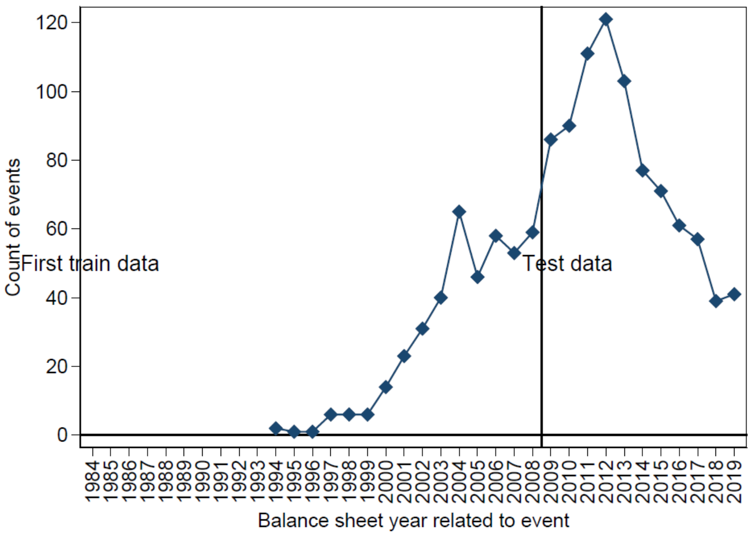
Quintile	Size	B2M	ROE	Investment	Age	SalesError
1	33.81	13.62	20.97	20.93	15.55	13.53
2	18.65	12.06	16.75	15.08	16.51	11.75
3	10.58	12.66	13.01	12.38	16.62	11.77
4	6.90	15.11	10.23	12.24	15.44	14.53
5	2.78	18.38	10.88	12.13	8.19	21.20

CAN PREDICTED FINANCIAL STATEMENTS ANTICIPATE ACCOUNTING IRREGULARITIES?

A: Severe acc. irreg. (AAER and BigR)



B: LittleR



OVERVIEW OF THE PREDICTION METHODS AND COLLABORATION

1. Raw data → **2. Prediction method** → **3. Process** → **4. Variable (scaling)**

1. Raw data

Compustat
28 Core Items

- Cleaning: $X_t^{initial}$
- $X_t^{initial}$ + augmentation
  + ma, cs, d_ma, d_cs
  + FF12
  = X_t
- Targets: Y_{t+1}
  = $X_{t+1}^{initial}$-$X_t^{initial}$

2. Prediction method

OLS

GBMs

3. Process

Parallel

- $X_t \to Y_{1,t+1}$
- $X_t \to Y_{2,t+1}$
- $X_t \to Y_{3,t+1}$
- .
- .
- .
- $X_t \to Y_{28,t+1}$

Chained (ordered by human or data)

- $X_t \to Y_{1,t+1}$
- $X_t + Y_{1,t+1} \to Y_{2,t+1}$
- $X_t + Y_{1,t+1} + Y_{2,t+1} \to Y_{3,t+1}$
- .
- .
- .
- $X_t + \cdots + Y_{27,t+1} \to Y_{28,t+1}$

4. Variable (scaling)

- Features: $X_t / Assets_{t-1}$
- Targets: $Y_{t+1} / Assets_t$

5. Training

(1st Training window: 1963-1982)

6. Prediction

(1st out-of-sample prediction: 1983)

Parallel

- $X_t \to \hat{Y}_{1,t+1}$
- $X_t \to \hat{Y}_{2,t+1}$
- $X_t \to \hat{Y}_{3,t+1}$
- .
- .
- .
- $X_t \to \hat{Y}_{28,t+1}$

Chained

- $X_t \to \hat{Y}_{1,t+1}$
- $X_t + \hat{Y}_{1,t+1} \to \hat{Y}_{2,t+1}$
- $X_t + \hat{Y}_{1,t+1} + \hat{Y}_{2,t+1} \to \hat{Y}_{3,t+1}$
- .
- .
- .
- $X_t + \cdots + \hat{Y}_{27,t+1} \to \hat{Y}_{28,t+1}$

Run	Method	Process	Variable	Order
1	OLS*	Parallel	Core	n/a
2	GBM	Parallel	Core	n/a
3	GBM	Chained	Core	Human
4	GBM	Chained	Scaled	Human
5	GBM	Chained	Scaled	Data

CLASSIC AI (TREE-BASED MODEL) USE CASE: EARLY-WARNING SIGNAL FOR CEO DISMISSALS

Predicting CEO dismissals using Artificial Intelligence

Henk Berkman,* Helen Lu† and Juebin Zeng‡

9th June 2025

Abstract

This study employs Artificial Intelligence (AI) to predict performance-induced CEO dismissals. In out-of-sample tests, our AI-based approach consistently outperforms classic benchmark models (Bushman, Dai and Wang, 2010; Jenter and Kanaan, 2015) across multiple metrics. Using over 70 candidate predictors, we find that firm-specific performance and risk measures account for a large portion of the model's predictive power, while industry-peer returns and risk measures make a limited contribution to the prediction. Dismissals deemed unexpected by the AI model are followed by significant negative stock price reactions, whereas those anticipated by the model do not trigger such responses. In contrast, the predictions of classic models do not align with market responses. CEOs that the AI predicts should be removed often remain in position and continue to deliver subpar performance. Firms led by CEOs with high AI-predicted dismissal probabilities experience significantly lower industry-adjusted return-on-assets over the following year. A trading strategy based on AI-generated signals earns abnormal returns exceeding 0.75 percentage points per month. These effects are concentrated in well-governed firms, suggesting that investors in such firms may underestimate frictions in the CEO labour market. [183 words]

Keywords: CEO turnovers; Corporate Governance; Artificial Intelligence; Machine Learning; Abnormal returns

JEL codes: C10, C45, G30, M12

CLASSIC AI (TREE-BASED MODEL) USE CASE: EARLY-WARNING SIGNAL FOR CEO DISMISSALS

Machine prediction

Reality

+

(OPTIONAL) CLASSIC AI USE CASE: *HOW MANY INVESTOR STRATEGIES? (CLUSTERING)*

Journal of Banking and Finance 119 (2020) 105934

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

The correlation structure of anomaly strategies☆

Paul Geertsema[a,b,*], Helen Lu[a,b]

[a] Department of Accounting and Finance, University of Auckland Business School, Owen G Glenn Building, 12 Grafton Road, Auckland, New Zealand
[b] The University of Auckland Business School, Private Bag 92019, Auckland 1142, New Zealand

ARTICLE INFO

Article history:
Received 9 September 2019
Accepted 12 August 2020
Available online 15 August 2020

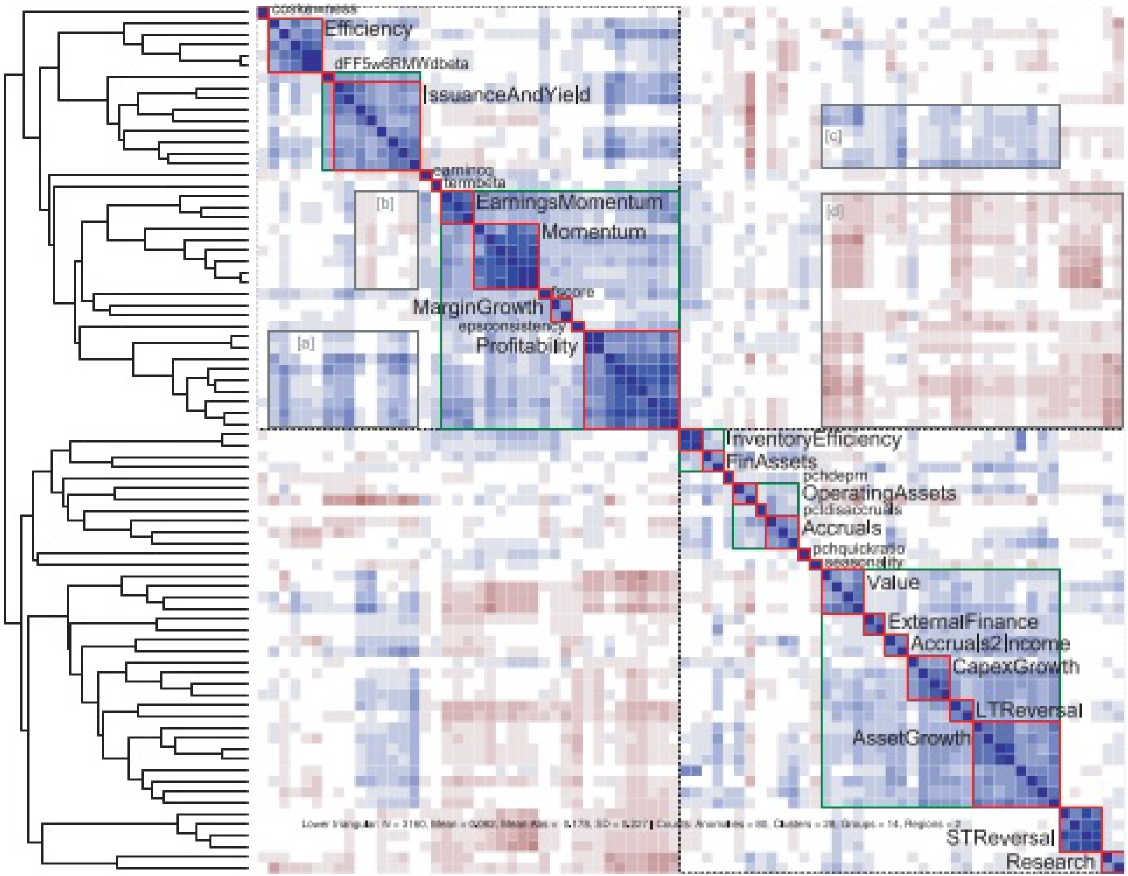
JEL classification:
G12
C38

Keywords:
Anomalies
Correlation
Cluster analysis
Machine learning
Asset pricing

ABSTRACT

We consolidate a large number of mean-significant anomalies into cluster portfolios. More than a third of cluster portfolios remain significant under the Hou et al. (2020) five-factor model – the best performing among six benchmark models tested. A best-first search yields nine factors that subsume *all* cluster portfolios as well as *all* significant anomalies, demonstrating the feasibility of a parsimonious description of average realised returns. The expected growth factor (EG) and a cluster portfolio linked to accruals are prominent factors that improve pricing performance. The search-generated model produces a monthly maximum squared Sharpe ratio of 0.51, considerably higher than current benchmark models.

CLASSIC AI (UNSUPERVISED) USE CASE:
HOW MANY INVESTOR STRATEGIES? (CLUSTERING)



(OPTIONAL) LLM USE CASE: THE VALUE OF BEING SPECIFIC

Specifics matter:

An analysis of mutual fund ESG disclosures

Huayu Shi[a], Xing Han[a], John B. Lee[a], and Helen Lu[b,c]

[a]Accounting and Finance Department, The University of Auckland
[b]Area Accounting and Finance, Vlerick Business School
[c]Department of Accounting, Finance and Insurance, Faculty of Economics and Business, KU Leuven

2nd June 2025

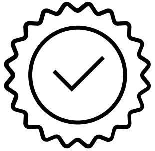
Abstract

This study employs a Large Language Model to quantify the specificity of mutual funds' Environmental, Social, and Governance (ESG) disclosures. We find that funds with more specific disclosures attract greater investor flows, particularly among institutional-oriented funds and during periods of heightened climate concern. Disclosure specificity is also positively associated with the ESG performance of future fund holdings. These findings suggest that specific ESG disclosures serve as credible signals of a fund's ESG commitment, enabling investors to make more informed, sustainability-conscious decisions. The results offer empirical support for recent regulatory efforts to promote more structured and detailed ESG reporting. [98 words]

JEL Codes: G11, G23
Keywords: Mutual Fund Disclosures; ESG; Fund Flows; LLM; Specificity

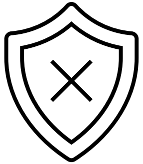
LLM USE CASE:
THE VALUE OF SPECIFIC ESG STRATEGIES

A specific ESG strategy



"…may invest in companies that focus on lowering the cost of healthcare, combatting the opioid epidemic, or offering ethically sourced products."

A generic ESG strategy



"The investment team may also consider the risks and return potential presented by environmental, social, and governance (ESG) factors in investment decisions."

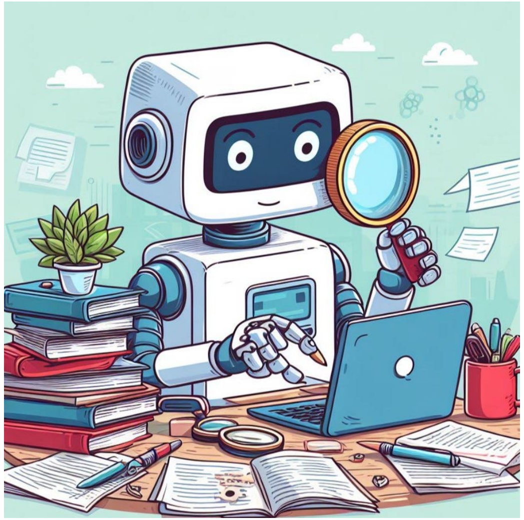
Shi, Han, Lee and Lu (2024)

SELF-SUPERVISED LEARNING: HOW DOES AI READ AND WRITE?

- He loves **durian** juice.
- He loves **orange** juice.
- He loves ____ juice.

	Fruit	**Human**	**Tropical**
Orange	1	0	0.1
Durian	1	0	0.8
Boy	0	1	0
Mother	0	1	0

PRETRAINED LLMS: MISSION IMPOSSIBLE BECOMES POSSIBLE



Young business man with his face passing through the screen of a laptop on binary code background
GETTY

Thank you!

helen.lu@vlerick.com

https://www.linkedin.com/in/helen-lu-52760a15/

Journal of Banking and Finance 119 (2020) 105934

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

The correlation structure of anomaly strategies☆

Paul Geertsema[a,b,*], Helen Lu[a,b]

[a] Department of Accounting and Finance, University of Auckland Business School, Owen G Glenn Building, 12 Grafton Road, Auckland, New Zealand
[b] The University of Auckland Business School, Private Bag 92019, Auckland 1142, New Zealand

ARTICLE INFO

Article history:
Received 9 September 2019
Accepted 12 August 2020
Available online 15 August 2020

JEL classification:
G12
C38

Keywords:
Anomalies
Correlation
Cluster analysis
Machine learning
Asset pricing

ABSTRACT

We consolidate a large number of mean-significant anomalies into cluster portfolios. More than a third of cluster portfolios remain significant under the Hou et al. (2020) five-factor model – the best performing among six benchmark models tested. A best-first search yields nine factors that subsume *all* cluster portfolios as well as *all* significant anomalies, demonstrating the feasibility of a parsimonious description of average realised returns. The expected growth factor (EG) and a cluster portfolio linked to accruals are prominent factors that improve pricing performance. The search-generated model produces a monthly maximum squared Sharpe ratio of 0.51, considerably higher than current benchmark models.

Journal *of* Accounting Research

DOI: 10.1111/1475-679X.12464
Journal of Accounting Research
Vol. No. October 2022
Printed in U.S.A.

CHICAGO BOOTH

Relative Valuation with Machine Learning

PAUL GEERTSEMA* AND HELEN LU*

Received 30 November 2020; accepted 8 September 2022

ABSTRACT

We use machine learning for relative valuation and peer firm selection. In out-of-sample tests, our machine learning models substantially outperform traditional models in valuation accuracy. This outperformance persists over time and holds across different types of firms. The valuations produced by machine learning models behave like fundamental values. Overvalued stocks decrease in price and undervalued stocks increase in price in the following month. Determinants of valuation multiples identified by machine learning models are consistent with theoretical predictions derived from a discounted cash flow approach. Profitability ratios, growth measures, and efficiency ratios are the most important value drivers throughout our sample period. We derive a novel method to express valuation multiples predicted by our machine learning models as weighted averages of peer firm multiples. These weights

HARVARD BUSINESS REVIEW

AI Can Make the Relative-Valuation Process Less Subjective

by Paul Geertsema, Helen Lu and Kristof Stouthuysen