ARTIFICIAL INTELLIGENCE, DATA SCIENCES, AND OPTIMIZATION IN ECONOMICS AND FINANCE

PANOS M. PARDALOS, DISTINGUISHED EMERITUS PROFESSOR UNIVERSITY OF FLORIDA

WWW.ISE.UFL.EDU/PARDALOS



What is (and is not) Al

Technology has allowed us to overcome the biological limits on the concentration of power inherent in the limits of the human eyes, ears and brains [microscopes, telescopes, sensors, computing machines, machine-brain interface, augmented/virtual reality, augmented intelligence, etc].

Artificial intelligence (A.I.) is a term used to describe machines performing human-like cognitive activities such as learning, understanding, reasoning, and interacting [prefer the term augmented intelligence].

As with Electricity is the field of fields (Thomas Edison 1901): "It holds the secrets which will reorganize the life of the world." Al is a new field of fields.

What is (and is not) Al

Al could generate incredible benefits for the world's economy and welfare. However, there are also concerns that Al—like any technology—could create new challenges and exacerbate existing problems

Al empires [https://science.sciencemag.org/content/372/6539/246/tab-pdf]

Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence (*Kate Crawford* Yale University Press, 2021)

A BRIEF HISTORY OF ARTIFICIAL INTELLIGENCE What It Is, Where We Are, and Where We Are Going Michael Wooldridge Flatiron Books (2021) https://us.macmillan.com/books/9781250770738

Status of Al-related **patents**:

World Intellectual Property Organization (<u>WIPO</u>): WIPO Technology Trends 2019 – Artificial Intelligence

https://www.wipo.int/publications/en/details.jsp?id=4386 Most filing patents are made in the USA and China

New trend- AI is a nationalized priority:

US, Canada, China, France, Europe, all have national strategies for AI Economy

USA government reports:

- 1. Artificial Intelligence, Automation, and the Economy;
- 2. Preparing for the Future of Artificial Intelligence;
- 3. The National Artificial Intelligence Research and Development Strategic Plan
- 4. National Security Commission on Artificial Intelligence (NSCAI) March 1, 2021 <u>https://www.nscai.gov/</u>

"The successful adoption of AI in adjacent fields and technologies will drive economies, shape societies, and determine which states exert influence and exercise power in the world."

Al is "the quintessential 'dual-use' technology," the report notes. "The ability of a machine to perceive, evaluate, and act more quickly and accurately than a human represents a competitive advantage in any field—civilian or military."

European Commission:

https://digital-strategy.ec.europa.eu/en/policies/strategy-artificialintelligence

The European Al Alliance is a forum engaged in a broad and open discussion of all aspects of Artificial Intelligence development and its impact.

https://digital-strategy.ec.europa.eu/en/policies/european-ai-alliance

Chinese government reports:

 Due to its national plan for AI domination and the billions of dollars the government has invested in the field, China is poised to take the lead in AI research and usage.

China introduced its Artificial Intelligence Development Plan (AIDP) in 2017 to make China an AI superpower by 2030, surpassing its rivals to become "the world's premier artificial intelligence innovation center".

See: <u>https://techwireasia.com/2021/03/is-china-beating-the-us-to-ai-supremacy/</u> (March 2021) See also article: <u>https://www.thenation.com/article/world/china-ai-military/</u>

World Intellectual Property Report 2022 (wipo.int)

Ben Buchanan and Andrew Imbrie

The New Fire

War, Peace, and Democracy in the Age of AI

RESTARTING THE FUTURE How to Fix the Intangible Economy JONATHAN HASKEL **STIAN WESTLAKE**

Digital Money?

What about cryptocurrencies?

AI and blockchain: Mining the future

Data is the future of business, but complex new technologies such as artificial intelligence and blockchain seem to muddy the waters. Are they really bringing benefits?

https://www.springerprofessional.de/en/introduction-to-blockchain-technology/19368450

How do you quantify risk?

Some of our related work

Advances in machine learning/data sciences and AI are progressing rapidly and demonstrating the potential to transform our lives.

The spectacular success of these areas relies in part in their sophisticated mathematical underpinnings (e.g. optimization techniques and operations research tools), even though this crucial aspect is often downplayed.

We worked in developing AI based approaches with applications in:

Agriculture Energy Medicine Manufacturing Dynamics of Financial Markets (market graph)

Objectives of this Research joint research with G. Adosoglou, G. Lombardo, and S. Park

- Can neural network embedding techniques provide a better way to detect semantic changes in companies' annual reports (10-Ks)?
- Do the detected changes have any implications for future firm returns?
- Do portfolios constructed using these signals yield statistically significant risk adjusted returns or are the returns explained by common risk factors?
- Can Network Science techniques be leveraged to also track relative changes and capture systemic signals along with idiosyncratic signals?



Contents

- Neural network embeddings on corporate annual filings for portfolio selection
- Lazy Networks: Avoiding economic shocks
 using word embedding based networks

Contents

- <u>Neural network embeddings on corporate</u> <u>annual filings for portfolio selection</u>
- Lazy Networks: Avoiding economic shocks
 using word embedding based networks

Companies' 10Ks have been increasing in complexity and length

700



Panel B: Textual Changes in 10-Ks



- Dyer et al. (2017) showed that the managers provide boilerplate information and avoid giving accurate signals of the company's status by extending the document length.
- Loughran and Mcdonald (2011) also showed that managers are incentivized to minimize the effect of negative news on stock prices
- The average publicly traded firm has their annual report downloaded from the SEC's website only 28.4. (Loughran and McDonald (2017))

Source: Lazy prices (L. Cohen, 2019)



Lazy prices (L. Cohen, 2019)

- **Context**: Public companies' management are often "lazy" when writing the annual 10K reports for the SEC
 - Routine Reporting: Often copy & paste while also attempting to hide negative information.
 - Investors are becoming increasingly less attentive to 10Ks.
 - Illustrated that tracking textual changes among companies' reports can be used as a signal for future performance
 - Documents are analyzed using the **Bag-of-words model (BoW)**
 - Long-only and long-short portfolios are built by buying each year companies featuring few changes and selling companies with major changes.

• Bag-of-words model (BoW) Limitations:

- BoW does not take into account semantics and words' order
- BoW is computationally expensive because we need a different model for each comparison
- A company could just use different words to say the same thing just to hide information in their reports.
- A change in CEO/CFO could mean a complete 10-K rewrite causing huge issues with the BOW model.

Neural network embeddings on corporate annual filings for portfolio selection

• Dataset:

- O Yearly 10Ks for 6.000 US companies (SEC) from Loughran-McDonald public dataset
- O Analyzed period: 1998-2018 About 120.000 documents (200 Gigabytes)
- O Past returns and companies capitalization from Bloomberg and Wharton research database

• Training:

- Word2vec model: Averaging word embedding to compute document embedding
- O Doc2vec model: Distributed memory mode (PV-DM) and Distributed Bag of Words version (PV-DBOW)
- O Bag-of-words: As a comparison with lazy prices

• Hyper-parameters selection:

- O Epochs and dimension according to Baldwin et al, 2016
- Window size, sampling threshold and mincount with grid search on a validation set of well-known cases of changes among documents (e.g., Baxter inc from lazy prices)

• Portfolios creation:

- Similar to Lazy prices, 3 long-only portfolios with BOW and the two modes of Doc2Vec (PV-DM,PV-DBOW)
- O A long-only portfolio that also incorporates a momentum (ret(t-12)) criterion
- Portfolios' performance comparison:
 - O Comparison of all 3 models in capturing the changes associated with future risk adjusted excess returns
 - (both equally and value weighted)
 - Comparison of the best performing model (PV-DM Doc2Vec) with the one that also incorporates momentum
 - O Comparison with the Lazy prices paper's performance

Representation Learning

- Machine learning algorithms learn from a set of observations $oldsymbol{\chi}$ in the form of real-valued vectors.
- For words and documents, being unstructured data, this representation is not always directly available and an encoding phase is required.

-> **Representation Learning**: Learning this encoding directly from data with neural networks

 \circ Training leads to a latent representation of $oldsymbol{\chi}$ at each layer

O Representation in the output layer is called: embedding



Source: Zhiyuan Liu 2020



Word2Vec and Doc2Vec

- *Word2Vec* is trained to complete surrounding words in corpus.
- In the figure: The skip-gram model of Word2Vec: Both the input vector x and the output y are one-hot encoded word representations. The hidden layer is the word embedding of size N. (T. Mikolov et al, 2013)
- In *Doc2Vec*, the paragraph vectors are also asked to contribute to the prediction task of the next word given many contexts sampled from the paragraph (L. Quoc et al, 2014)



Doc2Vec (PV-DM and PV-DBOW modes)



Distributed Bag ofWords version of Paragraph Vector (PV-DBOW)



The paragraph vector is trained to predict the words in a small window.

The concatenation or average of this vector with a context of three words is used to predict the fourth word.

Comparing similarity measures for Baxter's 2009-2010 10-Ks to Lazy Prices



- Baxter inc: Jaccard similarity between 2010 and 2009 10-Ks with BoW – Source: Lazy-Prices
- Issues with FDA reported in 10K long before it was published in the news



 Baxter inc: Cosine similarity among 2010-2009 with Doc2Vec PV-DM, our approach

Portfolio Construction

- Four portoflios were constructed:
- Two equally weighted and two capitalization weighted
- First two portfolios satisfy only one requirement:
 - <u>Cosine similarity >= 0.95 with the previous year 10K</u>
 - Named: Semantic Similarity Portfolio (SSP)
- The other two portfolio satisfy two requirements:
 - <u>Cosine similarity >= 0.95 with the previous year 10K</u>
 - <u>Momentum-related criterion: Returns(t-1) > 0</u>
 - Named Non-struggling Portfolio (NSP)
 - Attempt to avoid companies with persisting risks and difficulties that are documented in the 10-Ks and have not been removed in the most recent filing (leaving the 10-K semantically unchanged)

Portfolios' Risk Adjusted Returns and Factor Loadings

- Alphas: Portfolios' excess returns have been evaluated by running the following regressions:
 - CAPM (Capital Asset Pricing Model)
 - Fama & French Model (Incorporatea startegies based on Value & Size)
 - 5-factor Model (Incorporatea also Momentum & Liquidity startegies)

-> Each model returns an alpha (abnormal return) and statistical significance in terms of p-value and t-stat

 Factor Loadings' Analysis: Measuring the exposures to market, size, value, momentum, liquidity risks. Capital Asset Pricing Model (CAPM): $R_{t} - r_{t}^{f} = \alpha + \beta_{MKT} \left(R_{t}^{M} - r_{t}^{f} \right) + \varepsilon_{t}$ 3-Factor Fama and French Model: $R_{t} - r_{t}^{f} = \alpha + \beta_{MKT} \left(R_{t}^{M} - r_{t}^{f} \right) + \beta_{HML}HML_{t} + \beta_{SMB}SMB_{t} + \varepsilon_{t}$ 5-Factor Model: $R_{t} - r_{t}^{f} = \alpha + \beta_{MKT} \left(R_{t}^{M} - r_{t}^{f} \right) + \beta_{HML}HML_{t} + \beta_{SMB}SMB_{t} + \beta_{UMD}UMD_{t} + \beta_{PS_VWF}PS_VWF_{t} + \varepsilon_{t}$

Performance of the Portoflios in terms of Risk Adjusted Returns

	Value-Weighted			Equally-weighted		
Portfolio	CAPM	3-Factor	5-Factor	CAPM	3-Factor	5-Factor
	Alpha	Alpha	Alpha	Alpha	Alpha	Alpha
SSP	11.04**	7.93**	5.87	10.94**	6.60**	7.45**
t-stat	(2.75)	(2.27)	(1.59)	(2.56)	(2.50)	(2.69)
BoW	3.88	3.84	5.1	5.09	3.14	3.57
t-stat	(1.20)	(1.25)	(1.54)	(1.60)	(1.29)	(1.30)
NSP	11.11***	9.75***	8.45**	9.88**	7.40**	6.28**
t-stat	(3.30)	(3.24)	(2.61)	(2.57)	(2.45)	(2.15)
BoW-Mom	3.92	4.33	1.26	2.16	1.74	-1.35
t-stat	(1.51)	(1.56)	(0.44)	(0.70)	(0.52)	(0.71)

	Value-Weighted			Equally-weighted		
Portfolio	CAPM	3-Factor	5-Factor	CAPM	3-Factor	5-Factor
	Alpha	Alpha	Alpha	Alpha	Alpha	Alpha
Q5	12***	5.28**	5.16**	11.52***	2.88**	2.76**
t-stat	(3)	(2.78)	(2.81)	(3.05)	(2.76)	(2.7)
Q5-Q1	7.68***	8.88***	8.16***	1.92	2.88***	2.28**
t-stat	(3.55)	(4.17)	(3.77)	(1.5)	(2.82)	(2.24)

Cumulative Returns of the Portoflios



Factor Exposures

	Value-Weighted		Equally-Weighted		
Factors	3-Factor	5-Factor	3-Factor	5-Factor	
Intercept (α)	9.75***	8.45**	7.40**	6.28**	
t-stat	(3.24)	(2.61)	(2.45)	(2.15)	
MKTRF	0.48***	0.45**	0.61***	0.59***	
t-stat	(3.30)	(2.66)	(4.23)	(3.34)	
SMB	0.04	-0.19	0.31	0.10	
t-stat	(0.14)	(-0.62)	(1.08)	(0.30)	
HML	0.45***	0.34**	0.61***	0.49**	
t-stat	(2.96)	(1.74)	(3.94)	(2.48)	
UMD	-	-0.12	-	-0.12	
t-stat		(-0.50)		(-0.54)	
PS_VWF	-	0.49*	-	0.46	
t-stat		(1.84)		(1.69)	

- Table illustrates exposure to the market, size, value, momentum and liquidity risks.
- Small betas
- > Statistically significant exposure to HML factor \rightarrow Value tilt

Contents

- Neural network embeddings on corporate annual filings for portfolio selection
- Lazy Networks: Avoiding economic shocks
 <u>using word embedding based networks</u>

Lazy Networks: Avoiding economic shocks using word embedding based networks



The 10-K Network:

- A network constructed using the 10K embeddings for 2010
- Nodes: Companies
- Edges: cosine similarity between the Doc2Vec 10K embeddings
- Edge threshold = 0.6
- Colored by industries (Standard Industrial Classification – SIC)



Representing Economic Shocks by Network Structure

- Text-Based Network Industries and Endogenous Product Differentiation (Hoberg and Phillip, 2016)
- In text-based networks (from 10-K's using the BOW model)
 → Positive and negative demand shocks can be reflected in the total similarity of the companies that operate in specific industries (i.e. the weighted degree centrality or Strength)
- Positive demand shocks increase the similarity of rivals' products September 11 attacks → need to produce similar products -> increase in the total similarity of the firms' 10-Ks'
- <u>Negative demand shocks tend to make peers more dissimilar</u>
 2000 software market collapse → need to pivot and produce different products → decrease in the total similarity neighboring software firms

Q: How can we utilize neural networks and graph theory techniques to better measure the relative changes of companies in the markets? → Lazy Networks

Proposed Process of constructing 'Lazy Network'

Computing cosine similarity between Doc2Vec vectors for all companies Constructing the **10-K Network**, where edges represent the cosine similarities Constructing Lazy Network by taking the absolute difference between 10-K Networks for successive years

Computing centrality to measure economic shock Creating the portfolios selecting the most peripheral stocks in each case.

Choosing the Threshold (Slicing cut-off)



- Node degree distributions need to follow the power law
- Kolmogorov–Smirnov (KS) test: to choose lower threshold on node degree
- The resultant threshold chosen for the analysis = 0.4

Centrality Measures

Strength Centrality

The strength of node i, s_i , in the network is computed by summing the weights of its adjacent edges, namely:

$$s_i = \sum_{j \in N(i)} w_{ij}, \ \forall i \in V$$

where N(i) is the set of neighbor nodes of node *i*.

Closeness Centrality

Closeness centrality of node *i*, c_i , is thus defined as the reciprocal of the average shortest path distance to node *i* over all n - 1 reachable nodes:

$$c_i = \frac{n-1}{\sum_{j=1, j \neq i}^n d(i, j)}$$

d(i, j) is the shortest-path distance between node *i* and *j*.

Eigenvector Centrality (Bonacich, 1987)

The eigenvector centrality of a node *i* corresponds to the *i*th element of the eigenvector **v** that is associated with the maximum eigenvalue, λ , of the adjacency matrix $A \in \mathbb{R}^{n \times n}$. The elements of the adjacency matrix are the weighted edges between nodes. The eigenvector centrality is defined:

$$A\mathbf{v} = \lambda \mathbf{v}.$$

Betweenness Centrality (Freeman, 1978)

Betweenness centrality of node *i*, is defined as the total number of shortest paths between nodes that pass through node *i*:

$$b_i = \sum_{s,t \in V} \frac{\sigma(s,t \mid i)}{\sigma(s,t)}$$

where $\sigma(s, t)$ is the number of shortest paths between node *s* and *t*, and $\sigma(s, t|i)$ is the number of those paths passing through node *i*.

The Lazy Network Visualization



The Lazy Network for the year 2014

- Nodes represent companies
- Edges represent the year-on-year absolute change in the cosine similarity between the companies' 10-K vector representations
- Nodes are colored based on the different centrality measures
- Node size represents the degree.

Mean \pm Standard Deviation for the correlations of the network centrality measures for the years from 2000 - 2018.

Variable	Strength	Eigenvector	Closeness	Betweenness
Strength	1.000			
Eigenvector	0.81±0.07	1.000		
Closeness	0.81±0.04	0.88±0.06	1.00	
Betweenness	0.59±0.09	0.48±0.09	0.65±0.04	1.00

Portfolio Returns Comparison

	Equal-weighted			Value-weighted		
Portfolio	CAPM	3-Factor	5-Factor	CAPM	3-Factor	5-Factor
Fortiono	Alpha	Alpha	Alpha	Alpha	Alpha	Alpha
Strength	1.05***	0.89***	0.94***	0.84***	0.76***	0.87***
t-stat	(4.21)	(3.94)	(4.38)	(3.51)	(3.23)	(3.75)
Eigenvector	1.07***	0.92***	0.96***	0.64***	0.56**	0.62**
t-stat	(4.57)	(4.35)	(4.69)	(2.62)	(2.30)	(2.52)
Closeness	0.84***	0.70***	0.72***	0.94***	0.90***	0.92***
t-stat	(4.32)	(4.03)	(3.80)	(3.97)	(4.14)	(4.22)
Betweenness	0.71***	0.52***	0.51***	0.19	0.14	0.13
t-stat	(4.08)	(4.13)	(4.07)	(1.08)	(0.83)	(0.78)
Lazy Prices (Cohen et al. (2019))	0.96***	0.24***	0.23***	1.00***	0.44***	0.43***
t-stat	(3.05)	(2.76)	(2.70)	(3.00)	(2.78)	(2.81)
SSP (Adosoglou et al. (2021))	0.87**	0.53**	0.60**	0.87**	0.64**	0.48
t-stat	(2.56)	(2.50)	(2.69)	(2.75)	(2.27)	(1.59)

- All suggested portfolios yield large and statistically significant alphas
- All these portfolios use similar rebalancing as well as similar datasets for input.

Cumulative Returns



- Equally weighted (left) and value weighted (right) from 2000 2019
- Compared with buying and holding the S&P 500 index, as well as the whole universe of the available stocks portfolio (Universe of Stocks) and the CRSP market index.
- CRSP market index : market returns from WRDS (Wharton Research Data Services)



Factor Exposures

- For the long-only portfolio that goes long the 50 companies with the lowest closeness centrality in the Lazy Network
- We observed statistically significant low betas in all portfolios
- A value tilt across the portfolios (HML)
- A negative relation to the momentum factor (UMD)
 → lower risk: which is in line with expectations given the defensive nature of these strategies

	Equal-w	veighted	Value-weighted		
Factors	3-Factor	5-Factor	3-Factor	5-Factor	
Intercept (α)	0.70***	0.72***	0.90***	0.92***	
t-stat	(3.80)	(3.97)	(4.14)	(4.22)	
MKTRF	0.97***	0.90***	0.87***	0.81***	
t-stat	(21.71)	(18.92)	(16.55)	(14.30)	
SMB	0.52***	0.52***	-0.03	-0.03	
t-stat	(7.28)	(7.40)	(-0.35)	(-0.30)	
HML	0.26***	0.24***	0.17**	0.15**	
t-stat	(4.38)	(3.99)	(2.32)	(2.01)	
UMD	-	-0.13***	-	-0.11**	
t-stat		(-3.41)		(-2.38)	
PS_VWF	-	0.07	-	0.06	
t-stat		(1.32)		(0.97)	

Our strategies avoid high beta, high growth, high momentum stocks in rapidly changing industries.

Recent related papers

George Adosoglou, Gianfranco Lombardo, Panos M. Pardalos: Neural network embeddings on corporate annual filings for portfolio selection. Expert Syst. Appl. 164: 114053 (2021)

George Adosoglou, Seonho Park, Gianfranco Lombardo, Stefano Cagnoni, Panos M. Pardalos: Lazy Network: A Word Embedding-Based Temporal Financial Network to Avoid Economic Shocks in Asset Pricing Models. Complex. 2022: 9430919:1-9430919:12 (2022)

Gianfranco Lombardo, Mattia Pellegrino, George Adosoglou, Stefano Cagnoni, Panos M. Pardalos, Agostino Poggi: Machine Learning for Bankruptcy Prediction in the American Stock Market: Dataset and Benchmarks. Future Internet 14(8): 244 (2022)

Advantages of the Strategy

- The model can also be used as a signal or alert, helping identify emerging risk in specific sectors or industries
- Yields up to 95 basis points in monthly five factor alphas (over 12% annually)
- Slow strategy → difficult to be eroded/ arbitraged away
 i.e. not a strategy that depends on acting quickly.
- Can also be applied to:
 - o quarterly reports (10-Q)
 - quarterly earnings conference-call transcripts
 - foreign company's annual report (20-F)
 - proxy statements
- Lots of other approaches using neural networks and graph theory techniques could be utilized to improve performance

What is next for AI?

"It is expected in an AI-enhanced future, humans will become better at everything; they'll also become safer, less vulnerable to danger."

"Technology is transforming how humans and machines work together."

> Source: https://www.ibm.com/watson/advantage-reports/future-of-artificialintelligence.html

What is next for AI?

In 1951, <u>Alan Turing</u> wrote an article titled *Intelligent Machinery, A Heretical Theory*, in which he proposed that artificial general intelligences would likely "take control" of the world as they became more intelligent than human beings:

"Let us now assume, for the sake of argument, that [intelligent] machines are a genuine possibility, and look at the consequences of constructing them... There would be no question of the machines dying, and they would be able to converse with each other to sharpen their wits. At some stage therefore we should have to expect the machines to take control, in the way that is mentioned in Samuel Butler's "Erewhon"."

Are the risks underappreciated?

What is next for AI and the Economy?

<u>AI and the Economy</u> (by Jason Furman and Robert Seamans), Innovation Policy and the Economy, Vol 19 (2019), pp. 161-191, The University of Chicago Press.

"AI and other forms of advanced automation, including robots and sensors, can be thought of as a general purpose technology that enable lots of follow-on innovation that ultimately leads to productivity growth."

<u>The Wrong Kind of AI? Artificial Intelligence and the Future of Labor</u> <u>Demand</u>, (by Daron Acemoglu and Pascual Restrepo)

Digital Abundance and Scarce Genius: Implications for Wages, Interest Rates, and Growth, (by Seth Benzell and Erik Brynjolfsson)

A history of AI: <u>https://ahistoryofai.com/antiquity/</u>

What is next for AI and the Economy?

Fifth Industrial Revolution:

The combination of humans and machines in the workplace.

However, this is vastly oversimplified and does not even begin to explain the magnitude and complexity of the change.

Excitement and anxiety! The Fifth Industrial Revolution ha

The Fifth Industrial Revolution has the potential to initiate a **new socio-economic era**, creating infinite opportunities for humanity, and for a better planet.

Al in the economy of the fifth industrial revolution



AT THE CUSP OF THE 5TH INDUSTRIAL REVOLUTION

Recent rapid adoption and application of artificial intelligence algorithms — triggered by access to big data and better hardwareprocessing capabilities — are changing the face of blue and white collar jobs.



"Many fail to grasp what they have seen, and cannot judge what they have learned, although they tell themselves they know."

Heraclitus (born *c.* 540 BCE, <u>Ephesus</u>, died *c.* 480)

Questions?