Deep Learning in Real Estate Investment: Myth or Legend?

Richard Saldanha
r.saldanha@qmul.ac.uk

Trends in AI, real estate investing and financing

24 October 2025
CoreLab, Tampere University, Finland

Overview

- What is deep learning?
- Valuation and prediction
  - Property valuation
  - Macro trends
  - Economic forecasting
- Beware the hype
- Comments / Q&A

Source: Financial Times

What is deep learning?

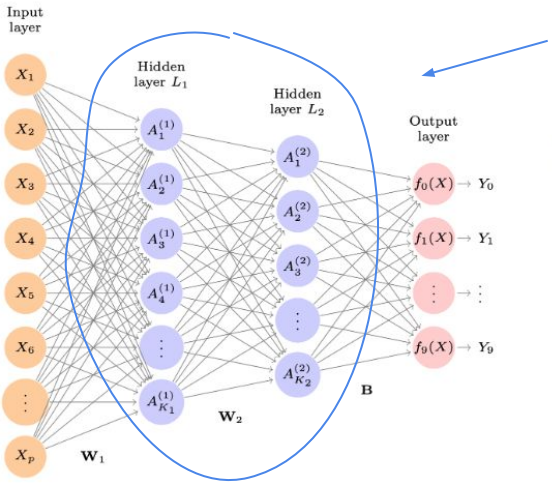


Figure 6.3: Feed-forward NN with two hidden layers and multiple outputs taken from (James *et al.*, 2021). The input layer has $p = 784$ units, the two hidden layers $K_1 = 256$ and $K_2 = 128$ units, and the output layer 10 units. Along with intercepts, constants referred to as 'biases' by NN practitioners, this network has 235,146 parameters or weights. The $\mathbf{W}_1$ (of dimension 785) and $\mathbf{W}_2$ (257) are matrices of weights feeding into the first and second hidden layers – L_1 and L_2, respectively. Finally, $\mathbf{B}$ (129 $\times$ 10) is another matrix of weights feeding into the output layer. (Note $p + 1$, $K_1 + 1$ and $K_2 + 1$ are the first dimensions of $\mathbf{W}_1$, $\mathbf{W}_2$ and $\mathbf{B}$ matrices, respectively, to include the biases.)

If the weights are near zero then the neural network is approximately a linear model. We usually choose starting values for the weights as random values close to zero. Thus, the model starts out nearly linear and becomes non-linear as the weights increase.

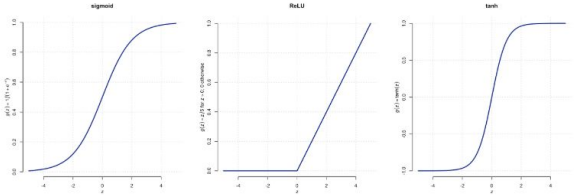
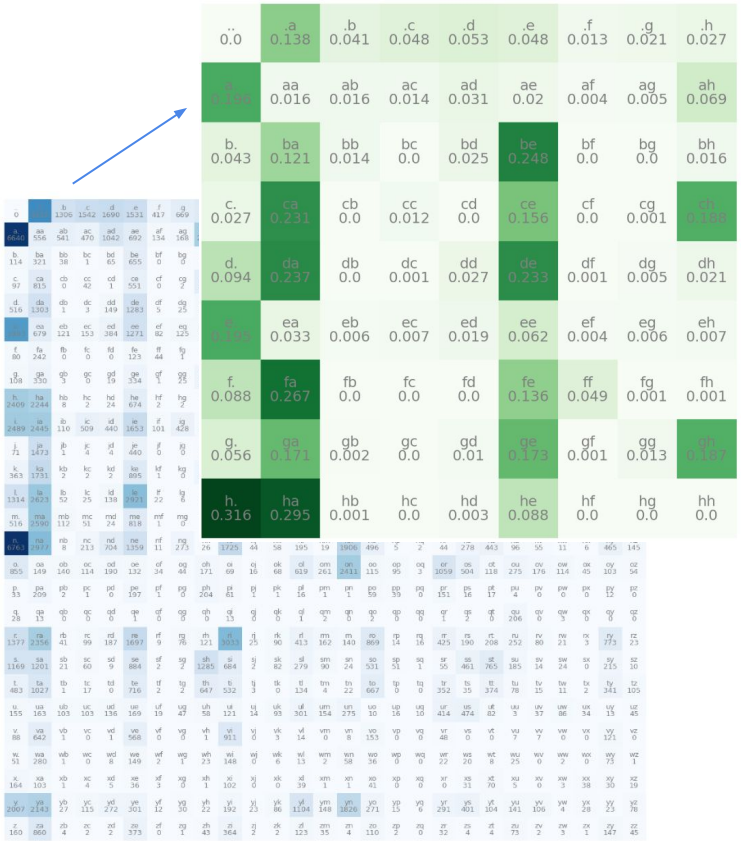


Figure 6.2: Typical NN activation functions: a sigmoid (various forms exist), Rectified Linear Unit (ReLU) scaled by a divisor of five, and tanh.

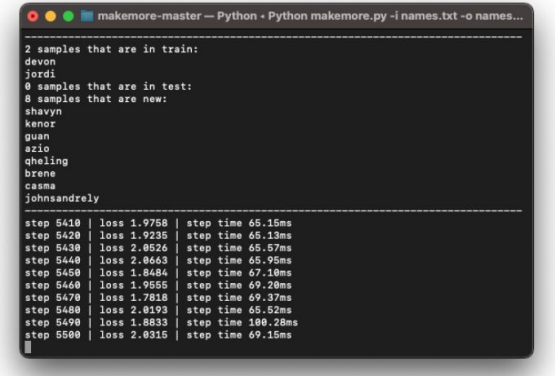
- **Deep learning** refers to the use of artificial neural networks with multiple layers between the input and output layers
- Depth of these models indicates the multiplicity of layers that allows NNs to learn hierarchical representations of data

Queen Mary University of London



Language has structure

- Examining two letter pairs within US baby names dataset
- Letter pair counts can be converted probabilities
- Fit probability model using standard statistical methods
- Fit equivalent and more sophisticated NN models



Queen Mary University of London

Caveat emptor

The problem with neural networks

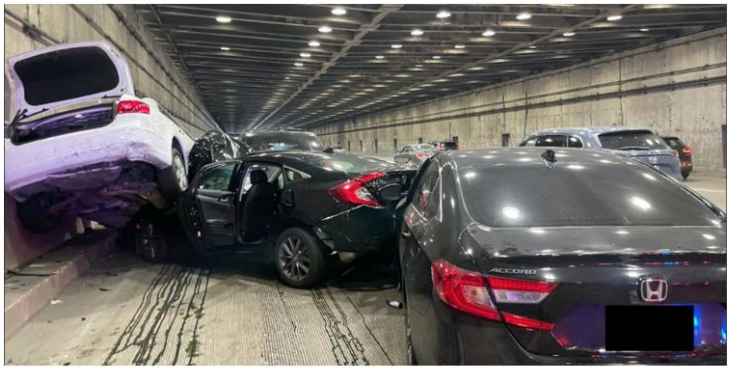


Figure 11.2: An eight-car pileup that occurred on the San Francisco to Oakland Bay Bridge in November 2022 caused by a Tesla in full self-driving mode, which abruptly changed several lanes and braked hard. Thankfully, nobody was seriously injured. Photograph: California Highway Patrol.

Neural-network-based self-driving systems are notorious for 'seeing' phantom objects (hallucinating) in the road; see Cummings (2023). This can be caused merely by shadows or slightly odd atmospheric conditions. When a self-driving car detects an object on the road ahead, even if it's just a shadow, it will apply the brakes. The neural network parameter estimation that goes into building LLMs is much the same as the image recognition system for a self-driving car. An autonomous car has no understanding of itself, the context, or any unobserved factors that a person would consider in a similar situation. Cummings suggests 'that while a language model may give you nonsense, a self-driving car can kill you'.

The downright deceitful

GOOGLE'S BARD

The earliest mention of artificial intelligence (AI) in The New York Times was in 1956, when the term was coined by John McCarthy [1] at a workshop at Dartmouth College. The workshop brought together some of the brightest minds in computer science to discuss the possibility of creating machines that could think like humans.

The Times reported on the workshop in a front-page article, [2] calling it "a new field of science" [3] that was "aimed at creating machines that can think like humans." The article quoted McCarthy as saying [4] that the goal of AI was to "build machines with minds, capable of learning and adapting to new situations."

NEW YORK TIMES ANALYSIS

1. Mr. McCarthy is credited with creating the term artificial intelligence in 1955, in a proposal for the Dartmouth conference.
2. The Times did not cover the Dartmouth workshop at all, let alone on the front page.
3. More recent online articles refer to the conference as starting "a new field of science."
4. The Times could not find a reference to Mr. McCarthy ever saying this phrase, though it includes common ways to describe A.I.

Figure 7.5: Analysis of Google Bard output in NY Times article (Weise and Metz, 2023, May 1).

From valuation to prediction

Property valuation

HouseCanary

- NN image analysis to analyse property photos for US for institutional investors, e.g. REIT providers and buyers
- Underlying NN-based Asset Valuation Model (AVM) assesses property value from condition and quality of finish from images
- Median Absolute Percentage Error (MDAPE) of 3% claimed

Walker & Dunlop

- Combination of NN image analysis and Natural Language Processing (NLP) with focus on commercial real estate
- Combines traditional data (e.g. rent, financials) with alternative data sources such as tenant credit ratings, foot traffic and satellite imagery
- Operates across US and Europe
- MDAPE of 7% claimed for multifamily properties

Zillow's Zestimate

- NN image analysis processes features for over 100 million homes in the US aimed at retail market
- Claims MDAPE error of 7% for off-market homes but more accurate (within 2%) for publicly listed properties
- Results used only as guide **not** as a basis for loan

Macro trends

Privateer and EarthDaily both supply satellite image data and analysis. For real estate investors, this can provide useful macro insights:

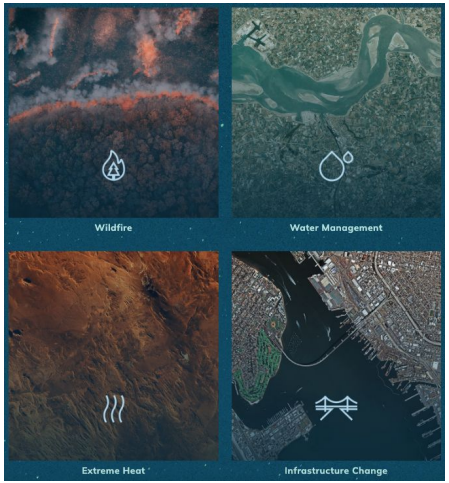


Image taken from EarthDaily website

Queen Mary University of London

- By identifying cranes, new building developments and supporting infrastructure the pace of development in a geography can be used to indicate future supply
- Counting cars parked at major retail centers or industrial parks over time can be used to measure consumer activity and commercial health
- Monitoring the expansion of residential areas to identify long-term growth corridors
- Evidence of climate change (e.g. fire, flooding risk) is an important element of investment decision making

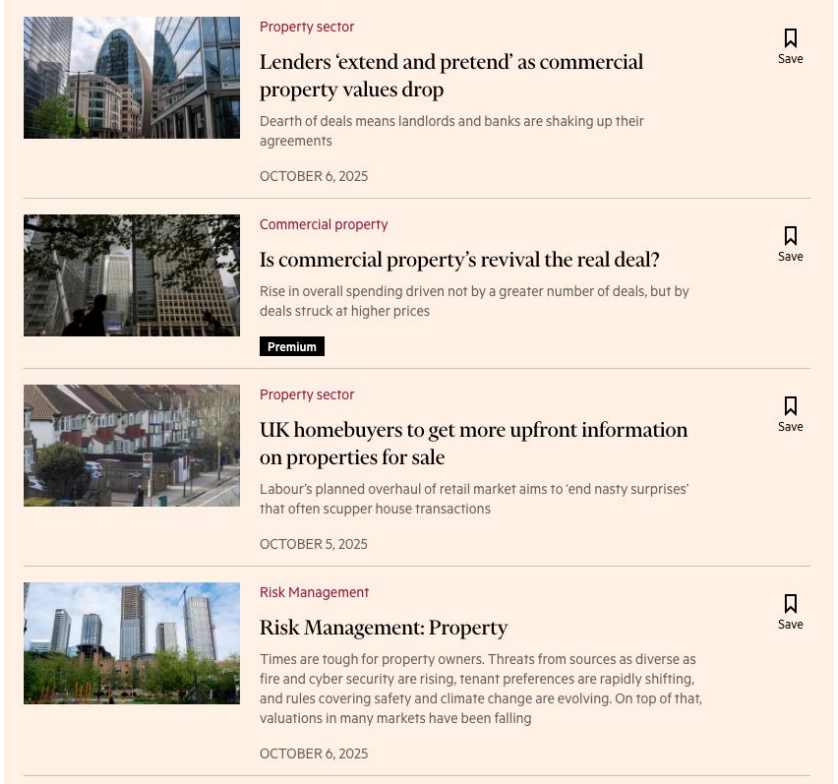
Placer.ai

- Concerned with foot traffic and human mobility analysis
- Anonymized location data from mobile devices, these platforms use machine learning to model population movement.
- This helps investors answer macro-level questions such as:
  - Which neighborhoods are attracting the most new residents?
  - Are suburban shopping centers gaining traffic at the expense of urban ones?

Economic forecasting

Deep learning – e.g. Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) models – often incorporated into large-scale econometric models

- NLP used to scan millions of news articles, planning documents and earnings call transcripts
- Aim to detect early signals of major infrastructure projects, corporate relocations, or shifts in area plans that impact future property values and rental demand
- Nowcasting is the prediction of the very recent past or the present, can take a *ragged edge* dataset (e.g. daily text sentiment, weekly satellite data, monthly employment figures) and combine it all to produce a single, real-time estimate of what some key financial variable is right now



Source: Financial Times

Queen Mary University of London

Beware the hype

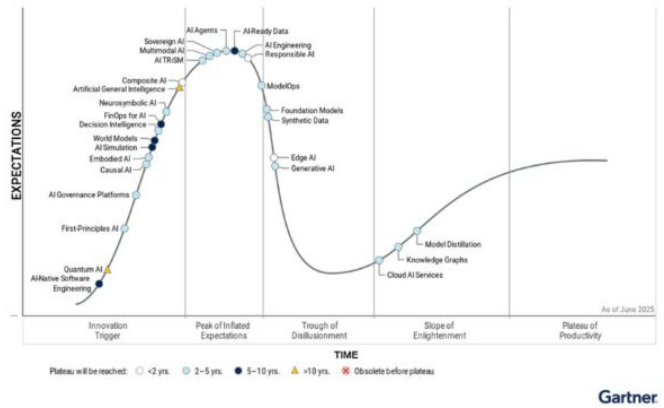
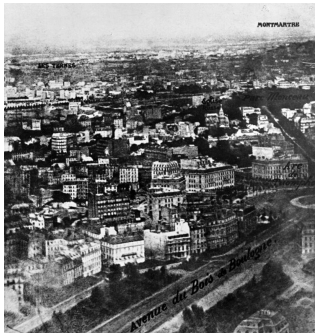


Figure 12.1: The hype cycle for AI in 2025. Source: Gartner August 2025.

- Classical maths/stats and computing still powerful today, it doesn't need to be nor should we pretend it's 'AI'
- Still live in a binary computing world (despite the promise of quantum or even analog

- LLMs are modern day NLP
- Remote sensing began with the first aerial photograph in 1858



French photographer and balloonist Gaspard-Félix Tournachon 'Nadar' first aerial photograph, taken in 1858 above the city of Paris
Source: https://www.onverticality.com/blog/nadar-and-the-aerial-perspective

- Any model can be 'brittle', economic models particularly so

Queen Mary University of London

Final comments / Q&A

- Still need to exercise our brains, i.e. human-in-the-loop is a powerful agent
- Some understanding of what underpins software (AI or other) is often a worthwhile exercise
- Data quality and integrity, as well as having the ‘right’ data are key as always
- Myth or legend – you decide?

kiitos

DISCLAIMER: The information and opinions contained in this document are for background purposes only, do not purport to be full or complete and do not constitute investment advice. No representation, warranty or undertaking, express or implied, is given as to the accuracy or completeness of the information or opinions contained in this document by Queen Mary University of London nor its directors or employees and no liability is accepted by such persons for the accuracy or completeness of any such information or opinions. Furthermore, the document does not constitute or form part of any offer to issue or sell, or any solicitation of any offer to subscribe or purchase, any investment nor shall it or the fact of its distribution form the basis of, or be relied on in connection with, any contract thereof.

Queen Mary University of London (QMUL) is a public research university with a legal structure that includes a governing body called Council, which is responsible for overseeing its strategic mission. The university is a member institution of the federal University of London and a member of the Russell Group of leading UK universities.

Queen Mary University of London
Mile End Road
London E1 4NS