Large Language Models

Strategic Innovation and Artificial Intelligence - Purple Edition

Dr. Philippe J.S. De Brouwer
Honorary Consul of Belgium in Kraków
guest professor at the UJ, AGH, UEK and UW
board member of AGH and ISK
SVP at HSBC in Kraków

2024-05-13

Contents

1 Artificial Intelligence

Artificial Intelligence

Artificial Intelligence (AI) is not new, however, we are in a critical phase where we have (a) computers that are fast enough to fit large models, and (b) we have enough digital content to allow machines to "learn".

An important class of AI models rely on learning from pre-defined good answers: "supervised learning." For example: we show the AI pictures and tell it when there is a cat in it. The AI learns to recognise pictures with cats. Given enough pictures, it seems like the machine recognises cats just like us.

2 Linear Regressions

Linear Regressions – Definition

Linear regressions also "learn" from data. However, there is no lengthy learning process because we can calculate the coefficients.

To fit the linear model

$$y = ax + b$$

we seek a and b so that the differences between the observation y_i and the estimate $\hat{y}_i$ are as small as possible: we minimise the sum of squares:

$$\min_{a,b}\left(\sum_{i=1}^{N}(y_i - \hat{y}_i)^2\right)$$

Linear Regressions – Definition

Solving this, we find the estimate for a and b to be as follows:

$$\hat{a} = \frac{\sigma_{x,y}}{\sigma_x^2} = \rho_{x,y}\frac{\sigma_y}{\sigma_x} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \hat{y})}{\sum_{i=1}^{n}(x - \bar{x})^2}$$
$$\hat{b} = \bar{y} - (\hat{a}\bar{x})$$

Linear Regressions – Example

For example, we want to predict fuel consumption as miles per gallon (`mpg`) for cars, based on their weight (`wt`), using the famous data set `mtcars`.

We can calculate the values of a and b for the linear regression.

$$mpg = a\ wt + b$$

The results can then be visualised as follows.

Linear Regressions – Example

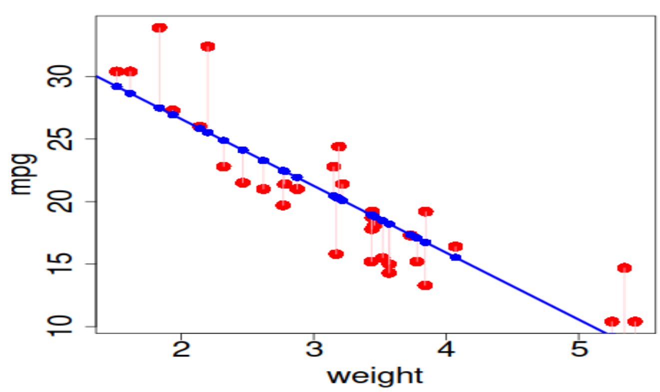

Figure 1: Fitting the model is finding the blue line that minimises (the square of) the distances between the red dots (observations) and blue dots (the estimates by the model).

Linear Regressions – Note

Linear regressions are not machine learning
Note that we don't call a linear regression "AI" or "machine learning". This is because we can calculate the a and b (from the formula $\hat{y} = ax + b$). To call something "machine learning" we want a long learning process instead of a calculation.

2.1 Minorities

Linear Regressions – Minorities

Models that are fitted so that total sum of deviations is minimal, do not necessarily work well for small sub-groups.

For example, in Figure 2 we show that small city cars see their fuel consumption on average overstated, and the fuel consumption is much more sensitive to the weight of a car (green) than the overall model (blue) would lead us to believe.

Linear Regressions – Minorities

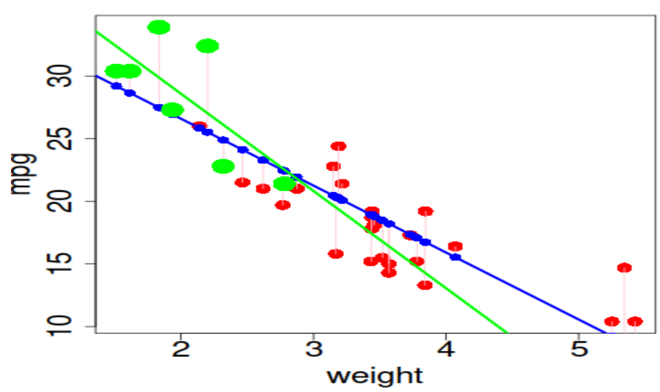

Figure 2: Small city cars (green) see their fuel consumption more over-stated than others. They are a minority and the model fails to capture their specific nature that is more sensitive to weight.

2.2 Multiple Linear Regression

Linear Regressions – Multiple explaining variables

To predict `mpg`, we can use more than one parameter. For example, we can use `wt` (weight), and `hp` (horsepower):

$$mpg = a_1 wt + a_2 hp + b$$

Also the parameters in this model can be calculated exactly and we do not consider this as AI.

Part V of "The Big R-Book: from data science learning machines and big data" Chapter 21 explains how linear regressions work and how to implement them in R.

3 Artificial Neural Networks

Make a symbolic representation of the linear regression

Let's start with visualising our linear regression

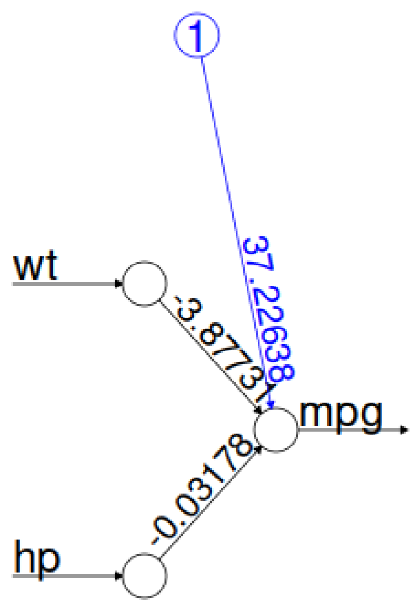

Figure 3: This model estimates `mpg` as the weight (`wt`) multiplied with −3.87751, then adds horsepower (`hp`) multiplied with −0.03178 and adds 37.22666 .

The deep learning Neural Network

Well, you have imagined a most simple neural network with just one neuron. Now imagine a network of regressions feeding into each other as follows:

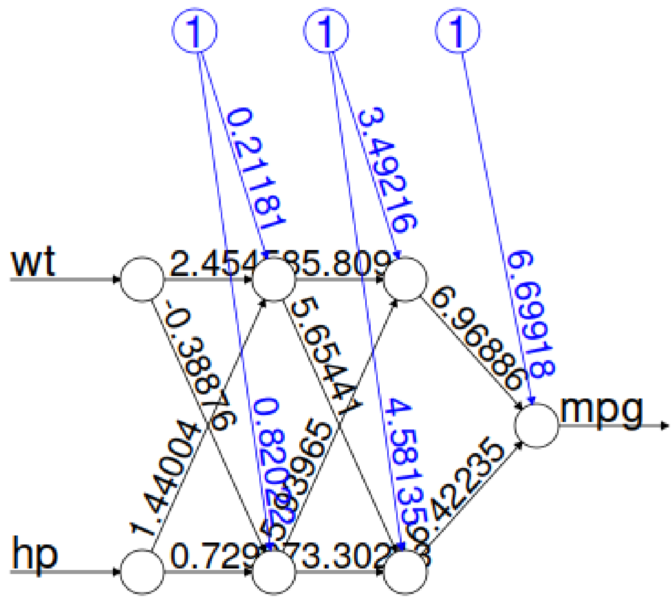

Figure 4: an Artificial Neural Network with 2 hidden layers of 2 neurons.

Notes on the deep learning ANN

Each circle is called a "node". Apart from the input layer (leftmost) we can imagine in each circle a regression that uses the output of the previous layer as input. All parameters are chosen so that the sum of squares of differences between observations and model has to be as small as possible.

We cannot calculate the parameters anymore, and hence we need an iterative algorithm: the model needs to "learn" by trying variations and subsequently keeping the best variations.

Congratulations, you have imagined an Artificial Neural Network (ANN)!

Your ANN is simple: has 2 input nodes `wt` (weight) and `hp` (horsepower). We can use more input nodes, more hidden layers, and more more neurons on those hidden layers.

A simple ANN

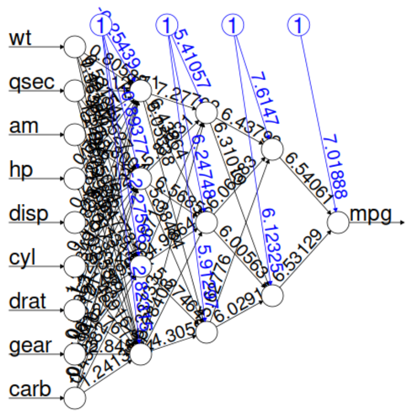

Figure 5: An Artificial Neural Network with 3 hidden layers of respectively 4, 3, and 2 neurons.

The complexity of an ANN

This simple model has 61 variables . . . and it becomes impossible to understand why a high or low `mpg` is predicted.

black box models

An ANN has so many parameters that it is impossible for us to understand the concept behind each node. Therefore we call it a “black box”: it is not transparent and we cannot be sure why a given decision is made.

Part V of “The Big R-Book: from data science learning machines and big data”
Artificial Neural Networks are described in my book in section V, chapter 23.

4 Large Language Models

4.1 The essence of LLMs

Large Language Models – the concept

Now imagine an ANN trained on all texts digitally available. Such model will have many internal layers and billions of parameters. The AI is trained to predict the next word (auto-regressive) or missing word (masked learning). Used in sequence this method creates sentences.

Allow the AI to learn on the data (unsupervised learning), and add some fine tuning (people providing the right answers and call this RHLF (reinforced learning from human feedback)).

Congratulations, you have imagined a large language model (LLM)!

4.2 The Transformer

The transformer

The English Language, has 147,000 words in use.Some sequences of words mean something else than the individual words, some must be understood together (grandfather in my mother's side, etc.)

A virtual space with many thousands of dimensions can encoude each word. Eg. the word sequence "brick house" could be:

$$[1034, 55446, 88, -8999, \ldots, 66676]$$

A "red house" could be:

$$[10, 554, 101, -8756, \ldots, 66170]$$

5 Generative AI

Generative AI

You also heard the term "generative AI". Well, that is the wider term for AI that is able to be creative in the sense that it can create things that didn't exist before. Well-known examples are structural design elements, images, and text (the LLMs).

An AI that generates pictures will be constructed and trained differently than an AI that provides conversations, but the basic principles are the same.

6 Add Some Magic

The magic of generative AI

If you believe now that AI does slavishly what it is taught to do, then read on.

6.1 Emergent Abilities

LLMs might acquire the ability to do something that they are not trained for. Given a critical level of complexity, they just can do it.

For example, LLMs seem to be able to speak Hinglish (combined English and Hindi), without being trained for it. Other emergent abilities are passing college-level exams, do multi-step arithmetic, identify the intended meaning of a word out of the context, etc.

Emergent Abilities
Already now –in the infancy of AI– we have models that display "emergent abilities": they can do things that they weren't really taught to do. Imagine what AI can do in 100 years from now.

6.2 Hallucinations

Hallicunating generative AI

Sometimes the LLM will confidentially assert something that does not follow from the data.

For example without fine tuning, most LLMs will answer "no" to the question "Can I teach an old dog new tricks." This is because of the popular expression "One cannot teach an old dog new tricks?"

hallucinations
LLMs can "hallucinate": they can make mistakes or show bias even when the data does not really have that mistake or bias. Those mistakes appear without warning!

7 Conclusions

Conclusions

- Large language models and generative AI use association (similar to a linear regression). There is no magic, nor conceptual understanding.
- Just as a linear regression, the LLM might miss-understand smaller groups and perpetuate bias (over-fitting).
- LLMs can gain abilities that we didn't directly teach them and they can hallucinate.
- Do use such LLMs, but always remain in the driving seat! Don't let your guard down.
- The rise of AI creates important ethical and societal issues.
- AI is a transformative technology and it will change our lives.
- AI is in it infancy.

8 Nomenclature

Nomenclature

Nomenclature

$\bar{x}$	the bar above a variable refers to the mean of the variable
$\hat{a}$	the estimate for the coefficient a
$\rho_{x,y}$	correlation between the variable x and y
$\sigma_{x,y}$	covariance between the variable x and y
σ_x	variance of the variable x
a	the coefficient in the linear regression $y = ax + b$
x	x is a variable (column name) that we use to explain y
x_i	the index i refers to the number of the row (observation i)
y	y is the variable (column name) that we want to predict using x
AI	Artificial Intelligence
ANN	artificial neural network
LLM	large language model
RHLF	reinforced learning from human feedback