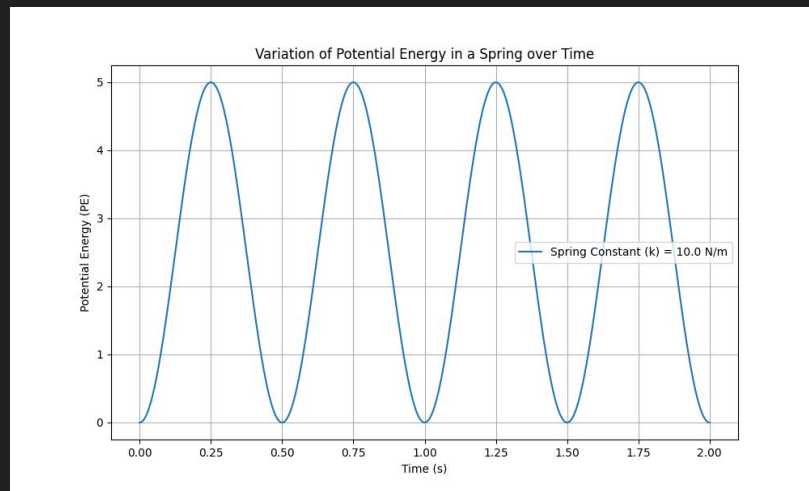
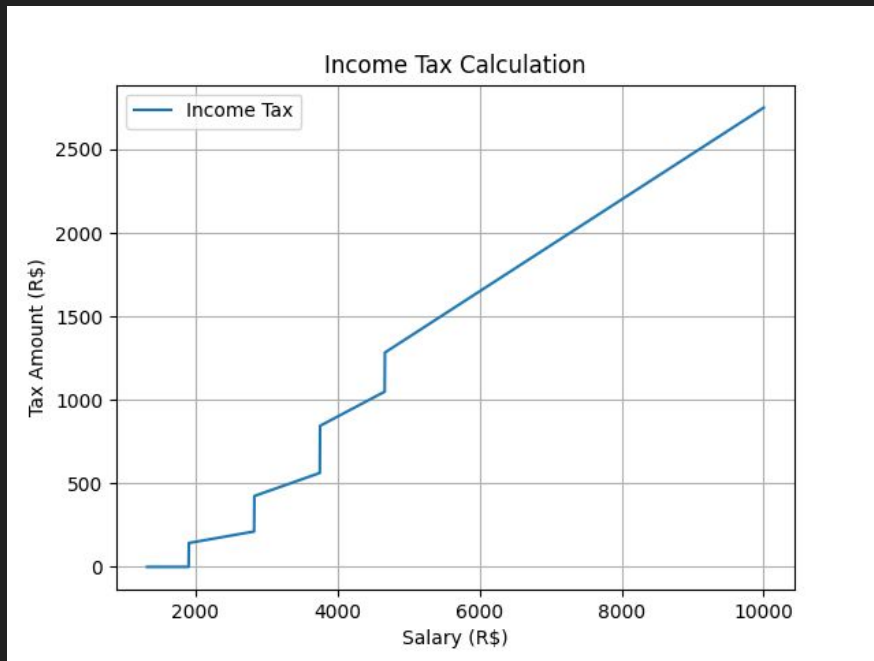


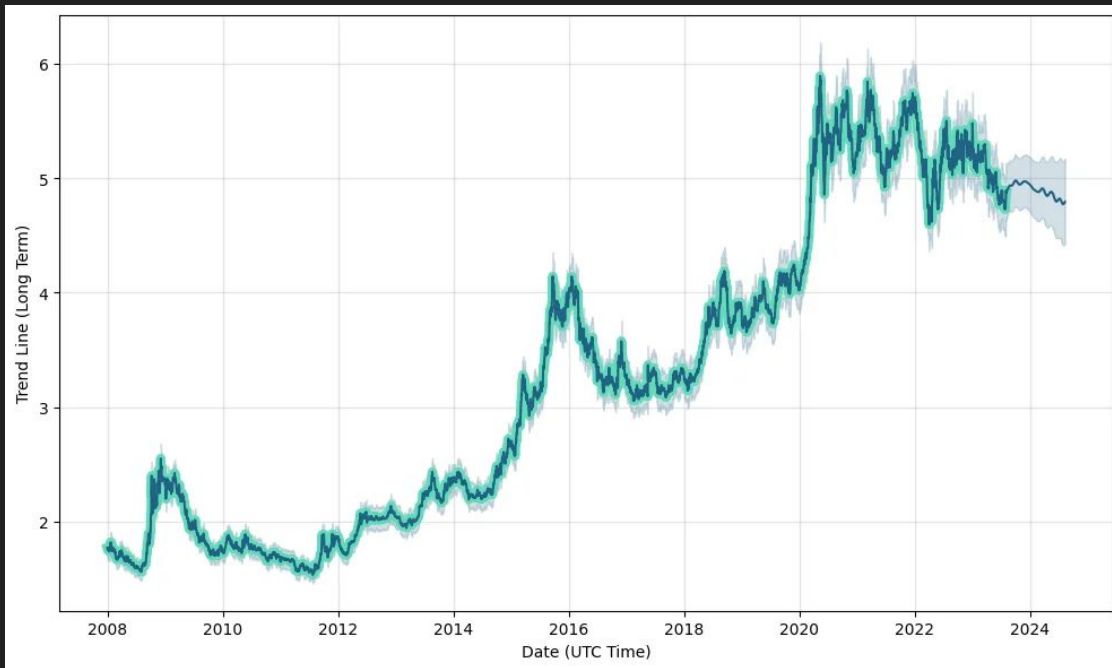
# Objetivos

1. Para que serve machine learning?
2. Como redes neurais funcionam?
3. O que são LLMs?

# O que é machine learning? | Funções

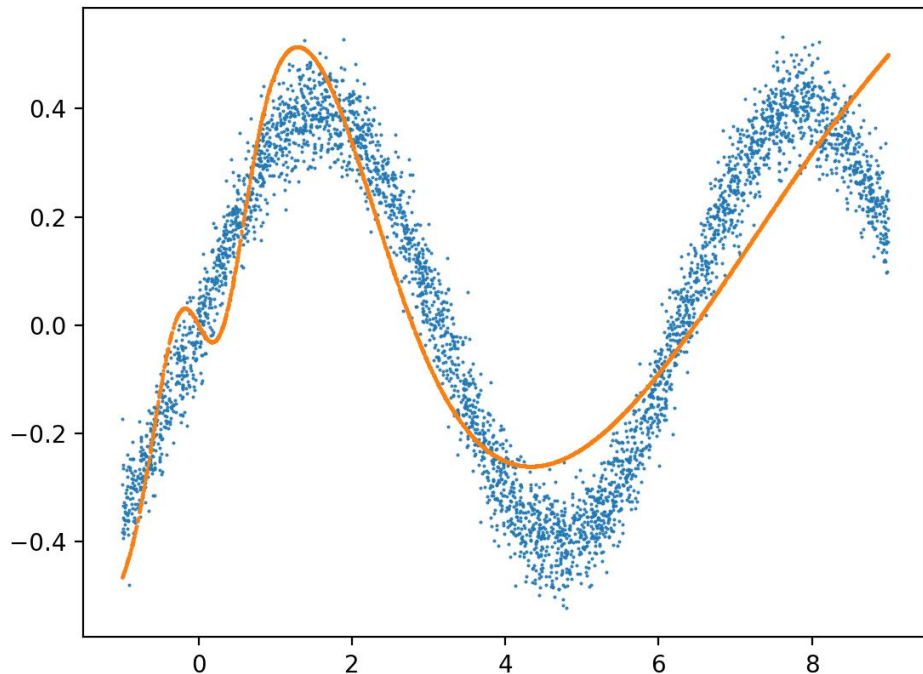


# O que é machine learning? | Funções



- Mercado Financeiro
- Clima
- Vendas
- Linguagem

# Qual o objetivo? | Aproximar funções



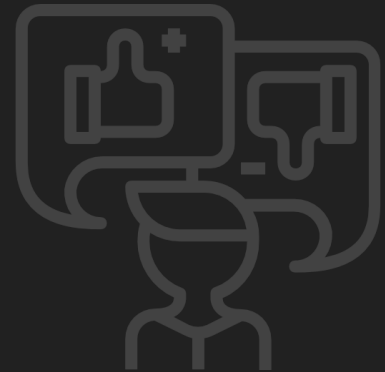
- Uma vez tendo dados de fenômenos passados, encontrar uma aproximação para uma função que descreve o comportamento d



Supervised  
Learning



Unsupervised  
Learning



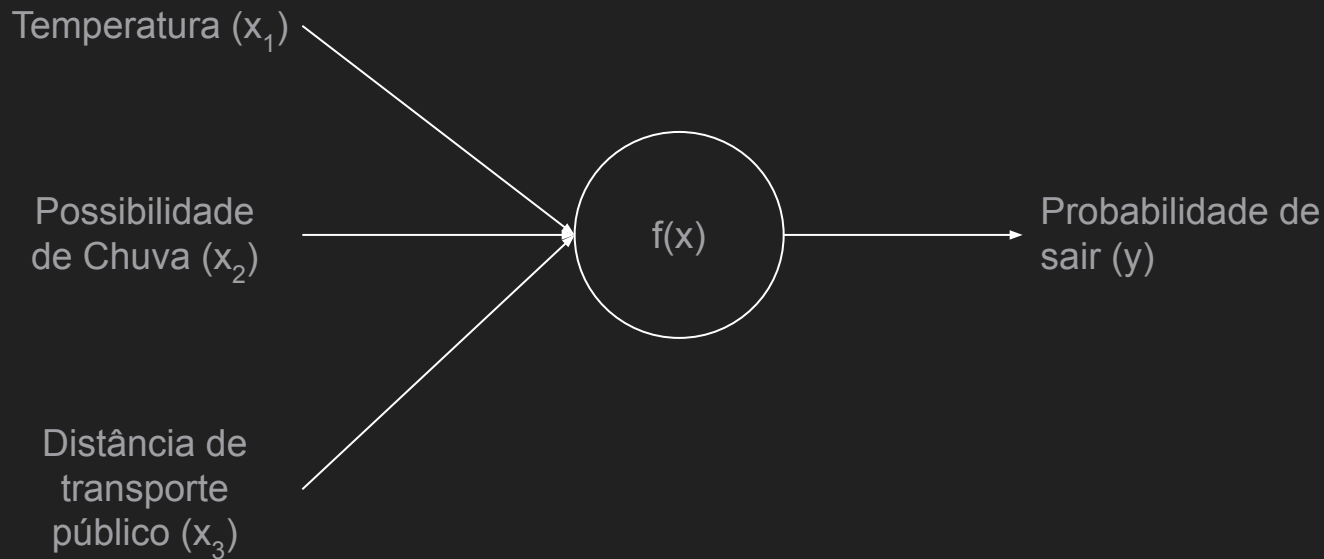
Reinforcement  
Learning

# O que são Redes Neurais? | Exemplo

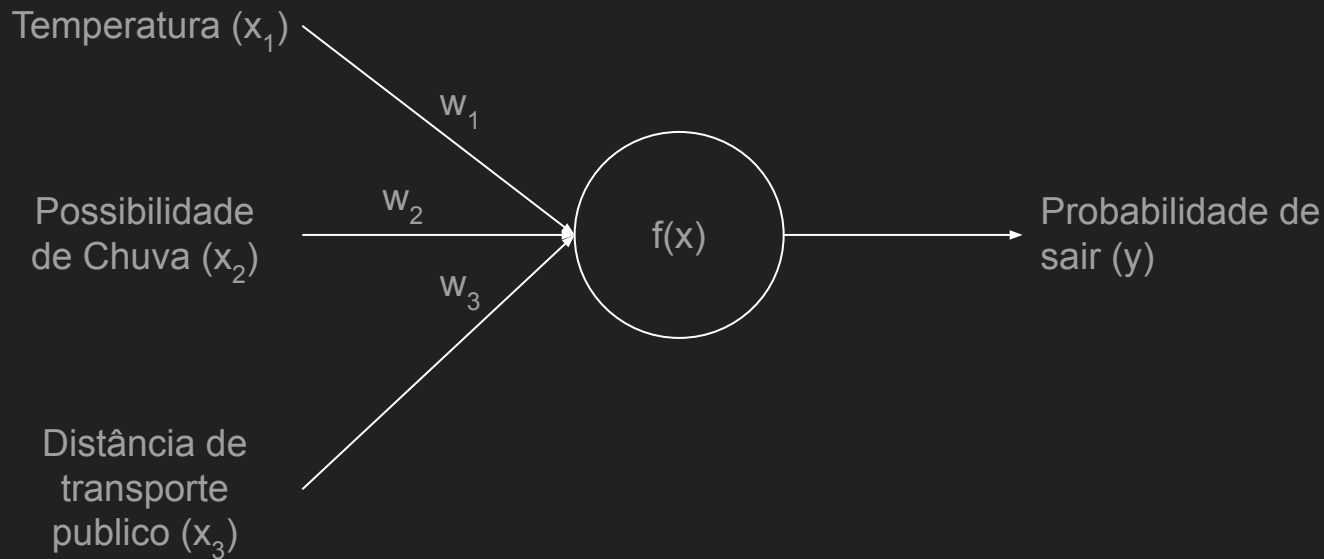
Probabilidade de sair de casa no final de semana para ir em algum local.

- Temperatura
- Probabilidade de Chuva
- Distância de transporte público

# O que são Redes Neurais? | Neurônio



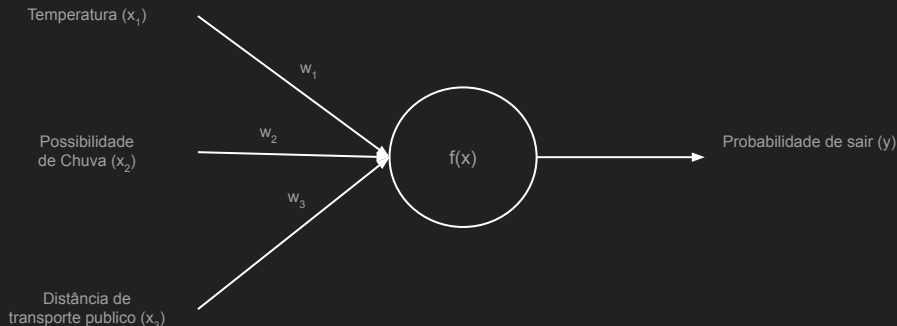
# O que são Redes Neurais? | Neurônio





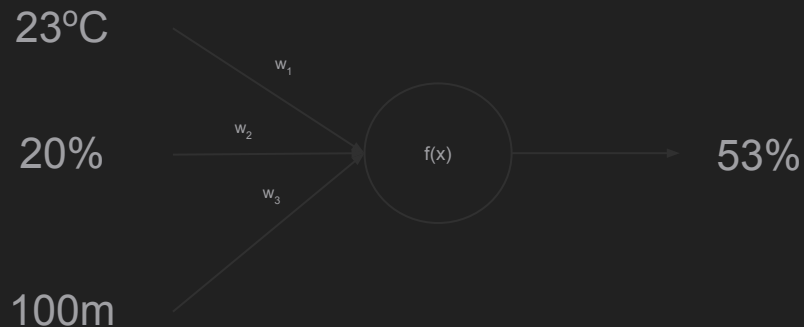
# O que são Redes Neurais? | Aprendizado

1.	$x_1 = 23^\circ\text{C}$	$x_2 = 20\%$	$x_3 = 100\text{m}$	$\Rightarrow y = 96\%$
2.	$x_1 = 10^\circ\text{C}$	$x_2 = 40\%$	$x_3 = 1000\text{m}$	$\Rightarrow y = 26\%$
3.	$x_1 = -1^\circ\text{C}$	$x_2 = 5\%$	$x_3 = 50\text{m}$	$\Rightarrow y = 2\%$
4.	$x_1 = 18^\circ\text{C}$	$x_2 = 85\%$	$x_3 = 10\text{m}$	$\Rightarrow y = 64\%$



# O que são Redes Neurais? | Aprendizado

1.	$x_1 = 23^{\circ}\text{C}$	$x_2 = 20\%$	$x_3 = 100\text{m}$	$\Rightarrow y = 96\%$
2.	$x_1 = 10^{\circ}\text{C}$	$x_2 = 40\%$	$x_3 = 1000\text{m}$	$\Rightarrow y = 26\%$
3.	$x_1 = -1^{\circ}\text{C}$	$x_2 = 5\%$	$x_3 = 50\text{m}$	$\Rightarrow y = 2\%$
4.	$x_1 = 18^{\circ}\text{C}$	$x_2 = 85\%$	$x_3 = 10\text{m}$	$\Rightarrow y = 64\%$



# O que são Redes Neurais? | Aprendizado

1.	$x_1 = 23^{\circ}\text{C}$	$x_2 = 20\%$	$x_3 = 100\text{m}$	$\Rightarrow y = 96\%$
2.	$x_1 = 10^{\circ}\text{C}$	$x_2 = 40\%$	$x_3 = 1000\text{m}$	$\Rightarrow y = 26\%$
3.	$x_1 = -1^{\circ}\text{C}$	$x_2 = 5\%$	$x_3 = 50\text{m}$	$\Rightarrow y = 2\%$
4.	$x_1 = 18^{\circ}\text{C}$	$x_2 = 85\%$	$x_3 = 10\text{m}$	$\Rightarrow y = 64\%$



# O que são Redes Neurais? | Aprendizado

1.	$x_1 = 23^\circ\text{C}$	$x_2 = 20\%$	$x_3 = 100\text{m}$	$\Rightarrow y = 96\%$
2.	$x_1 = 10^\circ\text{C}$	$x_2 = 40\%$	$x_3 = 1000\text{m}$	$\Rightarrow y = 26\%$
3.	$x_1 = -1^\circ\text{C}$	$x_2 = 5\%$	$x_3 = 50\text{m}$	$\Rightarrow y = 2\%$
4.	$x_1 = 18^\circ\text{C}$	$x_2 = 85\%$	$x_3 = 10\text{m}$	$\Rightarrow y = 64\%$



Atualização de pesos  
baseado no erro.

# O que são Redes Neurais? | Aprendizado

1.	$x_1 = 23^\circ\text{C}$	$x_2 = 20\%$	$x_3 = 100\text{m}$	$\Rightarrow y = 96\%$
2.	$x_1 = 10^\circ\text{C}$	$x_2 = 40\%$	$x_3 = 1000\text{m}$	$\Rightarrow y = 26\%$
3.	$x_1 = -1^\circ\text{C}$	$x_2 = 5\%$	$x_3 = 50\text{m}$	$\Rightarrow y = 2\%$
4.	$x_1 = 18^\circ\text{C}$	$x_2 = 85\%$	$x_3 = 10\text{m}$	$\Rightarrow y = 64\%$



Atualização de pesos  
baseado no erro.

# O que são Redes Neurais? | Aprendizado

1.	$x_1 = 23^\circ\text{C}$	$x_2 = 20\%$	$x_3 = 100\text{m}$	$\Rightarrow y = 96\%$
2.	$x_1 = 10^\circ\text{C}$	$x_2 = 40\%$	$x_3 = 1000\text{m}$	$\Rightarrow y = 26\%$
3.	$x_1 = -1^\circ\text{C}$	$x_2 = 5\%$	$x_3 = 50\text{m}$	$\Rightarrow y = 2\%$
4.	$x_1 = 18^\circ\text{C}$	$x_2 = 85\%$	$x_3 = 10\text{m}$	$\Rightarrow y = 64\%$



Atualização de pesos  
baseado no erro.

# O que são Redes Neurais? | Aprendizado

1.	$x_1 = 23^{\circ}\text{C}$	$x_2 = 20\%$	$x_3 = 100\text{m}$	$\Rightarrow y = 96\%$
2.	$x_1 = 10^{\circ}\text{C}$	$x_2 = 40\%$	$x_3 = 1000\text{m}$	$\Rightarrow y = 26\%$
3.	$x_1 = -1^{\circ}\text{C}$	$x_2 = 5\%$	$x_3 = 50\text{m}$	$\Rightarrow y = 2\%$
4.	$x_1 = 18^{\circ}\text{C}$	$x_2 = 85\%$	$x_3 = 10\text{m}$	$\Rightarrow y = 64\%$



# O que são Redes Neurais? | Aprendizado

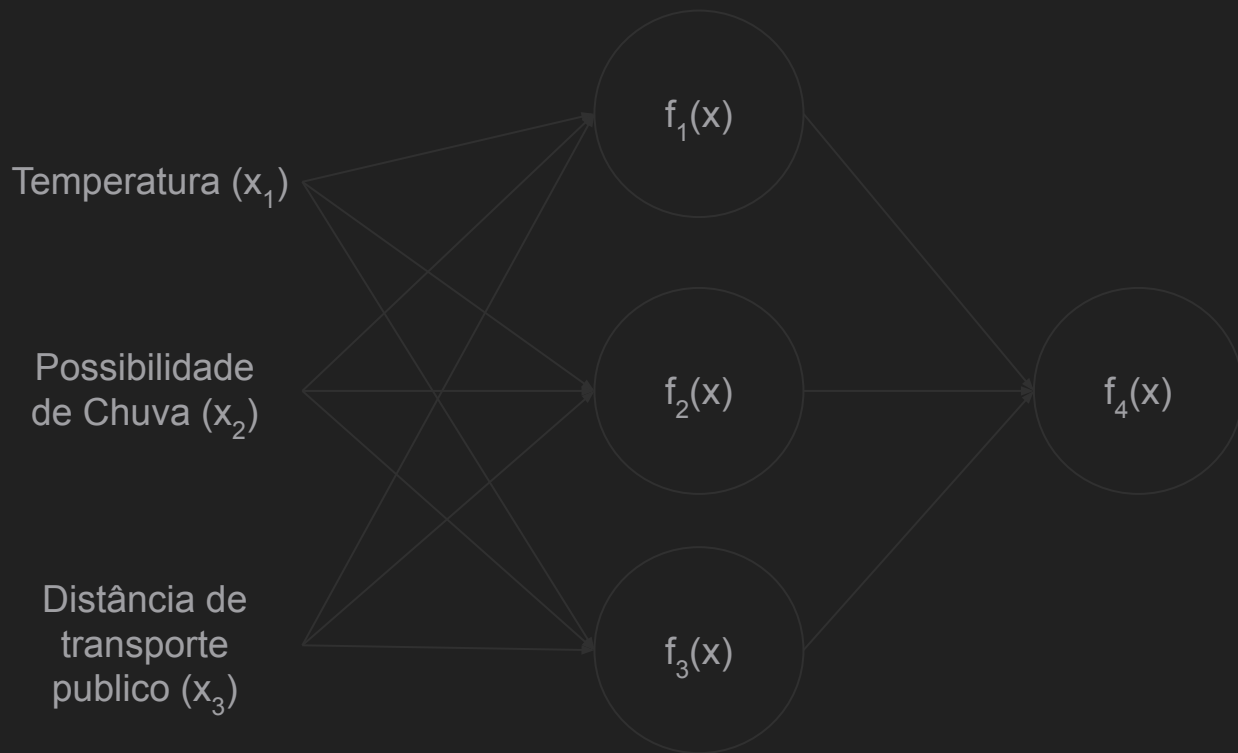
- |    |                          |              |                      |                        |
|----|--------------------------|--------------|----------------------|------------------------|
| 1. | $x_1 = 23^\circ\text{C}$ | $x_2 = 20\%$ | $x_3 = 100\text{m}$  | $\Rightarrow y = 96\%$ |
| 2. | $x_1 = 10^\circ\text{C}$ | $x_2 = 40\%$ | $x_3 = 1000\text{m}$ | $\Rightarrow y = 26\%$ |
| 3. | $x_1 = -1^\circ\text{C}$ | $x_2 = 5\%$  | $x_3 = 50\text{m}$   | $\Rightarrow y = 2\%$  |
| 4. | $x_1 = 18^\circ\text{C}$ | $x_2 = 85\%$ | $x_3 = 10\text{m}$   | $\Rightarrow y = 64\%$ |

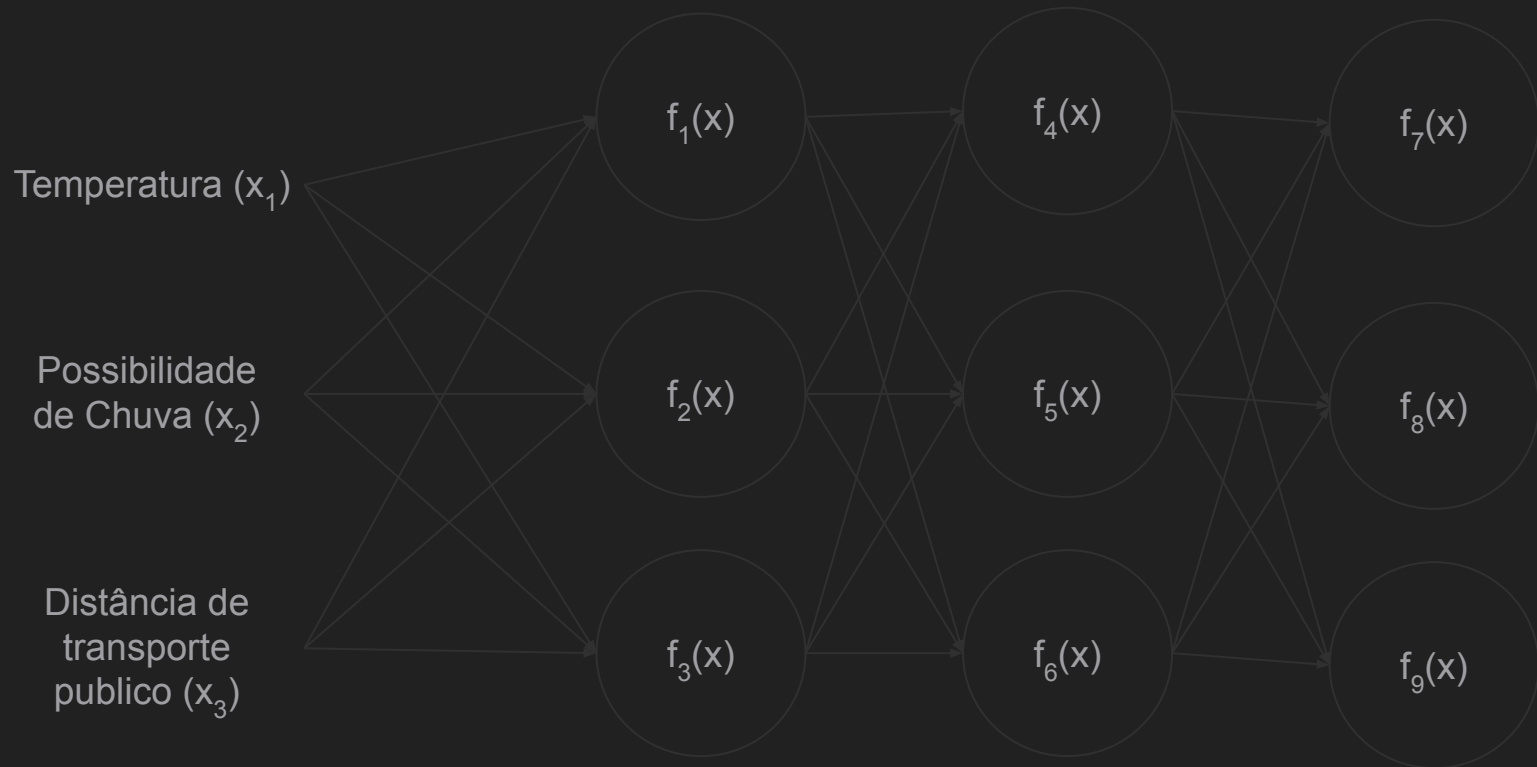


Aproxima a função que descreve o comportamento a partir do ajuste de pesos da rede.

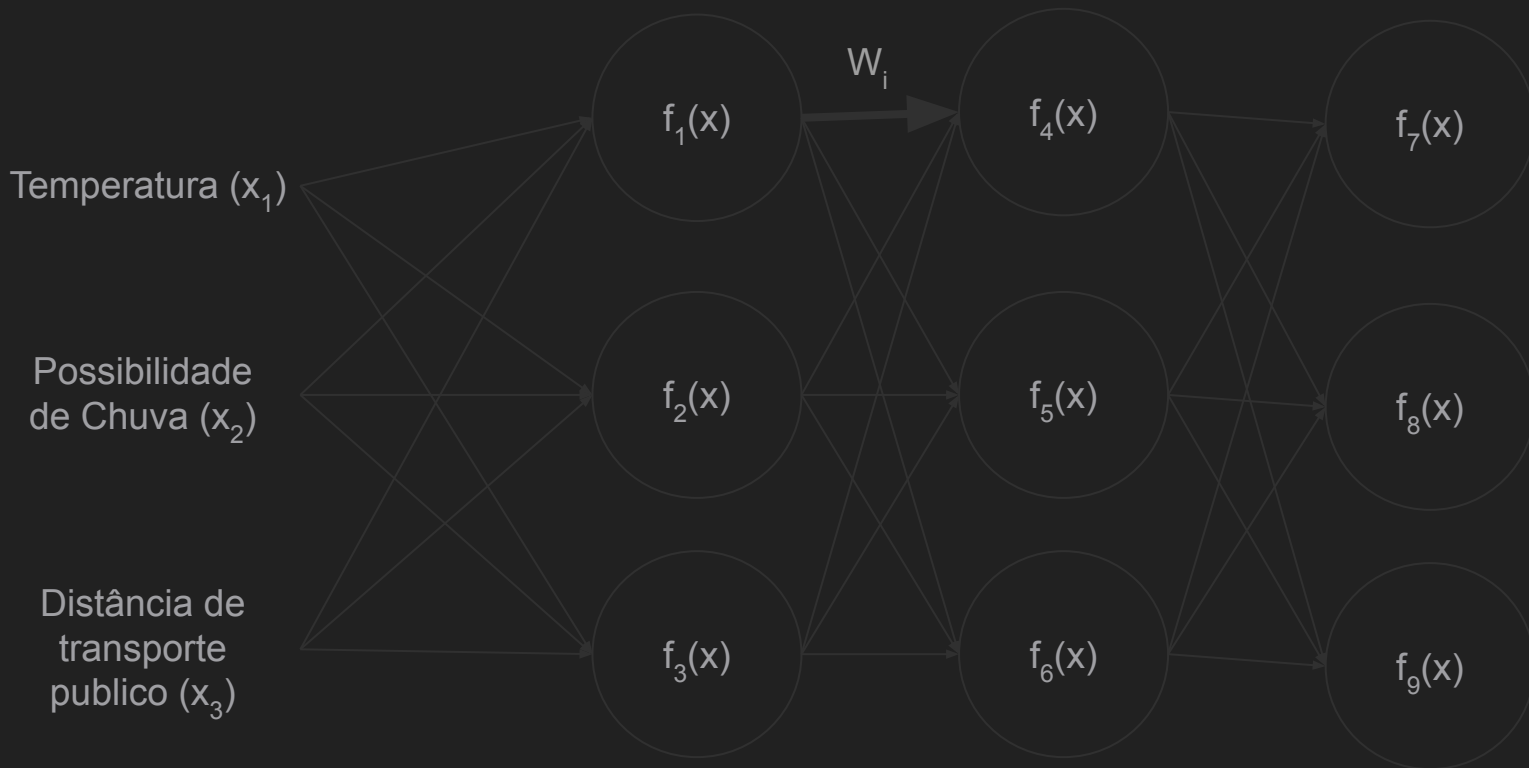


# Como captar mais detalhes? | Adicionando Neurônios

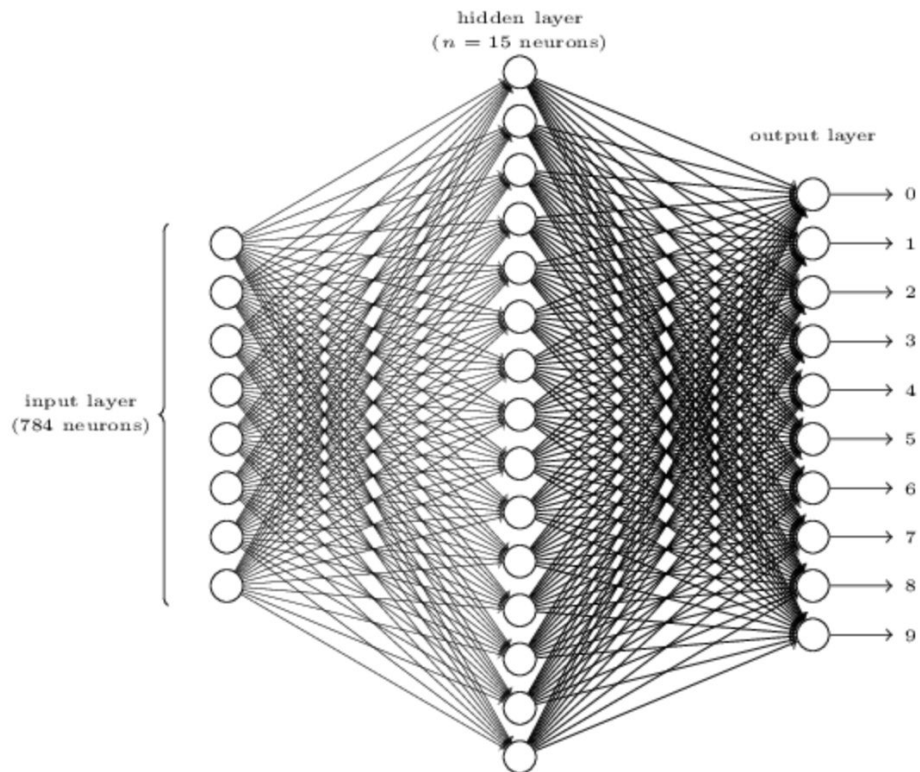




# Como captar mais detalhes? | Adicionando layers



# Deep Learning | Layers



# Tamanho da rede | Quanto mais neurônios e layers melhor?

- Muitos neurônios e layers: O treinamento torna a rede totalmente acoplada aos exemplos prévios, tornando-a incapaz de prever cenários nunca antes vistos.
  - Overfitting
- Poucos neurônios e layers: O treinamento captura poucas variáveis dos dados de exemplo, sendo mais difícil de prever outros cenários e captar pequenas variações.
  - Underfitting

# O que é um LM? | Modelagem estatística da linguagem

Hoje o [redacted] foi oficialmente <??>



$\max(P(\text{próximo termo} \mid \text{termos anteriores})) = \text{"lançado"}$

Qual o elemento químico com número atômico 6?

$\max(P(\text{resposta} \mid \text{contexto})) = \text{"carbono"}$

# Qual o primeiro LM? | Claude Shannon (1948)

## 3. THE SERIES OF APPROXIMATIONS TO ENGLISH

To give a visual idea of how this series of processes approaches a language, typical sequences in the approximations to English have been constructed and are given below. In all cases we have assumed a 27-symbol “alphabet,” the 26 letters and a space.

1. Zero-order approximation (symbols independent and equiprobable).

XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGHYD QPAAMKBZAACIBZLHJQD.

2. First-order approximation (symbols independent but with frequencies of English text).

OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL.

3. Second-order approximation (digram structure as in English).

ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TU COOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE.

4. Third-order approximation (trigram structure as in English).

IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE.

5. First-order word approximation. Rather than continue with tetragram,  $\dots$ ,  $n$ -gram structure it is easier and better to jump at this point to word units. Here words are chosen independently but with their appropriate frequencies.

REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME CAN DIFFERENT NATURAL HERE HE THE A IN CAME THE TO OF TO EXPERT GRAY COME TO FURNISHES THE LINE MESSAGE HAD BE THESE.

6. Second-order word approximation. The word transition probabilities are correct but no further structure is included.

THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHARACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED.

# Qual o primeiro LM?

Claude Shannon (1948)

TABLE OF RANDOM DIGITS

136

06750	40432	31880	52885	30223	92287	52845	72912	71573	27159	76506
06751	48004	14884	50400	09870	20851	17022	16982	80061	80455	15722
06752	11180	09129	91287	82537	90548	17022	58746	24623	31181	80569
06753	82316	88002	52608	21824	61884	40446	67587	12139	24643	91291
06754	20992	65043	13874	73883	89008	82312	58892	47055	09491	44984
06755	24968	82203	07446	09796	90568	23508	10526	22527	25212	62816
06756	78646	21817	12718	43803	16018	18449	24842	64963	53132	76509
06757	01862	03535	89018	82360	64912	77086	50164	20086	03053	60214
06758	83042	15129	52628	28010	03278	05591	08366	73136	99716	80124
06759	08952	76364	68801	09765	39840	01087	98243	58755	32656	40209
06760	21237	22621	91884	23885	47087	26389	86074	26921	48873	15794
06761	21088	57629	10453	60319	35885	01726	40186	78713	77972	72632
06762	24547	72653	81245	21210	80410	33461	29868	94900	59069	91882
06763	12412	40248	81006	49025	08355	40869	61323	81854	00899	98146
06764	98040	86212	80518	47006	80136	41384	13425	73058	51519	86061
06765	50083	22052	27529	81547	78065	49154	80614	32645	22883	38408
06766	54788	10191	23487	18088	00891	06446	34130	05300	16027	30587
06767	37495	48786	58995	05658	12747	76745	34801	78810	18451	82931
06768	82905	52190	84878	16027	78878	81016	90237	78868	78128	21645
06769	15560	90174	97458	92248	34055	18345	51795	62616	30585	24708
06770	42381	37908	16691	12867	84321	84545	40413	55342	10325	85335
06771	18179	14799	94883	42271	28489	93064	53877	38895	44225	23571
06772	85082	16059	16047	44624	81222	63172	63284	21951	73674	78621
06773	84123	18717	88091	66883	82569	07282	10120	30736	04284	78621
06774	28810	68101	84605	78688	42811	92969	72307	01124	00641	44763
06775	58980	54463	02362	08004	57984	23092	70100	97659	56582	40247
06776	17806	47752	45484	28901	13851	72419	18105	30875	52358	30648
06777	61977	07645	10518	41203	97170	29580	80885	91897	89552	31836
06778	02106	92145	62292	15861	38012	17554	24054	92442	43104	70414
06779	98042	51223	05400	98231	06874	52607	23763	57348	82806	47774
06780	02692	55557	25865	39051	50011	76284	25735	72568	80939	08573
06781	70215	12348	5285	72208	21960	24888	38705	77197	38140	45481
06782	47602	84646	82683	47416	55353	08879	26957	77705	92882	78477
06783	84685	98266	90190	27385	32574	89909	78316	64473	60334	90244
06784	17039	40915	24846	16023	03535	36957	85772	90294	78213	83497
06785	30413	82644	23975	02831	08846	26510	21129	46076	90114	63945
06786	22962	56376	87732	80887	21801	78766	16952	81621	84846	80776
06787	81513	21655	00846	43652	17218	64303	66857	01991	35811	46848
06788	57406	27107	48628	29909	98800	22244	68883	10168	41073	78807
06789	27585	04368	07030	34288	34506	86213	78454	55339	78807	80774
06790	11993	08814	93328	99860	62997	97263	60519	76764	68087	93475
06791	30212	55553	74872	12650	85751	65979	29912	62519	51192	50298
06792	54729	99291	57363	46271	84806	60639	15890	13972	43023	46880
06793	47992	54709	28119	28103	68883	10168	41073	78807	13511	37268
06794	72146	85466	46108	23860	34074	39988	56710	45873	16957	98374
06795	48415	90116	38077	99005	35727	76348	09565	28933	19189	75033
06796	22964	19356	90118	85916	61104	62840	48490	09771	21305	07161
06797	08807	97755	80595	27623	71777	67791	79800	87781	18560	40816
06798	43442	89977	71011	20791	15869	47973	80593	44571	97167	88123
06799	49556	69880	21299	60015	24505	21857	78919	05031	74892	80904

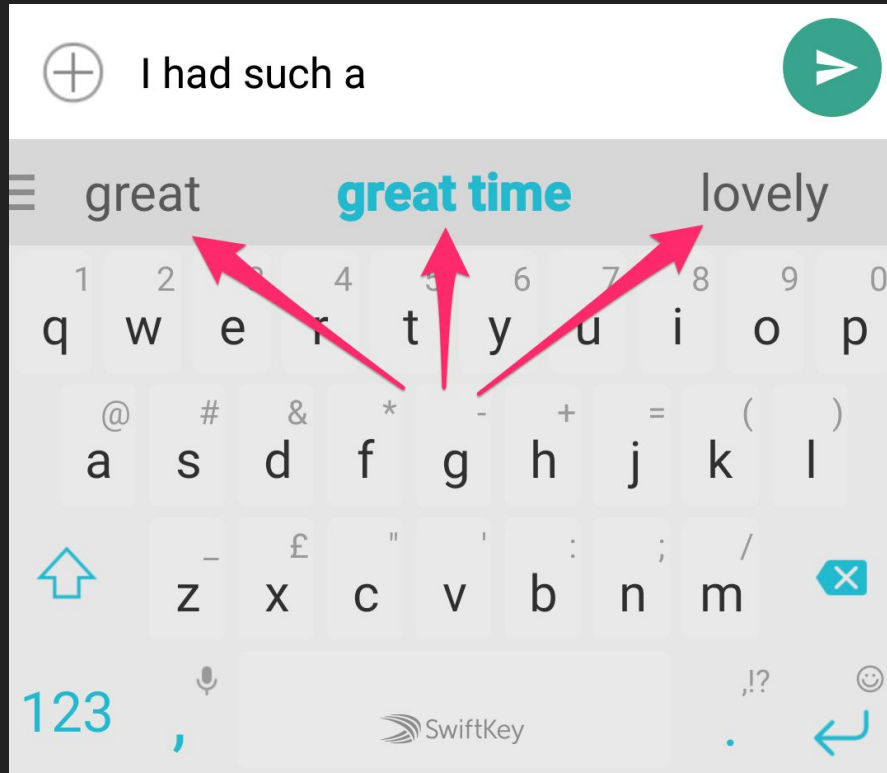
TABLE OF RANDOM DIGITS

137

06800	21827	80191	43585	20270	74558	48861	49781	43152	80052	05750
06801	92204	68587	84725	31701	74558	48861	49781	43152	80052	05750
06802	72608	47719	94413	07081	49781	43152	80052	05750	83718	27893
06803	11181	34112	21804	22884	38705	77197	38140	45481	44440	61844
06804	90738	58381	02501	10840	07605	80601	68272	37402	50299	44941
06805	97806	62153	46984	88450	86772	21601	60123	11842	19098	02527
06806	64801	15291	70325	24453	4742	33354	17051	01497	07764	59507
06807	51517	20148	58482	85851	4742	33354	17051	01497	07764	59507
06808	86035	09025	34342	01026	01900	87885	34416	18097	05218	13502
06809	40704	13850	78982	10013	72214	98454	38445	21181	94549	35512
06810	09130	82082	16513	04318	48844	42677	33651	92884	60732	83781
06811	71235	76894	81382	48674	23672	77020	32251	70463	66312	25038
06812	13116	64816	14165	01982	19943	11088	32449	54013	76240	91860
06813	97227	69784	76413	30598	82123	74699	05519	22601	23968	19257
06814	55499	59591	92600	73882	25113	99288	43688	22501	31577	50791
06815	68114	65784	03503	02342	33465	78067	62339	87327	50998	48084
06816	10844	70223	87979	08189	81988	79313	41960	38484	44654	81809
06817	63563	42765	58463	28808	7994	03555	85149	88776	05904	85119
06818	50696	67283	15472	16243	06020	77514	02533	41561	86849	63729
06819	38018	81790	07801	50846	59924	61794	38229	18692	74217	87486
06820	29833	65742	94607	41151	40162	56513	17119	69344	05109	06509
06821	81722	86318	35863	02825	03242	02124	32181	36576	86228	92081
06822	76893	26845	98750	07129	26694	31274	96539	03974	52918	79355
06823	49035	20556	69838	18267	50454	68776	00591	81476	80160	32618
06824	21626	25025	16717	87974	58254	09438	16845	70276	45279	47740
06825	21413	49624	17413	92485	88605	17066	49553	43121	83541	34840
06826	20882	36088	10376	13157	22479	29798	08852	98103	43943	44438
06827	41284	09786	32189	23352	72568	43449	41222	91977	73728	40282
06828	17486	24548	43744	28648	82119	17327	20886	12025	48181	21522
06829	84534	85028	24040	92081	32144	90247	36813	37041	52031	40205
06830	84770	38718	43464	28531	51519	96267	20105	08067	75599	63821
06831	71312	02987	91514	47178	77587	99400	56878	54742	89001	75477
06832	64826	46172	01831	08463	06071	32967	80075	48441	73480	88646
06833	70411	41223	97763	52366	06071	32967	80075	48441	73480	88646
06834	32545	35103	84863	56129	91052	58833	91861	35813	28319	18439
06835	89904	07019	11723	27044	91405	04809	38411	56070	03870	31461
06836	79183	25657	77892	13003	64816	80819	24709	71826	82401	30778
06837	48882	86644	41008	32795	31584	38421	12552	71153	24882	83679
06838	70037	31852	87418	96027	09894	43826	80380	52418	12767	67027
06839	17517	48860	09293	41303	00117	13912	46878	38007	08537	27885
06840	83388	12308	01113	21707	18677	80807	22543	08868	21672	39023
06841	55719	09726	72750	18180	31068	77780	14917	80518	86410	87268
06842	24838	18018	05772	71842	34325	02884	78088	00340	15052	71153
06843	54899	97283	82385	34763	52021	81298	14532	32583	42666	00135
06844	10078	53685	81841	14489	40858	34363	00072	28619	03788	70842
06845	30880	53378	41676	67807	03084	17931	03034	80627	44152	21253
06846	44114	33580	28817	75858	82315	28283	90403	22371	27823	34423
06847	31995	38893	57776	70414	14885	19011	44858	36850	30843	31087
06848	75124	01815	79153	32913	14171	14944	09944	77418	33862	86228
06849	48839	38487	43058	78326	31584	38421	86315	14236	40858	80874



# Autocomplete | Language Model Comum



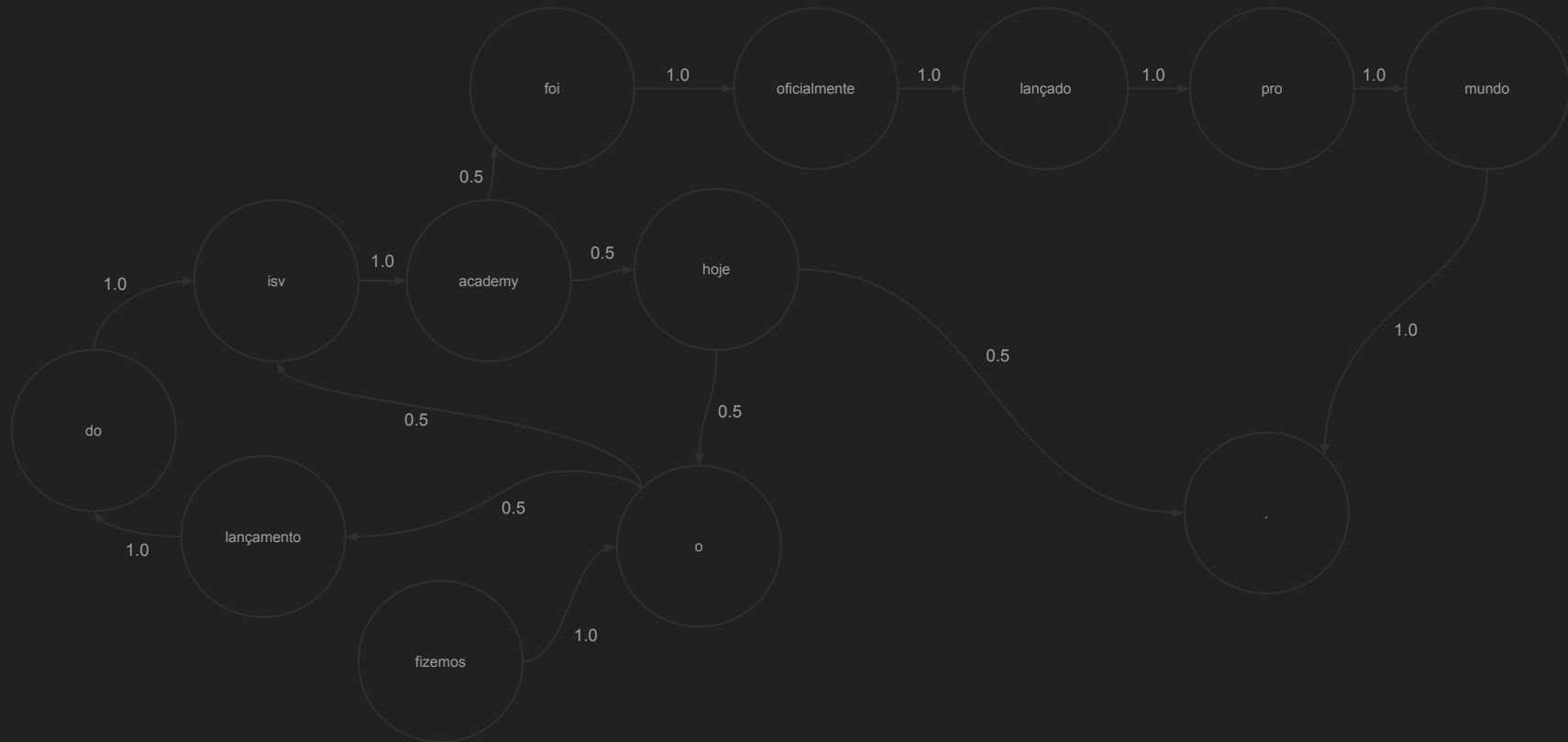
# Modelagem Estatística | Como construímos esse modelo?

Universo  
de  
Sentenças

Hoje o [redacted] foi oficialmente lançado pro mundo.

Fizemos o lançamento do [redacted] hoje.

# Cadeia de Markov | Language Model Comum



# Matriz de Probabilidades | Language Model Comum

	academy	fizemos	hoje	isv	o
academy	0	0	0.5	0	0
fizemos	0	0	0	0	1.0
hoje	0	0	0	0	0.5
isv	1.0	0	0	0	0
o	0	0	0	0.5	0

...

⋮

# Problema

Apenas utilizando o termo anterior

	academy	fizemos	hoje	[redacted]	o
academy	0	0	0.5	0	0
fizemos	0	0	0	0	1.0
hoje	0	0	0	0	0.5
[redacted]	1.0	0	0	0	0
o	0	0	0	0.5	0

...

⋮

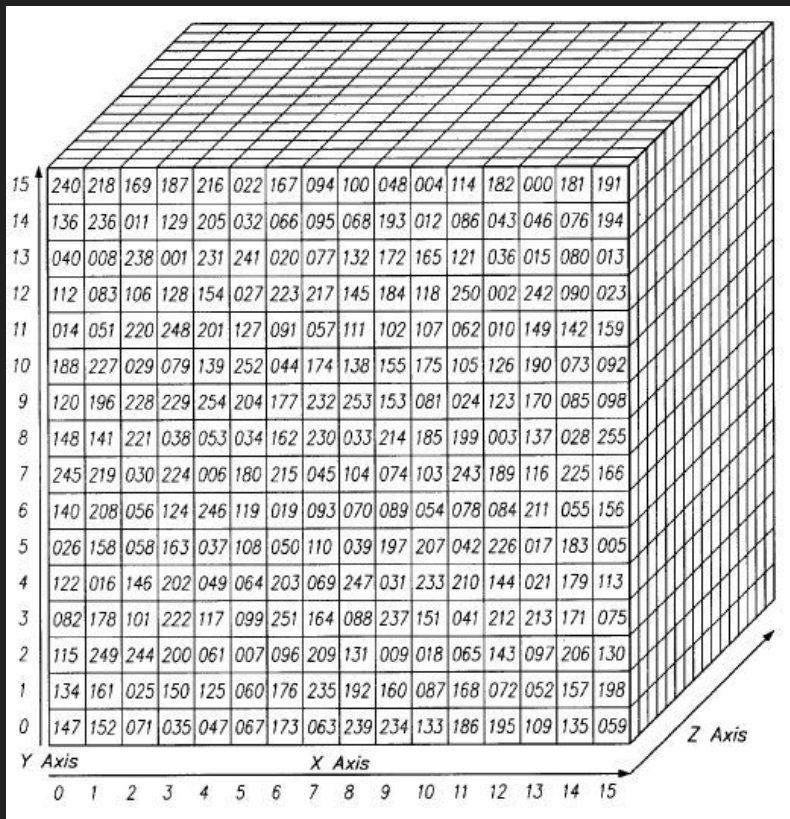
# Contexto

Hoje o [redacted] foi oficialmente <??>

No modelo atual: Hoje o [redacted] foi oficialmente <??>

No modelo ideal: Hoje o [redacted] foi oficialmente <??>

# Matriz de Probabilidades | Contexto



Número de palavras na língua portuguesa: 370 000

- 1 palavra:  $(370\,000)^2 = 1 * 10^{11}$
- 2 palavras:  $(370\,000)^3 = 5 * 10^{16}$
- 3 palavras:  $(370\,000)^4 = 2 * 10^{22}$

Mais do que 5 palavras:

- Número de elementos maior que o número de átomos no universo ( $7 * 10^{27}$ )

---

### Attention Is All You Need

---

**Ashish Vaswani\***  
Google Brain  
avaswani@google.com

**Noam Shazeer\***  
Google Brain  
noam@google.com

**Niki Parmar\***  
Google Research  
nikip@google.com

**Jakob Uszkoreit\***  
Google Research  
usz@google.com

**Llion Jones\***  
Google Research  
llion@google.com

**Aidan N. Gomez\* †**  
University of Toronto  
aidan@cs.toronto.edu

**Łukasz Kaiser\***  
Google Brain  
lukaszkaizer@google.com

**Illia Polosukhin\* ‡**  
illia.polosukhin@gmail.com

#### Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions



# Atenção

Foco em termos específicos do input

The FBI is chasing a criminal on the run .

The FBI is chasing a criminal on the run .

The FBI is chasing a criminal on the run .

The FBI is chasing a criminal on the run .

The FBI is chasing a criminal on the run .

The FBI is chasing a criminal on the run .

The FBI is chasing a criminal on the run .

The FBI is chasing a criminal on the run .

The FBI is chasing a criminal on the run .

The FBI is chasing a criminal on the run .

# Treinamento | Como treinar o modelo?

Hoje o [redacted] foi <??> lançado pro mundo.

Fizemos o <??> do [redacted] hoje.

Hoje o [redacted] foi <??> lançado pro mundo.

Fizemos o <??> do [redacted] hoje.



Hoje o [redacted] foi **jorge** lançado pro mundo.

Fizemos o **lançamento** do [redacted] hoje.

Hoje o [redacted] foi <??> lançado pro mundo.

Fizemos o <??> do [redacted] hoje.



Hoje o [redacted] foi **jorge** lançado pro mundo.

Fizemos o **lançamento** do [redacted] hoje.

Hoje o [redacted] foi <??> lançado pro mundo.

Fizemos o <??> do [redacted] hoje.



Hoje o [redacted] foi **jorge** lançado pro mundo.

Fizemos o **lançamento** do [redacted] hoje.

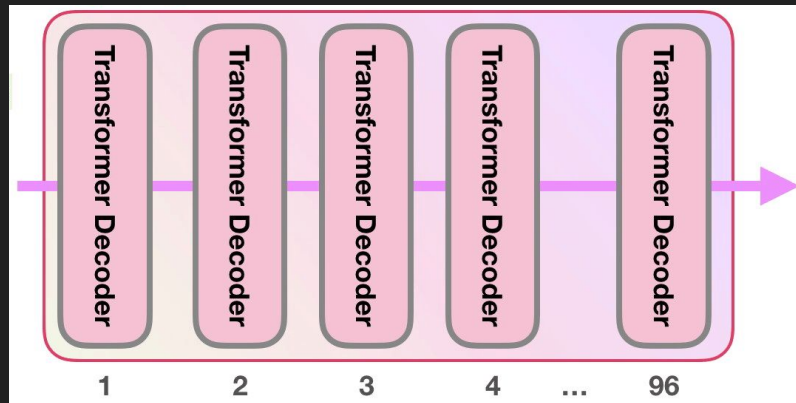
Modelo

Atualização  
do pesos

- CommonCrawl (60%)
- WebText2 (22%)
- Wikipedia (3%)
- Livros (15%)

Embedding

0.35  
0.24  
.  
.  
.  
0.87

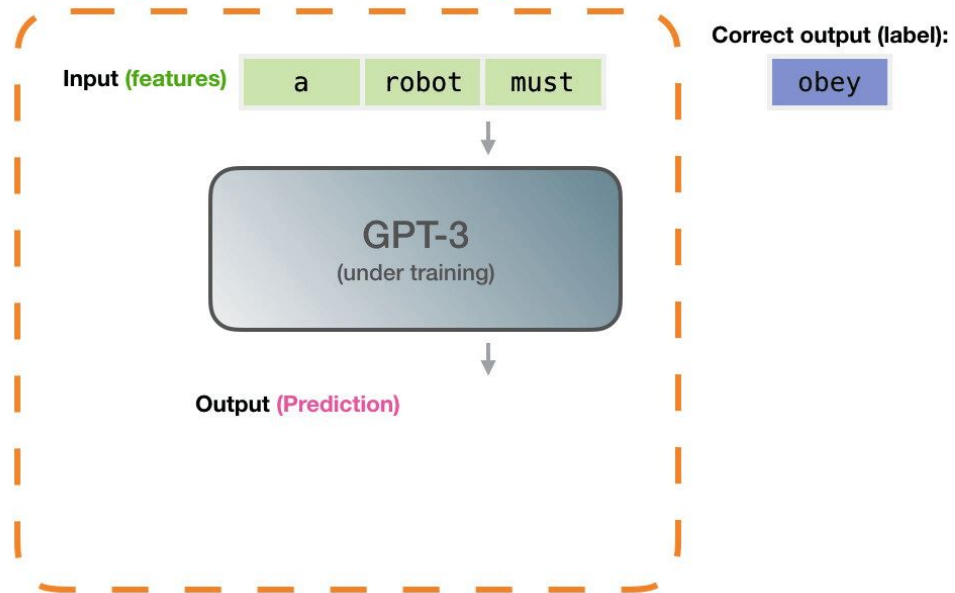


Rede Neural com  
96 Layers

0.18  
0.43  
.  
.  
.  
0.07

"Academy"

## Unsupervised Pre-training



1 2 3 4 5 6

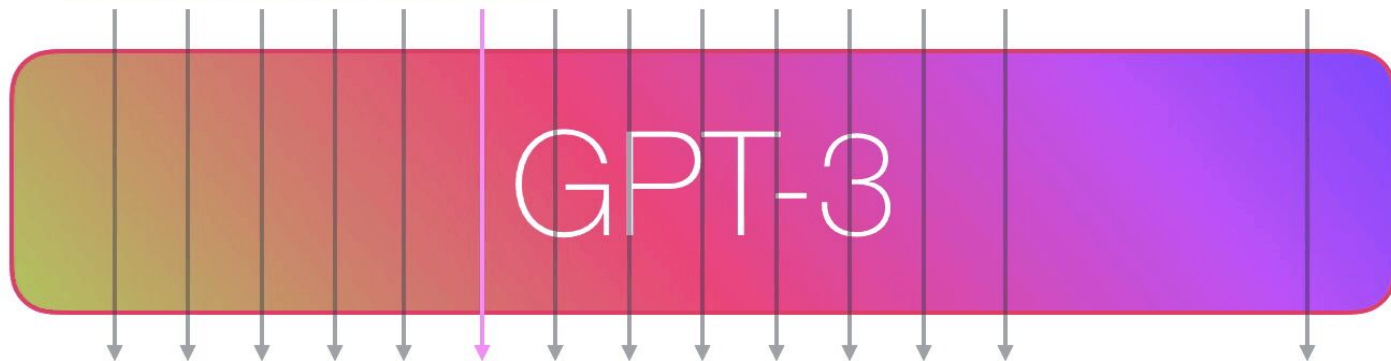
... 2048

Input

Recite the first law of robotics

GPT-3

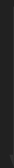
Output:





- O português possui 360 mil palavras. Isso gera um vocabulário enorme.
- Quanto maior o tamanho do vocabulário, maior será o número de conexões necessárias.
- Palavras praticamente iguais como "garoto" e "garotos" recebem tokens completamente diferentes.
- Palavras digitadas erradas contam como um token inválido (fora do vocabulário). Ex: "principlamente" ("principalmente").

Hoje o [redacted] foi oficialmente lançado pro mundo.



principlamente

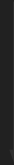
principlamente

principalmente

principalmente

- 2 tokens iguais
- 3 muito similares

Hoje o [redacted] foi oficialmente lançado pro mundo.



- A separação em tokens limita o número de inputs e conexões da rede neural.
- Facilita a predição em múltiplas linguagens.
- Performa melhor em cenários com palavras nunca antes vistas e erros de ortografia.

# Tokens | GPT

Texto	Tokens / palavra
Wikipedia (Filosofia - Inglês)	1.46
Wikipedia (Filosofia - Italiano)	2.48
Wikipedia (Filosofia - Polonês)	8.67
Wikipedia (Filosofia - Chinês)	13.13

# Como o modelo evolui? | Ajuste de pesos

100 iterações

tyntd-iafhatawiaoihrdemot lytdws e ,tfti, astai f ogoh eoase rrranbyne 'nhthnee e  
plia tklrge t o idoe ns,smtt h ne etie h,hregtrs nigtike,aoaenns lng

300 iterações

"Tmont thithey" fomesscerliund  
Keushey. Thom here  
sheulke, anmerenith ol sivh l lalterthend Bleipile shuwyl on aseterlome  
coaniogennc Phe lism thond hon at. MeiDimorotion in ther thize."

# Como o modelo evolui? | Ajuste de pesos

500 iterações

we counter. He stutn co des. His stanted out one ofler that concossions and was  
to gearang reay Jotrets and with fre colt off paitt thin wall. Which das stimn

700 iterações

Aftair fall unsuch that the hall for Prince Velzonski's that me of  
her hearly, and behs to so arwage fiving were to it beloge, pavu say falling  
misfort  
how, and Gogition is so overelical and ofter.

# Como o modelo evolui? | Ajuste de pesos

1200 iterações

"Kite vouch!" he repeated by her  
door. "But I would be done and quarts, feeling, then, son is people...."

2000 iterações

"Why do what that day," replied Natasha, and wishing to himself the fact the  
princess, Princess Mary was easier, fed in had opened him.  
Pierre asking his soul came to the packs and drove up his father-in-law women.

# Reinforcement Learning | Refinamento do modelo

