

Generativní umělá inteligence

Díl druhý: Umělý neuron Warrena

McCullocha a Waltera Pittse

EDUARD BARTL

Přírodovědecká fakulta UP, Olomouc

Velké jazykové modely jsou postavené na hlubokých neuronových sítích, které využívají vícehlavový attention mechanismus. Co tyto „komplikovaně znějící“ pojmy znamenají se postupně dozvíme v jednotlivých dílech série článků věnované generativní umělé inteligenci. Začneme pěkně od začátku – v tomto dílu se budeme věnovat základní stavební jednotce hlubokých neuronových sítí, které se dnes říká *umělý neuron*. Konkrétně se budeme zabývat historicky prvním a zároveň nejjednodušším umělým neuronem, takzvaným MCP neuronem.

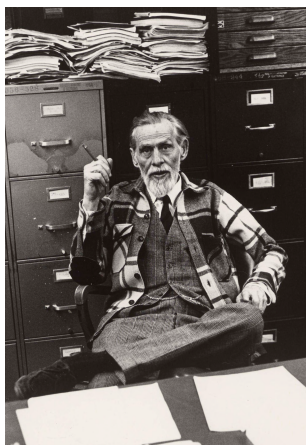
1. Dva géniové

Velké jazykové modely, například GPT od OpenAI, Gemini od Googlu, LLaMA od Mety nebo rodina velkých jazykových modelů Claude od Anthropicu, patří mezi nejmodernější přístupy v oblasti zpracování přirozeného jazyka.¹⁾ Na mnoho lidí, zejména ze starší generace, musí tyto nástroje působit futuristickým dojmem – jako by patřily spíše do sci-fi literatury než do dnešního světa. Tento dojem navíc umocnil překotný nástup velkých jazykových modelů. Pro laickou veřejnost se totiž objevily takřka přes noc, když 30. listopadu 2022 společnost OpenAI představila ChatGPT. Skutečnost je ovšem taková, že tomuto okamžiku předcházelo mnoho desítek let vývoje zejména v oblasti informatiky a informačních technologií. První krok k dnešním velkým jazykovým modelům byl učiněn již v roce 1943 dvěma poněkud excentrickými americkými vědci, Warrenem McCullochem a Walterem Pittsem.

¹⁾ Zpracování přirozeného jazyka (v angličtině se používá zkratka NLP – Natural Language Processing) je obor umělé inteligence, jenž se zabývá interpretací a generováním textu napsaného v přirozeném jazyce, jakým je například čeština nebo angličtina. Mezi typické úlohy NLP patří strojový překlad, sumarizace textu, rozpoznávání řeči a podobně.

Warren Sturgis McCulloch (1898–1969) byl vskutku renesanční osobností. Vystudoval filozofii a psychologii, velmi se zajímal o fungování lidského mozku, úzce spolupracoval se zakladatelem kybernetiky, Norbertem Wienerem. V mládí ho však lákalo studium teologie a chtěl se stát básníkem; je autorem několika sonetů. Byl také velmi zručný a prakticky orientovaný, co prokázal mimo jiné tím, že na svém statku v Connecticutu navrhl a postavil několik budov a malou přehradu.

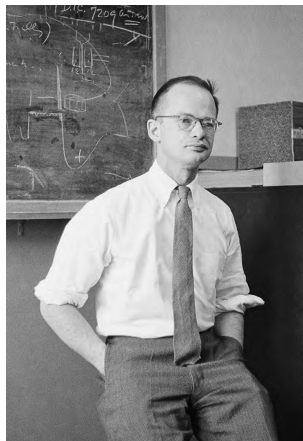
McCulloch úzce spolupracoval se zakladatelem kybernetiky, Norbertem Wienerem. Byl předsedou takzvaných *Macy Conferences on Cybernetics*, což byla setkání předních učenců z různých disciplín uskutečňovaných v letech 1946 až 1953, která měla zásadní vliv na vznik a rozvoj nejen kybernetiky, ale také umělé inteligence, systémových a kognitivních věd.



Obr. 1 Warren Sturgis McCulloch (zdroj: mitmuseum.mit.edu).

Walter Harry Pitts (1923 – 1969) byl neméně zajímavou osobností. Byl zázračným dítětem, celý život se zabýval matematickou logikou a neurovědami. Logiku však nestudoval na žádné škole, byl samoukem. V literatuře věnované Pittsovi je často zmiňovaná historka z jeho dětství, kdy jako dvanáctiletý přečetl za tři dny celou třísvazkovou *Principia Mathematica* od Bertranda Russella. Poté Russelovi napsal dopis, v němž poukázal na závažné nedostatky, které objevil v prvním svazku tohoto významného díla. Russell se zachoval velkoryse, námitky mladého Waltera ocenil a pozval ho ke studiu na Cambridge University. Pitts však tuto nabídku odmítl s tím, že se chce plně věnovat studiu logiky.

Osobní život Waltera Pittse však šťastný nebyl. Nikdy se neoženil a začátkem padesátých let začal ztrácet zájem o svojí vědeckou práci a postupně se propadal do sociální izolace. Vše vyvrcholilo rozkoem s Norbertem Wienerem, se kterým (stejně jako McCulloch) spolupracoval a měl k němu velmi přátelský vztah. Poté Pitts spálil všechny své nepublikované vědecké práce včetně rozepsané dizertace, která se věnovala neuronovým sítím. Zemřel ve věku 46 let na zdravotní potíže, které pravděpodobně souvisely s alkoholismem.



Obr. 2 Walter Harry Pitts (zdroj: Wikipedia).

2. MCP neuron

McCulloch a Pitts ve svém článku nazvaném *A Logical Calculus of Ideas Immanent in Nervous Activity* [1] vycházejí z předpokladu, že nervová aktivita má binární charakter,²⁾ to znamená, že biologický neuron v daném okamžiku vyše plný akční potenciál (takzvaně zahoří), nebo nevyšle vůbec žádný potenciál (nezahoří). V jeho aktivitě tedy neexistuje žádný mezistav, něco na způsob částečného zahoření. Nervovou aktivitu je proto možné přesně popsat formálními prostředky výrokové logiky. McCulloch a Pitts v článku dokazují, že chování každé sítě neuronů je možné popsat pomocí formule výrokové logiky a naopak, každou formuli je možné reprezentovat pomocí sítě neuronů.

²⁾V abstraktu článku autoři přímo mluví o „all-or-none character of nervous activity, neural events and the relations among them“.

Pro důkaz svých tvrzení vymysleli McCulloch a Pitts *umělý neuron*, který modeloval fungování neuronu biologického. Tomuto modelu dnes říkáme *McCullochův–Pittsův umělý neuron* (zkráceně *MCP neuron*), občas také *prahová logická jednotka* (tento název je velmi výstižný, jak můžeme vidět z dalšího popisu). Princip jeho fungování je velmi jednoduchý. MCP neuron má n binárních vstupů $x_1, x_2, \dots, x_n$ a právě jeden binární výstup y , je tedy pouze realizací booleovské funkce n proměnných

$$\{0, 1\}^n \rightarrow \{0, 1\}.$$

Každý vstup neuronu je buď *excitační*, nebo *inhibiční* a předpokládejme, že vždy existuje alespoň jeden excitační a alespoň jeden inhibiční vstup.³⁾ Excitační vstupy budeme značit pomocí symbolů $x_1^e, x_2^e, \dots, x_p^e$ a inhibiční vstupy pomocí symbolů $x_1^i, x_2^i, \dots, x_r^i$, přičemž zřejmě platí

$$p + r = n.$$

Výstup neuronu budeme označovat symbolem y .

Neuronu je navíc přiřazena prahová hodnota $\theta \in \{0, 1, 2, \dots\}$, která spolu se vstupními hodnotami ovlivňuje jeho výstup, a to následujícím způsobem: Na výstupu neuronu je logická jednička, pokud excitační buzení vstupů převáží nad inhibičním. Přesněji řečeno, výstup je roven logické jedničce tehdy, je-li počet excitačních vstupů nastavených na jedničku alespoň roven prahové hodnotě θ a zároveň jsou všechny inhibiční vstupy nastaveny na nulu. Jinými slovy, $y = 1$, jsou-li současně splněny následující dvě podmínky:

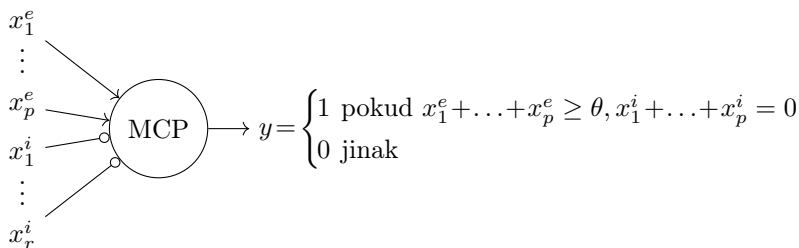
$$x_1^e + \dots + x_p^e \geq \theta, \tag{1}$$

$$x_1^i + \dots + x_r^i = 0. \tag{2}$$

Pokud není splněna alespoň jedna z těchto podmínek, výstup neuronu je nastaven na logickou nulu.

Schéma MCP neuronu a jeho fungování je znázorněno na obr. 3. Kvůli přehlednosti odlišujeme v tomto a dalších obrázcích excitační a inhibiční vstupy různými šípkami. Konkrétně excitační vstup označíme šípkou s vyplněným kolečkem, inhibiční vstup naopak šípkou s prázdným kolečkem.

³⁾ Tato podmínka je zavedena z čistě technických důvodů tak, abychom si mírně zjednodušili definici výstupní hodnoty neuronu.



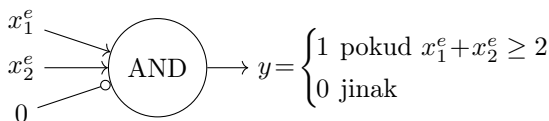
Obr. 3 McCullochův–Pittsův umělý neuron.

3. Booleovské operace pomocí MCP neuronu

Základní booleovské operace – tedy operace konjunkce (logický součin, AND), disjunce (logický součet, OR) a negace (NOT) – je nyní možné vyjádřit pomocí vhodného nastavení MCP neuronu. Konkrétně operaci konjunkce vyjádříme pomocí MCP neuronu s pouze dvěma excitačními vstupy x_1^e a x_2^e , jedním inhibičním vstupem, který je konstantně roven 0 a s prahem $\theta = 2$. Zdůvodnění je skutečně jednoduché. Pokud je totiž jediný inhibiční vstup stále nulový, pak je podmínka (2) vždy splněna. Podmínka (1), v tomto případě $x_1^e + x_2^e \geq 2$, je splněna pouze tehdy, když současně platí $x_1^e = 1$ a $x_2^e = 1$ (viz také tabulka 1). Tento neuron, jehož schéma je uvedeno na obr. 4, tak skutečně realizuje na svých excitačních vstupech operaci konjunkce.

x_1^e	x_2^e	$x_1^e + x_2^e$	y
0	0	0	0
0	1	1	0
1	0	1	0
1	1	2	1

Tabulka 1 Hodnoty excitačních vstupů a výstupu MCP neuronu, který realizuje operaci konjunkce.

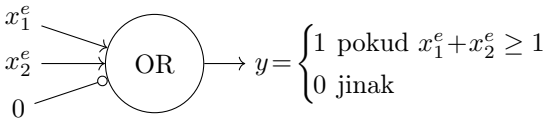


Obr. 4 Operace konjunkce vyjádřená pomocí McCullochova–Pittsova umělého neuronu.

Operaci disjunkce získáme malou úpravou předchozího nastavení. Změníme pouze prahovou hodnotu – tentokrát bude platit $\theta = 1$; vše ostatní zůstává stejné. Toto nastavení prahové hodnoty způsobí, že podmínka (1), tedy $x_1^e + x_2^e \geq 1$, je splněna v případech, že alespoň jeden excitační vstup nabývá hodnoty 1 (viz tabulka 2). Schéma tohoto neuronu je uvedeno na obr. 5.

x_1^e	x_2^e	$x_1^e + x_2^e$	y
0	0	0	0
0	1	1	1
1	0	1	1
1	1	2	1

Tabulka 2 Hodnoty excitačních vstupů a výstupu MCP neuronu, který realizuje operaci disjunkce.

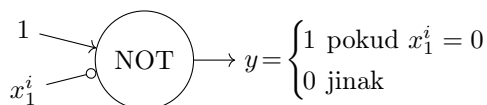


Obr. 5 Operace disjunkce vyjádřená pomocí McCullochova–Pittsova umělého neuronu.

Operaci negace vyjádříme také velmi jednoduše. Použijeme jeden excitační vstup konstantně nastavený na logickou nulu a jeden inhibiční vstup x_1^i . Prahová hodnota θ bude rovna nule. Při tomto nastavení si v jistém smyslu podmínky (1) a (2) vymění svoji roli: podmínka (1), tedy $x_1^e \geq 0$, bude vždy splněna a výstup neuronu tak bude určovat splnění nebo nesplnění podmínky (2), jak můžeme vidět v tabulce 3. Schéma tohoto MCP neuronu je znázorněno na obr. 6.

x_1^i	y
0	1
1	0

Tabulka 3 Hodnoty inhibičního vstupu a výstupu MCP neuronu, který realizuje operaci negace.



Obr. 6 Operace negace vyjádřená pomocí McCullochova–Pittsova umělého neuronu.

Je dobře známým faktem, že je možné každou formuli výrokové logiky převést do konjunktivní nebo disjunktivní normální formy, to znamená do tvaru, který využívá pouze konjunkci, disjunkci a negaci. Vhodným spojením MCP neuronů proto můžeme navrhnout *umělou neuronovou síť*, která je schopná vyjádřit libovolnou formuli výrokové logiky. Navzdory své jednoduchosti jsou umělé neuronové sítě založené na MCP neuronech mocným výpočetním nástrojem. Pokud je totiž vybavíme nekonečnou parametřovou páskou, mohou dokonce simulovat činnost libovolného Turingova stroje [2].

4. Rosenblattovo vylepšení

Řekli jsme, že MCP neuron je velmi jednoduchý model biologického neuronu. McCulloch a Pitts svou práci vnímali čistě teoreticky jako prostředek k pochopení výpočetních vlastností nervových systémů. Nezabývali se proto fyzickou realizací svého modelu a nepředpokládali jeho technické využití.

Příští díl série se bude věnovat rozšířené variantě MCP neuronu, takzvanému *perceptronu*, jehož autorem je americký psycholog Frank Rosenblatt. Perceptron je biologickému neuronu bližší, na rozdíl od neuronu McCullocha a Pittse není statický – je schopný se jistým způsobem učit. Brzy proto našel významné praktické uplatnění a je základní výpočetní jednotkou dnešních hlubokých neuronových sítí včetně těch, na kterých jsou postaveny velké jazykové modely GPT, Gemini a další.

Literatura

- [1] *McCulloch, W. S., Pitts, W.*: A Logical Calculus of Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics, roč. 5 (1943), č. 4, 115–133.
- [2] *Michael Sipser*: Introduction to the Theory of Computation. 3. vydání, Cengage Learning, 2012.