Umělá inteligence, rizika, příležitosti

Martina Šimůnková

KMA FP TUL

6. června 2025

Program

1. Co je strojové učení a hluboké neuronové sítě.
   - Strojové učení (machine learning)
   - Hluboké neuronové sítě
   - Studium: 3Blue1Brown
2. Naše digitální stopa, její rizika a zamyšlení jak jim předcházet.
   - Center for Human Technology, The Social Dilemma
   - Juval Noach Harari
   - Mindf*ck, Christopher Wylie
3. Zdroje ke studiu.

Program

1. Co je strojové učení a hluboké neuronové sítě.
   - Strojové učení (machine learning)
   - Hluboké neuronové sítě
   - Studium: 3Blue1Brown
2. Naše digitální stopa, její rizika a zamyšlení jak jim předcházet.
   - Center for Human Technology, The Social Dilemma
   - Juval Noach Harari
   - Mindf*ck, Christopher Wylie
3. Zdroje ke studiu.

Program

1. Co je strojové učení a hluboké neuronové sítě.
   - Strojové učení (machine learning)
   - Hluboké neuronové sítě
   - Studium: 3Blue1Brown
2. Naše digitální stopa, její rizika a zamyšlení jak jim předcházet.
   - Center for Human Technology, The Social Dilemma
   - Juval Noach Harari
   - Mindf*ck, Christopher Wylie
3. Zdroje ke studiu.

Strojové učení

Ke strojovému učení (machine learning) je potřeba

- Velký objem dat
- Výkonné počítače pro trénování modelu (pro práci modelu pak stačí mnohem menší výkon)
- Algoritmy strojového učení

Strojové učení

Ke strojovému učení (machine learning) je potřeba

- Velký objem dat
- Výkonné počítače pro trénování modelu (pro práci modelu pak stačí mnohem menší výkon)
- Algoritmy strojového učení

Strojové učení

Ke strojovému učení (machine learning) je potřeba

- Velký objem dat
- Výkonné počítače pro trénování modelu (pro práci modelu pak stačí mnohem menší výkon)
- Algoritmy strojového učení

Strojové učení

Ke strojovému učení (machine learning) je potřeba

- Velký objem dat
- Výkonné počítače pro trénování modelu (pro práci modelu pak stačí mnohem menší výkon)
- Algoritmy strojového učení

Metoda nejmenších čtverců – formulace problému

Popis problému:
Daty $\{(x_i, y_i)\}_{i=1}^n$ prokládáme přímku $y = ax + b$. Zvolíme koeficienty a, b, pro které má minimální hodnotu cenová funkce c

$$c(a, b) = \sum_{i=1}^{n} (ax_i + b - y_i)^2$$

Metoda nejmenších čtverců – výpočet rovnice přímky

Přímý výpočet a, b**:** Z podmínky $\nabla c = 0$ dostaneme soustavu

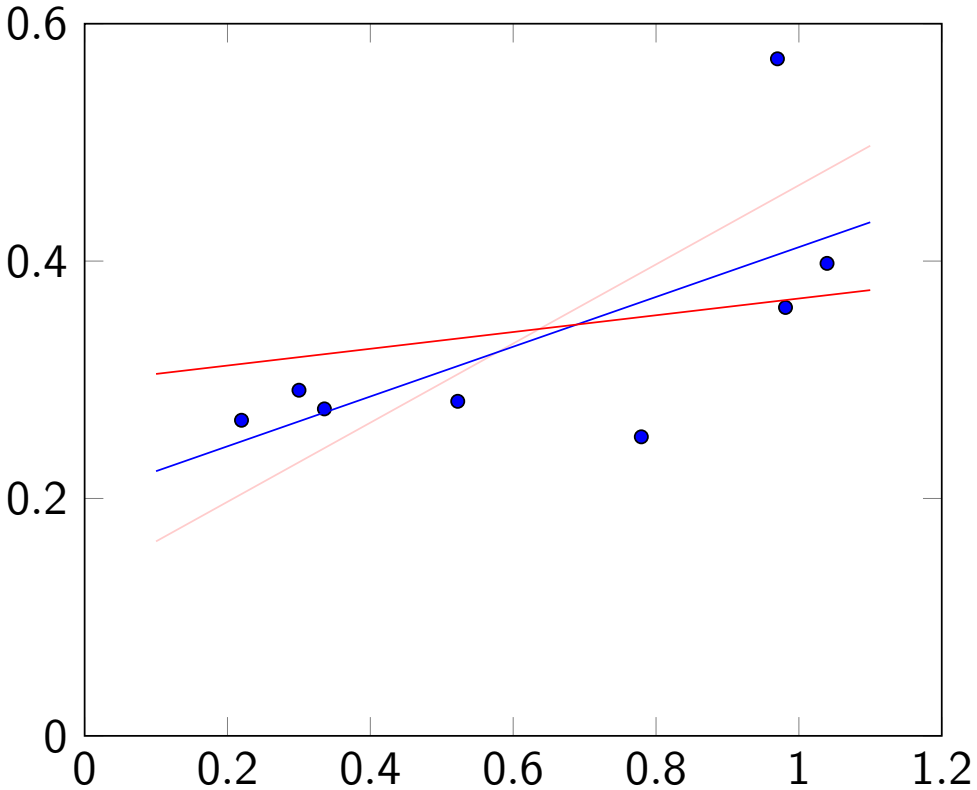
$$\begin{aligned} a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i &= \sum_{i=1}^{n} x_i y_i \\ a\sum_{i=1}^{n} x_i + nb &= \sum_{i=1}^{n} y_i \end{aligned}$$

Ilustrace strojového učení (machine learning), aneb, jak by bylo použito k výpočtu koeficientů a, b modelu: Koeficienty a, b zvolíme náhodně a postupně je pomocí dat aktualizujeme metodou gradient descent, α je vhodně zvolené malé kladné číslo

$$(a, b)_{new} = (a, b) - \alpha\nabla c$$

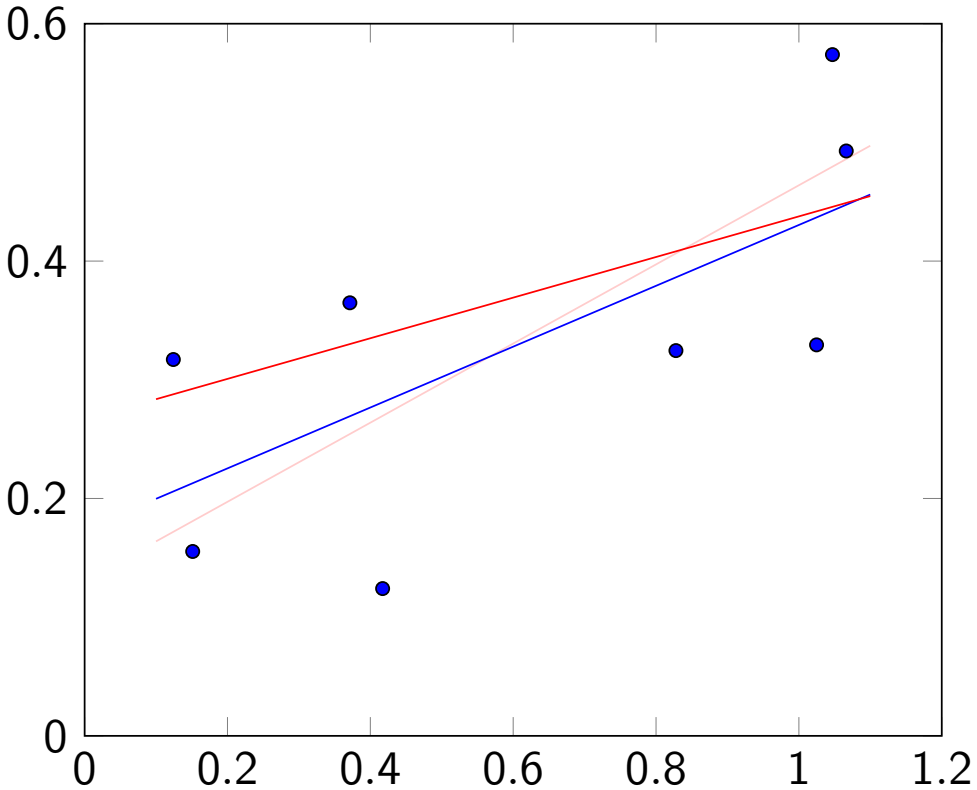
Metoda nejmenších čtverců – gradient descent method

Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



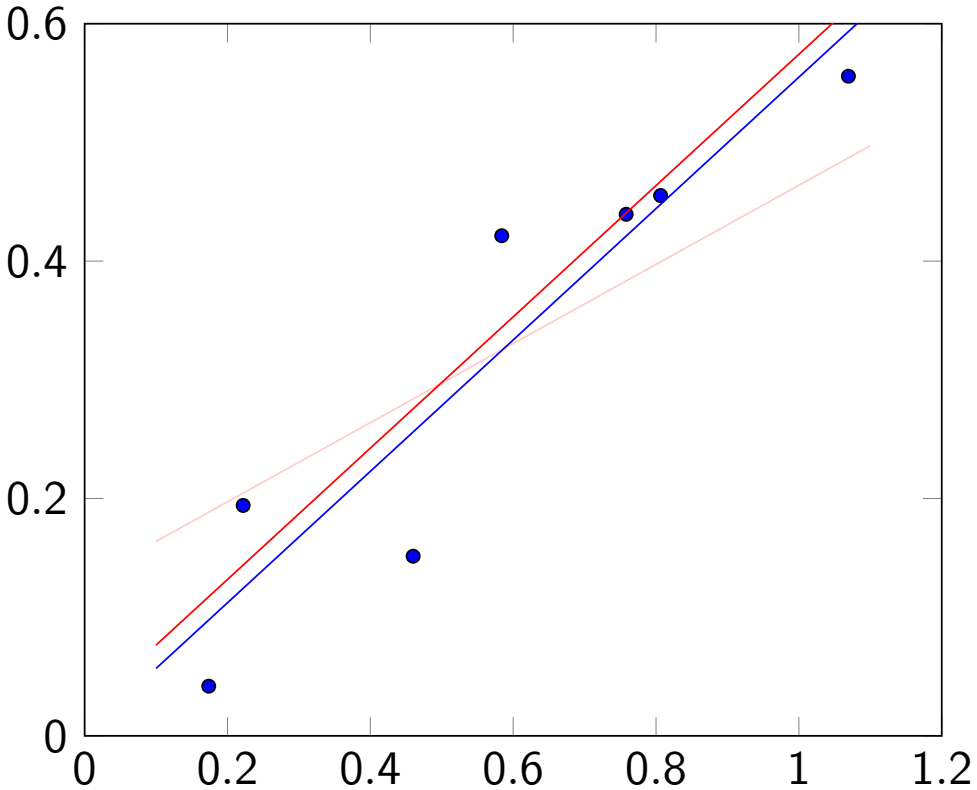
Metoda nejmenších čtverců – gradient descent method

Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



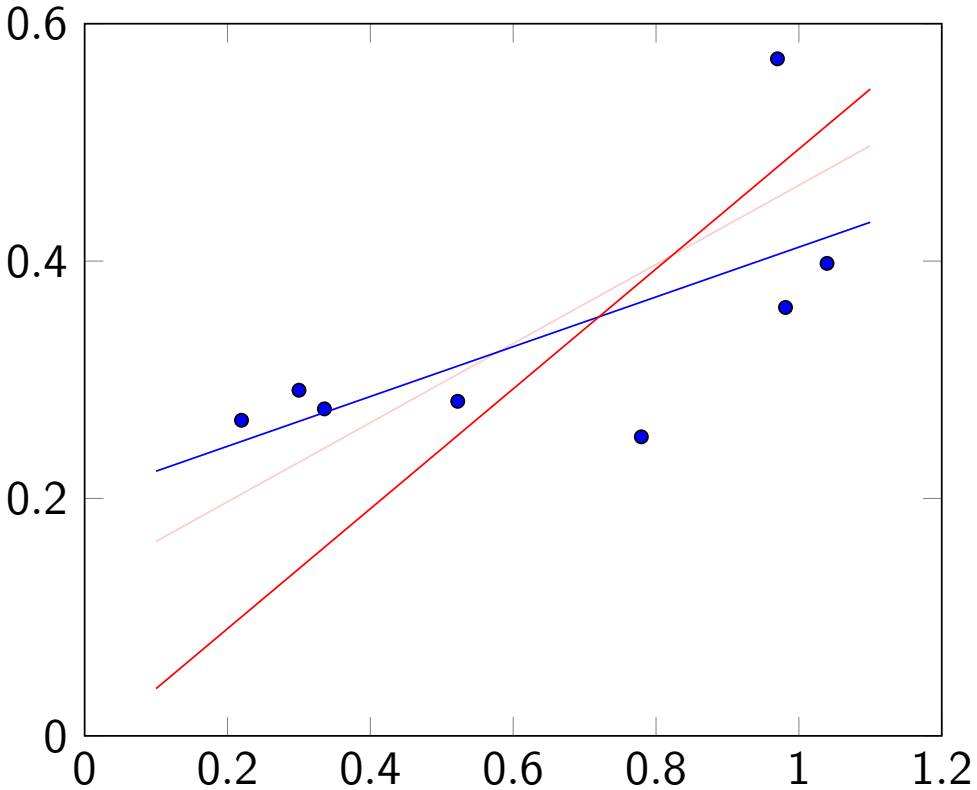
Metoda nejmenších čtverců – gradient descent method

Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



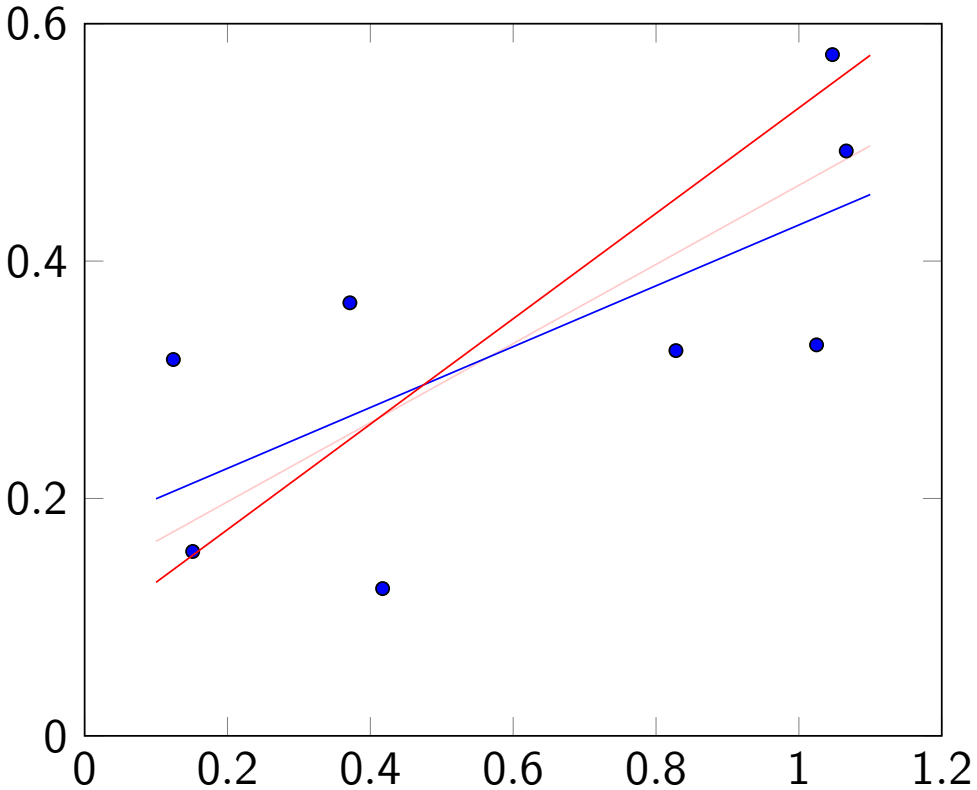
Metoda nejmenších čtverců – gradient descent method

Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



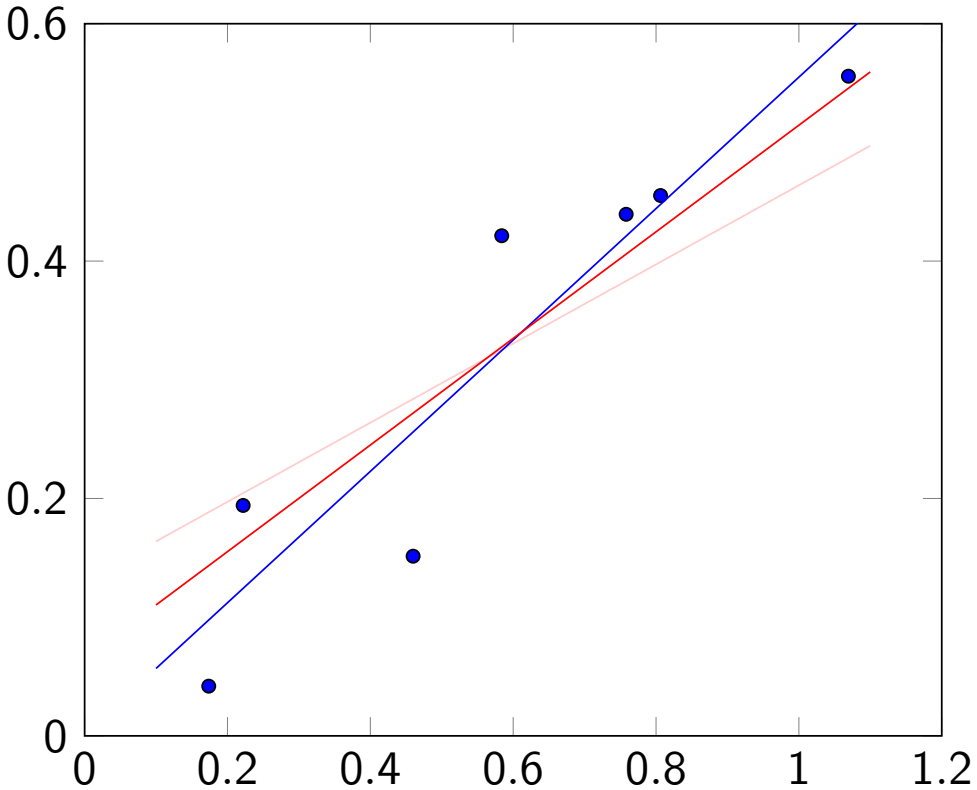
Metoda nejmenších čtverců – gradient descent method

Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



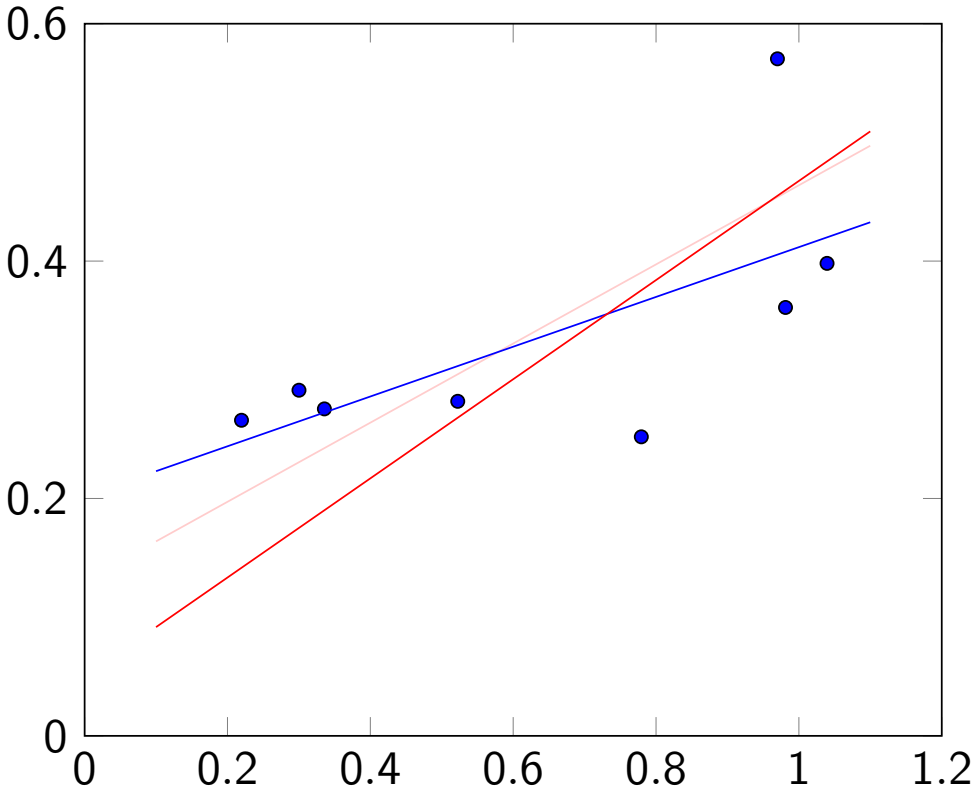
Metoda nejmenších čtverců – gradient descent method

Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



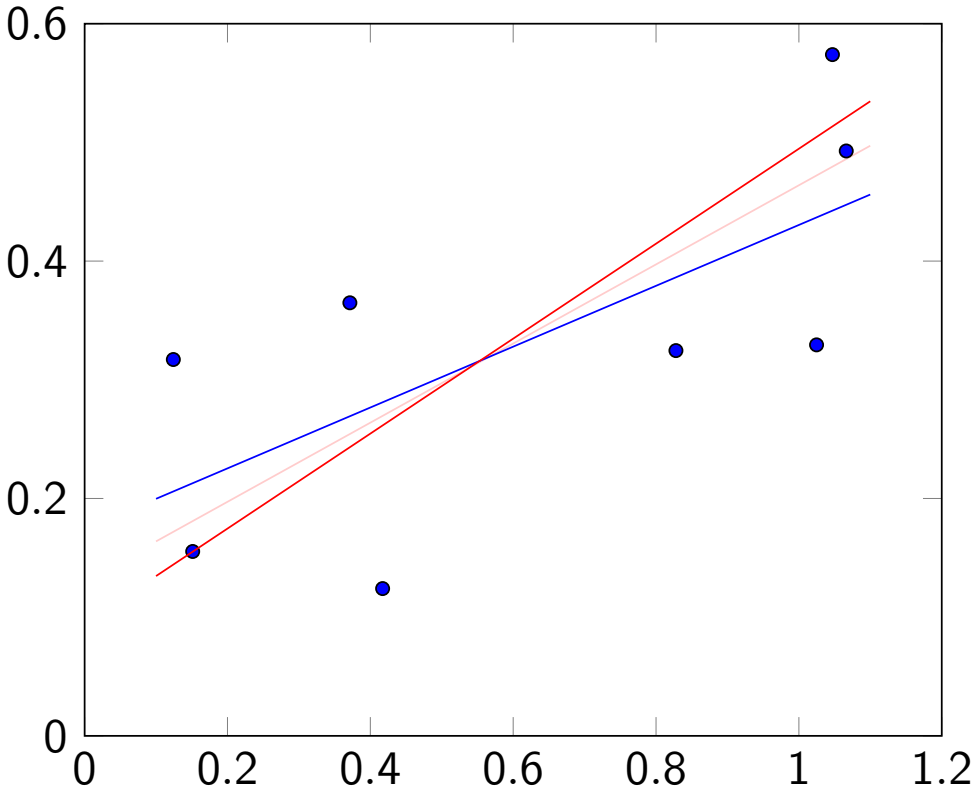
Metoda nejmenších čtverců – gradient descent method

Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



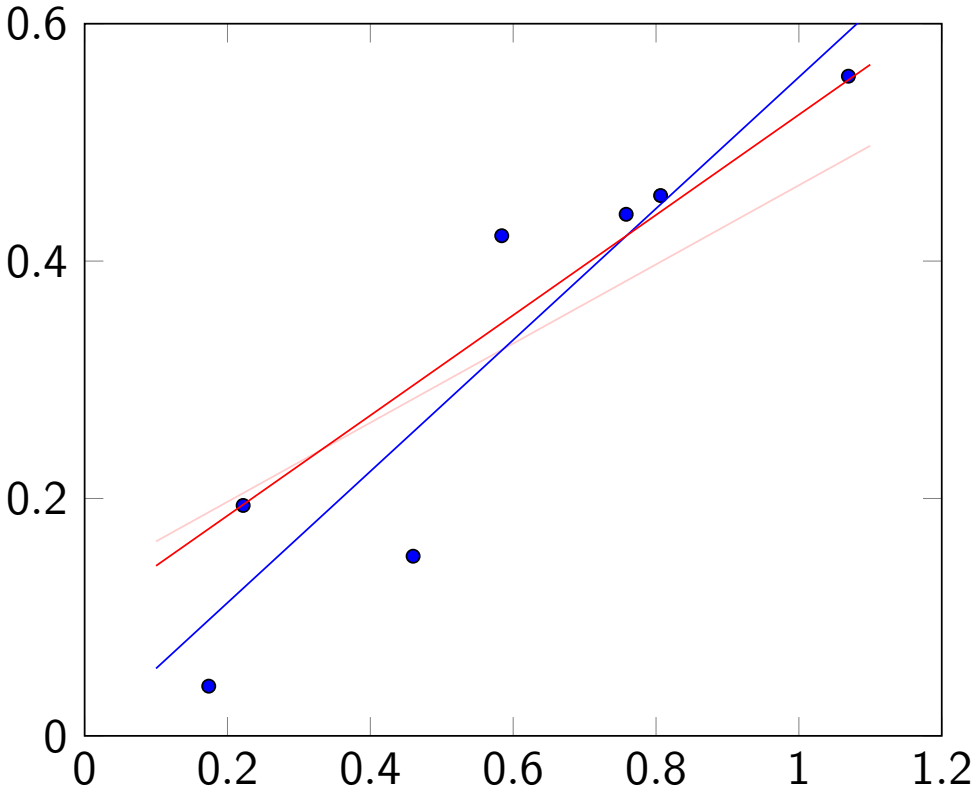
Metoda nejmenších čtverců – gradient descent method

Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



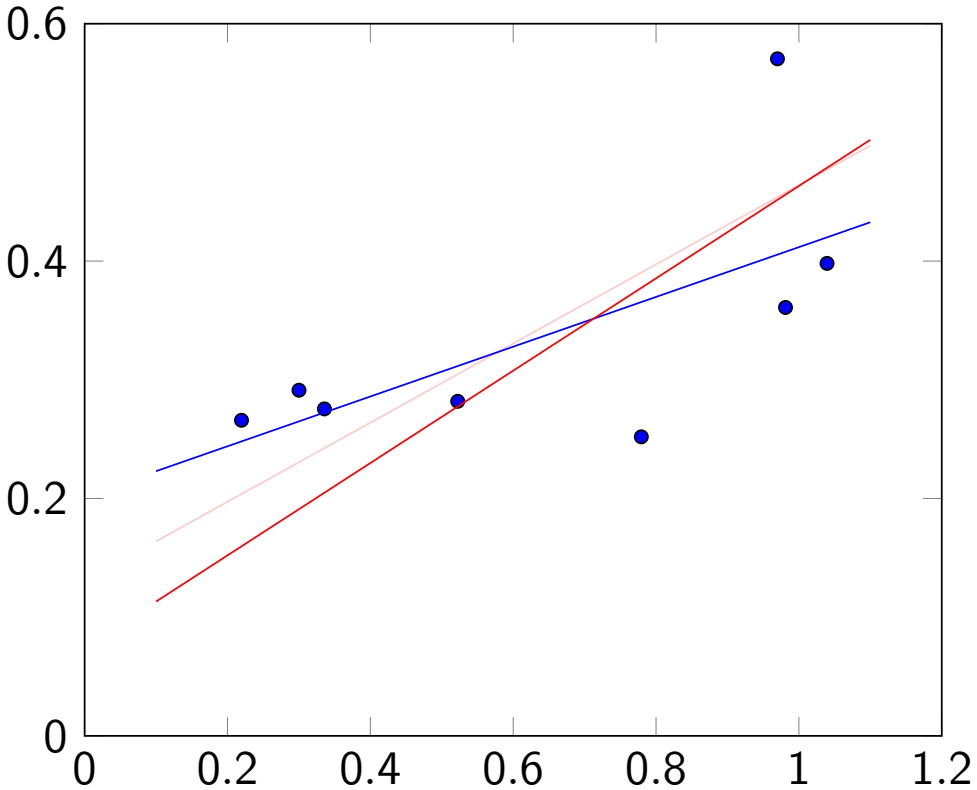
Metoda nejmenších čtverců – gradient descent method

Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



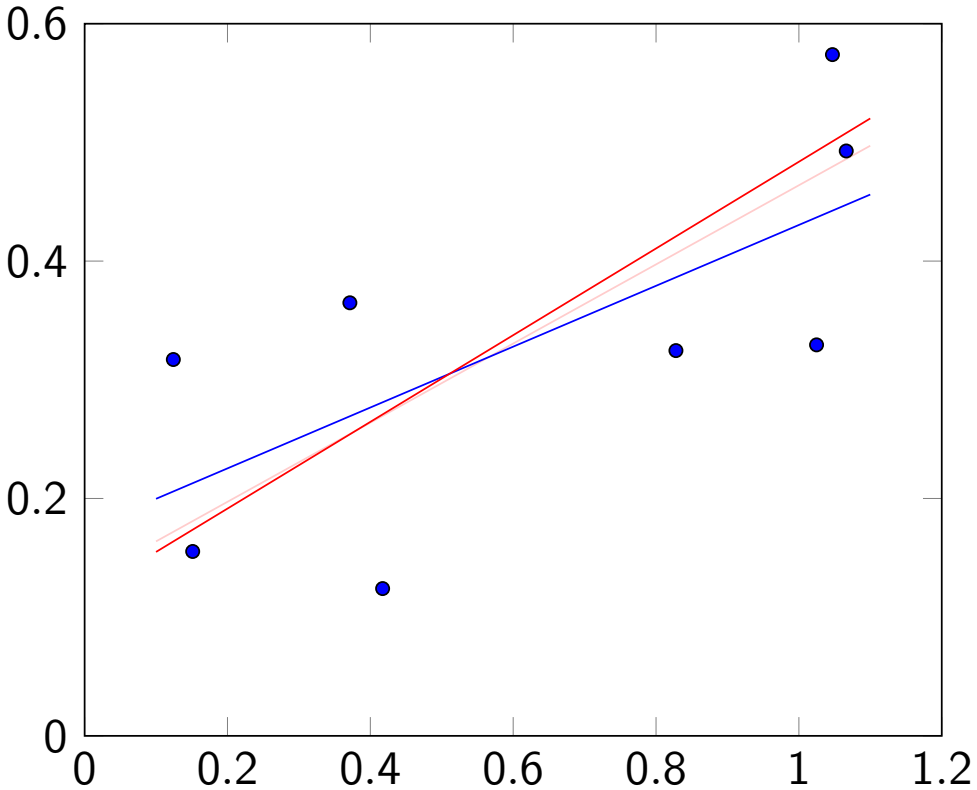
Metoda nejmenších čtverců – gradient descent method

Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



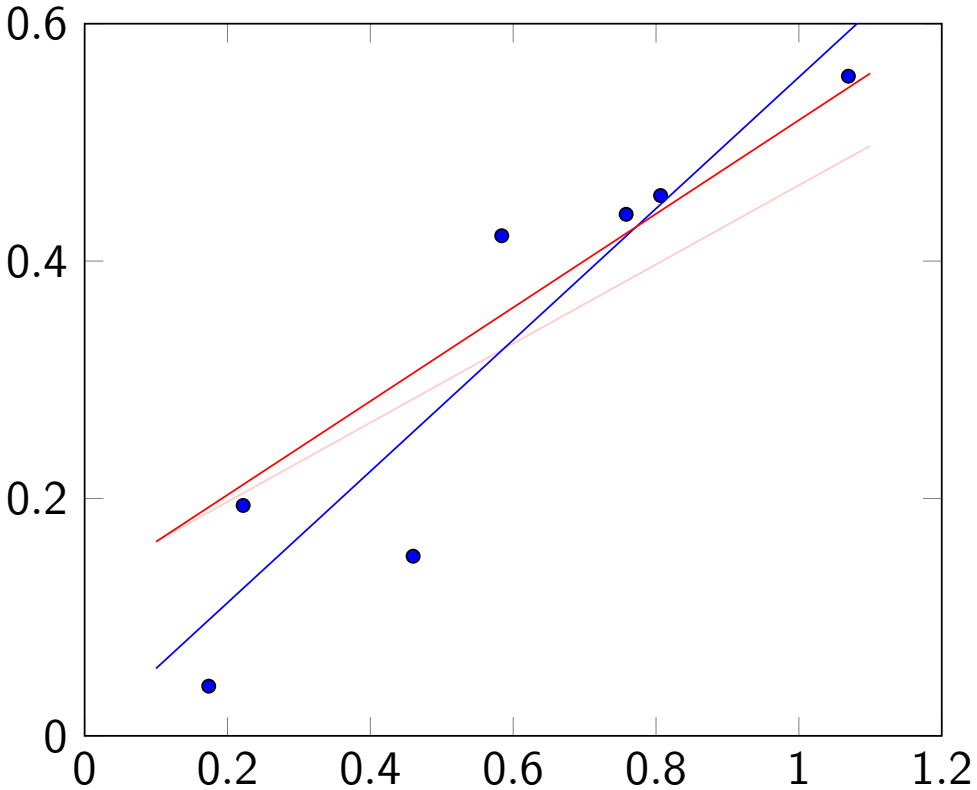
Metoda nejmenších čtverců – gradient descent method

Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



Metoda nejmenších čtverců – gradient descent method

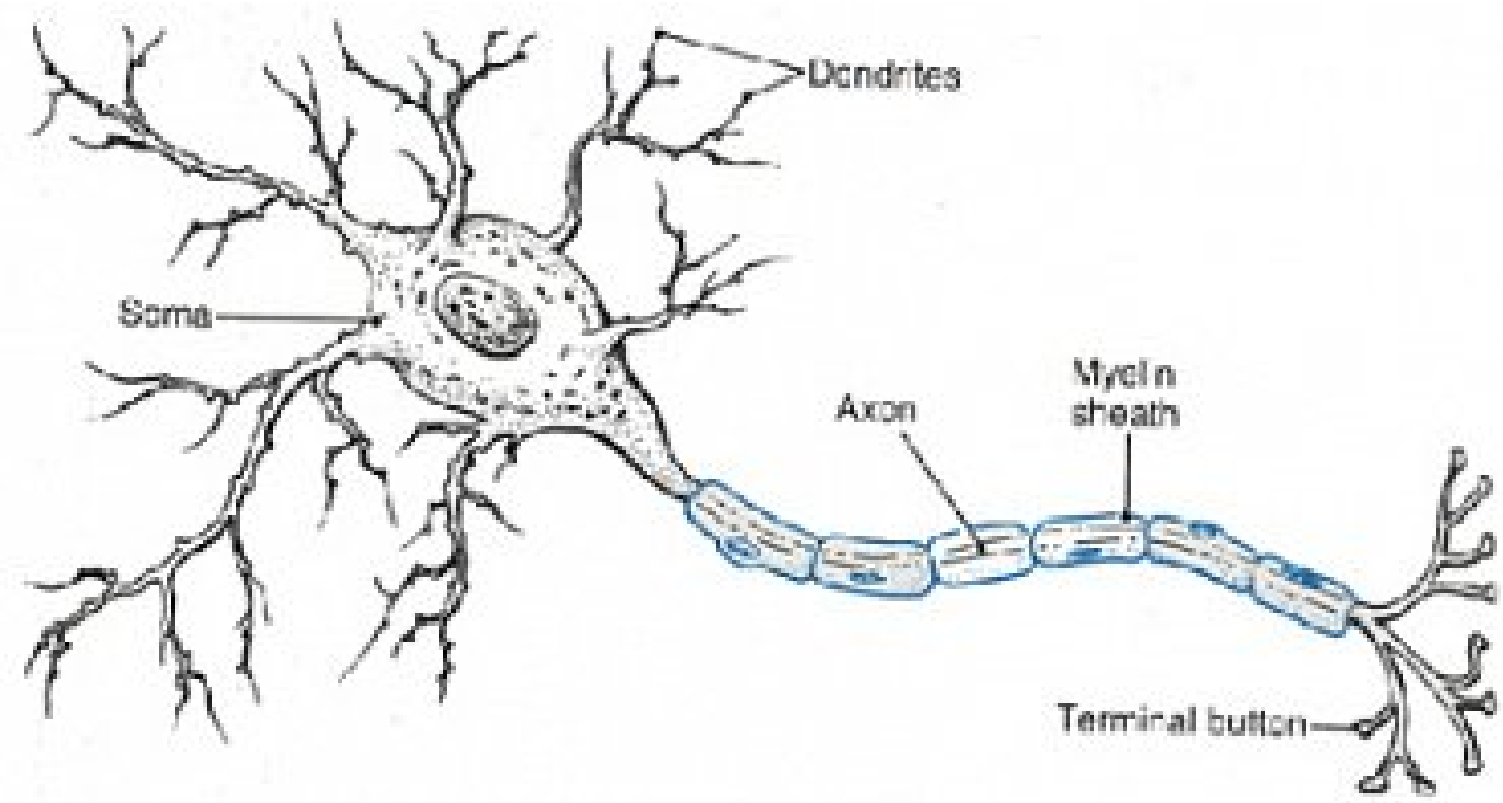
Ilustrace strojového učení bez neuronové sítě:
Sytě červená přímka je zvolena náhodně a postupně je daty upravovaná metodou gradient descent.
Slabě červená přímka je proložená všemi daty. Modrá aktuálními daty.



Program

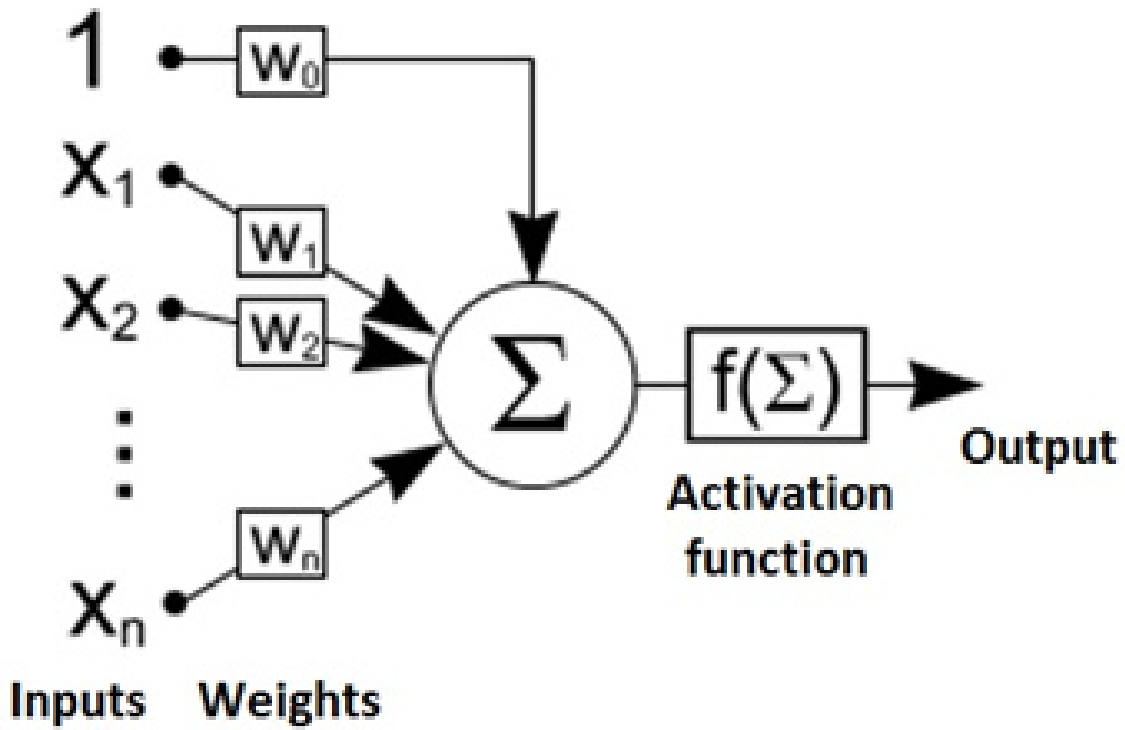
1. Co je strojové učení a hluboké neuronové sítě.
   - Strojové učení (machine learning)
   - Hluboké neuronové sítě
   - Studium: 3Blue1Brown
2. Naše digitální stopa, její rizika a zamyšlení jak jim předcházet.
   - Center for Human Technology, The Social Dilemma
   - Juval Noach Harari
   - Mindf*ck, Christopher Wylie
3. Zdroje ke studiu.

Nervová buňka



https://www.root.cz/clanky/biologicke-algoritmy-4-neuronove-site

Umělý neuron (perceptron, 1957)

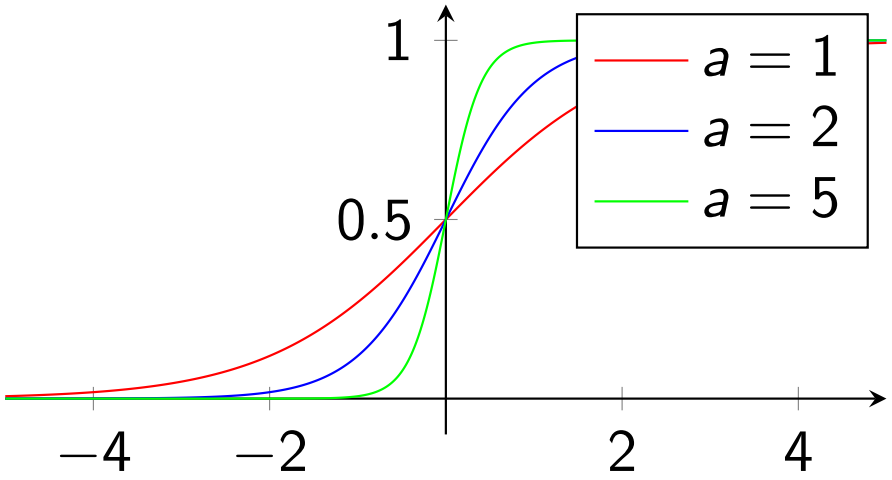


https://www.root.cz/clanky/biologicke-algoritmy-4-neuronove-site

$$\Sigma = w_0 + \sum_{k=1}^{n} w_k x_k$$

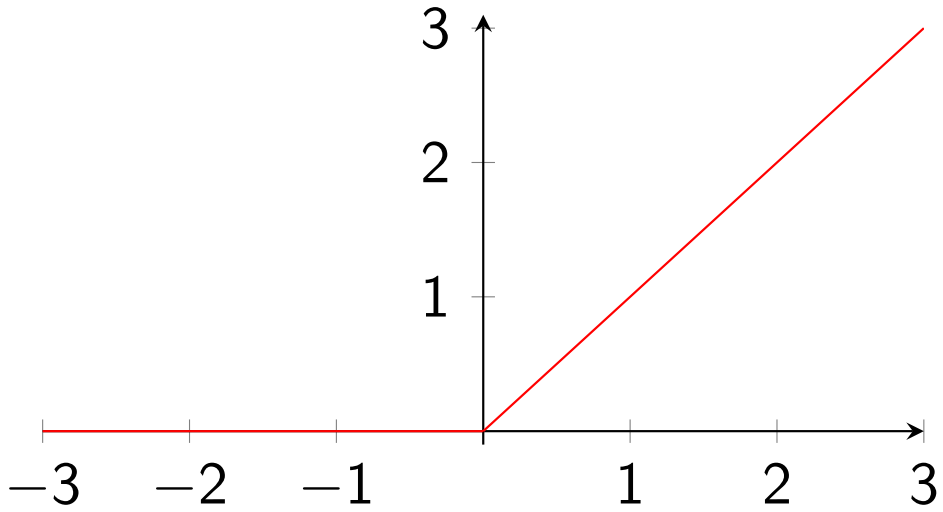
Příklady aktivačních funkcí

Sigmoid



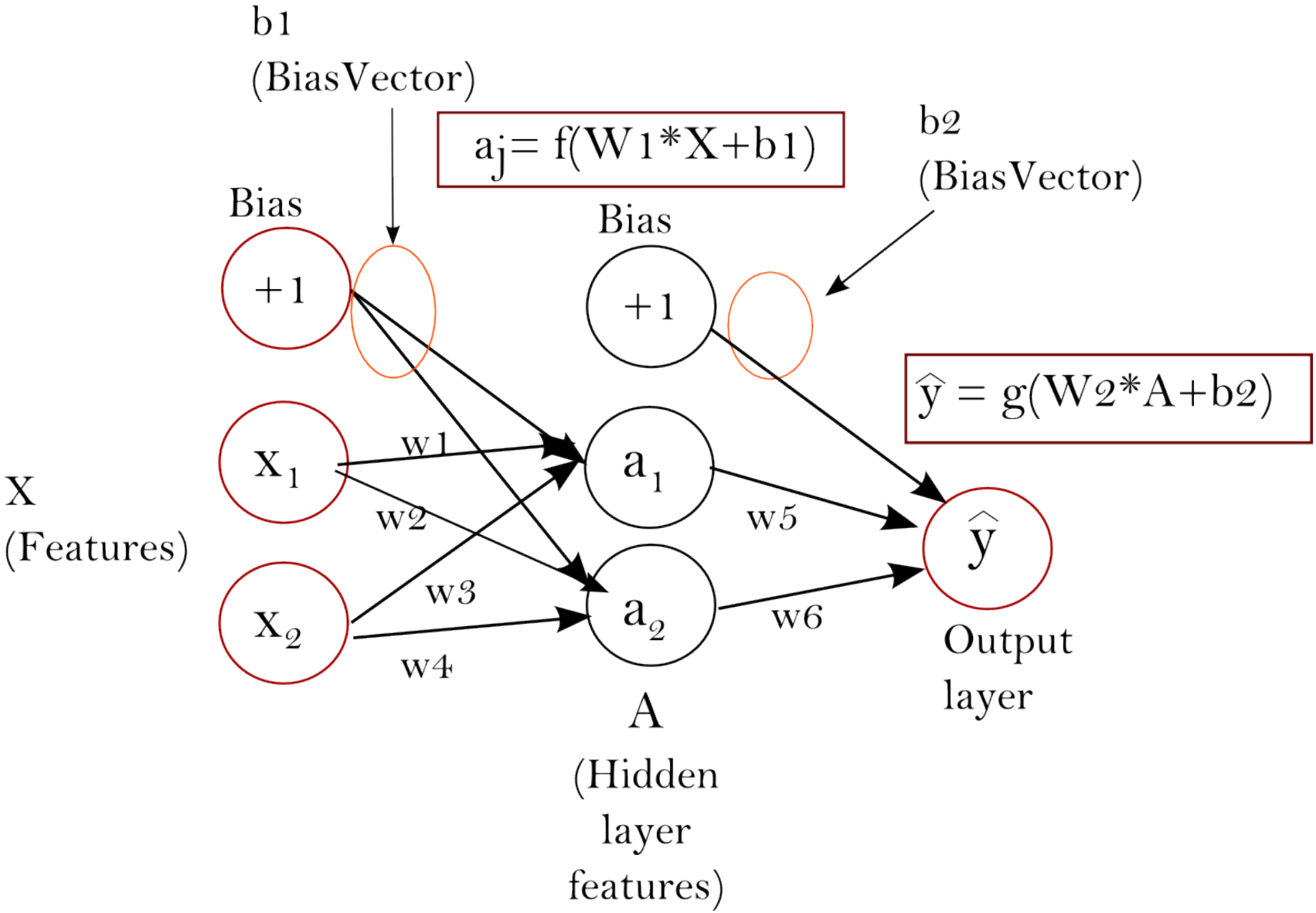
$$f(\Sigma) = \frac{1}{1 + \exp(-a\Sigma)}$$

ReLU (rectified linear unit)



$$f(\Sigma) = \max(0, \Sigma)$$

Hluboká neuronová síť (alespoň jedna skrytá vrstva)



https://i.sstatic.net/5SkqL.png

Vstup (input) a výstup (output) hluboké sítě

- Vstup: Informace o pixelech obrázku.
  Výstup: Pravděpodobnost, že na obrázku je ...
- Vstup: Text.
  Výstup: Pravděpodobnostní rozložení pokračujícího slova/tokenu textu.
- Vstup: Pozice v partii deskové hry (například šachy, go).
  Výstup: Pravděpodobnostní rozložení dalšího tahu, pravděpodobnost výhry.
- Vstup: Informace o žadateli o půjčku.
  Výstup: Pravděpodobnost, že žadatel bude půjčku řádně splácet.

Supervised learning

1. Sestavíme hlubokou síť a nastavíme náhodně její váhy w_k.
2. Sestavíme sadu dat. Ke každé položce v datech máme výstup v_0 nebo sadu výstupů (v_k).
3. Sadu rozdělíme na trénovací a testovací část.
4. Síti předložíme jeden soubor trénovacích dat a podle výstupů sítě (y_k) upravíme váhy sítě w_i tak, aby se snížila hodnota odchylky

$$\sum_k (v_k - y_k)^2$$

   Váhy se přenastavují pomocí algoritmu zpětného šíření chyby (error backpropagation).
5. Opakujeme krok 4 pro další soubor dat.
6. Kroky 4, 5 můžeme zopakovat několikrát pro celou datovou sadu.
7. Pomocí testovací sady dat natrénovanou síť otestujeme.

Reinforcement learning

Neuronová síť řídí chování agenta v prostředí, reaguje na změny v prostředí.

Chování agenta je po určitém počtu akcí ohodnoceno.

Na základě ohodnocení jsou přenastaveny váhy sítě.

Příklady:
Samořiditelná auta: počet kilometrů bez nehody.
Odpověď chatbota: ohodnocení uživatele (například počtem hvězdiček).
Konverzace chatbota s uživatelem: $+-$ ohodnotí člověk v roli hodnotitele (anotátora).
Desková hra: výhra, prohra.
Pohyb robota: zda a jak rychle splní úkol.

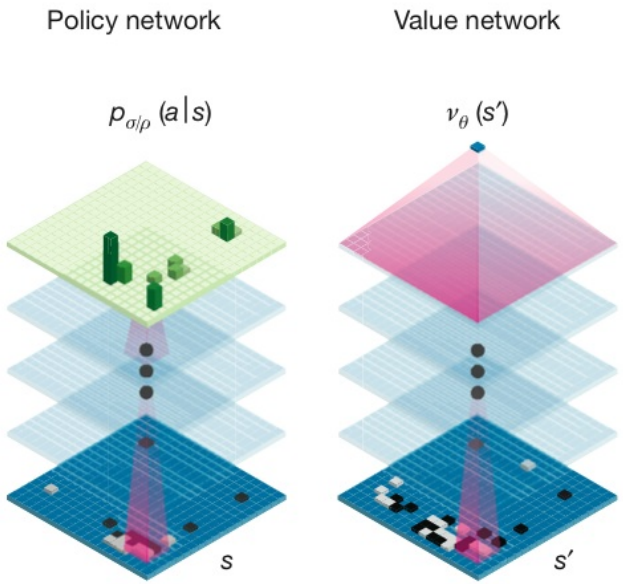
Video (72s) o tom, jak se roboti učí hrát fotbal:
`https://www.youtube.com/watch?v=qrvK_KuIeJk&t=181s`

Deskové hry

DeepBlue
V roce 1997 porazilo Garriho Kasparova.
Uvažuje a propočítává všechny možné tahy do určité hloubky.

AlphaGo
V roce 2016 porazilo I Se-tola.
Uvažuje jen tahy vybrané neuronovou sítí, která je natrénovaná primárně na partiích lidí a následně na hře proti sobě.

Video (necelých pět minut) o jednom neočekávaném tahu robota.



AlphaZero
Systém, který se učí výhradně hraním sám proti sobě, na vstupu má jen pravidla. Otestován na hrách šachy, go, šógi.

Program

1. Co je strojové učení a hluboké neuronové sítě.
   - Strojové učení (machine learning)
   - Hluboké neuronové sítě
   - Studium: 3Blue1Brown
2. Naše digitální stopa, její rizika a zamyšlení jak jim předcházet.
   - Center for Human Technology, The Social Dilemma
   - Juval Noach Harari
   - Mindf*ck, Christopher Wylie
3. Zdroje ke studiu.

Zdroje o neuronových sítích

Výuková videa na kanálu 3Blue1Brown
`https://www.3blue1brown.com/topics/neural-networks`

Program

1. Co je strojové učení a hluboké neuronové sítě.
   - Strojové učení (machine learning)
   - Hluboké neuronové sítě
   - Studium: 3Blue1Brown
2. Naše digitální stopa, její rizika a zamyšlení jak jim předcházet.
   - Center for Human Technology, The Social Dilemma
   - Juval Noach Harari
   - Mindf*ck, Christopher Wylie
3. Zdroje ke studiu.

Program

1. Co je strojové učení a hluboké neuronové sítě.
   - Strojové učení (machine learning)
   - Hluboké neuronové sítě
   - Studium: 3Blue1Brown
2. Naše digitální stopa, její rizika a zamyšlení jak jim předcházet.
   - Center for Human Technology, The Social Dilemma
   - Juval Noach Harari
   - Mindf*ck, Christopher Wylie
3. Zdroje ke studiu.

Center for Human Technology

Center for Human Technology
https://www.humanetech.com/who-we-are

The Social Dilemma, film z roku 2020, trailer
https://www.imdb.com/video/vi303154457/?ref_=ttvg_vi_1

Vyhlížíme okamžik, kdy technologie překonají lidstvo . . . technologie dříve využijí lidské slabosti.
https://www.youtube.com/watch?v=iYVVgGWUKKg&t=315s

Jsou lidé tak zlí?
https://www.youtube.com/watch?v=iYVVgGWUKKg&t=977s

Program

1. Co je strojové učení a hluboké neuronové sítě.
   - Strojové učení (machine learning)
   - Hluboké neuronové sítě
   - Studium: 3Blue1Brown
2. Naše digitální stopa, její rizika a zamyšlení jak jim předcházet.
   - Center for Human Technology, The Social Dilemma
   - Juval Noach Harari
   - Mindf*ck, Christopher Wylie
3. Zdroje ke studiu.

Z knihy Nexus

Za etnické čistky není nikdy odpovědná jen jedna strana. Vinu lze rozdělit mezi celou řadu aktérů. Jisté je, že nenávist vůči Rohingům existovala už před příchodem Facebooku do Myanmaru a že největší podíl na zvěrstvech z let 2016–2017 nesou lidé jako Wirathu a myanmarští vojenští velitelé, a stejně tak vůdci ARSA z řad Rohingů, kteří tuto vlnu násilí vyvolali. Určitá odpovědnost padá také na programátory a vedoucí pracovníky Facebooku, kteří algoritmy nakódovali, dali jim příliš velkou moc a nesnažili se je moderovat. *Zásadní však je, že část viny padá i na algoritmy samotné.* Metodou pokus–omyl zjistily, že rozhořčení vede k aktivitě uživatelů, a bez výslovného příkazu shora se rozhodly tuto emoci podporovat. ...

Z knihy Nexus

... Tím se dostáváme k definujícímu rysu umělé inteligence: schopnosti učit se a samostatně jednat. I kdybychom algoritmům přisoudili jen jedno procento viny, stále se jedná o první etnickou čistku v historii částečně způsobenou rozhodnutími jiné než lidské inteligence. A s velkou pravděpodobností ne poslední – především proto, že algoritmy už nezůstávají jen u zviditelňování fake news a konspiračních teorií vytvořených extremisty z masa a kostí. Na začátku dvacátých let už algoritmy fake news a konspirační teorie samy vytvářejí.

Harari z rozhovoru v deníku N

Osobně mám informační dietu stejně, jako mají lidé dietu potravinovou.
Informační dietu mohu každému jen doporučit, protože jsme zaplaveni příliš velkým množstvím informací a většina z nich je balast. Stejně jako si lidé dávají velký pozor na to, co jedí, měli by si dávat velký pozor i na to, jaké informace konzumují a v jakém množství.
Mám tendenci číst dlouhé knihy, a ne krátké tweety. Pokud chci opravdu pochopit, co se děje na Ukrajině nebo co se odehrává třeba v Libanonu, přečtu si o tom v závislosti na tématu několik knih, přijde na to, jestli jde o LLM, nebo o Římskou říši, o historii, o biologii, nebo informatiku.
Také si dávám informační půst. Protože většina informací je balast a informace je potřeba zpracovávat, pouhý přísun dalších informací do hlavy nijak nezaručuje, že budete chytřejší nebo moudřejší. . . .

Harari, rozhovor v deníku N

. . . Jen si zaplníte hlavu balastem.
Takže stejně jako je důležité informace přijímat, potřebujeme také čas na jejich strávení a detoxikaci hlavy. Proto denně dvě hodiny medituju. Každý rok odjíždím na třicet až šedesát dní do ústraní, kde už nepřijímám žádné nové informace. Panuje tam ticho. V meditačním centru s ostatními lidmi ani nemluvíte, jen zpracováváte, trávíte, detoxifikujete všechno, co jste během roku nashromáždili.
Vím, že to většina lidí bude pokládat za extrém. Většina lidí na to navíc nemá čas ani prostředky. *Ale přesto si myslím, že by se každý měl více zamyslet nad svým informačním jídelníčkem a také si alespoň jednou týdně třeba na den od informací odpočinout. Nebo si vyhradit pár hodin denně, kdy už další informace nebude přijímat.*

Program

1. Co je strojové učení a hluboké neuronové sítě.
   - Strojové učení (machine learning)
   - Hluboké neuronové sítě
   - Studium: 3Blue1Brown
2. Naše digitální stopa, její rizika a zamyšlení jak jim předcházet.
   - Center for Human Technology, The Social Dilemma
   - Juval Noach Harari
   - Mindf*ck, Christopher Wylie
3. Zdroje ke studiu.

Další četba z Mindf*ck

Sociální média a internetové platformy nejsou služby - jsou to architektury a infrastruktury. Tím, že své architektury označují za „služby“, snaží se společnosti přenést zodpovědnost na zákazníky, protože odklikli „souhlas“. Ale v žádném jiném sektoru podnikání se takovéto břemeno na zákazníky neklade. Pasažéři leteckých společností nejsou žádáni, aby „odsouhlasili“ konstrukci letadel, po hotelových hostech nikdo nežádá, aby „odsouhlasili“ počet únikových východů z budovy, nikdo po nás nechce, abychom „odsouhlasili“ čistotu vody, kterou nám servírují k pití. A jako dřívější návštěvník klubů vás můžu ujistit, že když je kapacita baru nebo koncertu překročena a podmínky začnou být evidentně nebezpečné, požární inspektoři nařídí návštěvníkům, aby budovu opustili.

Další četba z Mindf*ck

Historie stavebních norem jde až do roku 64 n. l., kdy Nero po devastujícím požáru Říma, který řádil devět dní, omezil výšku domů, šířku ulic a zásobování obyvatel vodou. I když požár Bostonu v roce 1631 přiměl město zakázat dřevěné komíny a doškové střechy, první moderní stavební normy se objevily až po zničujícím požáru Londýna v roce 1666. Tak jako v Bostonu byly londýnské domy postaveny většinou ze dřeva a měly doškové střechy, takže se požár rychle rozšířil a trval čtyři dny. Schořelo 13 200 domů, 84 kostelů a téměř všechny vládní budovy. V reakci na požár král Karel II. nařídil, že nikdo „nevztyčí dům či budovu, velkou nebo malou, aniž by byla z cihel nebo kamene“. Jeho výnos rovněž nařídil rozšířit průběžné ulice, aby oheň nemohl přeskočit z jedné strany na druhou. Po dalších velkých požárech v devatenáctém století následovala příklad Londýna mnohá další města, až byli nakonec jmenováni inspektoři s pravomocí soukromé stavby prohlížet a vydávat úřední souhlas, že jsou bezpečné pro jejich obyvatele i pro veřejnost. . . .

Další četba z Mindf*ck

. . . Objevila se nová pravidla a nakonec se termíny „bezpečnostní předpisy a normy“ staly všezahrnujícím principem, který dokázal zabránit nebezpečným nebo nepovoleným projektům, bez ohledu na přání majitelů pozemků nebo na souhlas obyvatel. Platforma jako Facebook zažívá svoje požáry už řadu let – Cambridge Analytica, ruské vměšování, myanmarská etnická čistka, hromadná střelba na Novém Zélandě – a tak jako následovaly změny po Velkém požáru Londýna, musíme dohlédnout dál než politici a zabývat se problémy digitální architektury, která ohrožuje sociální harmonii a blaho občanů.

Další četba z Mindf*ck

Internet zahrnuje nespočet různých druhů architektur, s nimiž lidé interagují každý den, někdy i několikrát denně. A jak se digitální a fyzický svět prolínají, tyto digitální architektury mají stále větší a větší vliv na naše životy. *Právo na soukromí* je fundamentální lidské právo a jako takové by mělo být hodnoceno. Jenomže soukromí je příliš často narušeno pouhým kliknutím na „souhlasím“ pod nekonečným seznamem všeobecných podmínek. Tento jakoby souhlas neustále umožňuje velkým technologickým platformám obhajovat svoje manipulativní praktiky neupřímným odvoláním se na „souhlas uživatele“. A nás to staví do pozice, kdy se nezabýváme designem – a designéry – těchto cinknutých architektur, ale neplodně se místo toho soustřeďujeme na aktivitu uživatele, který nechápe, jak byl systém navržen, a ani nemá možnost do ničeho mluvit. Nedáváme lidem možnost „vybrat“ si budovy, které mají vadnou elektroinstalaci nebo nemají nouzový východ. Bylo by to nebezpečné – a žádné podmínky připíchnuté na dveřích nepomohou architektovi, aby dostal svolení nebezpečnou stavbu realizovat. Proč by inženýři a architekti softwaru a internetových platforem měli mít výjimku?

Další četba z Mindf*ck

Tak jako v případě tradičních stavebních předpisů by ústředním rysem digitálních stavebních předpisů byl princip *bezpečnosti stavby pro uživatele*. Ten bude vyžadovat, aby platformy a jejich aplikace prošly auditem jejich zneužitelnosti a sadou bezpečnostních testů *předtím*, než bude produkt uvolněn a masově rozšířen. Břemeno důkazu bude spočívat na technologických společnostech – to ony musí dokázat, že jejich produkty jsou bezpečné pro masové použití veřejností. V souladu s tím bude zakázáno využívat veřejnost k experimentům testujícím ve velkém měřítku nové prvky, aby se občané nemohli stát pokusnými králíky. To pomůže zabránit případům jako Myanmar, kdy Facebook předem nestudoval nebezpečí, zda jeho platforma nemůže spolupracovat na zažehnutí násilí v oblasti etnického konfliktu.

Další četba z Mindf*ck

Do našich životů dnes ve velkém měřítku proniká umělá inteligence, digitální ekosystémy a software vůbec, a přitom ti, kdo tyto každodenně používané přístroje a programy vyrábějí, nejsou povinni dodržovat žádné zákony nebo vynutitelné normy, které by je přiměly pečlivě zvážit etické dopady na uživatele nebo i celou společnost. Softwarové inženýrství jako profese má vážné etické problémy, které je třeba řešit. Technologické společnosti své problematické a nebezpečné platformy nekouzlí ze vzduchu – mají zaměstnance, kteří je konstruují. A problém je v tom, že softwaroví inženýři a datoví analytici nenesou za své produkty žádnou osobní zodpovědnost. Když zaměstnavatel požádá svého inženýra, aby vytvořil systémy, které manipulují lidmi, jsou eticky pochybné nebo jsou ledabyle implementované a neberou v potaz bezpečí uživatele, neexistují žádné předpisy opravňující inženýra, aby takovou práci odmítl. Za současného stavu by takovým odmítnutím riskoval postih nebo i propuštění. . . .

Další četba z Mindf*ck

. . . I kdyby se později prokázalo, že design je neetický a odporuje zákonům, společnost stejně absorbuje případný postih, zaplatí pokuty a lidé, kteří danou technologii vyrobili, neponesou žádné následky – na rozdíl od lékaře nebo právníka, který vážným způsobem poruší etické standardy své profese. To je zvrácený stav věcí, jaký neexistuje v žádné jiné profesi. Kdyby zaměstnavatel požadoval od právníka nebo zdravotní sestry, aby provedli něco neetického, mají povinnost to odmítnout, v opačném případě jim hrozí ztráta licence. Jinými slovy, mají vážné důvody, aby se zaměstnavateli vzepřeli.

Další četba z Mindf*ck

Pokud softwaroví inženýři a datoví analytici mají být profesionály, kteří jsou hodni své pověsti a vysokých platů, musí tomu odpovídat i příslušná povinnost jednat eticky. Zákony svazující technologické společnosti nebudou zdaleka tak účinné, jak by mohly být, jestliže nebudeme vyžadovat osobní zodpovědnost lidí uvnitř těchto společností. Potřebujeme přimět inženýry, aby se začali zajímat o to, co postavili. Odpolední workshop nebo semestrální kurz o zásadách etického jednání jsou naprosto nedostatečná řešení problémů, které nám předkládají nové technologie. Nemůžeme prodlužovat současný stav technologického paternalismu, v němž šéfové ze slonovinové věže Silicon Valley vytvářejí typ nebezpečných odborníků, kteří se nehodlají zabývat škodami, jež jejich práce může potenciálně způsobit.

Další četba z Mindf*ck

Potřebujeme profesní etický kodex, na jehož plnění bude dohlížet profesní komora, jako to existuje v případě stavebních inženýrů a architektů v mnoha zemích. V těchto předpisech musejí být konkrétní tresty pro softwarové inženýry nebo datové analytiky, kteří využívají svůj talent a znalosti ke konstrukci nebezpečných a neetických technologií. Tento kodex nesmí být formulován volně, ale musí jasně, konkrétně a jednoznačně stanovit, co je přijatelné a co ne. Musí obsahovat povinnost respektovat autonomii uživatele, identifikovat a specifikovat rizika a musí projít recenzním řízením. Tento etický kodex by měl také zahrnovat požadavek vzít v úvahu důsledky provedené práce na zranitelnou část populace, včetně nepřiměřeného účinku na uživatele různých ras, pohlaví, schopností či sexuálních orientací. A když po důkladném zvážení dojde inženýr k závěru, že je žádost jeho zaměstnavatele postavit jistý prvek designu neetická, bude jeho povinností práci odmítnout a celou věc nahlásit, a pokud to neudělá, bude to mít vážné profesní důsledky. Proto musí být odmítnutí práce a hlášení chráněno zákonem před perzekucí ze strany zaměstnavatele.

Další četba z Mindf*ck

Ze všech možných druhů regulace zabrání etický kodex softwarových inženýrů největšímu množství škod, jelikož přímo přinutí stavitele, aby uvažovali o své práci *předtím, než je distribuována* na veřejnosti, a neumožňuje je zbavit se morální zodpovědnosti výmluvou, že stavitelé jen poslouchali příkazy. Technologie je často odrazem hodnot, jež společnost vyznává, takže zavedení etiky do chování technologických společností je životně důležité, pokud má naše společnost stále více záviset na jejích výtvorech. Pokud budou softwaroví inženýři skutečně osobně zodpovědni za svou práci, stanou se naší nejlepší obranou proti budoucímu zneužití technologií. *A budeme-li v roli softwarových inženýrů, měli bychom se všichni snažit, abychom si při vytváření nových architektur důvěru veřejnosti v naši práci zasloužili.*

Program

1. Co je strojové učení a hluboké neuronové sítě.
   - Strojové učení (machine learning)
   - Hluboké neuronové sítě
   - Studium: 3Blue1Brown
2. Naše digitální stopa, její rizika a zamyšlení jak jim předcházet.
   - Center for Human Technology, The Social Dilemma
   - Juval Noach Harari
   - Mindf*ck, Christopher Wylie
3. Zdroje ke studiu.

Zdroje

Juval Noach Harari: Nexus (stručné dějiny informačních sítí od doby kamenné k umělé inteligenci)

Juval Noach Harari: rozhovor v deníku N
https://denikn.cz/1615227/nemeli-bychom-se-bat-terminatoru-ale-ai-byrokratu-rika-histor?ref=inc&cst=2b2bfd880bb2c8ed2999c0c4b757b3e73139c2d065d9b7a0c3b64dbd8810fe

Mimo naše téma, přesto podnětné pro učitele a rodiče

BBC: will Australia's social media ban for under 16-s work?
https://www.youtube.com/watch?v=-FNIAzYLCQA