

Neuronové sítě

(11. přednáška)

O co jde?

Machine Learning — Naučit stroje se učit

O co jde?

Máme **model výpočtu**

O co jde?

Máme **model výpočtu** (t.j. výpočetní postup jednoznačně daný vstupy a nějakými parametry)

O co jde?

Máme **model výpočtu** (t.j. výpočetní postup jednoznačně daný vstupy a nějakými parametry), chceme najít vhodné nastavení parametrů, aby postup (model) dával řešení našeho problému.

O co jde?

Máme **model výpočtu** (t.j. výpočetní postup jednoznačně daný vstupy a nějakými parametry), chceme najít vhodné nastavení parametrů, aby postup (model) dával řešení našeho problému.

Příklad

Chceme převést stupně Celsia na stupně Farenheita.

O co jde?

Máme **model výpočtu** (t.j. výpočetní postup jednoznačně daný vstupy a nějakými parametry), chceme najít vhodné nastavení parametrů, aby postup (model) dával řešení našeho problému.

Příklad

Chceme převést stupně Celsia na stupně Farenheita. Modelem výpočtu bude lineární funkce $f(x) = \alpha x + b$

O co jde?

Máme **model výpočtu** (t.j. výpočetní postup jednoznačně daný vstupy a nějakými parametry), chceme najít vhodné nastavení parametrů, aby postup (model) dával řešení našeho problému.

Příklad

Chceme převést stupně Celsia na stupně Farenheita. Modelem výpočtu bude lineární funkce $f(x) = \alpha x + b$, která je daná dvěma parametry — α, b .

O co jde?

Máme **model výpočtu** (t.j. výpočetní postup jednoznačně daný vstupy a nějakými parametry), chceme najít vhodné nastavení parametrů, aby postup (model) dával řešení našeho problému.

Příklad

Chceme převést stupně Celsia na stupně Farenheita. Modelem výpočtu bude lineární funkce $f(x) = \alpha x + b$, která je daná dvěma parametry — α, b . Vhodné nastavení parametrů v tomto případě je $\alpha = 9/5$ a $b = 32$.

O co jde?

Máme **model výpočtu** (t.j. výpočetní postup jednoznačně daný vstupy a nějakými parametry), chceme najít vhodné nastavení parametrů, aby postup (model) dával řešení našeho problému.

Příklad

Chceme převést stupně Celsia na stupně Farenheita. Modelem výpočtu bude lineární funkce $f(x) = \alpha x + b$, která je daná dvěma parametry — α, b . Vhodné nastavení parametrů v tomto případě je $\alpha = 9/5$ a $b = 32$.

Při strojovém učení zvolíme dostatečně silný model

O co jde?

Máme **model výpočtu** (t.j. výpočetní postup jednoznačně daný vstupy a nějakými parametry), chceme najít vhodné nastavení parametrů, aby postup (model) dával řešení našeho problému.

Příklad

Chceme převést stupně Celsia na stupně Farenheita. Modelem výpočtu bude lineární funkce $f(x) = \alpha x + b$, která je daná dvěma parametry — α, b . Vhodné nastavení parametrů v tomto případě je $\alpha = 9/5$ a $b = 32$.

Při strojovém učení zvolíme dostatečně silný model a nastavení parametrů chceme nalézt automaticky na základě vzorových řešení daného problému

O co jde?

Máme **model výpočtu** (t.j. výpočetní postup jednoznačně daný vstupy a nějakými parametry), chceme najít vhodné nastavení parametrů, aby postup (model) dával řešení našeho problému.

Příklad

Chceme převést stupně Celsia na stupně Farenheita. Modelem výpočtu bude lineární funkce $f(x) = \alpha x + b$, která je daná dvěma parametry — α, b . Vhodné nastavení parametrů v tomto případě je $\alpha = 9/5$ a $b = 32$.

Při strojovém učení zvolíme dostatečně silný model a nastavení parametrů chceme nalézt automaticky na základě vzorových řešení daného problému (např. víme, že $0^{\circ}C = 32^{\circ}F$ a $100^{\circ}C = 212^{\circ}F$).

O co jde?

Máme **model výpočtu** (t.j. výpočetní postup jednoznačně daný vstupy a nějakými parametry), chceme najít vhodné nastavení parametrů, aby postup (model) dával řešení našeho problému.

Příklad

Chceme převést stupně Celsia na stupně Farenheita. Modelem výpočtu bude lineární funkce $f(x) = \alpha x + b$, která je daná dvěma parametry — α, b . Vhodné nastavení parametrů v tomto případě je $\alpha = 9/5$ a $b = 32$.

Při strojovém učení zvolíme dostatečně silný model a nastavení parametrů chceme nalézt automaticky na základě vzorových řešení daného problému (např. víme, že $0^{\circ}C = 32^{\circ}F$ a $100^{\circ}C = 212^{\circ}F$).

Slogan

Slogan

Chceme, aby se stroj naučil řešit zadaný problém na základě
vzorových řešení

Slogan

Chceme, aby se stroj naučil řešit zadaný problém na základě vzorových řešení

- řešení je příliš komplikované

Slogan

Chceme, aby se stroj naučil řešit zadaný problém na základě vzorových řešení

- řešení je příliš komplikované
- problém se často mění, vyvíjí

Slogan

Chceme, aby se stroj naučil řešit zadaný problém na základě vzorových řešení

- řešení je příliš komplikované
- problém se často mění, vyvíjí
- lidská práce je drahá (v porovnání se strojovou)

- Rozpoznávání vzorců

- Rozpoznávání vzorců
 - věci na fotkách

Strojové učení — typické aplikace

- Rozpoznávání vzorců
 - věci na fotkách
 - osoby na fotkách

- Rozpoznávání vzorců
 - věci na fotkách
 - osoby na fotkách
 - výrazy tváře

- Rozpoznávání vzorců
 - věci na fotkách
 - osoby na fotkách
 - výrazy tváře
 - mluvená slova

- Rozpoznávání vzorců
 - věci na fotkách
 - osoby na fotkách
 - výrazy tváře
 - mluvená slova
- Rozpoznávání anomálií

Strojové učení — typické aplikace

- Rozpoznávání vzorců
 - věci na fotkách
 - osoby na fotkách
 - výrazy tváře
 - mluvená slova
- Rozpoznávání anomálií
 - netypické sekvence finančních transakcí

- Rozpoznávání vzorců
 - věci na fotkách
 - osoby na fotkách
 - výrazy tváře
 - mluvená slova
- Rozpoznávání anomálií
 - netypické sekvence finančních transakcí
 - netypická data přicházející ze senzorů v atomové elektrárně

- Rozpoznávání vzorců
 - věci na fotkách
 - osoby na fotkách
 - výrazy tváře
 - mluvená slova
- Rozpoznávání anomálií
 - netypické sekvence finančních transakcí
 - netypická data přicházející ze senzorů v atomové elektrárně
- Předpovídání
 - vývoj ceny akcií na burze / vývoj měnového kurzu

- Rozpoznávání vzorců
 - věci na fotkách
 - osoby na fotkách
 - výrazy tváře
 - mluvená slova
- Rozpoznávání anomálií
 - netypické sekvence finančních transakcí
 - netypická data přicházející ze senzorů v atomové elektrárně
- Předpovídání
 - vývoj ceny akcií na burze / vývoj měnového kurzu
 - jaké filmy bude mít daný člověk rád

Rozpoznávání věcí na fotkách

(zdroj: G. Hinton, Neural Networks for ML)

Rozpoznávání věcí na fotkách

(zdroj: G. Hinton, Neural Networks for ML)



40 vydra

15 křepelka

7 tetřev

6 koroptev

Rozpoznávání věcí na fotkách

(zdroj: G. Hinton, Neural Networks for ML)



40 vydra

15 křepelka

7 tetřev

6 koroptev

Rozpoznávání věcí na fotkách

(zdroj: G. Hinton, Neural Networks for ML)



- 40 vydra
- 15 křepelka
- 7 tetřev
- 6 koroptev



- 85 sněžný pluh
- 6 vrtná plošina
- 6 záchranný člun
- 2 popelářské auto

Rozpoznávání věcí na fotkách

(zdroj: G. Hinton, Neural Networks for ML)



40 vydra

15 křepelka

7 tetřev

6 koroptev



85 sněžný pluh

6 vrtná plošina

6 záchranný člun

2 popelářské auto

Rozpoznávání věcí na fotkách

(zdroj: G. Hinton, Neural Networks for ML)



40 vydra

15 křepelka

7 tetřev

6 koroptev



85 sněžný pluh

6 vrtná plošina

6 záchranný člun

2 popelářské auto



15 žížala

12 gilotina

7 orangutan

6 koště

Výpočetní modely

Výpočetní modely

- Rozhodovací stromy (decision trees)

Výpočetní modely

- Rozhodovací stromy (decision trees)
- Bayesovské sítě

Výpočetní modely

- Rozhodovací stromy (decision trees)
- Bayesovské sítě
- Support Vector Machines

Výpočetní modely

- Rozhodovací stromy (decision trees)
- Bayesovské sítě
- Support Vector Machines
- Genetické algoritmy

Výpočetní modely

- Rozhodovací stromy (decision trees)
- Bayesovské sítě
- Support Vector Machines
- Genetické algoritmy
- Neuronové sítě

Výpočetní modely

- Rozhodovací stromy (decision trees)
- Bayesovské sítě
- Support Vector Machines
- Genetické algoritmy
- **Neuronové sítě**

Proč neuronové sítě?

- studium mozku

Proč neuronové sítě?

- studium mozku
- lidského myšlení

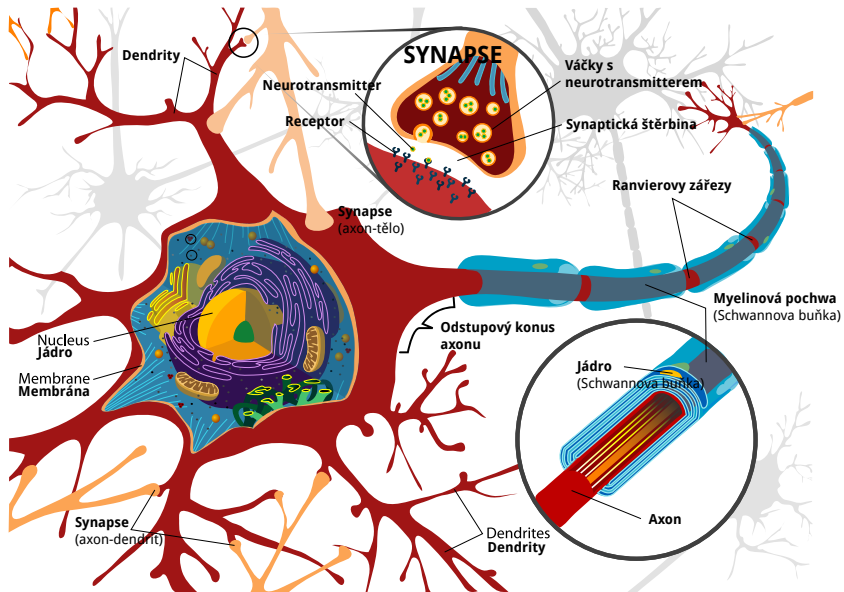
Proč neuronové sítě?

- studium mozku
- lidského myšlení
- zajímavý výpočetní model

Proč neuronové sítě?

- studium mozku
- lidského myšlení
- zajímavý výpočetní model

Biologie neuronu



Matematický neuron — $N : X_1 \times \cdots \times X_n \rightarrow Y$

Matematický neuron — $N : X_1 \times \cdots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

Matematický neuron — $N : X_1 \times \cdots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony

Matematický neuron — $N : X_1 \times \cdots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony
 - McCulloch & Pitts, 1943

Matematický neuron — $N : X_1 \times \cdots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony
 - McCulloch & Pitts, 1943
 - spike \approx pravdivostní hodnota výroku

Matematický neuron — $N : X_1 \times \cdots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony
 - McCulloch & Pitts, 1943
 - spike \approx pravdivostní hodnota výroku
 - neuron počítá se pravdivostní hodnotu jiného výroku ze vstupních výroků

Matematický neuron — $N : X_1 \times \cdots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony
 - McCulloch & Pitts, 1943
 - spike \approx pravdivostní hodnota výroku
 - neuron počítá se pravdivostní hodnotu jiného výroku ze vstupních výroků
- Lineární prahové neurony

Matematický neuron — $N : X_1 \times \cdots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony
 - McCulloch & Pitts, 1943
 - spike \approx pravdivostní hodnota výroku
 - neuron počítá se pravdivostní hodnotu jiného výroku ze vstupních výroků
- Lineární prahové neurony
- Sigmoidy (logistická funkce, příp. tanh)

Matematický neuron — $N : X_1 \times \cdots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony
 - McCulloch & Pitts, 1943
 - spike \approx pravdivostní hodnota výroku
 - neuron počítá se pravdivostní hodnotu jiného výroku ze vstupních výroků
- Lineární prahové neurony
- Sigmoidy (logistická funkce, příp. tanh)

$$N(X) = \frac{1}{1 + e^{-z}}$$

Matematický neuron — $N : X_1 \times \dots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony
 - McCulloch & Pitts, 1943
 - spike \approx pravdivostní hodnota výroku
 - neuron počítá se pravdivostní hodnotu jiného výroku ze vstupních výroků
- Lineární prahové neurony
- Sigmoidy (logistická funkce, příp. tanh)

$$N(X) = \frac{1}{1 + e^{-z}}, z = b + \sum_{i=0}^n x_i w_i$$

Matematický neuron — $N : X_1 \times \dots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony
 - McCulloch & Pitts, 1943
 - spike \approx pravdivostní hodnota výroku
 - neuron počítá se pravdivostní hodnotu jiného výroku ze vstupních výroků
- Lineární prahové neurony
- Sigmoidy (logistická funkce, příp. tanh)

$$N(X) = \frac{1}{1 + e^{-z}}, z = b + \sum_{i=0}^n x_i w_i$$

- mají hezké derivace

Matematický neuron — $N : X_1 \times \dots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony
 - McCulloch & Pitts, 1943
 - spike \approx pravdivostní hodnota výroku
 - neuron počítá se pravdivostní hodnotu jiného výroku ze vstupních výroků
- Lineární prahové neurony
- Sigmoidy (logistická funkce, příp. tanh)

$$N(X) = \frac{1}{1 + e^{-z}}, z = b + \sum_{i=0}^n x_i w_i$$

- mají hezké derivace
- proto se s nimi dobře pracuje

Matematický neuron — $N : X_1 \times \cdots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony
 - McCulloch & Pitts, 1943
 - spike \approx pravdivostní hodnota výroku
 - neuron počítá se pravdivostní hodnotu jiného výroku ze vstupních výroků
- Lineární prahové neurony
- Sigmoidy (logistická funkce, příp. tanh)

$$N(X) = \frac{1}{1 + e^{-z}}, z = b + \sum_{i=0}^n x_i w_i$$

- mají hezké derivace
 - proto se s nimi dobře pracuje
- Stochastické neurony

Matematický neuron — $N : X_1 \times \cdots \times X_n \rightarrow Y$

- Lineární neurony

$$N(X) = b + \sum_{i=0}^n x_i w_i$$

- Binární prahové neurony
 - McCulloch & Pitts, 1943
 - spike \approx pravdivostní hodnota výroku
 - neuron počítá se pravdivostní hodnotu jiného výroku ze vstupních výroků
- Lineární prahové neurony
- Sigmoidy (logistická funkce, příp. tanh)

$$N(X) = \frac{1}{1 + e^{-z}}, z = b + \sum_{i=0}^n x_i w_i$$

- mají hezké derivace
 - proto se s nimi dobře pracuje
- Stochastické neurony
 - výstup se interpretuje jako pravděpodobnost

Jednoduché rozpoznávání znaků

- dvouvrstvá síť

Jednoduché rozpoznávání znaků

- dvouvrstvá síť
- vstupní neurony = pixely

Jednoduché rozpoznávání znaků

- dvouvrstvá síť
- vstupní neurony = pixely
- výstupní neurony = jednotlivé znaky

Jednoduché rozpoznávání znaků

- dvouvrstvá síť
- vstupní neurony = pixely
- výstupní neurony = jednotlivé znaky
- pixel může hlasovat pokud je zabarvený

Jednoduché rozpoznávání znaků

- dvouvrstvá síť
- vstupní neurony = pixely
- výstupní neurony = jednotlivé znaky
- pixel může hlasovat pokud je zabarvený
- pixel může hlasovat pro víc znaků

Jednoduché rozpoznávání znaků

- dvouvrstvá síť
- vstupní neurony = pixely
- výstupní neurony = jednotlivé znaky
- pixel může hlasovat pokud je zabarvený
- pixel může hlasovat pro víc znaků
- znak s největším počtem hlasů vyhrává

Jednoduché rozpoznávání znaků

- dvouvrstvá síť
- vstupní neurony = pixely
- výstupní neurony = jednotlivé znaky
- pixel může hlasovat pokud je zabarvený
- pixel může hlasovat pro víc znaků
- znak s největším počtem hlasů vyhrává

Proces učení

Jednoduché rozpoznávání znaků

- dvouvrstvá síť
- vstupní neurony = pixely
- výstupní neurony = jednotlivé znaky
- pixel může hlasovat pokud je zabarvený
- pixel může hlasovat pro víc znaků
- znak s největším počtem hlasů vyhrává

Proces učení

V každém kroku:

Jednoduché rozpoznávání znaků

- dvouvrstvá síť
- vstupní neurony = pixely
- výstupní neurony = jednotlivé znaky
- pixel může hlasovat pokud je zabarvený
- pixel může hlasovat pro víc znaků
- znak s největším počtem hlasů vyhrává

Proces učení

V každém kroku:

- zvyš váhy z aktivních pixelů do správné třídy

Jednoduché rozpoznávání znaků

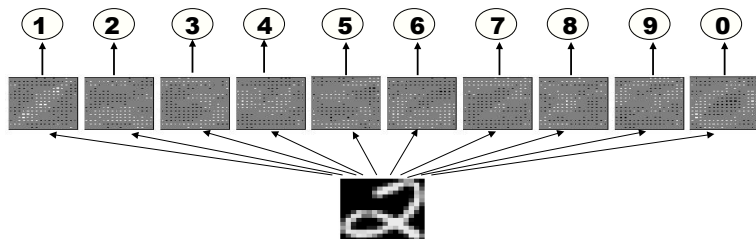
- dvouvrstvá síť
- vstupní neurony = pixely
- výstupní neurony = jednotlivé znaky
- pixel může hlasovat pokud je zabarvený
- pixel může hlasovat pro víc znaků
- znak s největším počtem hlasů vyhrává

Proces učení

V každém kroku:

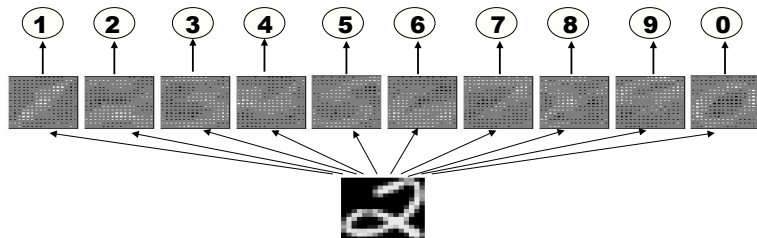
- zvyš váhy z aktivních pixelů do správné třídy
- následně sniž váhy z aktivních pixelů do aktuálně uhádnuté třídy

Rozpoznávání znaků — učení v obrázcích



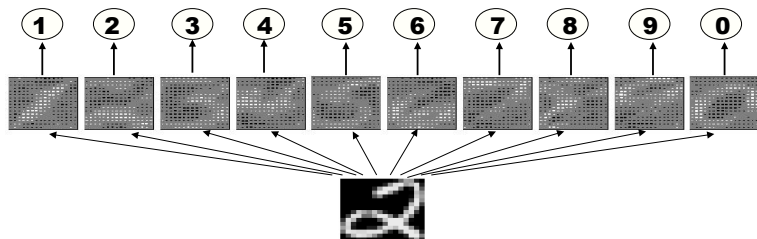
(zdroj: G. Hinton, Neural Networks for ML)

Rozpoznávání znaků — učení v obrázcích



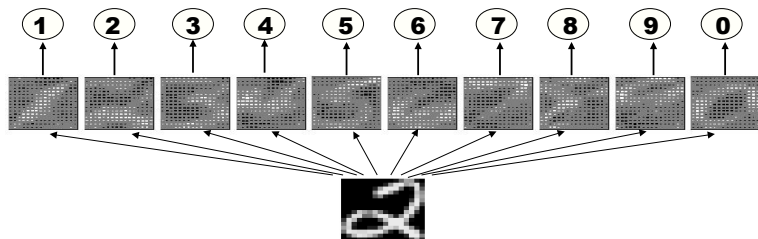
(zdroj: G. Hinton, Neural Networks for ML)

Rozpoznávání znaků — učení v obrázcích



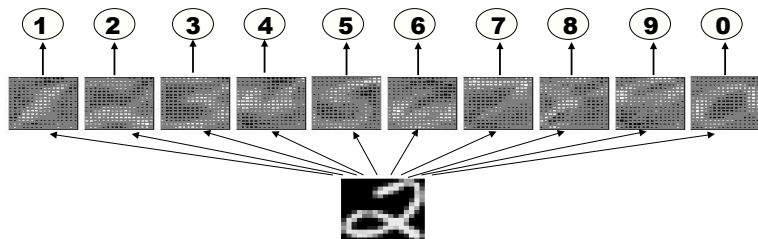
(zdroj: G. Hinton, Neural Networks for ML)

Rozpoznávání znaků — učení v obrázcích



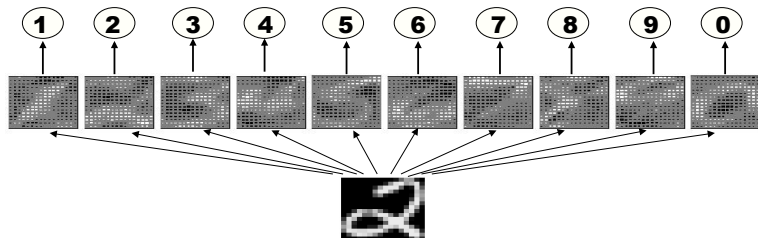
(zdroj: G. Hinton, Neural Networks for ML)

Rozpoznávání znaků — učení v obrázcích



(zdroj: G. Hinton, Neural Networks for ML)

Rozpoznávání znaků — učení v obrázcích



(zdroj: G. Hinton, Neural Networks for ML)

Typy učících úloh

- řízené učení (supervised learning)

Typy učících úloh

- řízené učení (supervised learning)
 - regrese

Typy učících úloh

- řízené učení (supervised learning)
 - regrese
 - klasifikace

Typy učících úloh

- řízené učení (supervised learning)
 - regrese
 - klasifikace
- reinforcement learning

Typy učících úloh

- řízené učení (supervised learning)
 - regrese
 - klasifikace
- reinforcement learning
- neřízené učení (unsupervised learning)

Typy neuronových sítí

- feed forward

Typy neuronových sítí

- feed forward (hluboké — >2 vnitřní vrstvy)

Typy neuronových sítí

- feed forward (hluboké — >2 vnitřní vrstvy)
- rekurentní neuronové sítě

Typy neuronových sítí

- feed forward (hluboké — >2 vnitřní vrstvy)
- rekurentní neuronové sítě
- symetrické (Hopfieldova síť)

I. Sutskever (2011) trénoval Neuronovou síť tak, aby uhodla následující znak v posloupnosti znaků.

I. Sutskever (2011) trénoval Neuronovou síť tak, aby uhodla následující znak v posloupnosti znaků. Trénovací data:

I. Sutskever (2011) trénoval Neuronovou síť tak, aby uhodla následující znak v posloupnosti znaků. Trénovací data: Wikipedie.

I. Sutskever (2011) trénoval Neuronovou síť tak, aby uhodla následující znak v posloupnosti znaků. Trénovací data: Wikipedie.

Výsledek (generovaný po jednom znaku)

I. Sutskever (2011) trénoval Neuronovou síť tak, aby uhodla následující znak v posloupnosti znaků. Trénovací data: Wikipedie.

Výsledek (generovaný po jednom znaku)

In 1974 Northern Denver had been overshadowed by CNL, and several Irish intelligence agencies in the Mediterranean region. However, on the Victoria, Kings Hebrew stated that Charles decided to escape during an alliance. The mansion house was completed in 1882, the second in its bridge are omitted, while closing is the proton reticulum composed below it aims, such that it is the blurring of appearing on any well-paid type of box printer.

(zdroj: G. Hinton, Neural Networks for ML)

Perceptron

Hopfieldova síť

<https://class.coursera.org/neuralnets-2012-001/>