



PLIN037 Sémantika a počítače

Zuzana Nevěřilová
2024

Strojové učení

- Počítačový program se nemění
- Výpočet se mění na základě dat
- Z minulých příkladů může program
 - Kategorizovat (klasifikovat)
 - Predikovat
 - Spojité hodnoty
 - Diskrétní hodnoty
- Existuje podobnost mezi minulými příklady a budoucími daty

statistika

Tradiční strojové učení

Neuronové sítě

Proč „učení“?

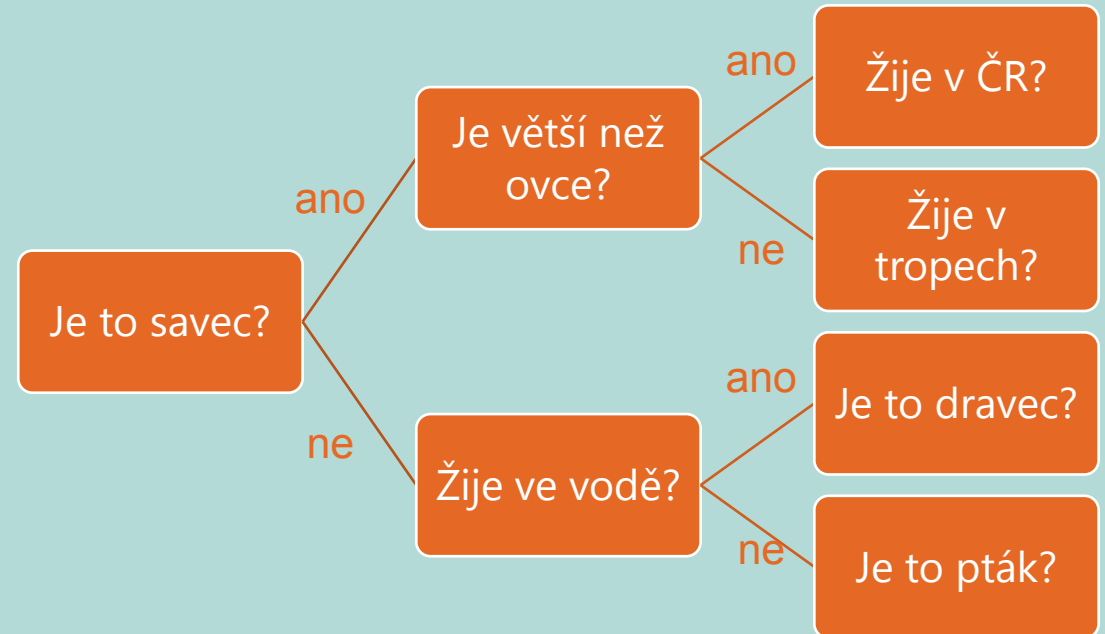
- Iterativní proces
- Zpětná vazba (účelová funkce, loss function)

Příklad: hra „Myslím si zvíře“

Znalosti (data)

Pořadí otázek

- Otázku klademe tak, aby odpověď rychle redukovala počet možných řešení (ideální by bylo půlení)



Příklad: hra „Myslím si zvíře“

	Savec	>ovce	V ČR	voda	dravec	pták
liška	ano	ne	ano	ne	ano	ne
kapr	ne	ne	ano	ano	ne	ne
zebra	ano	ano	ne	ne	ne	ne
?	ano	ano	ano	ne	ne	ne



- Data o všech zvířatech
- Nepřesné odpovědi
- Neúplné odpovědi
- Vývoj v čase (nová zvířata)

Příklad: hra „Myslím si zvíře“

	Savec	>ovce	V ČR	voda	dravec	pták
liška	ano	ne	ano	ne	ano	ne
kapr	ne	ne	ano	ano	ne	ne
zebra	ano	ano	ne	ne	ne	ne
?	ano	ano	ano	ne	ne	ne

Je to savec?



Strojové učení:

Existuje algoritmus, který se naučí klasifikovat neznámé zvíře na základě zjištěných rysů?

Algoritmy strojového učení

Rozhodovací stromy a rozhodovací seznamy

- If ryba & dravec then pstruh

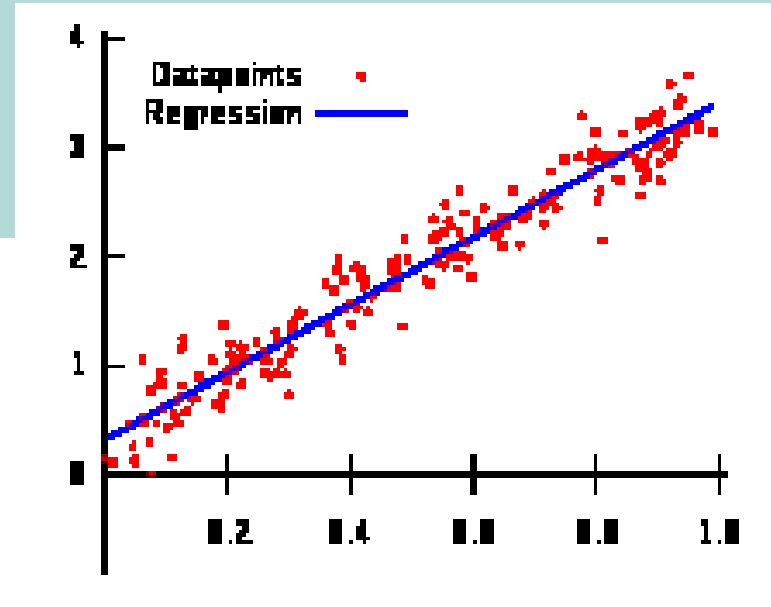
Lineární regrese

- Jaká (rovná) čára nejlépe rozdělí prvky do dvou tříd?

Naivní Bayesovský klasifikátor

- $P(pstruh|ryba\&dravec) = \frac{P(pstruh)P(ryba\&dravec|pstruh)}{P(ryba\&dravec)}$
- Jakou třídu nejpravděpodobněji pozorujeme?

Mnoho dalších...



model

Apriorní pravděpodobnost

Rozhodovací strom

Každý uzel obsahuje

- Prvky uzlu
- Jeden příznak
- Dělicí hodnotu

Giniho nečistota (Gini Impurity)

- Jaká je pravděpodobnost, že náhodně zvolený prvek uzlu bude zvolený špatně?

	Savec	> ovce	V ČR	voda	dravec	pták
liška	ano	ne	ano	ne	ano	ne
kapr	ne	ne	ano	ano	ne	ne
zebra	ano	ano	ne	ne	ne	ne

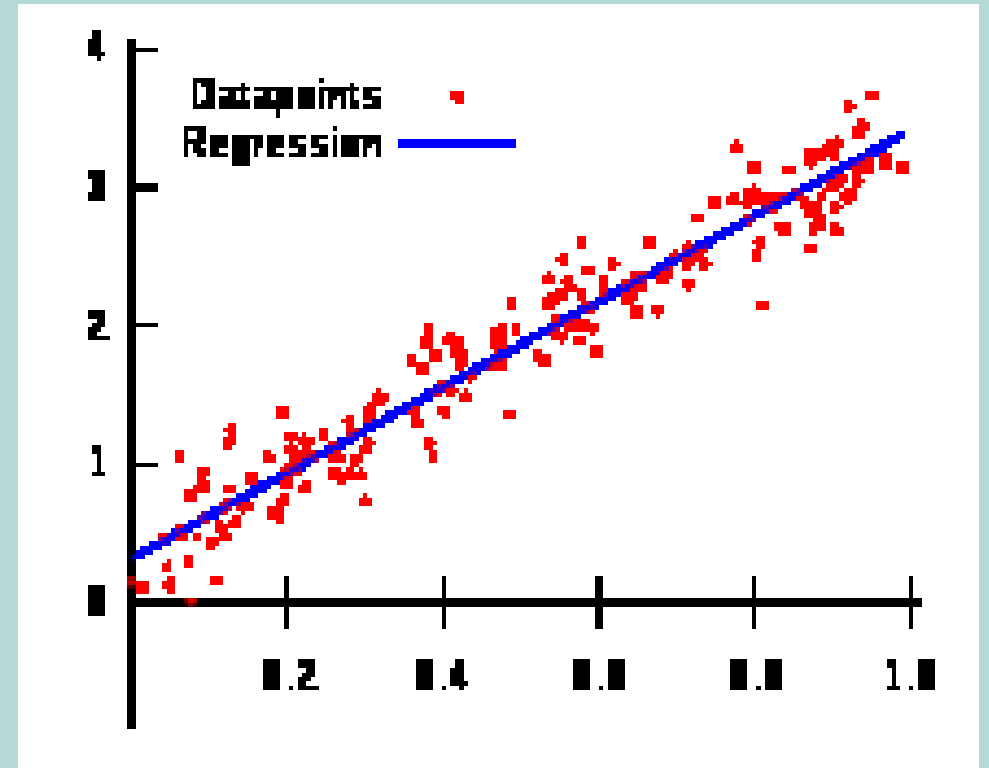


Lineární regrese

$$y = ax + b$$

Vzdálenost od optimální hodnoty:

- Součet nejmenších čtverců odchylek od přímky
- Naivní přístup – zkusit každé a i b
- 2D příklad: cena piva vs. objem
- Vícedimenzionální příklad: cena domu vs. velikost pozemku, vzdálenost od MHD, počet místností, stáří budovy, velikost obce, znečištění, ...



Naivní Bayesovský klasifikátor

Zvíře	Velikost	Barva	Potrava
Slon	Velký	Šedý	Býložravec
Slon	Střední	Šedý	Býložravec
Kráva	Velký	Strakatý	Býložravec
Kráva	Střední	Strakatý	Býložravec
Kráva	Velký	Strakatý	Býložravec
Vlk	Malý	Černý	Dravec
Kráva	Velký	Černý	Býložravec
Vlk	Velký	Šedý	Dravec

$$\text{posterior} = \frac{\text{prior} \cdot \text{likelihood}}{\text{evidence}}$$

$$P(C|F_1, F_2, \dots, F_n) = \frac{P(C) \cdot P(F_1, F_2, \dots, F_n|C)}{P(F_1, F_2, \dots, F_n)}$$

Minulá data

Nové pozorování:
Velký černý býložravec

Nepodmíněná pravděpodobnost:

- 50 % kráva
- 25 % vlk
- 25 % slon

Naivní Bayesovský klasifikátor

Nepodmíněná pravděpodobnost:

- 50 % kráva
- 25 % vlk
- 25 % slon

Podmíněné pravděpodobnosti:

$$P(\text{černý}|\text{slon}) = 0$$

$$P(\text{černý}|\text{vlk}) = 0.5$$

$$P(\text{černý}|\text{kráva}) = 0.25$$

$$P(\text{velký}|\text{slon}) = 0.5$$

$$P(\text{velký}|\text{vlk}) = 0.5$$

$$P(\text{velký}|\text{kráva}) = 0.75$$

$$P(\text{býložravec}|\text{slon}) = 1$$

$$P(\text{býložravec}|\text{vlk}) = 0$$

$$P(\text{býložravec}|\text{kráva}) = 1$$

Zvíře	Velikost	Barva	Potrava
Slon	Velký	Šedý	Býložra
Slon	Střední	Šedý	Býložra
Kráva	Velký	Strakatý	Býložr
Kráva	Střední	Strakatý	Býložr
Kráva	Velký	Strakatý	Býložr
Vlk	Malý	Černý	Dravec
Kráva	Velký	Černý	Býložr
Vlk	Velký	Šedý	Dravec

Naivní Bayesovský klasifikátor

Zvíře	Velikost	Barva	Potrava
Slon	Velký	Šedý	Býložravec
Slon	Střední	Šedý	Býložravec
Kráva	Velký	Strakatý	Býložravec
Kráva	Střední	Strakatý	Býložravec
Kráva	Velký	Strakatý	Býložravec
Vlk	Malý	Černý	Dravec
Kráva	Velký	Černý	Býložravec
Vlk	Velký	Šedý	Dravec

$$P(\text{velký}|\text{slon}) = 0.5, P(\text{velký}|\text{vlk}) = 0.5,$$
$$P(\text{velký}|\text{kráva}) = 0.75,$$
$$P(\text{býložravec}|\text{slon}) = 1, P(\text{býložravec}|\text{vlk}) = 0,$$
$$P(\text{býložravec}|\text{kráva}) = 1$$

$$P(\text{slon}|\text{velký, černý, býložravec})$$
$$= P(\text{slon}) P(\text{velký}|\text{slon}) P(\text{černý}|\text{slon})$$
$$P(\text{býložravec}|\text{slon}) = 0.25 \cdot 0.25 \cdot 0 \cdot 1 = 0$$
$$P(\text{vlk}|\text{velký, černý, býložravec})$$
$$= P(\text{vlk}) P(\text{velký}|\text{vlk}) P(\text{černý}|\text{vlk})$$
$$P(\text{býložravec}|\text{vlk}) = 0.25 \cdot 0.5 \cdot 0.5 \cdot 0 = 0$$

$$P(\text{býložravec}|\text{kráva}) = 0.5 \cdot 0.75 \cdot 0.25 \cdot 1$$
$$= 0.09375$$

(tradiční) strojové učení

Uživatelské strojové učení
Klasifikace a regrese

Pozorování dat
Určení rysů
Volba algoritmu



ron

Tradiční strojové učení

Pozorování dat
Architektura neuronové sítě

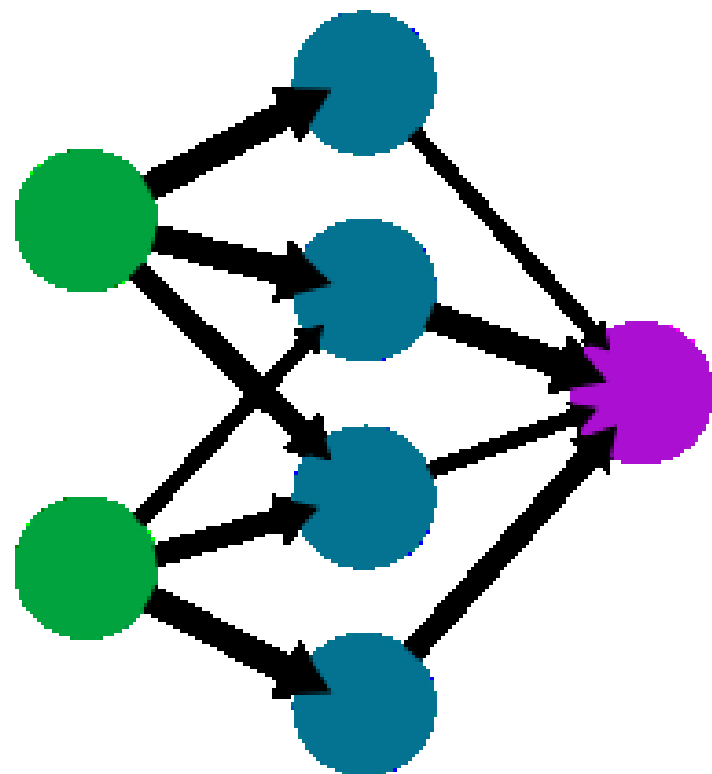
Neuronové sítě

Výpočetní model:

- Lineární složka $y = ax + b$
- Nelineární složka (nelineární funkce)
- Dopředný výpočet (feed forward)
- Zpětná propagace (back propagation)
- Výpočet hodnot lineární složky v každém neuronu
- Obrovské množství výpočtů = náročné na paměť počítače

A simple neural network

input layer hidden layer output layer



Neuronové sítě

Terminologie

$$y = ax + b$$

- a , b – parametry modelu
- V každém neuronu, takže spíš a_i a b_j

Layer = vrstva

- Vstupní vrstva, skrytá vrstva, výstupní vrstva

Hloubka sítě (depth) = počet skrytých vrstev L + 1 (výstupní)

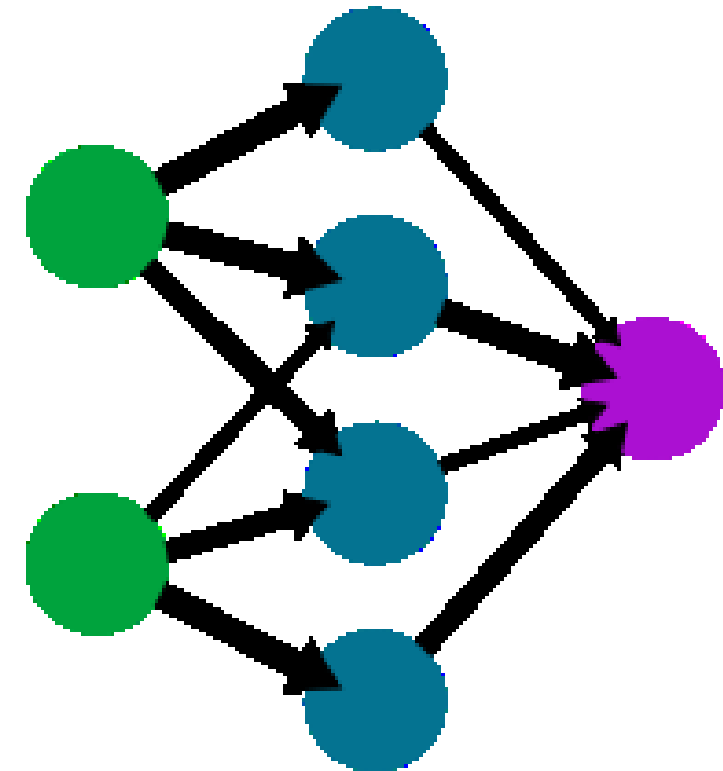
- V každém neuronu v každé vrstvě, takže spíš a_i^k a b_j^k

A simple neural network

input
layer

hidden
layer

output
layer



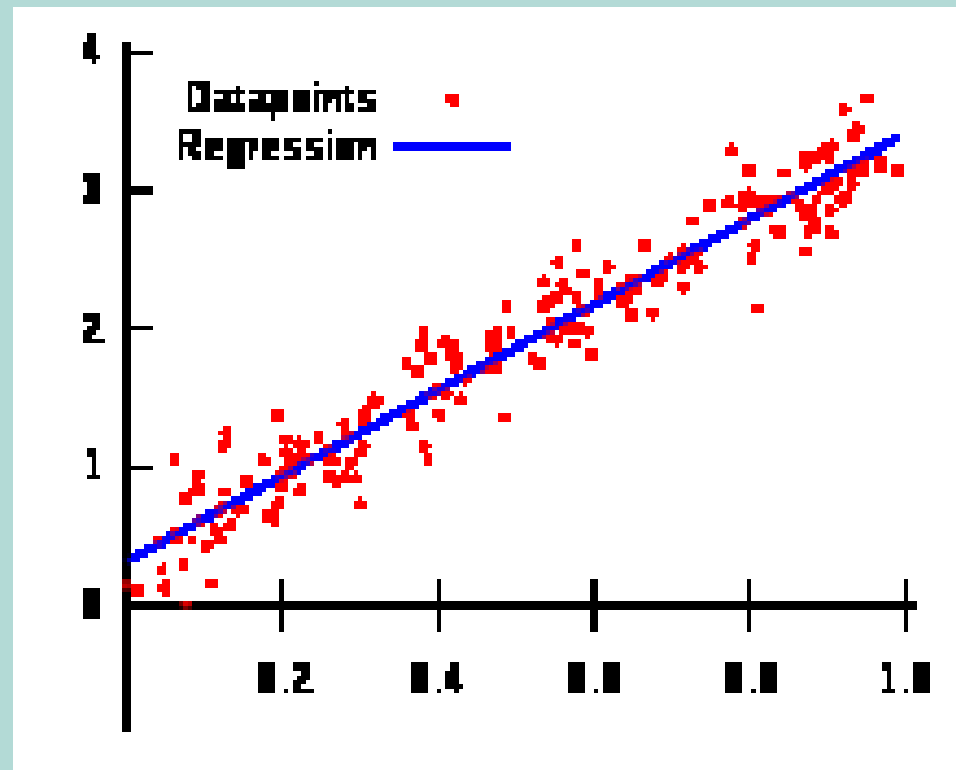
Neuronové sítě

Výpočet

$$y = ax + b$$

Nějak tipneme a a b

- **Dopředný výpočet (feed forward)**
 - Spočítáme $y = ax + b$
 - Spočítáme aktivační funkci (nelineární)
 $\hat{y} = \sigma(ax + b)$
- **Zpětná propagace (back propagation)**
 - Spočítáme, o kolik jsme se spletli (loss function, loss, účelová funkce)
 - Navrhujeme, jak parametry a a b upravit



$$y = ax + b$$

místo a používáme w – weight (váhy)

b – bias (zkreslení)

Výpočet $y = wx + b$ běží souběžně v neuronech jedné vrstvy, díky propojení se váhy neuronů ovlivňují.

Neuronové sítě

Word Embeddings

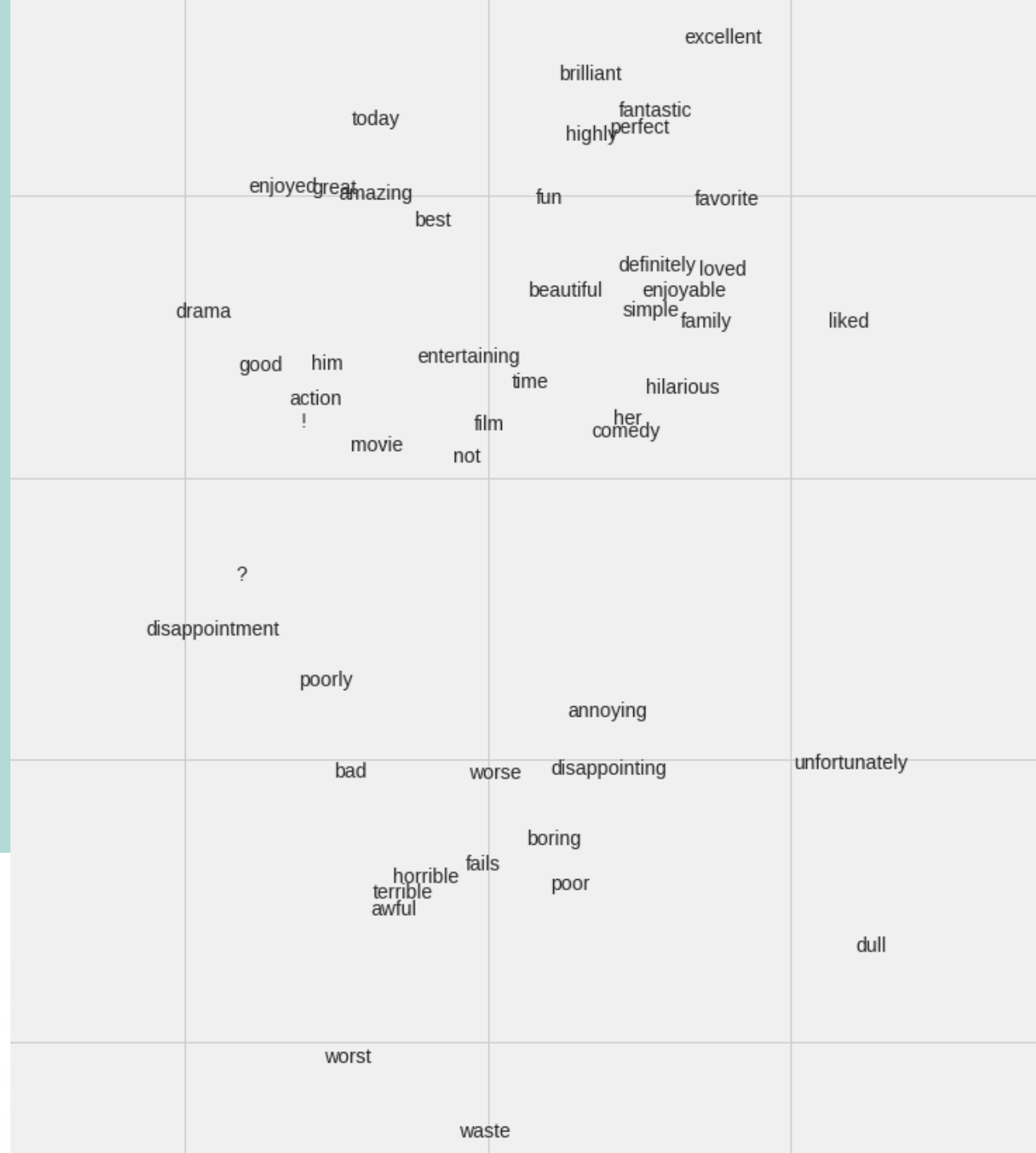
Modelování diskretních objektů (slova)
pomocí spojitých hodnot (vektory reálných
čísel)

Vektory kódující slova tak, aby:

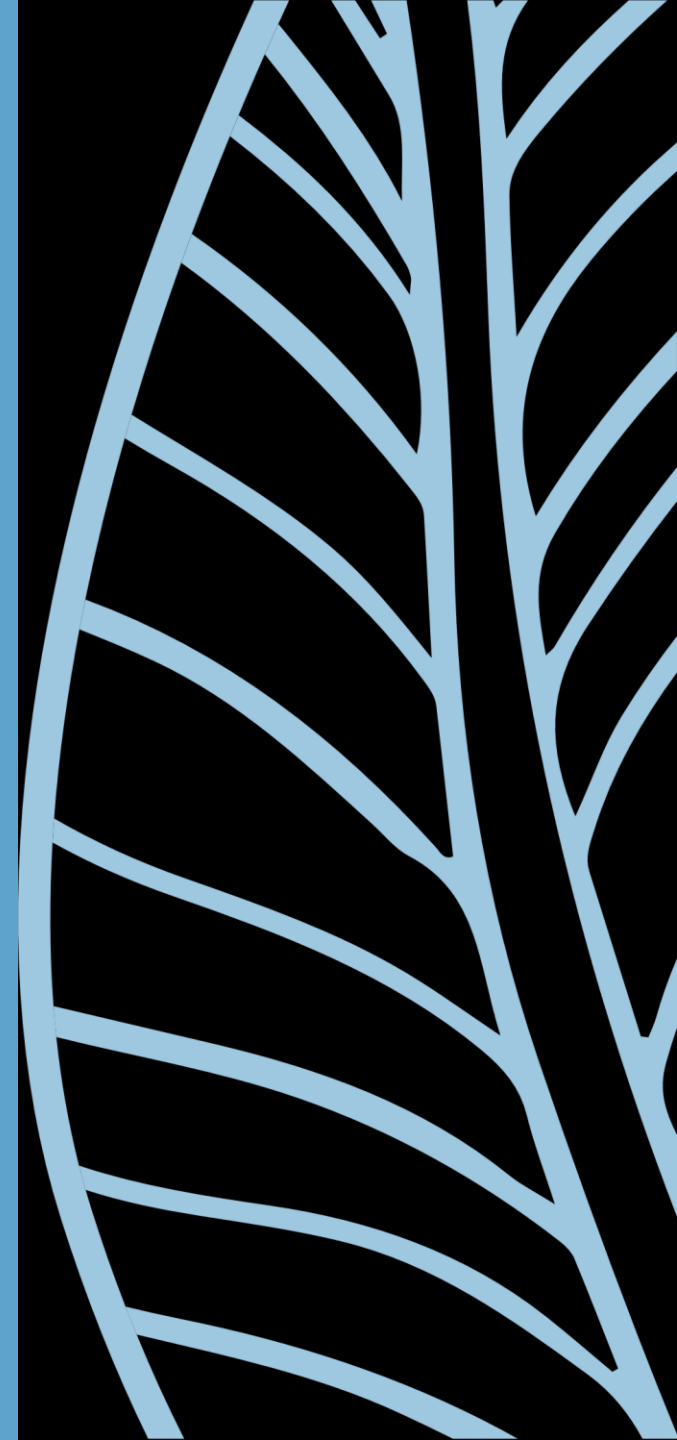
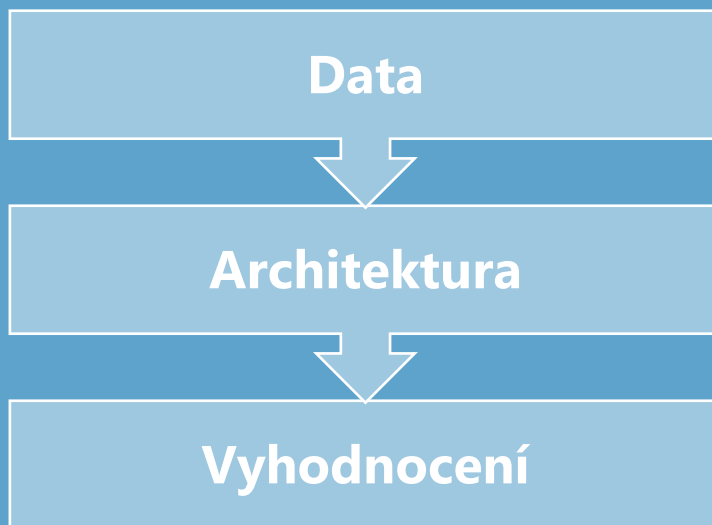
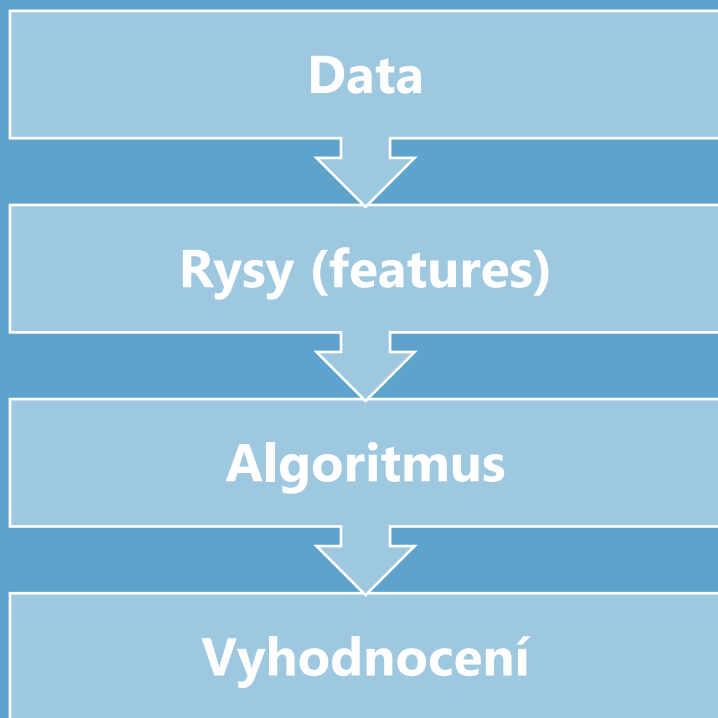
- Počet dimenzí nebyl moc vysoký
- Vektory významově blízkých slov svíraly malý úhel

Výpočet:

- Predikce chybějícího slova ve větě



Strojové učení a neuronové sítě





Literatura

- Will Koehrsen: **Neural Network Embeddings Explained. How deep learning can represent War and Peace as a vector.** Towards Data Science. Oct 2, 2018:8. <https://towardsdatascience.com/neural-network-embeddings-explained-4d028e6f0526>
- Wikipedia contributors. (2021, April 13). **Simple linear regression.** In Wikipedia, The Free Encyclopedia. Retrieved 00:39, May 4, 2021, from https://en.wikipedia.org/w/index.php?title=Simple_linear_regression&oldid=1017542751
- Wikipedia contributors. (2021, May 3). **Naive Bayes classifier.** In Wikipedia, The Free Encyclopedia. Retrieved 00:40, May 4, 2021, from https://en.wikipedia.org/w/index.php?title=Naive_Bayes_classifier&oldid=1021201855
- Wikipedia contributors. (2021, April 23). **Neural network.** In Wikipedia, The Free Encyclopedia. Retrieved 00:44, May 4, 2021, from https://en.wikipedia.org/w/index.php?title=Neural_network&oldid=1019519177
- Avinash Navlani. **Decision Tree Classification in Python.** Datacamp Tutorial. December 28th, 2018. <https://www.datacamp.com/community/tutorials/decision-tree-classification-python>
- Ria Kulshrestha: **Yes, you should listen to Andrej Karpathy, and understand Backpropagation.** Towards Data Science. 2020. <https://towardsdatascience.com/back-propagation-721bfcc94e34>