



Vilnius universitetas
Matematikos ir informatikos fakultetas
Duomenų mokslo ir skaitmeninių
technologijų institutas

Mašininis mokymasis

prof. dr. Olga Kurasova
Olga.Kurasova@mif.vu.lt

Šių dienų raktažodžiai

- **Didieji duomenys** (angl. *big data*)
- **Dirbtinis intelektas** (angl. *artificial intelligence*)
- **Mašininis mokymasis** (angl. *machine learning*)
- **Gilusis mokymasis** (angl. *deep learning*)
- **Gilieji neuroniniai tinklai** (angl. *deep neural networks*)

Mašininis mokymasis

- **Mašininis mokymasis** (angl. *machine learning*) – tai informatikos sritis, suteikianti kompiuteriams galimybę mokytis be aiškių instrukcijų.
- **Machine learning** is a field of computer science that gives computers the ability to learn without being explicitly programmed (*Wikipedia*).
- **Machine learning** is the use of artificial intelligence to automate algorithms that can learn without explicit instructions (*Intro to Machine Learning Course / Udacity*)

AI vs ML vs DL

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



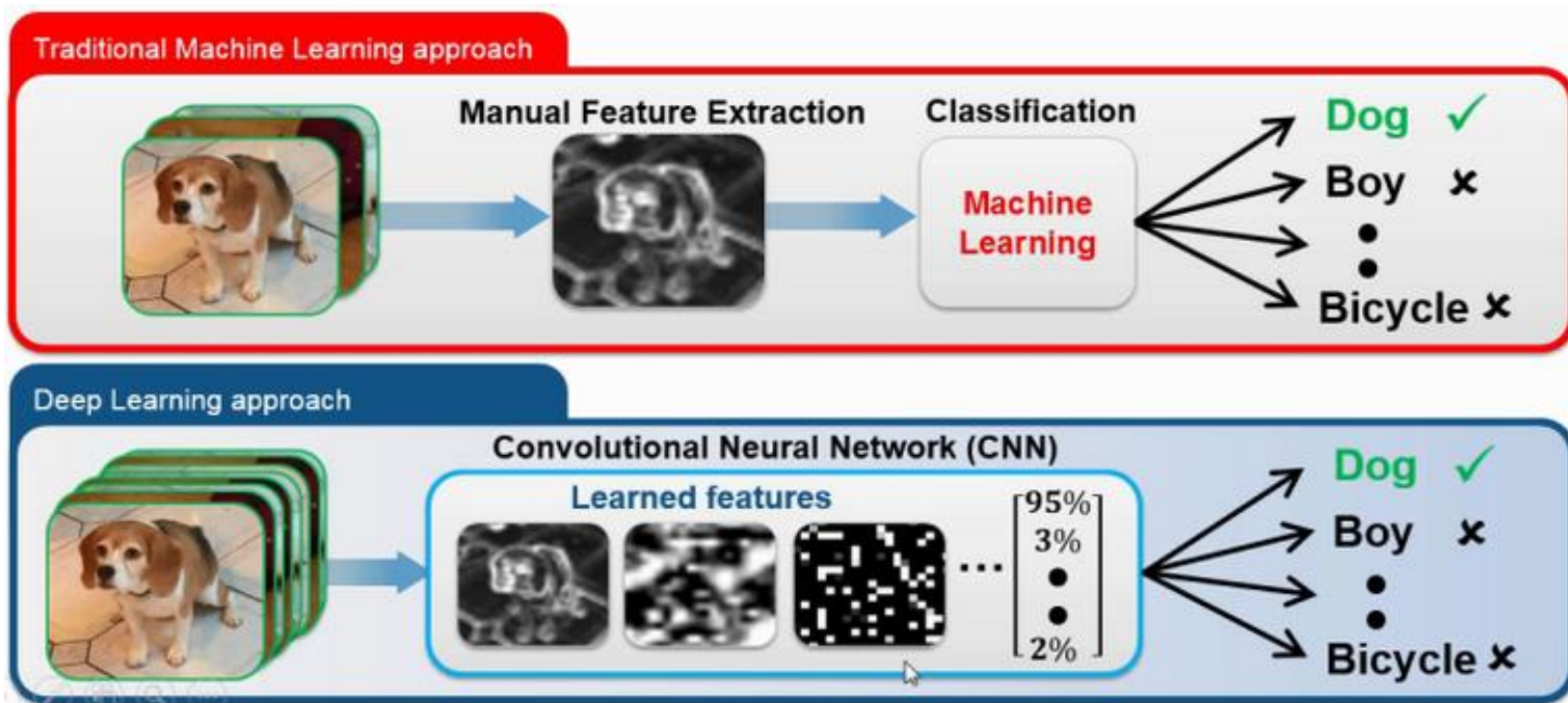
1950's 1960's 1970's 1980's 1990's 2000's 2010's

Skirtumai

- **Dirbtinis intelektas** – tai sistema, kuri gali mąstyti, jausti, veikti ir prisitaikyti.
- **Mašininis mokymasis** – tai algoritmai, kurių veikimas gerėja, kai jie gauna daugiau duomenų.
- **Gilusis mokymasis** – tai mašininio mokymosi poaibis, kuris efektyvus mokant naudojant didžiulius duomenų kiekius.



Mašininis vs gilusis mokymasis



P. S. Vėliau bus pateikta daugiau informacijos apie gilųjį mokymąsi.

https://leonardoaraujosantos.gitbooks.io/artificial-intelligence/content/deep_learning.html

Mašininis mokymasis

Sinonimai lietuvių kalba:

- Automatinis mokymasis
- Kompiuterių mokymasis
- Sistemų mokymasis

Mašininis mokymasis

- Algoritmas/programa **išmoksta** iteratyviai pateikiant **daug duomenų** be iš anksto suprogramuotų aiškių instrukcijų.



Brain Differences

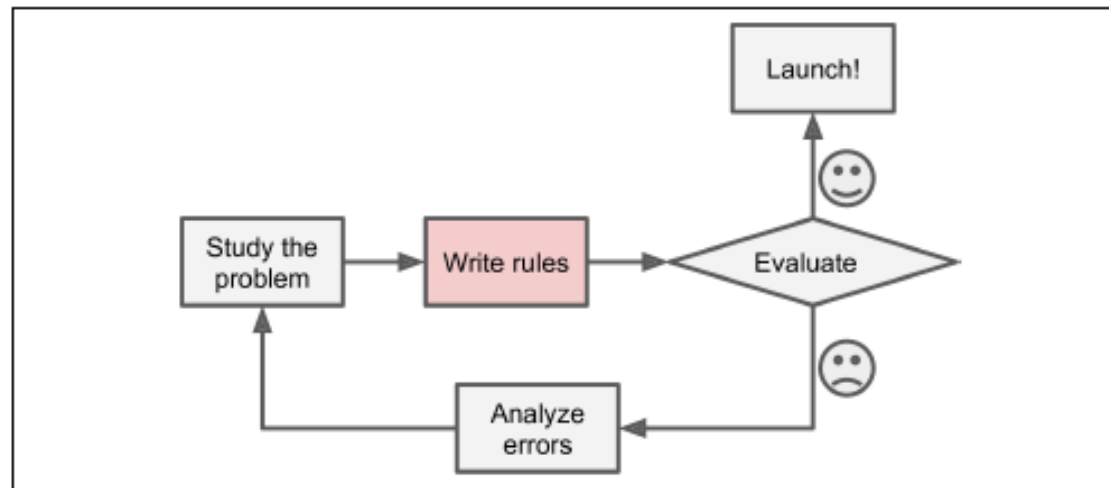
Now wait before you start thinking that you can just create a huge neural network and call strong AI, there are some few points to remember:

Just a list:

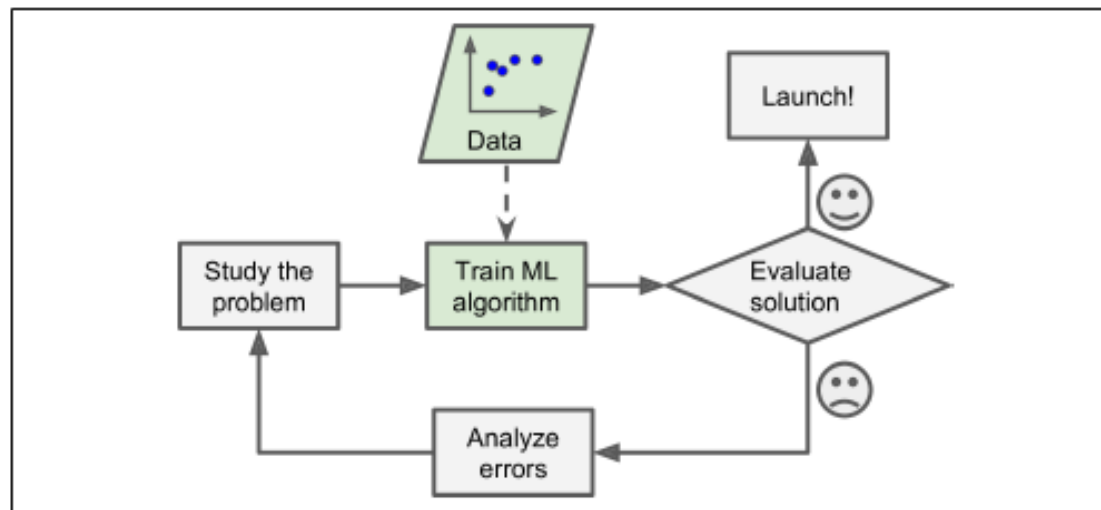
- The artificial neuron fires totally different than the brain
- A human brain has 100 billion neurons and 100 trillion connections (synapses) and operates on 20 watts(enough to run a dim light bulb) - in comparison the biggest neural network have 10 million neurons and 1 billion connections on 16,000 CPUs (about 3 million watts)
- The brain is limited to 5 types of input data from the 5 senses.
- Children do not learn what a cow is by reviewing 100,000 pictures labelled "cow" and "not cow", but this is how machine learning works.
- Probably we don't learn by calculating the partial derivative of each neuron related to our initial concept. (By the way we don't know how we learn)

https://leonardoaraujosantos.gitbooks.io/artificial-intelligence/content/neural_networks.html

Tradīcinis būdas



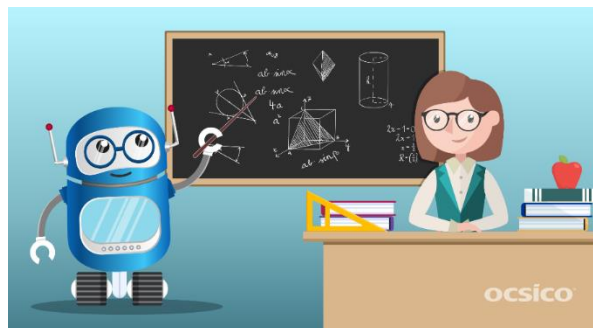
Mašīniniu mokymusi grīstas būdas



Géron, A. (2017). *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, Inc.

Mašininio mokymosi tipai

- **Su mokytoju / Prižiūrimas** mokymasis (*supervised*),
- **Be mokytojo / Neprižiūrimas** mokymasis (*unsupervised*),
- **Dalinei prižiūrimas** mokymasis (*semi-supervised*),
- **Sustiprintas** mokymasis (*reinforcement*).

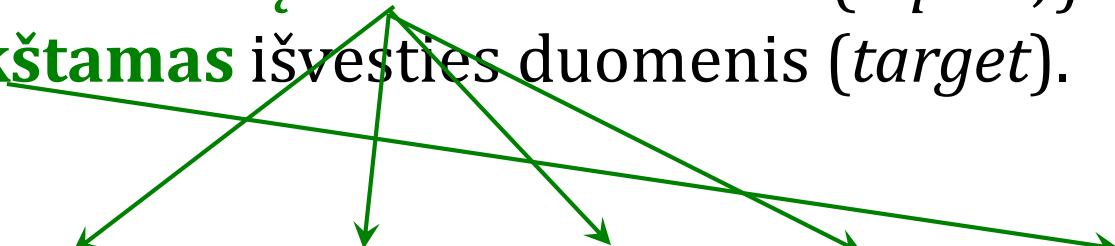


Prižiūrimas mokymasis

- Būtina žinoti ir **įvesties** duomenis (*inputs, features*) ir **trokštamą** išvesties duomenis (*target*).

Prižiūrimas mokymasis

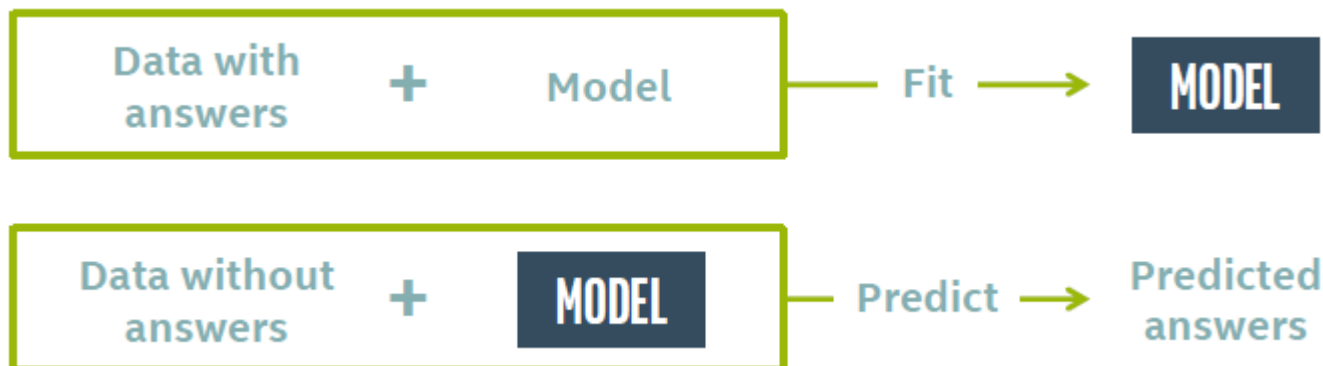
- Būtina žinoti ir **įvesties** duomenis (*inputs, features*) ir **trokštamą** išvesties duomenį (*target*).



	x_1	x_2	x_3	x_4	Klasė
X_1	85	0,001	24,1	1025	sveikas
X_2	77	0,002	21,3	2036	sveikas
X_3	68	0,015	35,8	1059	sveikas
...
X_{101}	101	0,001	22,4	3011	serga
X_{102}	95	0,001	28,0	2645	serga
...
X_{201}	86	0,002	30,1	2987	???
X_{202}	72	0,010	19,5	1259	???

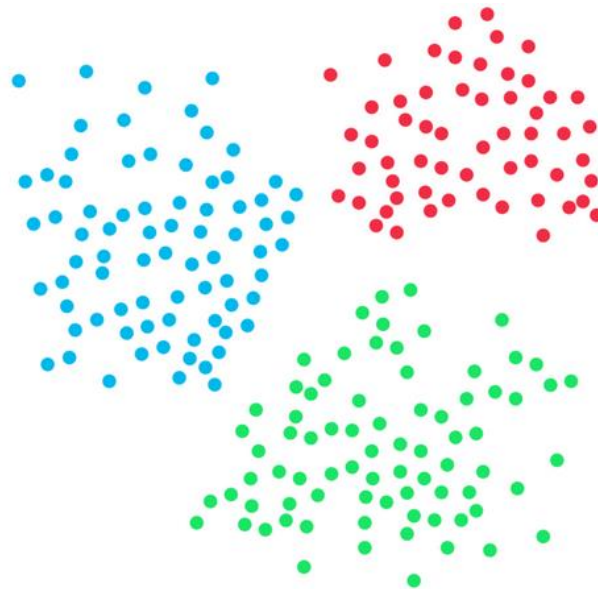
Prižiūrimas mokymasis

- Mokymosi eigoje **ieškomi ryšiai** tarp įvesties duomenų ir trokštamų išvesties duomenų.
- Tam yra taikomi **įvairūs algoritmai**.
- Nustačius tokius ryšius, algoritmas gebės **automatiškai nustatyti išvesties reikšmes** naujiems duomenimis, kurie nebuvo naudojami mokyme ir kurių išvestis nėra žinoma.



Neprižiūrimas mokymasis

- Algoritmai **ieško ryšių** tarp įvesties duomenų nežinodamas jokių trokštamų išvesčių.



Dalinai prižiūrimas mokymasis

- **Dalinai prižiūrimo** (*semi-supervised*) mokymosi algoritmai naudoja dalį duomenų, kurių klasės (ar kitos žymės) žinomos ir dalį duomenų su nežinomomis klasėmis.
- Jis svarbus, kai **mokymui reikia daug duomenų**, o anototų (sužymėtų) yra nedaug.

Charakteringų uždavinių klasės

- **Prižiūrimo mokymosi** algoritmais sprendžiami uždaviniai:
 - klasifikavimas,
 - prognozavimas,
- **Neprižiūrimo mokymosi** algoritmais sprendžiami uždaviniai:
 - klasterizavimas.

Mašininio mokymosi taikymai

- Automated theorem proving^{[39][40]}
- Adaptive websites^[citation needed]
- Affective computing
- Bioinformatics
- Brain–machine interfaces
- Cheminformatics
- Classifying DNA sequences
- Computational anatomy
- Computer vision, including object recognition
- Detecting credit-card fraud
- General game playing^[41]
- Information retrieval
- Internet fraud detection^[28]
- Linguistics
- Marketing
- Machine learning control
- Machine perception
- Medical diagnosis
- Economics
- Insurance
- Natural language processing
- Natural language understanding^[42]
- Optimization and metaheuristic
- Online advertising
- Recommender systems
- Robot locomotion
- Search engines
- Sentiment analysis (or opinion mining)
- Sequence mining
- Software engineering
- Speech and handwriting recognition
- Financial market analysis
- Structural health monitoring
- Syntactic pattern recognition
- Time series forecasting
- User behavior analytics
- Translation^[43]

Mašininio mokymosi iššūkiai

- **Mažas/nepakankamas duomenų kiekis**: reikia turėtų tūkstančius ar dar daugiau duomenų.
- Mokymo duomenys **nėra reprezentatyvūs**.
- **Blogos kokybės** duomenys: klaidos, triukšmas, išskirtys, trūkstami įrašai ir kt.
- **Netinkami duomenų požymiai**: būtinas tinkamas požymių atrinkimas iš esamų (*feature selection*), naujų sukūrimas iš esamų (*feature extraction*).
- **Modelio permokymas** (*overfitting*): gerai veikia mokymo duomenims, bet blogai su naujais duomenimis.

Gilusis mokymasis

- **Gilusis mokymasis** (angl. *deep learning*) – tai plačios mašininio mokymosi metodų šeimos dalis pagrįsta duomenų pateikimo (reprezentatyvumo) mokymusi, priešingai nei yra algoritmuose konkretiems uždaviniams spręsti.
- **Deep learning** (also known as deep structured learning or hierarchical learning) is part of a broader family of machine learning methods based on **learning data representations**, as opposed to task-specific algorithms.

Dimensiškumo prakeiksmas

- Mašininio mokymosi uždaviniuose labai dažna vadinamojo **dimensiškumo prakeiksmo** (angl. *curse of dimensionality*) problema.
- Jei nagrinėjamus duomenis apibūdina daug požymių, tokių duomenų **dimensija bus didelė**.
- Tuomet objektų (duomenų įrašų) turi **būti labai daug**.

Dimensiškumo prakeikimas

- Tarkime turime vienmatį atvejį: vienetinio ilgio atkarpoje atidėtus **10 taškus**.
- Norint, kad toks pats tankis išliktų dvimatėje erdvėje, reikia **$10^2=100$ taškų**, trimatėje – **$10^3=1000$ taškų**, n-matėje – **10^n taškų**. Priešingu atveju, duomenis yra **labai reti** (angl. sparse).
- Pavyzdžiui, jei duomenis apibūdina 20 požymių, norint kad algoritmas gerai apsimokytų, reikia jį mokyti **10^{20} įrašais**. Tiek įrašų turėti arba visai **neįmanoma**, arba mokymosi procesas truks **labai ilgą laiką**.
- Tikslinga pasitelkti **dimensijos mažinimo metodus**.

Gilusi mokymasis

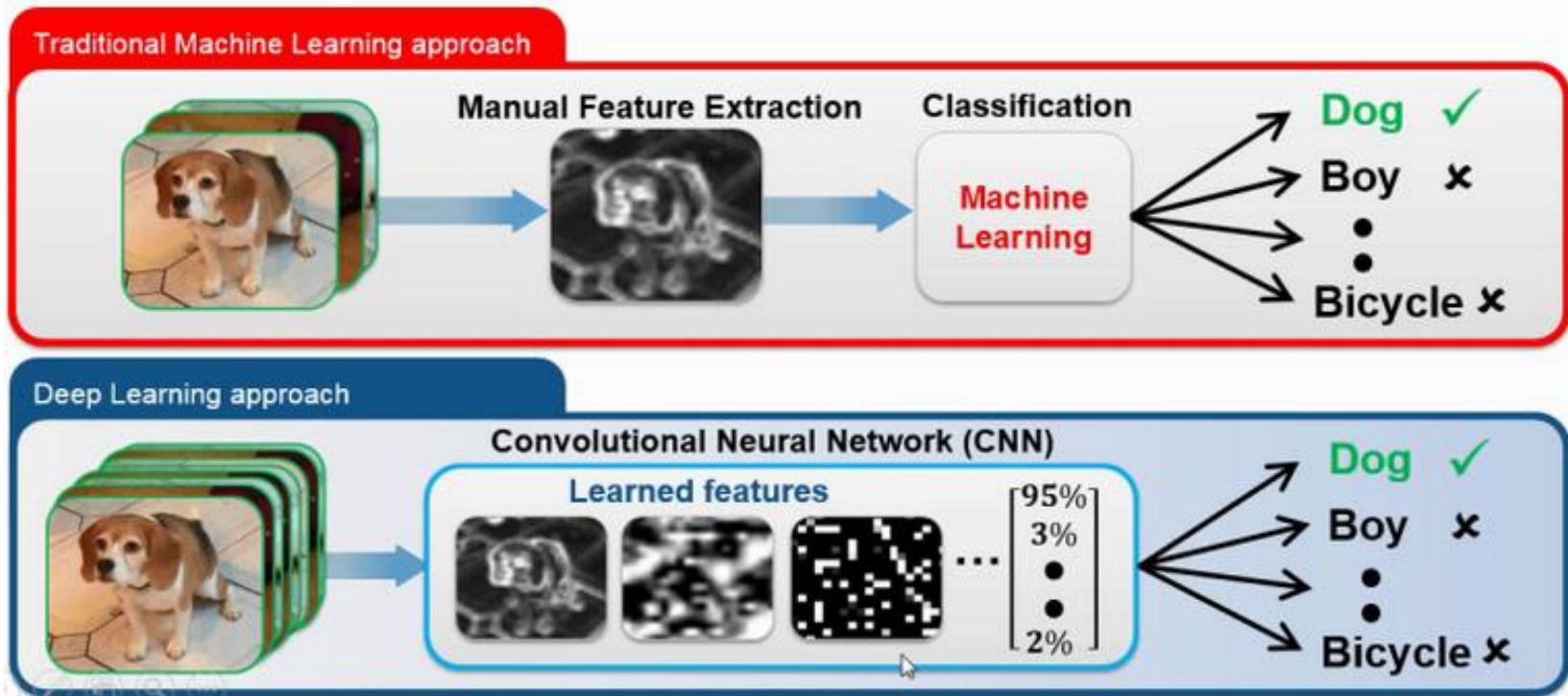
Deep learning is a class of machine learning algorithms that:

- use a **cascade of multiple layers** of nonlinear processing units for feature extraction and transformation. Each successive layer uses the output from the previous layer as input.
- learn in **supervised** (e.g., classification) and/or **unsupervised** (e.g., pattern analysis) manners.
- learn **multiple levels of representations** that correspond to different levels of abstraction; the levels form a hierarchy of concepts.
- use some form of **gradient descent** for training via backpropagation.

Duomenų reprezentavimas (pateikimas)

- **Duomenų reprezentavimas (pateikimas)** yra labai svarbus aspektas.
- Būtina atsakyti į klausimą, **kokie požymiai geriausiai reprezentuoja** analizuojamus objektus.
- Tai ypač aktualu tokiose srityse, kaip **šnekos, vaizdo** ir **garso** atpažinimas ir pan., kur būtina išgauti tinkamiausius požymius iš turimo signalo ar vaizdo.

Mašininis vs gilusis mokymasis



https://leonardoaraujosantos.gitbooks.io/artificial-intelligence/content/deep_learning.html

Duomenų reprezentavimo (pateikimo) pavyzdys

- Pavyzdžiui, įprastai, žmogui atlikti skaičių, pateiktų **arabiškais skaitmenimis**, aritmetinius veiksmus yra nepalyginamai **paprasciau** nei tą patį atlikti su skaičiais, pateiktais **romėniškais skaitmenimis**.

$$\text{III} + \text{VII} =$$

$$\text{II} + \text{V} =$$

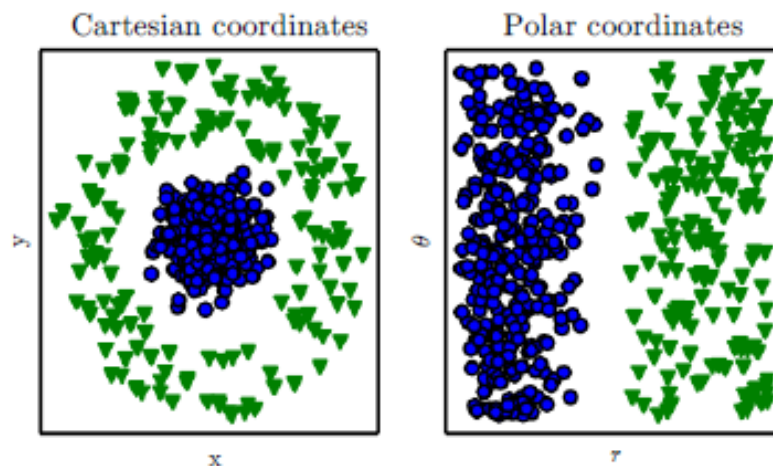
$$23 + 32 =$$

$$42 + 21 =$$

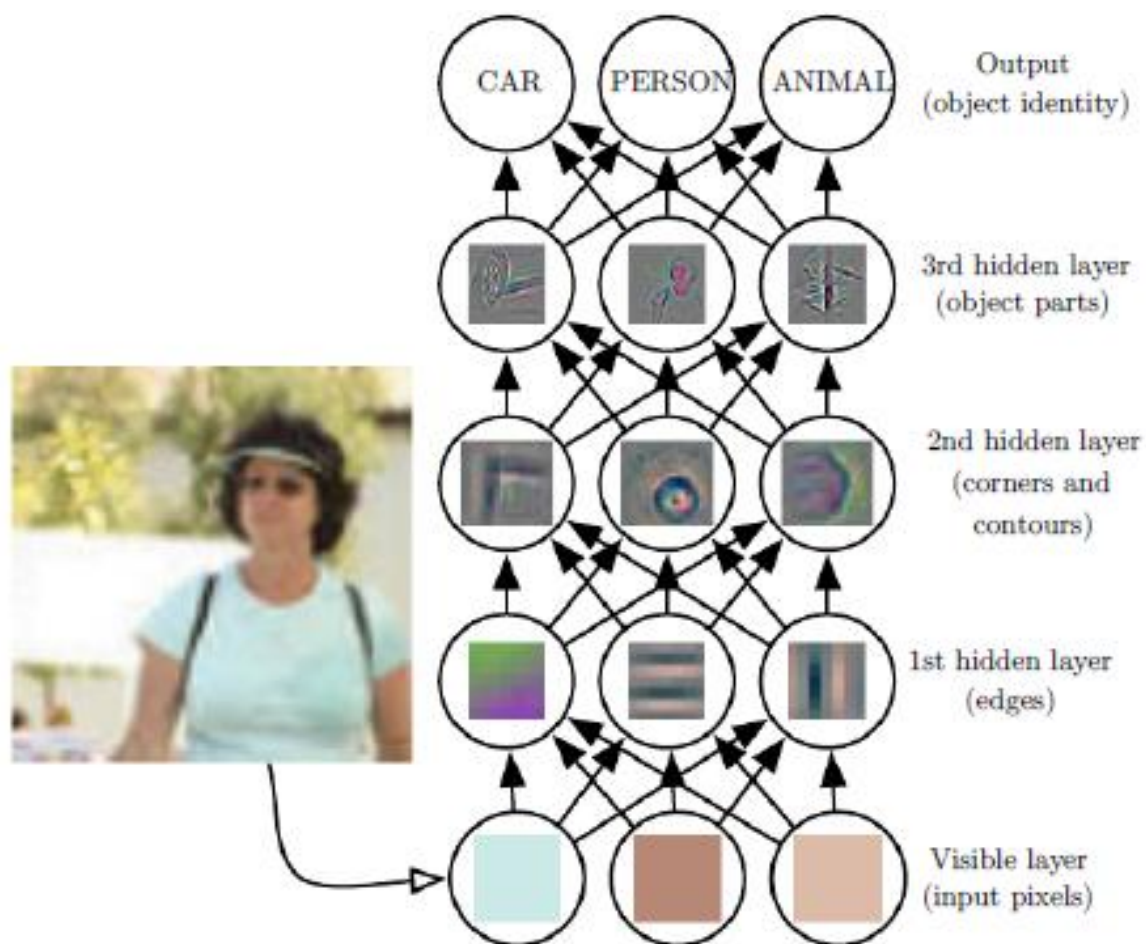
$$55 + 12 =$$

Duomenų reprezentavimo (pateikimo) pavyzdys

- Pavyzdžiui, pateikus tuos pačius duomenis Dekarto ar polinėje koordinačių sistemoje, iš esmės pasikeičia **klasių tiesinio atskyrimo** problema.



Gilusis mokymasis



Duomenų reprezentavimas giliajame mokymesi

- Duomenų **reprezentavimas** giliajame mokymesi yra ypač svarbus, kadangi čia siekiama, kad **automatiškai** būtų parinkta **geriausia** duomenų reprezentacija.

Gilusi mokymasis

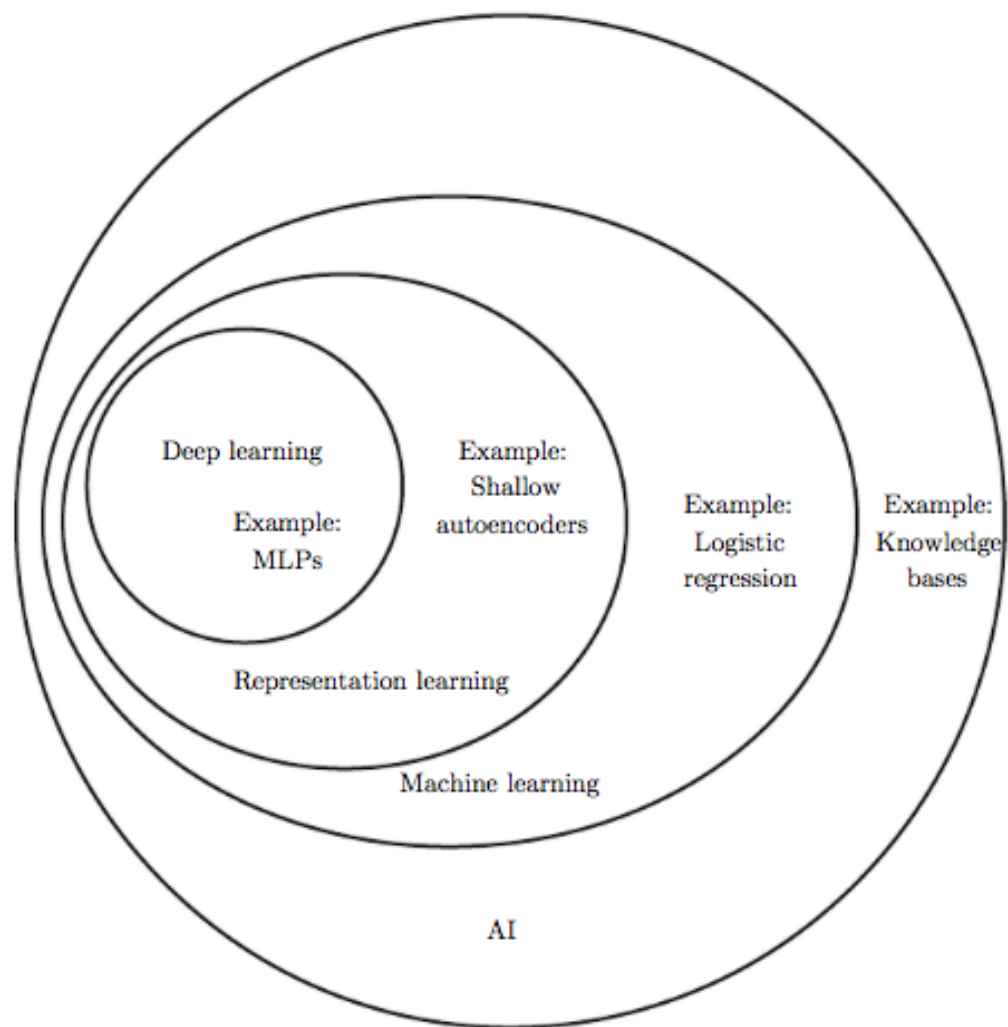
- **Deep learning** is a particular kind of machine learning that achieves **great power** and flexibility by representing the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts, and **more abstract representations** computed in terms of less abstract ones (Goodfellow et al., 2016)

Gilusis mokymasis

- Dėl savo savybių gilusis mokymasis dažniausiai taikomas
 - **Vaizdų** apdorojime ir analizėje
 - **Garsų** apdorojime ir analizėje (šnekos atpažinimas)
 - **Signalų** apdorojime ir analizėje (kalbos analizė)



Gilusis mokymasis ir dirbtinis intelektas



Gilusis mokymasis: istoriniai aspektai

- Giliojo mokymosi paradigma **nėra naujas dalykas**, bet įvairiais laikotarpiais turėjo įvairius vardus.
- Gilusis mokymasis tapo naudingas kai **išaugo mokymo duomenų kiekis**.
- Giliojo mokymosi modeliai tapo sudėtingesniais, kai **išsiplėtė skaičiavimų infrastruktūra** (tiek programinė, tiek techninė įranga). Tas leidžia spręsti sudėtingus taikomuosius uždavinius per sąlyginai trumpą laiką.

Gilusis mokymasis: istoriniai aspektai

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



1950's 1960's 1970's 1980's 1990's 2000's 2010's

Gilusis mokymasis: kodėl katinas?

- Šiais laikais **pirmasis pavyzdinis uždavinys** naudojant giliojo mokymo paradigmą buvo:
 - Iš YouTube vaizdų bibliotekos išrinkti **vaizdus su katinais**.



Gilusis mokymasis: šiandien

● deep learning
Paieškos terminas

+ Palyginti

Pasaulyje ▼

Pastarieji 5 metai ▼

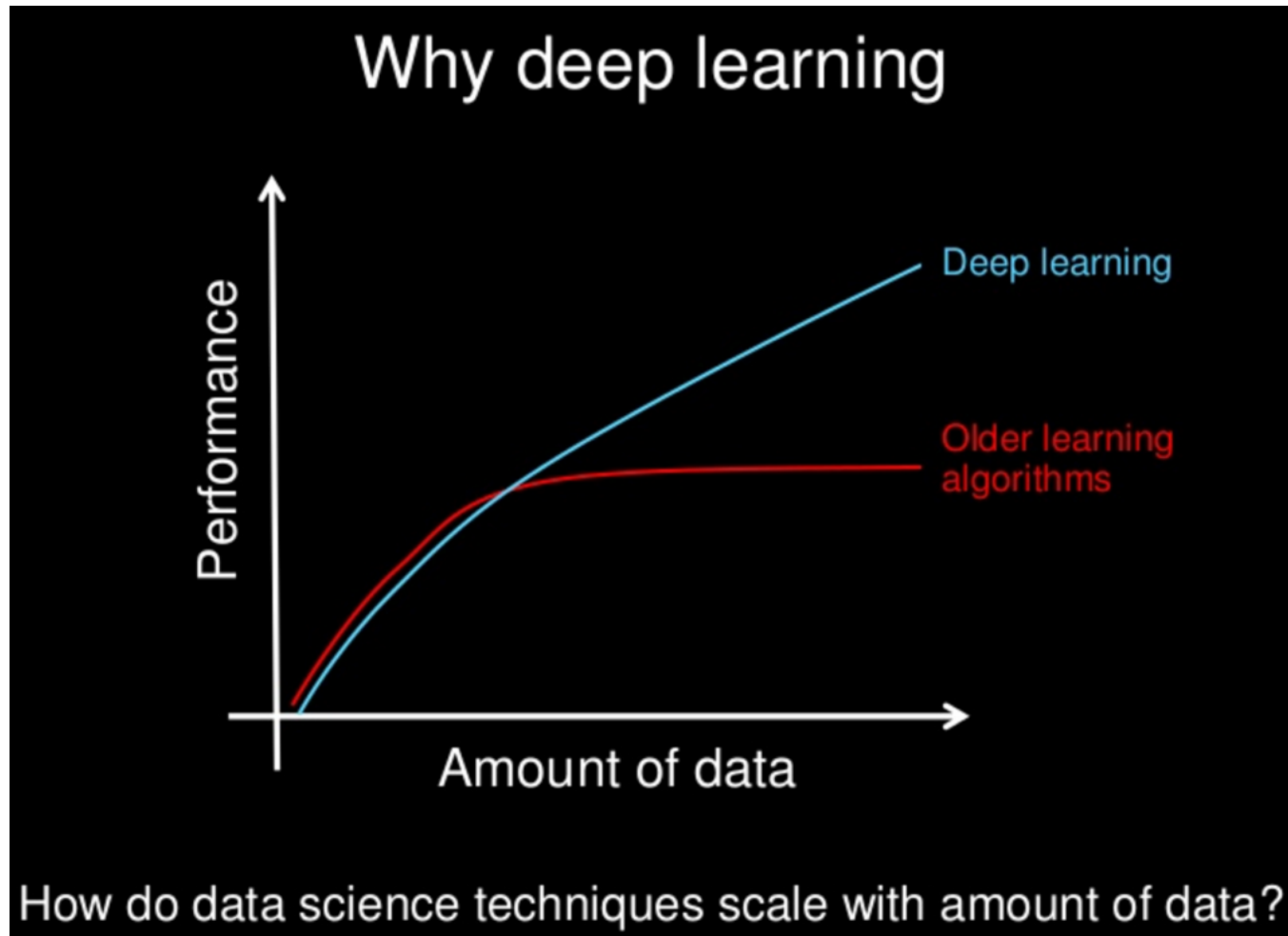
Visos kategorijos ▼

Žiniatinklio paieška ▼

Susidomėjimas per laikotarpį ?



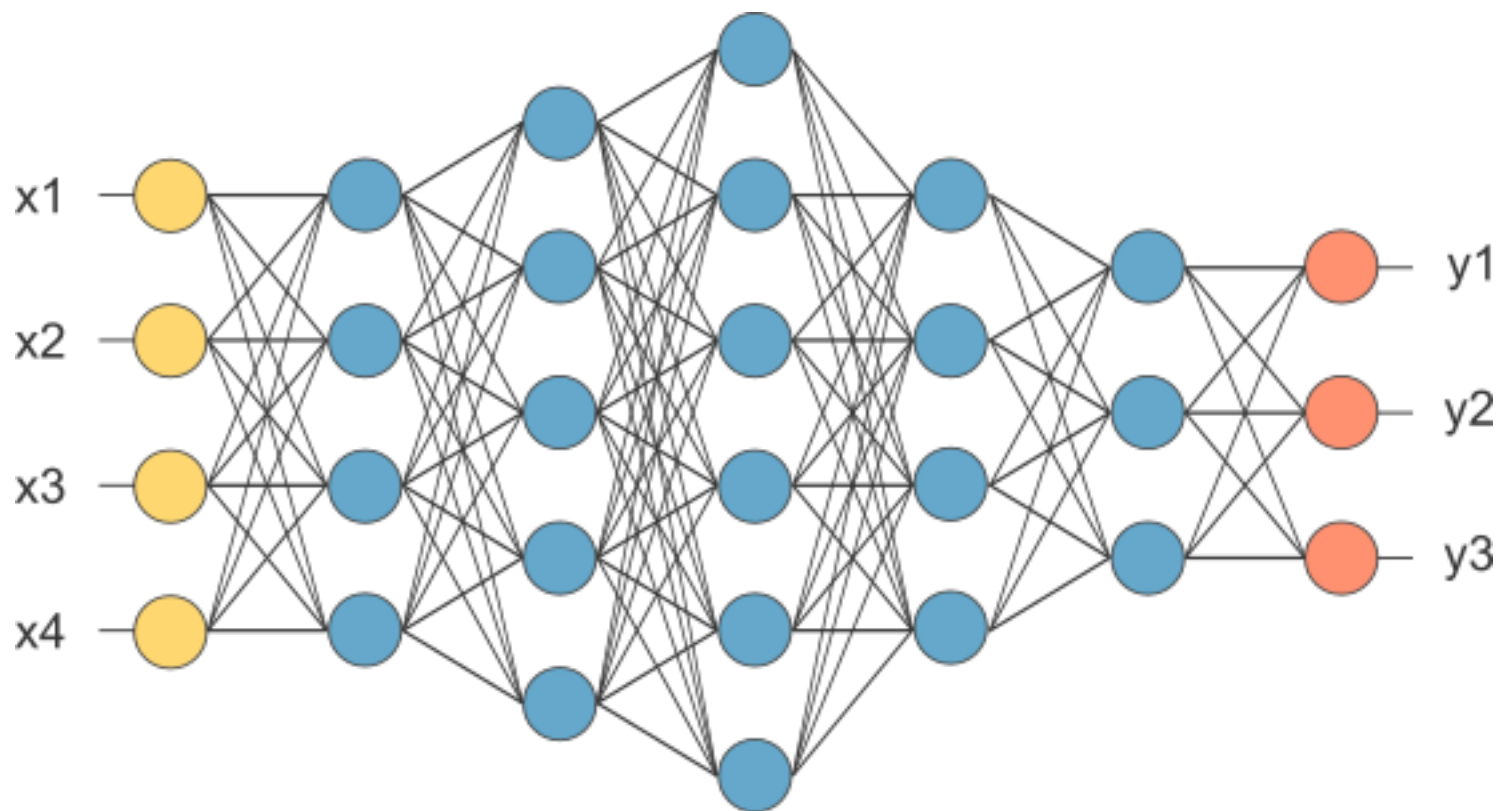
Giliojo mokymosi pajėgumai



Gilieji neuroniniai tinklai

- Nors **tiesioginio sklidimo neuroniniai tinklai** gali turėti norimą kiekį paslėptų sluoksnių ir neuronų skaičių juose, tačiau dėl buvusio skaičiavimų resursų ribotumo, dažnai buvo naudojami **tik vienas ar du paslėpti sluoksniai**, turintys po kelis neuronus.
- Šiuo metu plečiantis skaičiavimo resursų pajėgumams, atsirado galimybė naudoti **daugiau sluoksnių** ir **daugiau neuronų juose**.
- Taip išpopuliarėjo **gilieji neuroniniai tinklai**.

Giliųjų NN struktūra



Paveikslas iš <http://www.opennn.net>

Giliųjų DNT pradžia

- **Giliųjų DNT pradžia** laikoma **1965** m.
- Ivakhnenko, A. G. and Lapa, V. G. (1965). Cybernetic Predicting Devices. CCM Information Corporation.

