ICT academy

Uvod v strojno učenje

Univerza v Ljubljani
Fakulteta za elektrotehniko

IKT Katedra za informacijske in komunikacijske tehnologije

LTFE Laboratorij za telekomunikacije

LMMFE Laboratorij za multimedijo

LTFE ICT academy

www.ict-academy.eu

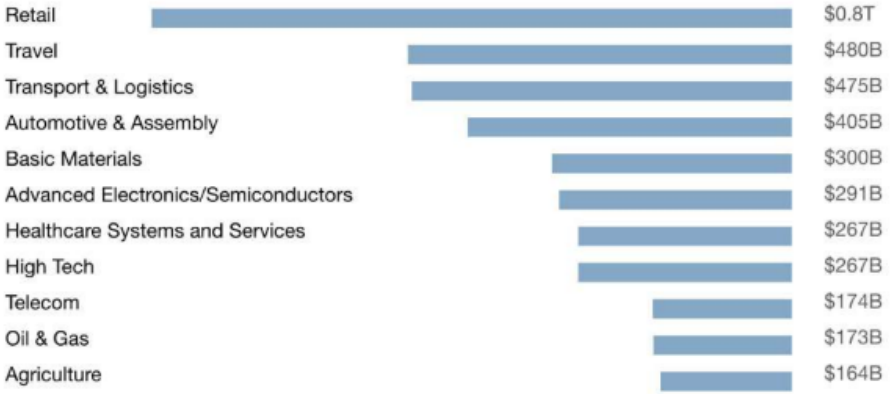
Kaj je umetna inteligenca (AI)?

Uvod

- Leta 2030 naj bi AI tehnologija proizvedla dodatnih 13 biljonov $ letno
- Vedno večja uporaba tudi v ne-tehnoloških sektorjih
  - Težko bo najti industrijo, ki ne uporablja AI tehnologije

AI value creation by 2030

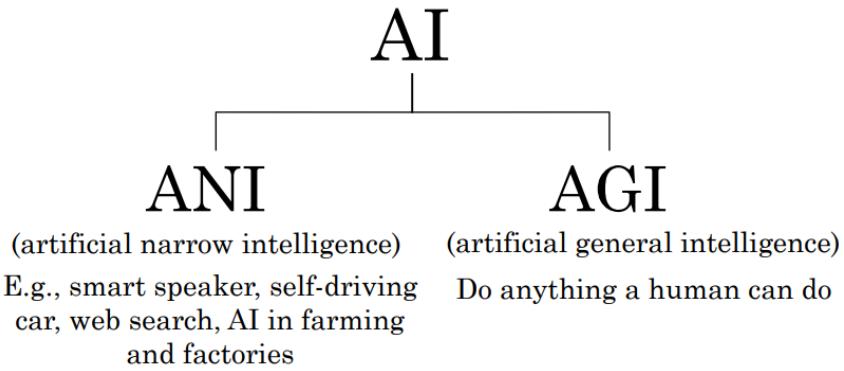
$13 trillion



[Source: McKinsey Global Institute.]

Tipi umetne inteligence

- ANI - Artificial Narrow Intelligence
- AGI - Artificial General Intelligence
- ASI - Artificial Super Intelligence



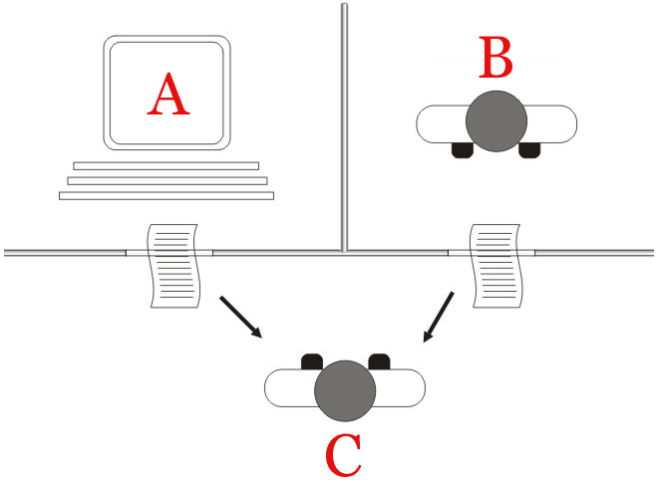
NARROW AI	ARTIFICIAL GENERAL INTELLIGENCE (AGI)
AI focused on a specific, singular or limited task	Not fully realized, with some developers questioning if it will be possible
Examples include image recognition, hyper-personalization, chatbots, predictive text	Seeks machines that can handle a range of cognitive tasks with little oversight
Trained on specific tasks by data scientists	The ability to learn, generalize, apply knowledge and plan for the future
Correlates questions or assignments to a specific data set to accomplish a task	Must consistently pass the Turing Test
No self-awareness, consciousness, ability to think	Single, general intelligence that possesses common sense and creativity and expresses emotions

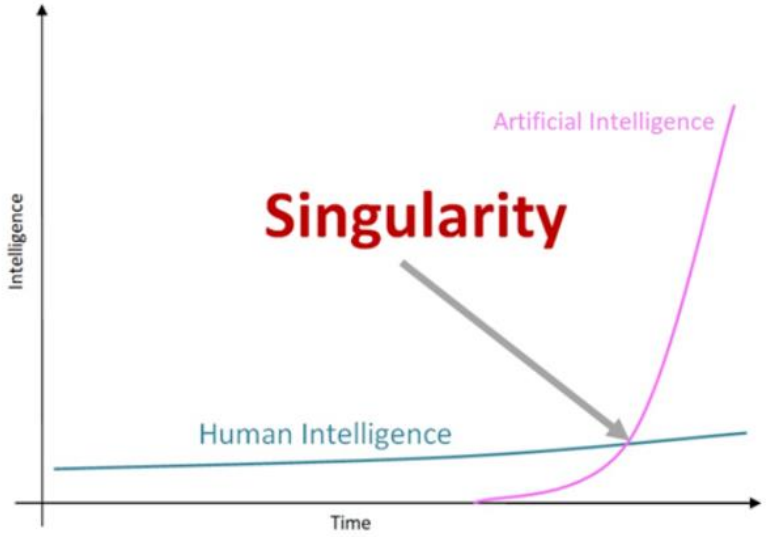
Source: https://www.techtarget.com/searchenterpriseai/definition/artificial-general-intelligence-AGI

- Razlika med ANI in AGI je torej v njihovi sposobnosti razmišljanja in učenja. ANI je omejena na določeno področje, medtem ko je AGI sposobna delovati na splošno, kot človek.
- Izumitelj in futurist Ray Kurzweil je napovedal, da bodo računalniki dosegli človeško raven inteligence do leta 2029
- Primeri ANI:
  - self-driving cars
  - customer service chatbots
  - voice assistants like Siri and Alexa
  - recommendation engines such as those Google, Netflix and Spotify use
  - facial recognition applications
- Primeri AGI (trenutno še ne obstaja, približki):
  - IBM Watson supercomputer
  - Language model Generative Pre-trained Transformer 3. GPT-3

Singularnost (Singularity)

- **Singularnost** je hipotetični trenutek v prihodnosti, ko bi lahko **umetna inteligenca dosegla ali presegla človeško inteligenco** in postala najmočnejša sila na svetu.
  - samostojno razvijala svoje sposobnosti
  - napredovala na način, ki bi presegel zmožnosti človeškega razumevanja
- Turingov test (Alan Turing, 1950)

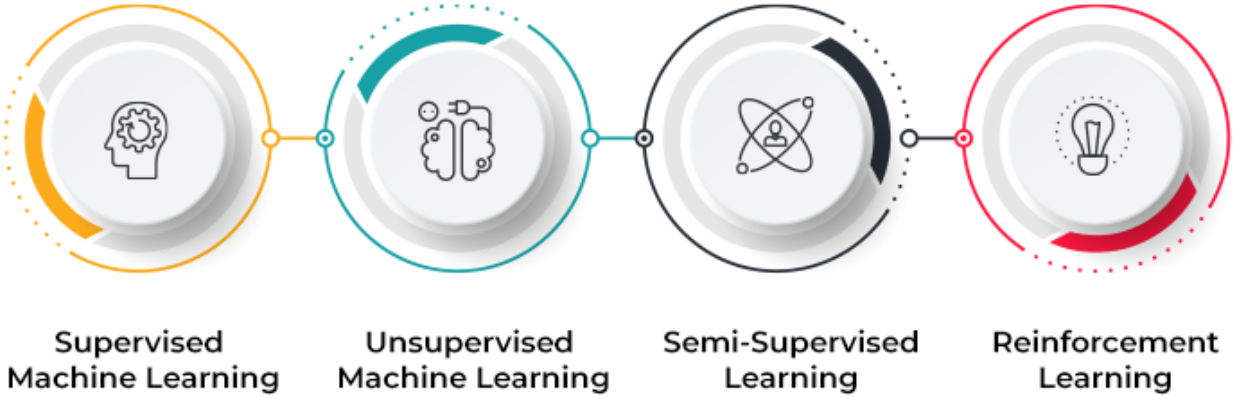




Strojno učenje

Kaj je strojno učenje?

- Strojno učenje je **področje umetne inteligence**, ki se ukvarja s tem, kako lahko računalniki samostojno pridobivajo nove sposobnosti in znanje iz podatkov, ne da bi jim bilo treba vnaprej določiti, kaj naj počnejo.
- **Pridobivanje znanja iz podatkov.**
- Uspeh AI-ja je močno odvisen od strojnega učenja.



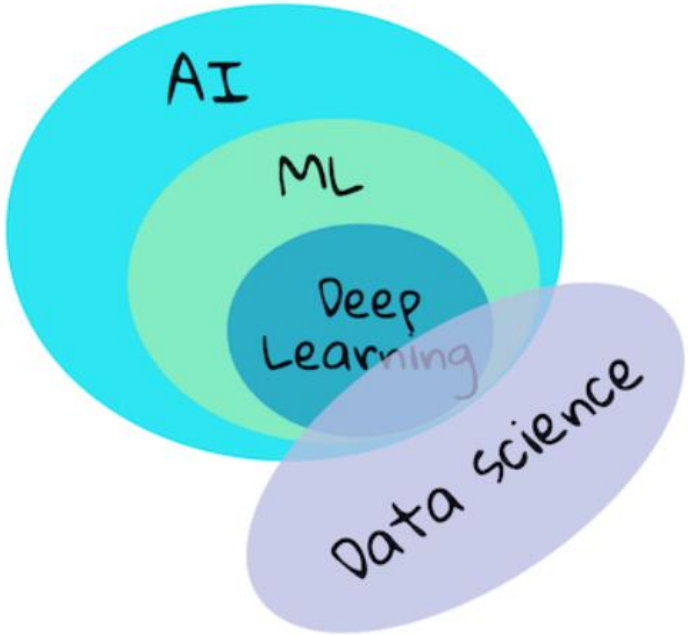
machine learning
Iskalni izraz

+ Primerjava

Ves svet | Od 2004 do zdaj | Vse kategorije | Spletno iskanje

Zanimanje skozi čas

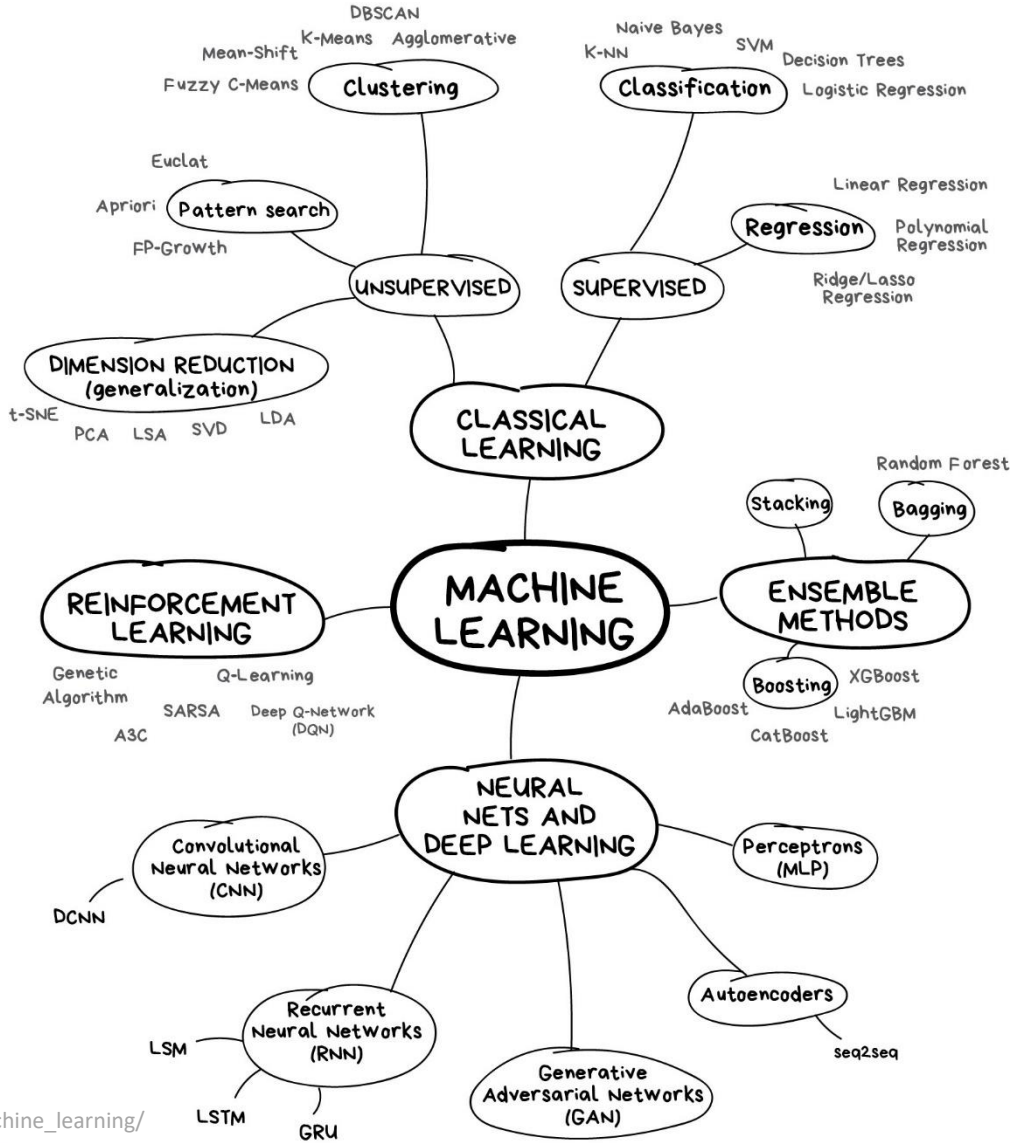
- **ML is concerned with using specialized algorithms to uncover meaningful information and find hidden patterns from perceived data to corroborate the rational decision-making process.**



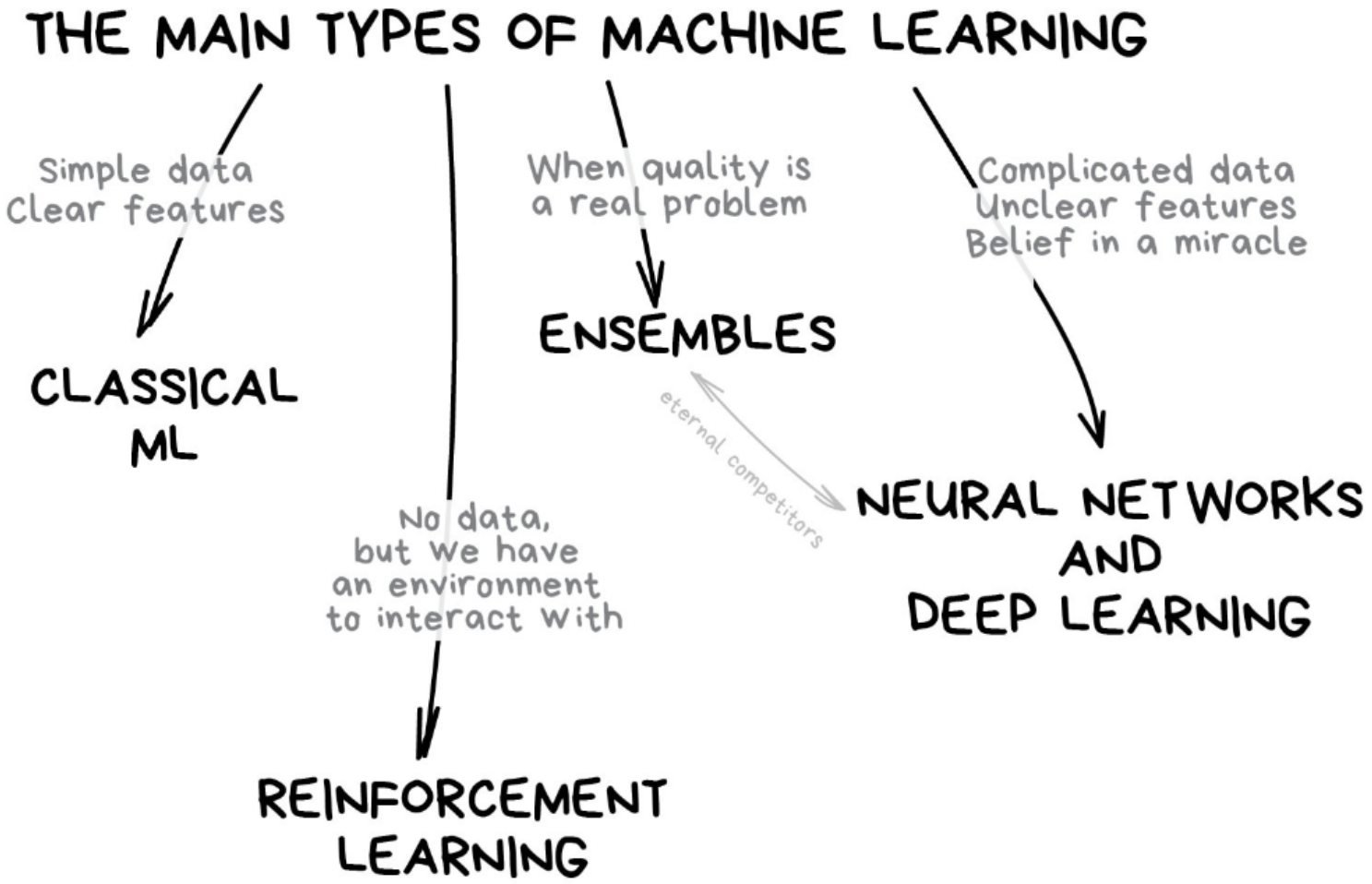
https://microsoft.github.io/ML-For-Beginners/1-Introduction/1-intro-to-ML/images/ai-ml-ds.png

Applications of machine learning

- Examples of applied ML:
  - To predict the likelihood of disease from a patient's medical history or reports.
  - To leverage weather data to predict weather events.
  - To understand the sentiment of a text.
  - To detect fake news to stop the spread of propaganda.
- Finance, economics, earth science, space exploration, biomedical engineering, cognitive science …



https://vas3k.com/blog/machine_learning/



https://vas3k.com/blog/machine_learning/

CLASSICAL MACHINE LEARNING

Data is pre-categorized or numerical

Data is not labeled in any way

SUPERVISED

Predict a category

Predict a number

CLASSIFICATION

«Divide the socks by color»

REGRESSION

«Divide the ties by length»

UNSUPERVISED

Divide by similarity

Find hidden dependencies

Identify sequences

CLUSTERING

«Split up similar clothing into stacks»

ASSOCIATION

«Find what clothes I often wear together»

DIMENSION REDUCTION (generalization)

«Make the best outfits from the given clothes»

https://vas3k.com/blog/machine_learning/

Podatki

- Kako jih pridobimo?
  - Ročno označevanje (Manual labeling)


cat


not cat


cat


not cat

  - Obnašanje uporabnikov

user ID	time	price ($)	purchased
4783	Jan 21 08:15.20	7.95	yes
3893	March 3 11:30.15	10.00	yes
8384	June 11 14:15.05	9.50	no
0931	Aug 2 20:30.55	12.90	yes

machine	temperature (◦C)	pressure (psi)	machine fault
17987	60	7.65	N
34672	100	25.50	N
08542	140	75.50	Y
98536	165	125.00	Y

  - Prenos obstoječih podatkov iz interneta

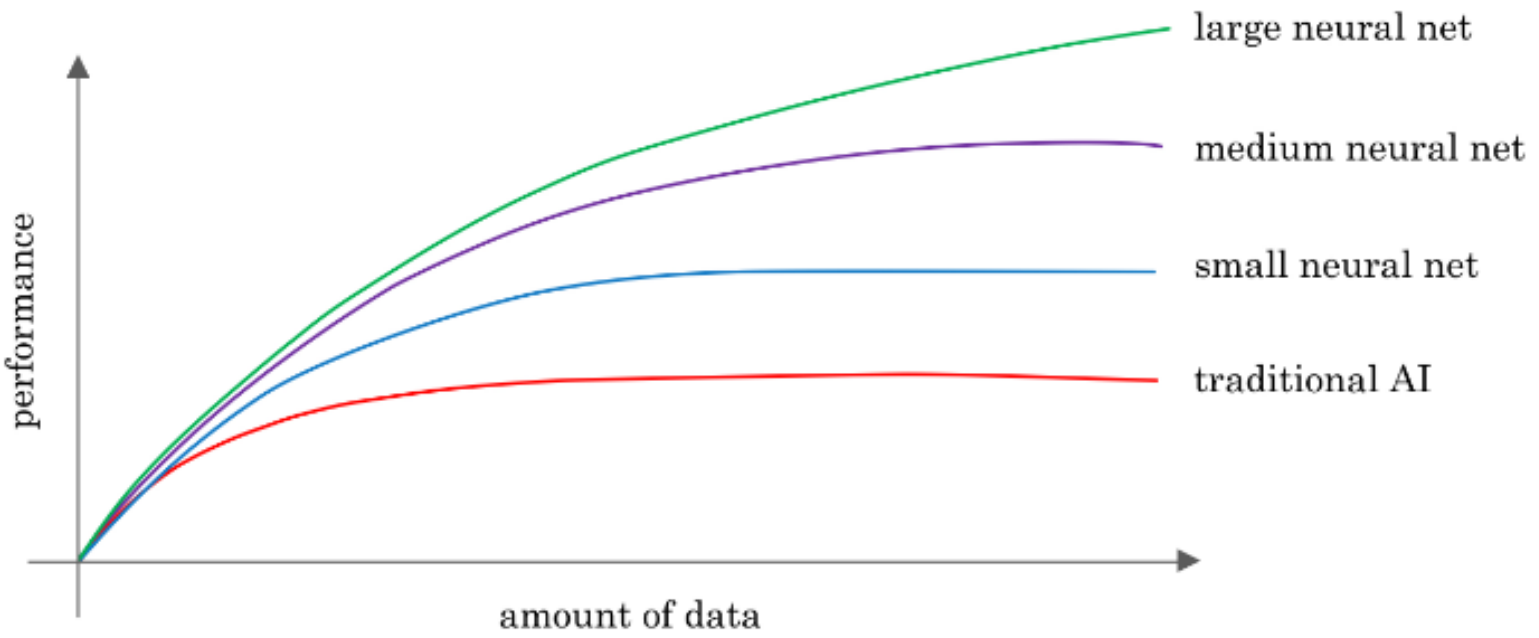
Advantages and Disadvantages of Machine Learning

- Easily identifies trends and patterns
- No human intervention needed (automation)
- Continuous Improvement
- Handling multi-dimensional and multi-variety data
- Wide Applications

- Data Acquisition
- Time and Resources
- Interpretation of Results
- High error-susceptibility

Workflow of a machine learning project

1. Decide on the question
2. Collect and prepare data
3. Choose a training method
4. Train model (iterate many times until good enough)
5. Evaluate the model
6. Parameter tuning
7. Predict
8. Deploy model
    1. Get data back
    2. Maintain / update model


large neural net
medium neural net
small neural net
traditional AI
performance
amount of data

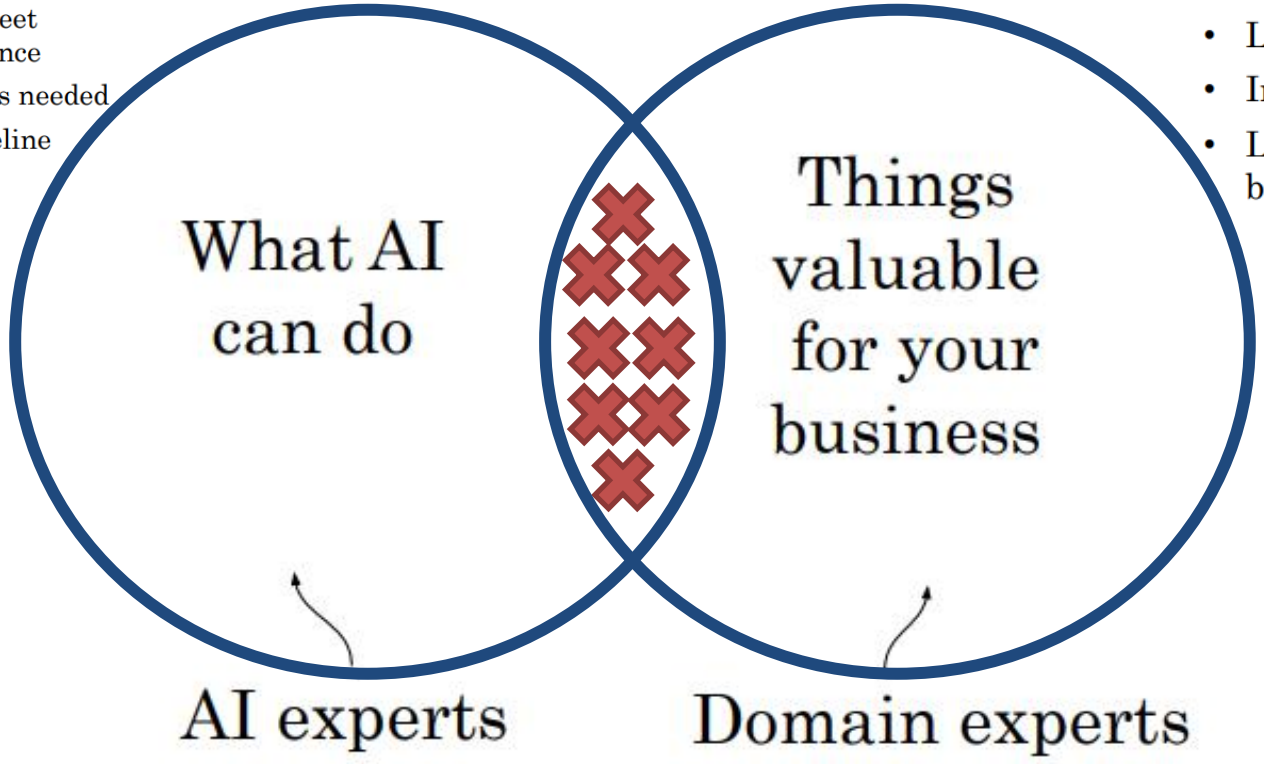
How to choose an AI project?

Technical diligence

- Can AI system meet desired performance
- How much data is needed
- Engineering timeline

Business diligence

- Lower costs
- Increase revenue
- Launch new product or business



Technical tools

- Machine learning frameworks:
  - TensorFlow
  - PyTorch
  - Keras
  - MXNet
  - CNTK
  - Caffe
  - PaddlePaddle
  - Scikit-learn
  - R
  - Weka

CPU vs. GPU

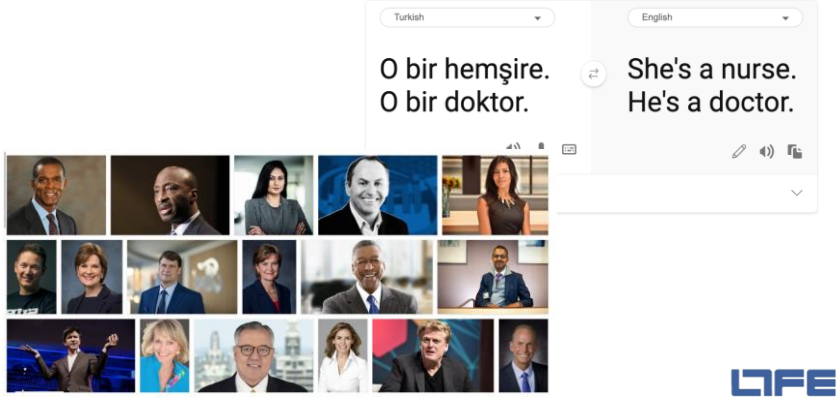
CPU: Computer processor (Central Processing Unit)

GPU: Graphics Processing Unit

Cloud vs. On-premises

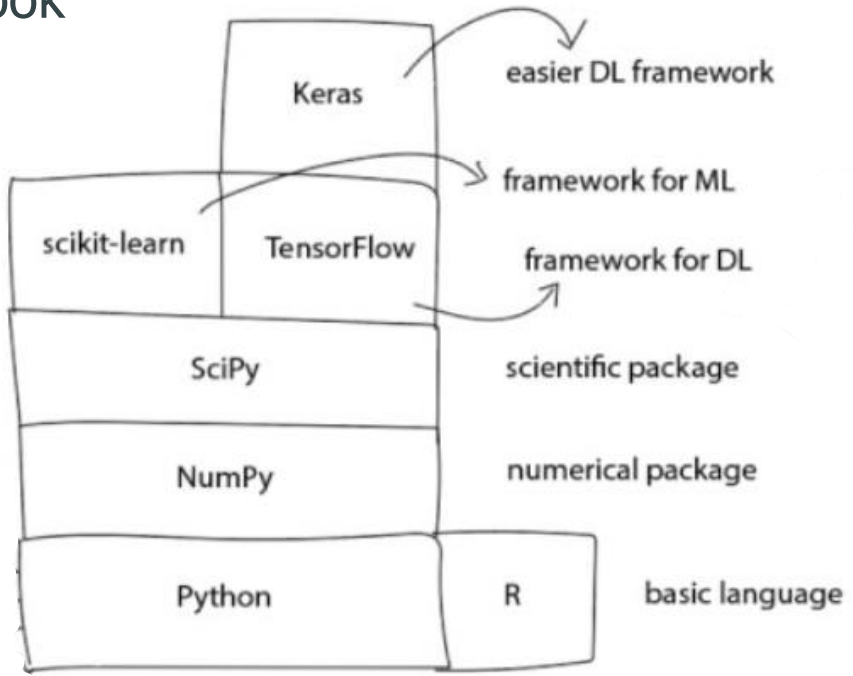
Fairness in Machine Learning

- *"If you torture the data long enough, it will confess to anything" - Ronald Coase*
- Data can be **manipulated** to support any conclusion
- Guaranteeing **fairness in AI** and machine learning remains a complex sociotechnical challenge
- Fairlearn is an open-source Python package that allows you to assess your systems' fairness and mitigate unfairness.
- The main fairness-related harms can be classified as:
  - Allocation
  - Quality of service
  - Stereotyping
  - Denigration
  - Over- or under- representation



Priprava okolja

- Python, VSCode, Jupyter Notebook



Hvala!

Leon Štefanič Južnič
leon.stefanic@ltfe.org

LTFE
Laboratorij za telekomunikacije

LMMFE
Laboratorij za multimedijo

IKT
Katedra za informacijske in komunikacijske tehnologije

Univerza v Ljubljani
Fakulteta za elektrotehniko