OEAW AI SUMMER SCHOOL

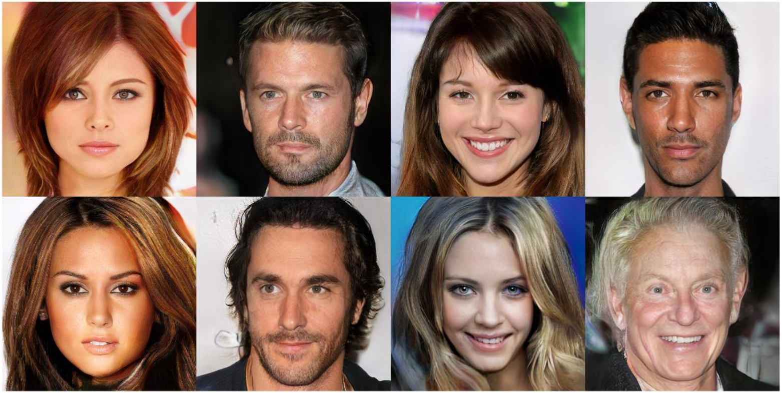
Introduction to Data Science

Johannes Brandstetter
Institute for Machine Learning

JKU
Institute for
Machine Learning

Welcome to Statistics, Machine Learning, Deep Learning, and ... Southern Styria

Which one of those is generated by AI?



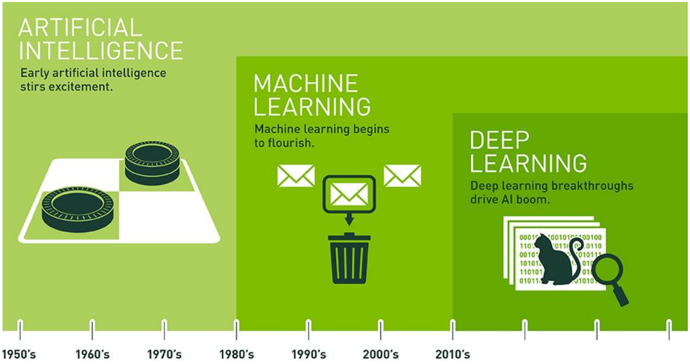
Content of this Summer School

- Bayes Statistics
  - Statistics based on the Bayesian interpretation of probability
  - Probability expresses a degree of belief in an event
  - No data science without statistics
- Advanced methods for classification and regression
  - Everything except Neural Networks
  - What you learn here will help you in many (scientific) daily-life problems
- Deep Learning
  - From simple Logistic Regression to Transformer Networks
  - How matrix multiplication and backpropagation changed the world

Why do we have an AI Summer School?

AI is more than Deep/Machine Learning (and it sounds much sexier)

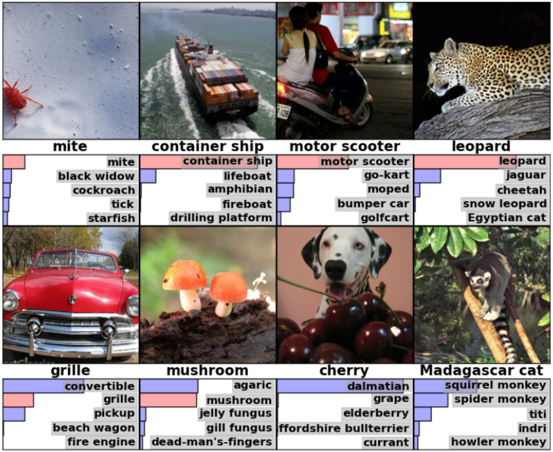
- The capability of a machine to imitate intelligent human behavior (Merriam Webster).
- The recent success of AI is mostly based on Deep Learning.


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

What is responsible for the AI boom?

- Starting point was around 2010.
- Boom mostly due to neural networks (NNs) together with recent availability of very fast computers (GPUs, TPUs) and massive data sets.
- Kurt Hornik in 1991: Neural network architecture itself gives NNs the potential of being universal function approximators.
  - □ Hornik K (1991). Approximation Capabilities of Multilayer Feedforward Networks, Neural Networks, 4(2), 251-257.
- First NN boom in the late 1980s and 1990s
  - □ AI winter due to lack of computational power
  - □ Research shift towards mathematically more profound methods: Support Vector Machines, various Unsupervised Learning methods, ...

Deep Learning methods classify images

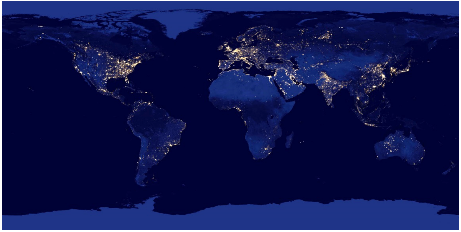


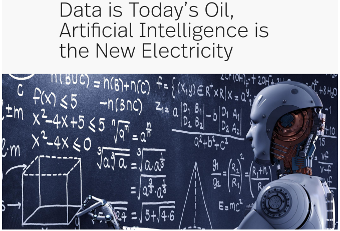
Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 1097-1105.
LeCun Y, Bottou L, Bengio Y, Haffner P (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE 86 (11), 2278-2324.

Deep Learning methods understand natural language

Hochreiter S, Schmidhuber J (1997). Long short-term memory. Neural computation 9 (8) 1735-1780.

Golden Age of AI

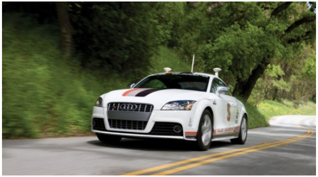




"It is a golden age. We are now solving problems with machine learning and artificial intelligence that were [...] in the realm of science fiction for the last several decades. [...] Machine Learning and AI is a horizontal enabling layer. It will empower and improve every business, every government organization, every philanthropy. Basically, there is no institution in the world that cannot be improved with machine learning."
(Jeff Bezos, 2017)

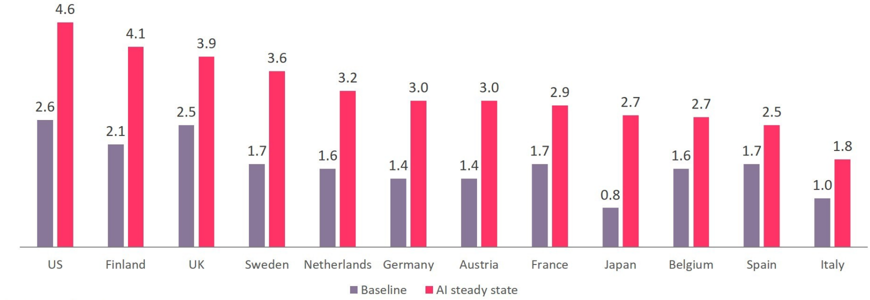
AI is ubiquitous

- AI pervades commercial applications in an unprecedented manner and is fundamentally changing how businesses operate across virtually all sectors
  - □ Information technology
  - □ Manufacturing and supply chains
  - □ Medicine and healthcare
  - □ Education
  - □ Financial, leagal and tax services
  - □ News and publishing
  - □ Transportation
  - □ ...
  - □ SCIENCE



Impact on the Economy

- Accenture (2016): AI doubles annual GDP growth rate until 2035 (similar studies: McKinsey, PwC, ...)
  - □ Austria without AI adoption: 1.4% growth rate
  - □ Austria with AI adoption: 3.0% growth rate



Source: Accenture and Frontier Economics

A word of caution

- (Deep) Machine Learning has the potential to revolutionize your field of science ...
- ... BUT (Deep) Machine Learning is no black box magic which always works:
  - □ Not every problem is a ML problem:
    - Sometime “simple” statistics is all you need
  - □ Not every ML problem is a Deep Learning problem:
    - You have to know your data
    - You have to know the statistics
    - You have to know what algorithm to use
    - You have to know how to control the beast (especially in Deep Learning)
- ... this is why we are here
- Keep in mind: Statistics is a (huge) field, (Deep) ML has become a (huge) field too! You cannot e.g. study physics in one week!

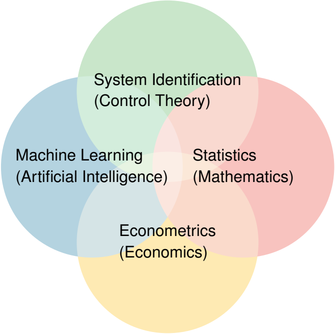
What is Machine Learning?

Machine Learning:

$$\text{data} + \text{model} \xrightarrow{compute} \text{prediction} \tag{1}$$

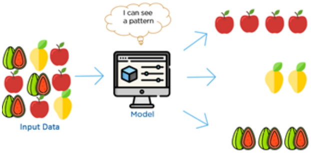
Goal of (supervised) ML is the minimization of Generalization Error:
Generalization error is a measure of how accurately an algorithm is able to predict values for unseen data.

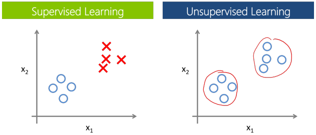
Machine Learning and Related Fields:

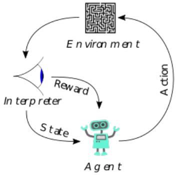


Machine Learning is a broad field

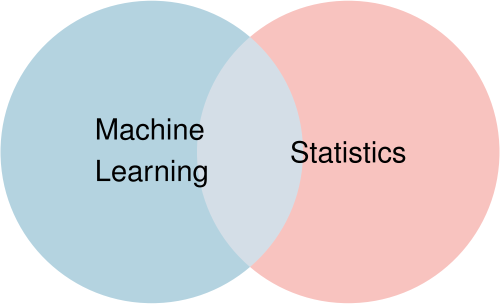
- Supervised Learning: data with labels
  - □ Quality of the predictive models depends on quality of labels.
  - □ Model can only learn “what’s in the data”.
- Unsupervised Learning: the world consists of lots of unlabeled data
  - □ PCA, ICA, FA
  - □ Projection and scaling methods
  - □ Clustering, Biclustering
  - □ Densitiy estimation
  - □ Generative models
- Reinforcement Learning







Machine Learning vs Statistics

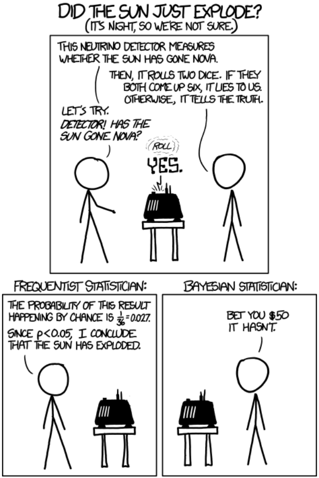


- Minimization of Generalization Error
- ML tries to make model predictions
- Statistical Learning Theory (Vapnik) is built on bias-variance tradeoff prediction.

- Parameter estimation and variance analysis
- Statistics tries to estimate parameters as precisely as possible.
- Statistics is built on bias-variance of parameter estimation.

Frequentists vs. Bayesians

- Information used:
  - Frequentists: data only
  - **Bayesians: data plus prior**
- Uncertainty measure:
  - Frequentists: confidence interval
  - Bayesians: credible interval
- Assessing significance:
  - Frequentists: hypothesis tests
  - Bayesians: direct interpretation of the posterior
- Basic concept:
  - Frequentists: relative frequency of an event
  - Bayesians: Bayes theorem


DID THE SUN JUST EXPLODE?
(IT'S NIGHT, SO WE'RE NOT SURE.)
THIS NEUTRINO DETECTOR MEASURES WHETHER THE SUN HAS GONE NOVA.
THEN, IT ROLLS TWO DICE. IF THEY BOTH COME UP SIX, IT LIES TO US. OTHERWISE, IT TELLS THE TRUTH.
LET'S TRY. DETECTOR! HAS THE SUN GONE NOVA?
(ROLL)
YES.
FREQUENTIST STATISTICIAN:
THE PROBABILITY OF THIS RESULT HAPPENING BY CHANCE IS $\frac{1}{36}$ = 0.027. SINCE $p < 0.05$, I CONCLUDE THAT THE SUN HAS EXPLODED.
BAYESIAN STATISTICIAN:
BET YOU $50 IT HASN'T.

Have fun @ the first OeAW AI Summer School