Artificial Intelligence and Forestry

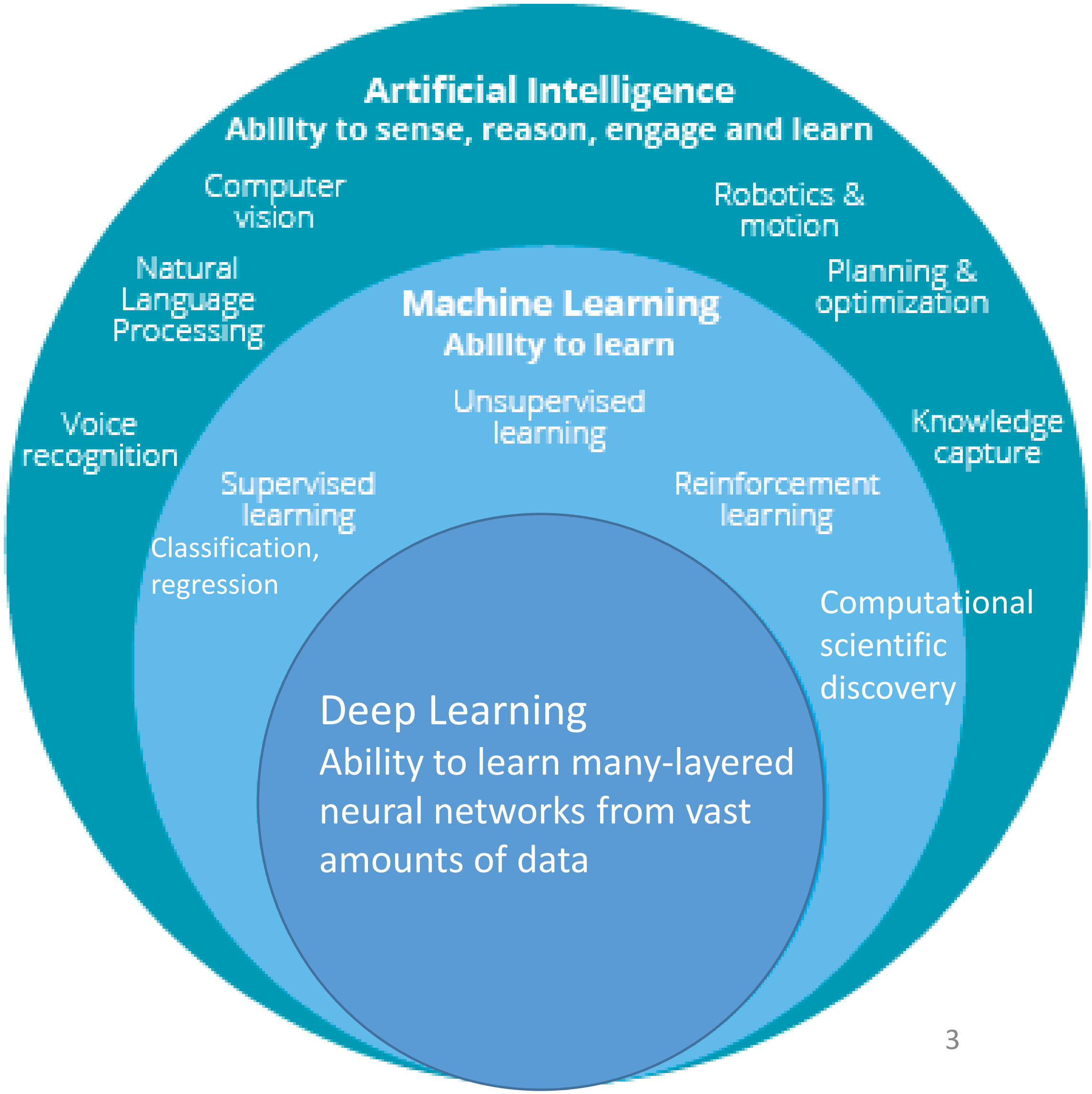
Sašo Džeroski
Jozef Stefan Institute (JSI), Ljubljana, Slovenia
Jozef Stefan International Postgraduate School (JS IPS)
European Space Agency (ESA/ESRIN), Frascati, Italy

Outline of the course

- Introduction to artificial intelligence
- Machine learning from big and complex data
- A gentle introduction to convolutional neural networks
- Some applications of machine learning in forestry (and broader)

Just what is AI?

AI is the science and engineering of making intelligent machines, especially intelligent computer programs (Alan Turing)



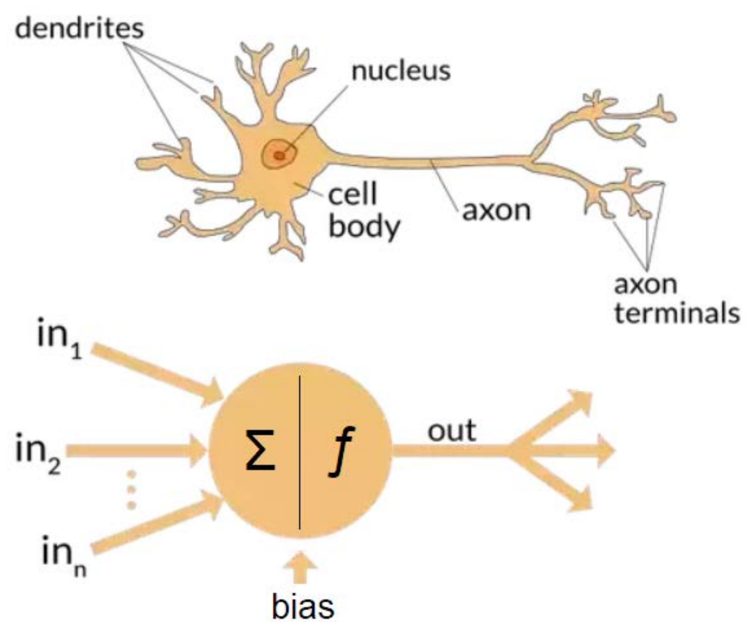
AI history & early optimism

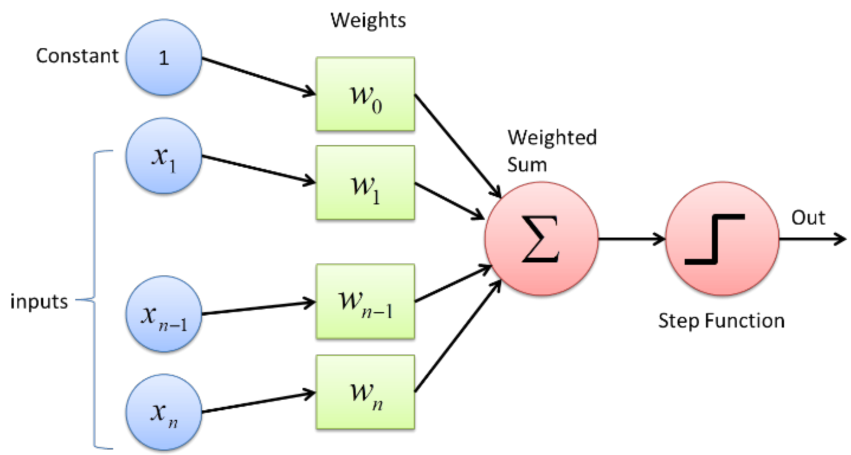
Start in the 1950s, Dartmouth conference in 1956: Big expectations, underestimating the difficulties that lie ahead

- 1958, H. A. Simon and Allen Newell: “… within ten years a digital computer will discover and prove an important new mathematical theorem.”
- 1965, H. A. Simon: “… machines will be capable, within twenty years, of doing any work a man can do.”
- 1967, Marvin Minsky: "Within a generation ... the problem of creating 'artificial intelligence' will substantially be solved.”
- 1970, Marvin Minsky: "In from three to eight years we will have a machine with the general intelligence of an average human being."

The perceptron: An artificial neuron

- The perceptron can be viewed as an artificial neuron

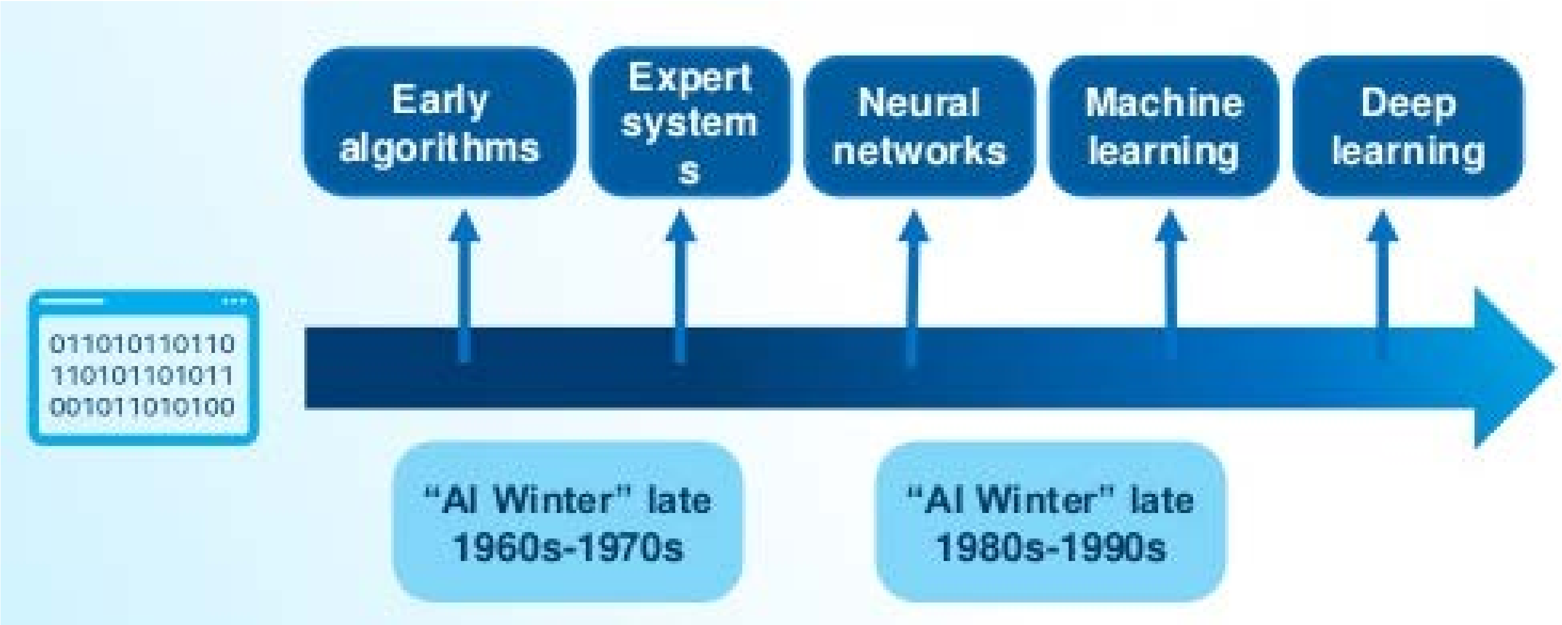




- Performs binary classification via linear discrimination: The basic unit of Artificial Neural Networks (Rosenblatt, 1958)
- Minsky and Papert (1969) show it cannot learn XOR function

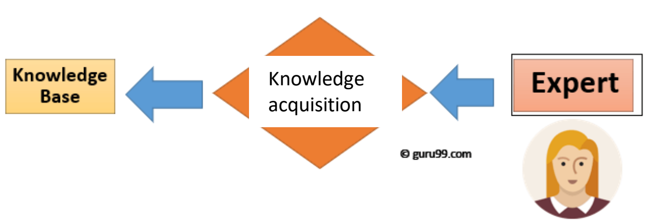
AI history

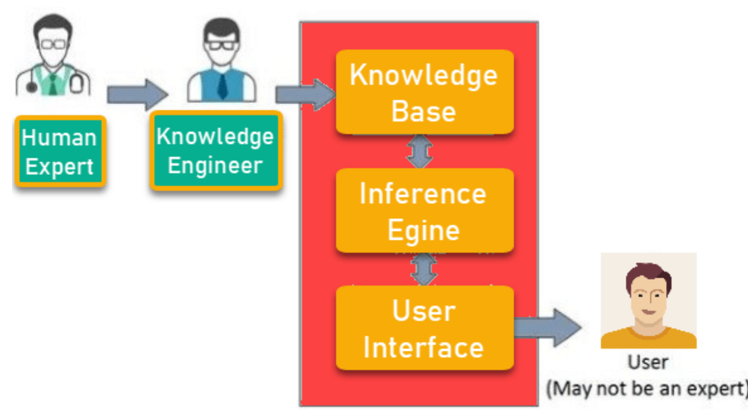
- Following the Minsky and Papert paper, the first AI winter ensues



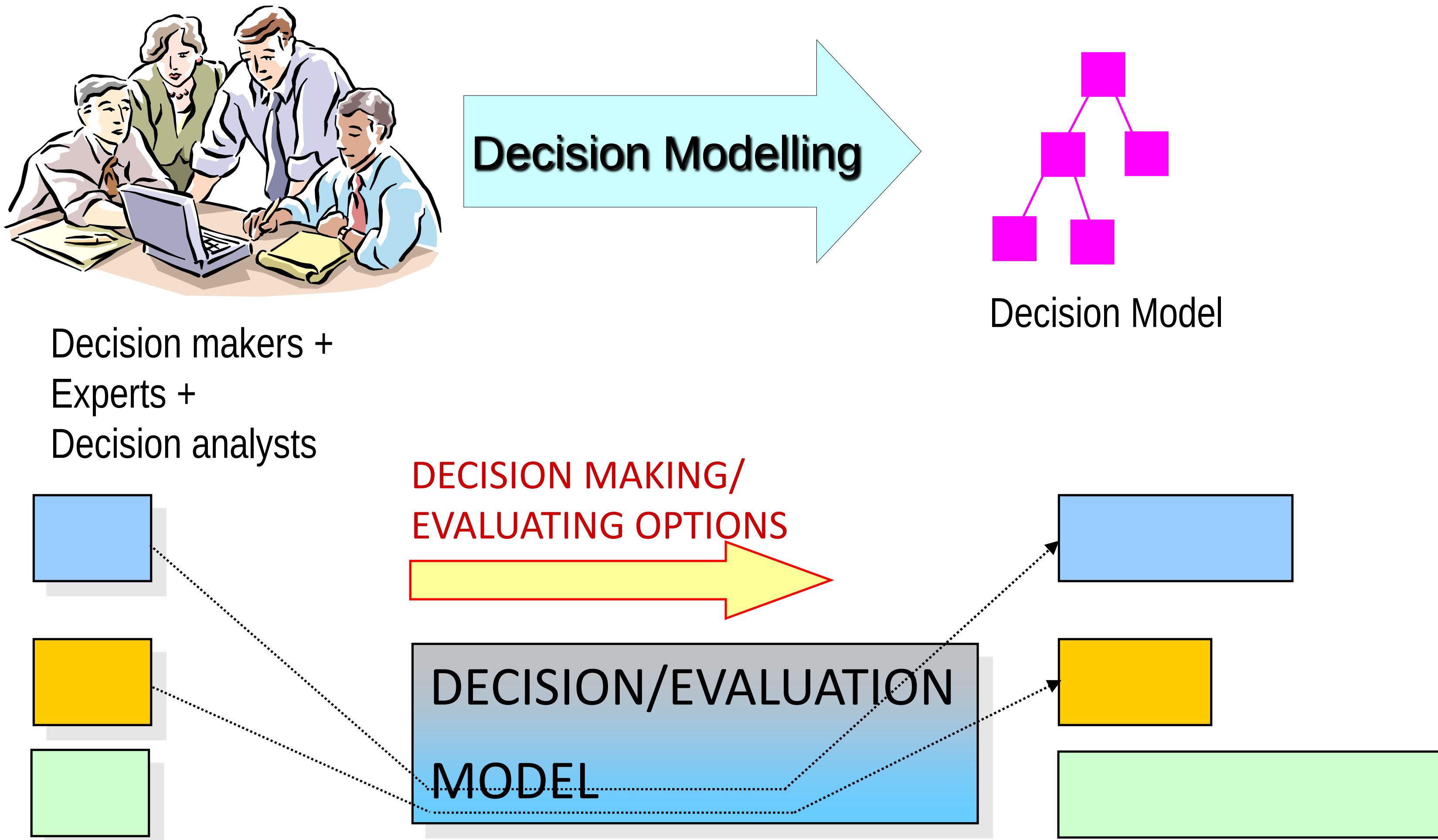
Expert systems

- An expert system is a computer system emulating the decision-making ability of a human expert
- Expert systems are designed to solve complex problems by reasoning over their knowledge bases, represented mainly as if–then rules
- Knowledge captured from human expert by a knowledge engineer, through a process known as knowledge acquisition

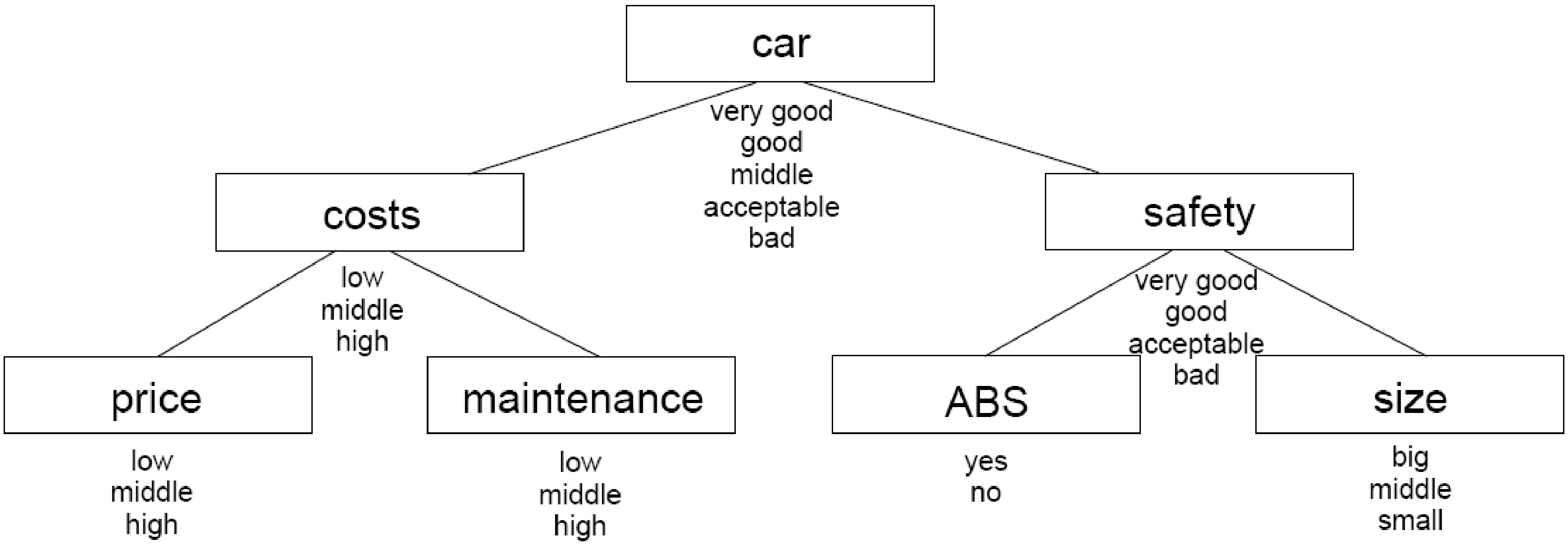




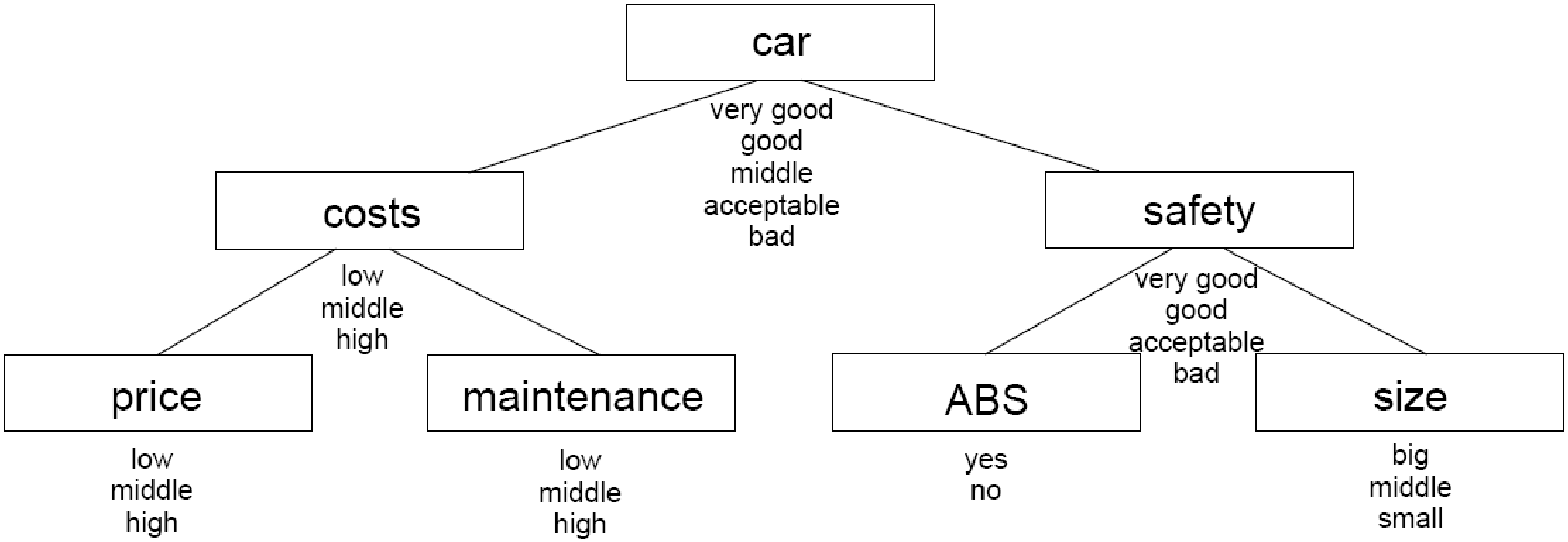
Expert Systems Here and Now: Decision Support Models



Decision Support Models: Hierarchical Multi-Attribute Models



Qualitative Multi-Attribute Models as Knowledge-Bases (DEXi)



costs	safety	car
low	v. good	v. good
low	good	good
low	acc.	mid.
low	bad	bad
mid.	v. good	good
mid.	good	good
mid.	acc.	acc.
mid.	bad	bad
high	v. good	good
high	good	mid.
high	acc.	bad
high	bad	bad

price	maint.	costs
low	low	low
low	mid.	low
low	high	mid.
mid.	low	low
mid.	mid.	mid.
mid.	high	high
high	low	mid.
high	mid.	high
high	high	high

IF ABS=no AND size=small
THEN safety = bad

ABS	size	safety
no	small	bad
no	mid.	acc.
no	big	good
yes	small	bad
yes	mid.	good
yes	big	v. good

The knowledge acquisition bottleneck

- To create an expert system, you need a knowledge-base.
- The knowledge needs to go from the experts to the computer.

Knowledge Engineering

"Critical scientific problem [...] successful applied AI requires that knowledge move from the heads of experts into programs"

FEIGENBAUM, E. A. (1984). Knowledge Engineering. Annals of the New York Academy of Sciences, 426: 91–107. doi: 10.1111/j.1749-6632.1984.tb16513.x

- Experts notoriously bad at explaining how they solve problems

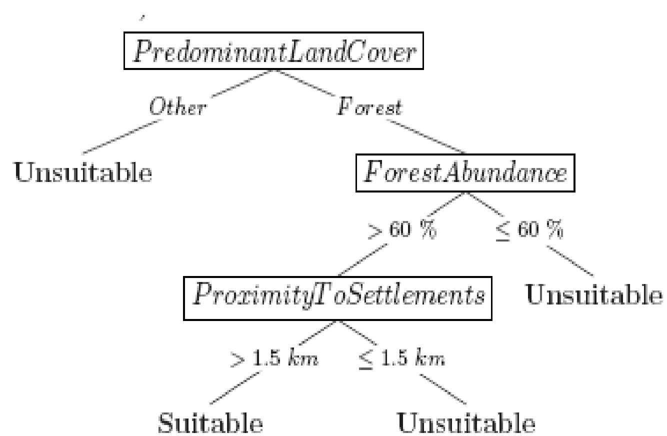
Machine Learning to the Rescue

- Input: Table of data

Location	PLC	FOREST-ABUNDANCE	PTS	OtherEnvVariables	BBH
l1	Forest	80	21.4	...	Yes
l2	Forest	66	13.9	...	Yes
l3	Forest	55	50.0	...	No
l4	Forest	72	1.2	...	No
l5	Grassland	6	19.1	...	No
l6	Grassland	0	11.4	...	No
l7	Wetland	3	5.8	...	No
l8	Water	0	3.9	...	No

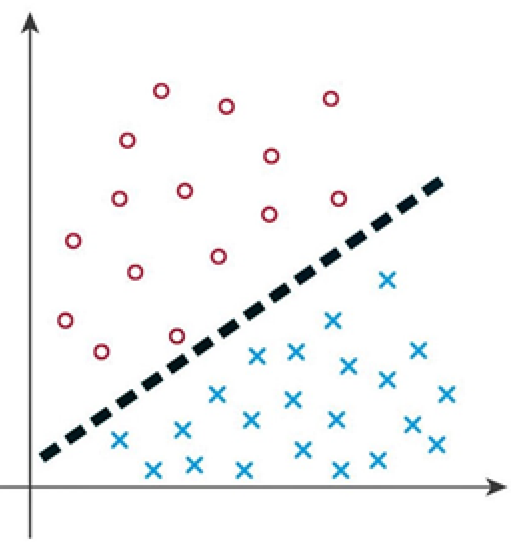
- Output: Models in the form of IF-THEN rules or decision trees

```
IF   PREDOMINANT-LAND-COVER = Forest
AND  FOREST-ABUNDANCE > 60%
AND  PROXIMITY-TO-SETTLEMENTS > 1.5 km
THEN BrownBearHabitat = Suitable
```

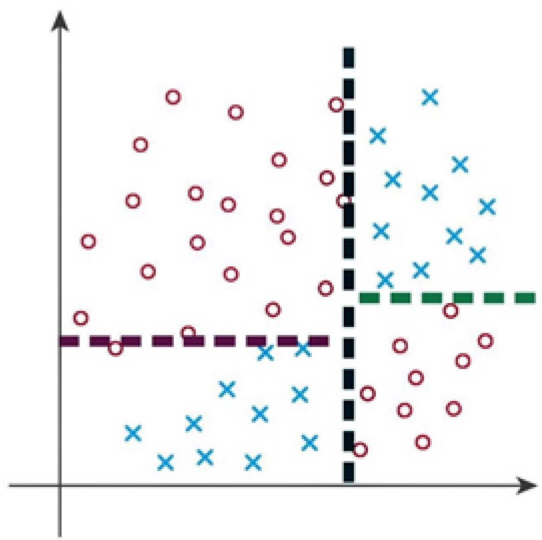


Decision Boundaries

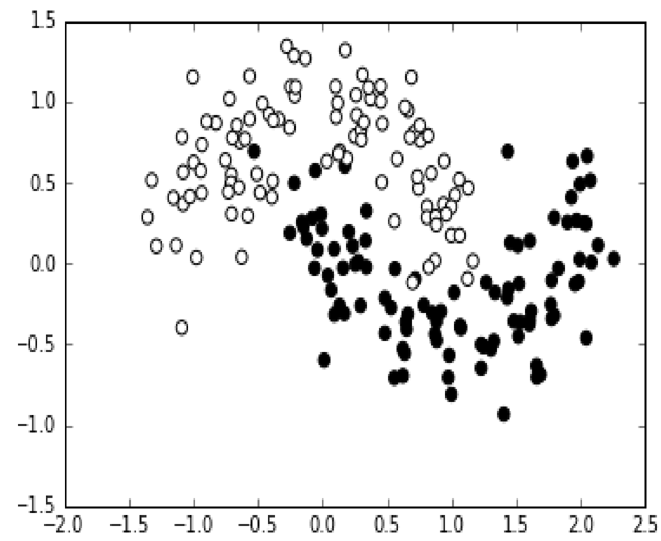
- Linear and nonlinear boundaries



Linearly separable dataset

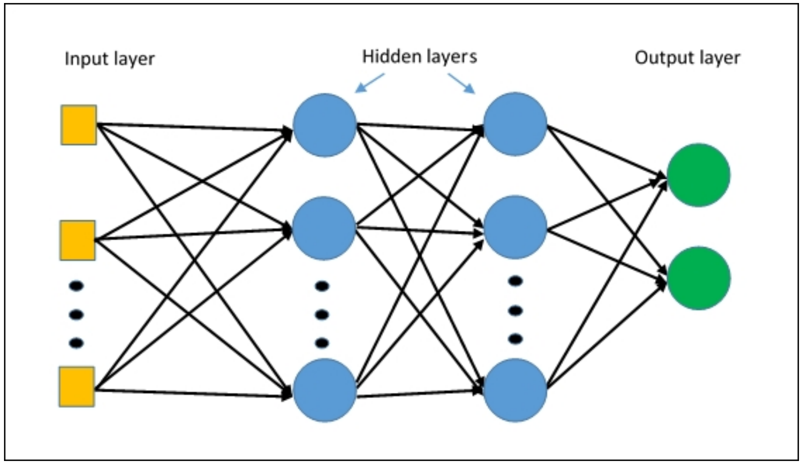


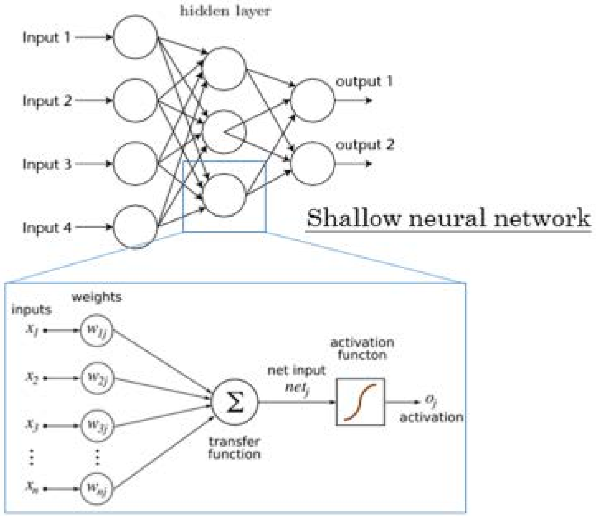
Linearly inseparable dataset



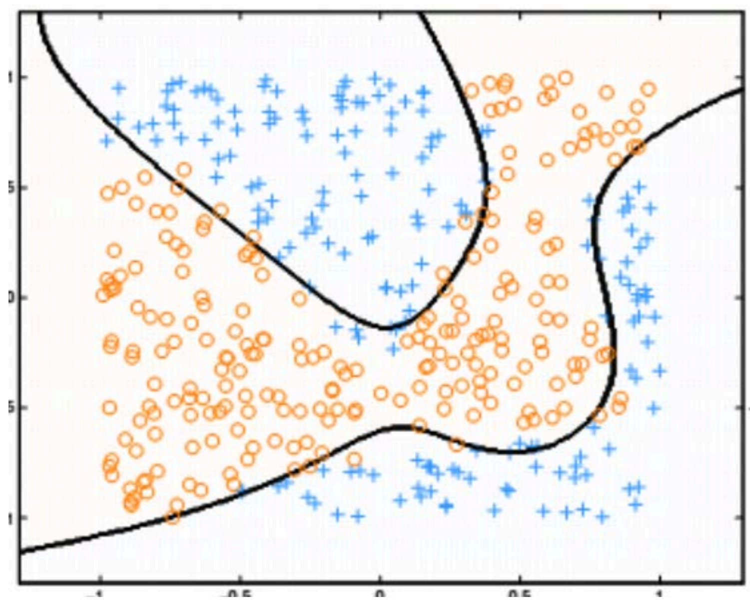
The multi-layer perceptron & NNs

- While a single perceptron cannot learn XOR, a multi-layer perceptron can. MLPs can approximate an arbitrary non-linear mapping between inputs & outputs
- Arbitrary depth or width may be needed
- Note the hidden layers: If only a few, shallow NNs



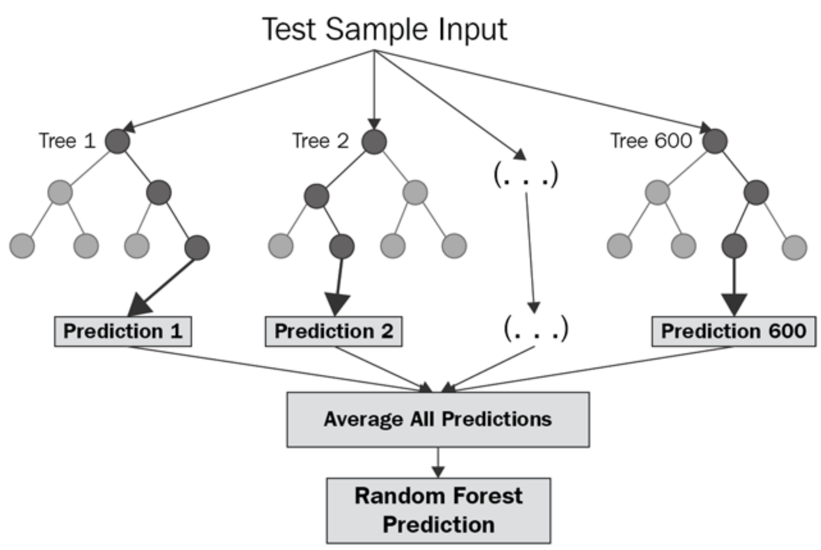


Neural Networks Decision Boundaries

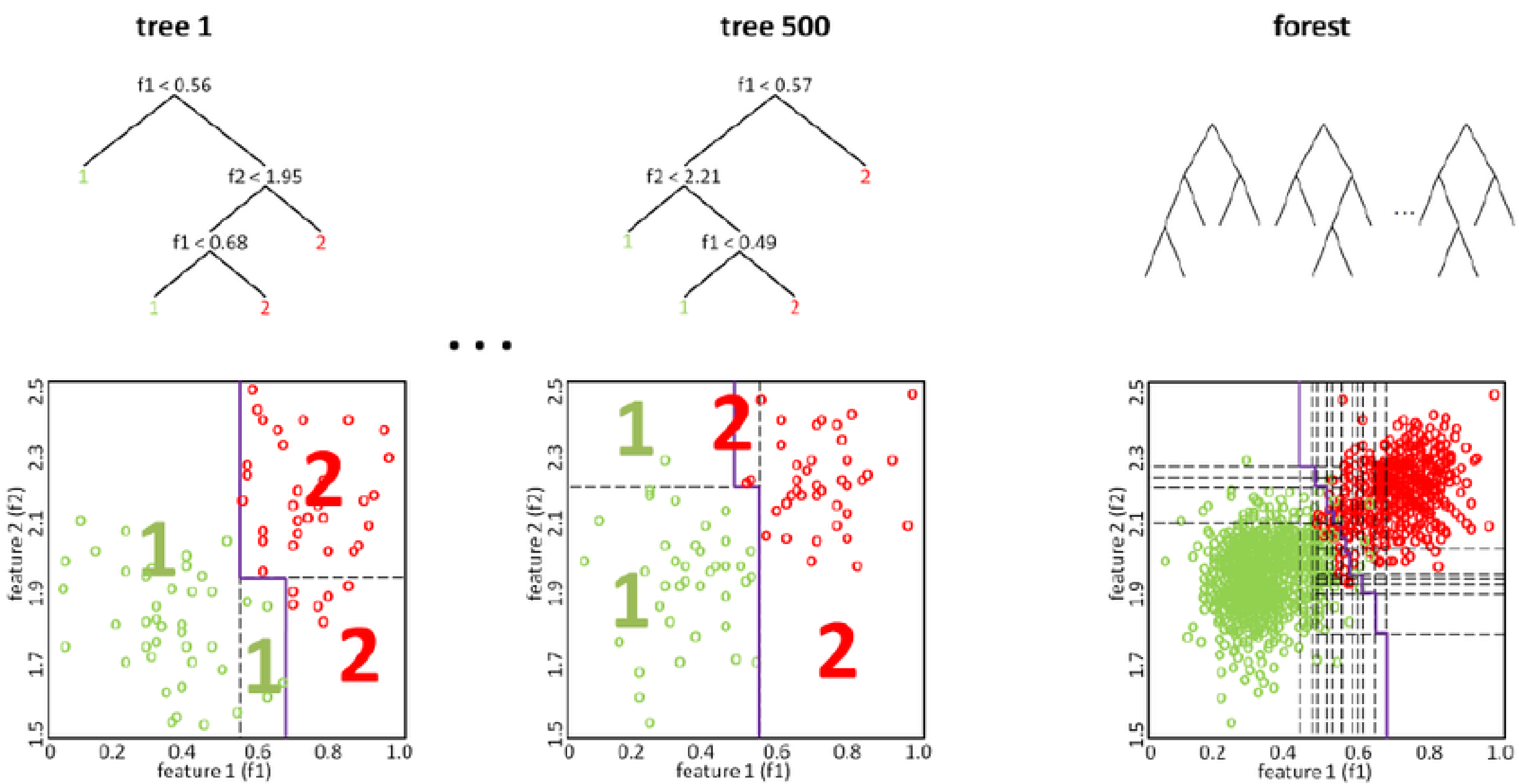


Ensemble methods: R. forests

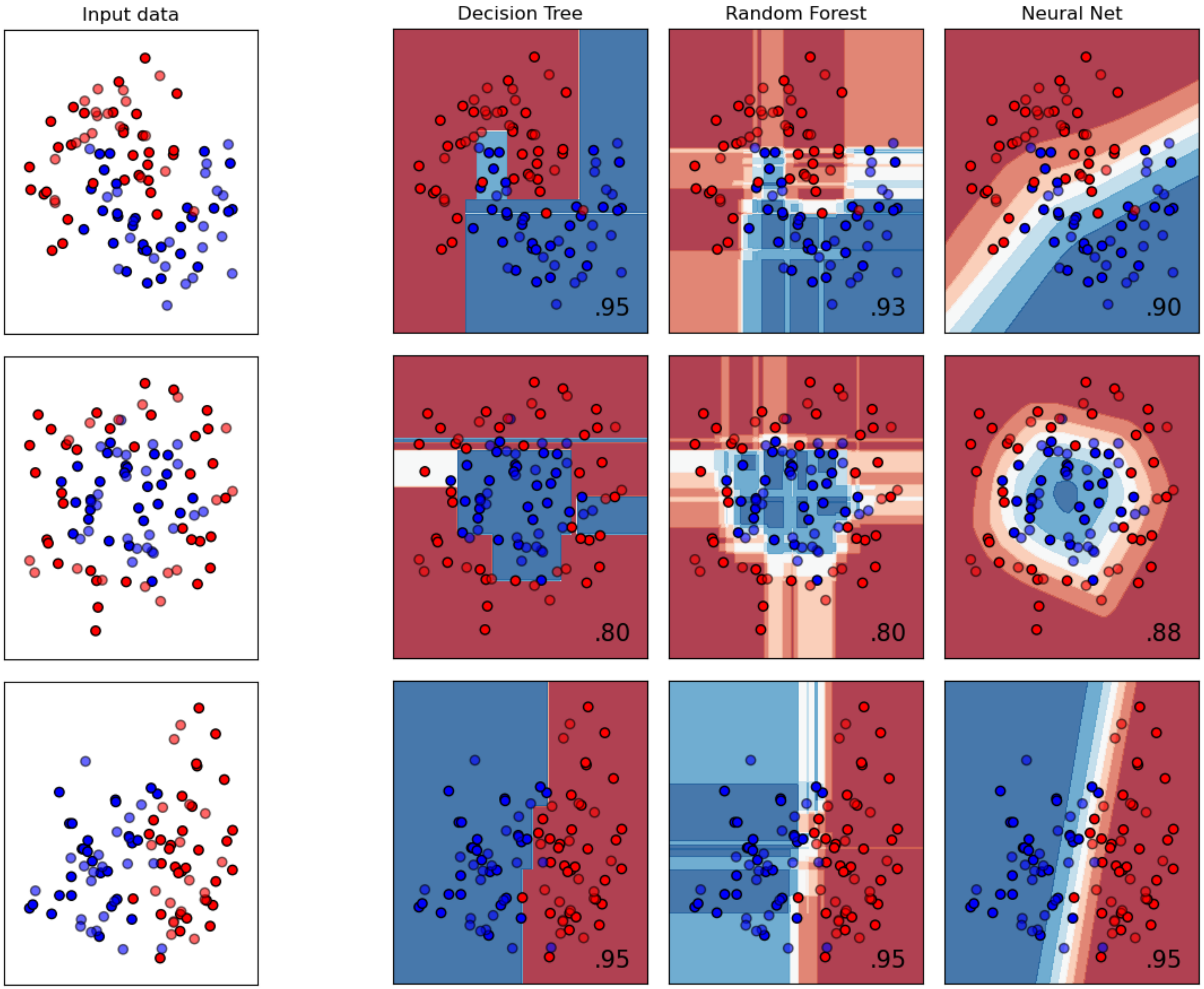
- Ensembles are collections of models
- Whose predictions are combined to obtain a final p.
- Can have much higher predictive performance



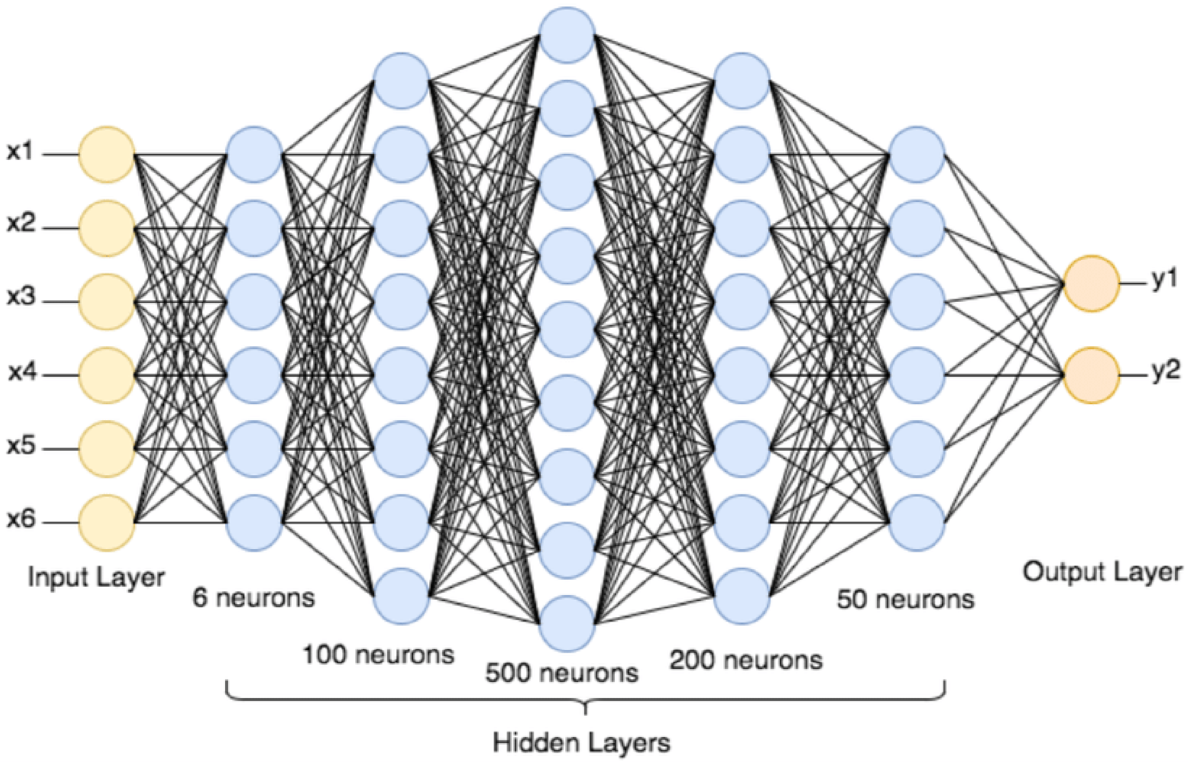
Decision Boundaries for Ensembles

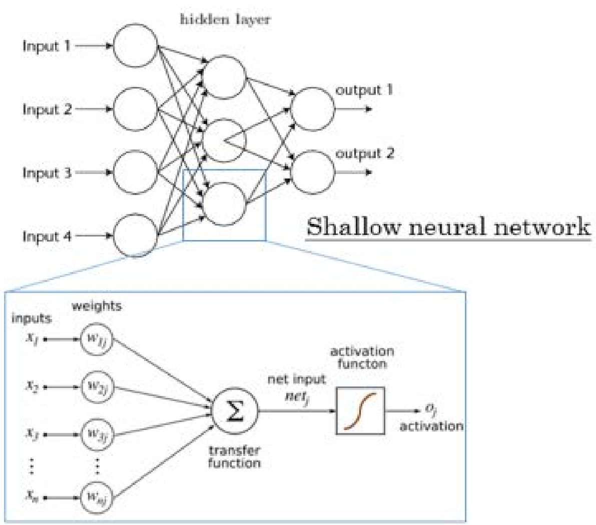


Decision Boundaries Illustrated



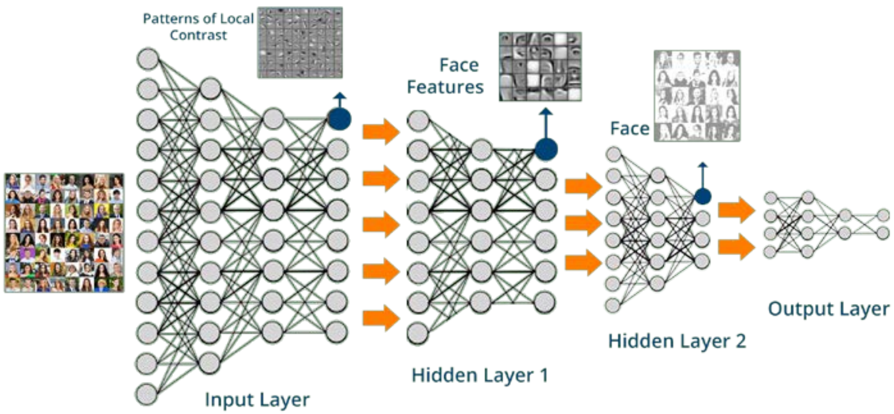
Deep and shallow neural nets: Many and few hidden layers



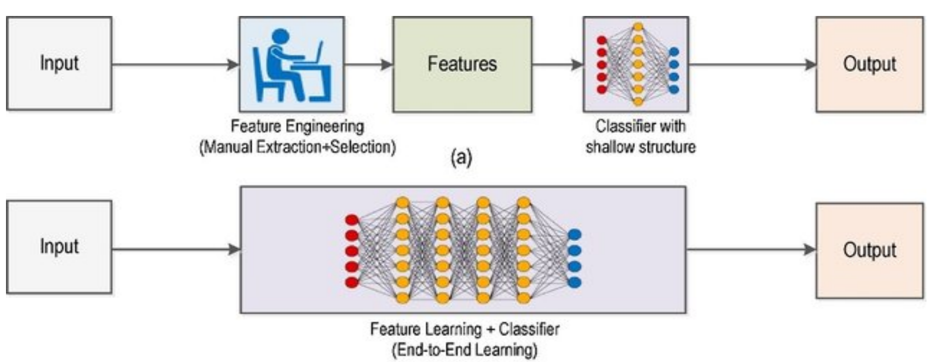


Deep Neural Networks: Advantages

CNNs are DNNs that include computer vision ideas (convolutional filters) and can learn features from images



This is the key to the success of NNs: End-to-end learning



Deep Neural Networks: Limits

But, neural networks **cannot 'explain their thinking'!**

- This is unacceptable in many areas (e.g., medicine)
- This also leaves the needs of scientists, whose very enterprise is founded on explanation, fundamentally unmet

Neural networks are data hungry

- They need lots of **labeled** training data, not easy to get

Neural networks are computationally demanding

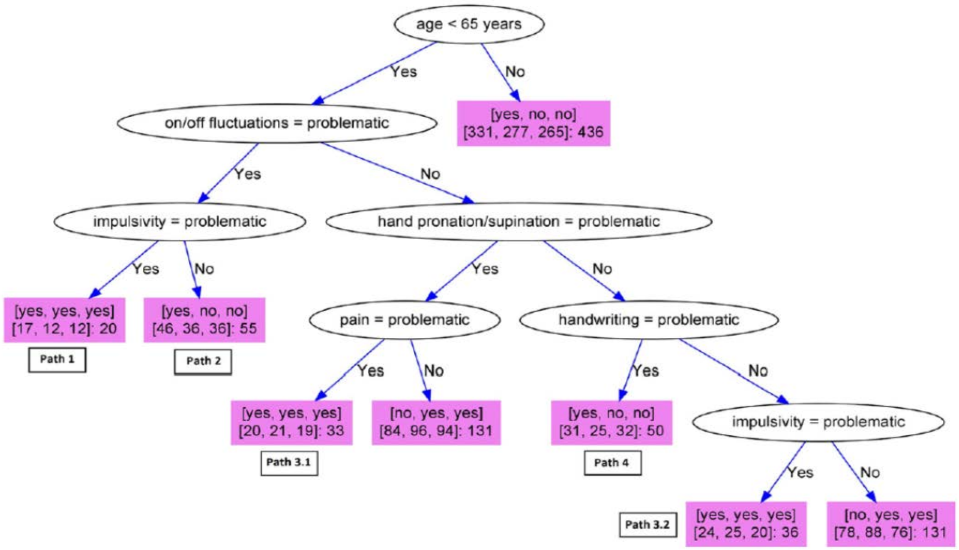
- In a recent study, DNN performance: 97% accuracy, 3300 hrs GPU, competing method 96% accuracy, 36 hrs CPU
- Electricity spent (rough estimate): 1KWh for competitor, 1 MWh for DNNs

There's more to ML than DNNs!

- Learning from data (inductive learning)
  - Unsupervised learning (clustering)
  - Supervised learning (predictive modeling)
    - Classification
    - Regression
- Learning from data (Data Mining)
  - Learning understandable/ explainable models
    - Trees & tree ensembles
    - Rules & rule ensembles
  - Learning neural networks/ Deep learning
- Reinforcement learning
- Computational scientific discovery

Explainable AI/ML

- Explaining predictions
- Explaining models
- Very important in medicine (changing therapy for Parkinson's)



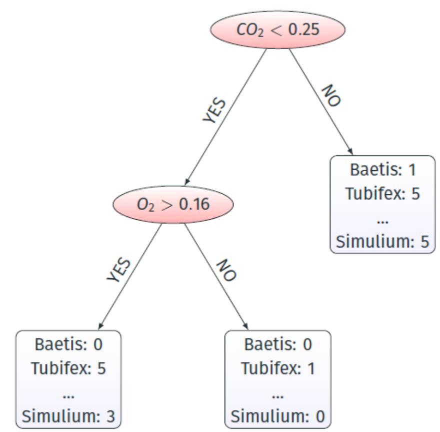
- Also very important in science

Ensembles: Accuracy/Expl. Trade-off

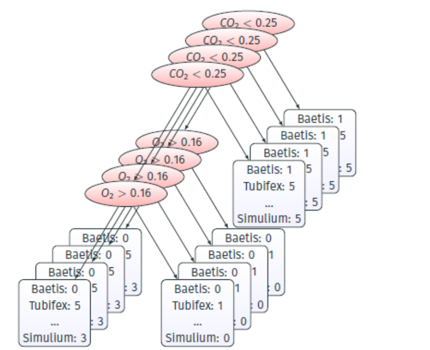
- Ensembles contain multiple-trees
- Predictions combined for better accuracy
- Provide feature importance estimates and thus some insight/ interpretation
- E.g., random forests
- Very efficient

#	Descriptive attributes				Target attributes			
	$KMnO_4$	CO_2	...	$K_2Cr_2O_7$	Baetis	Tubifex	...	Simulium
1	0.66	0.15	...	2.7	3	0	...	3
2	2.05	0.56	...	2.8	0	0	...	5
⋮	⋮	⋮		⋮	⋮	⋮		⋮
1060	1.3	1.23	...	1.1	5	3	...	1

A single decision tree

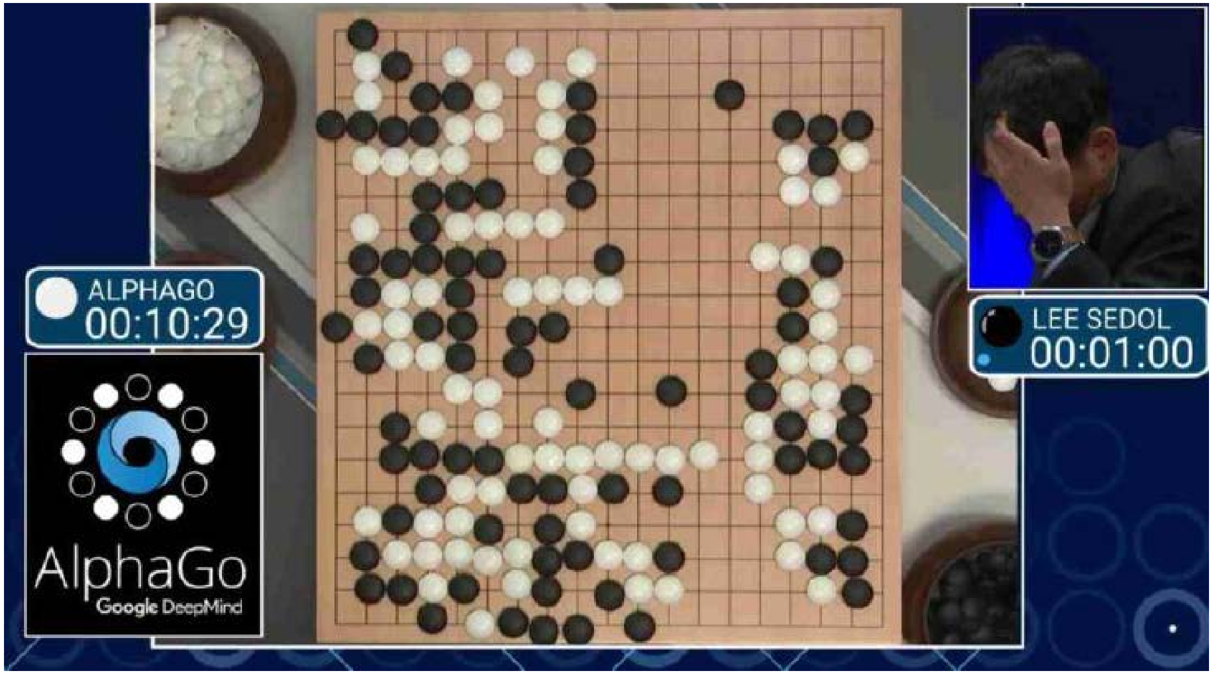


An ensemble of decision trees



Reinforcement learning

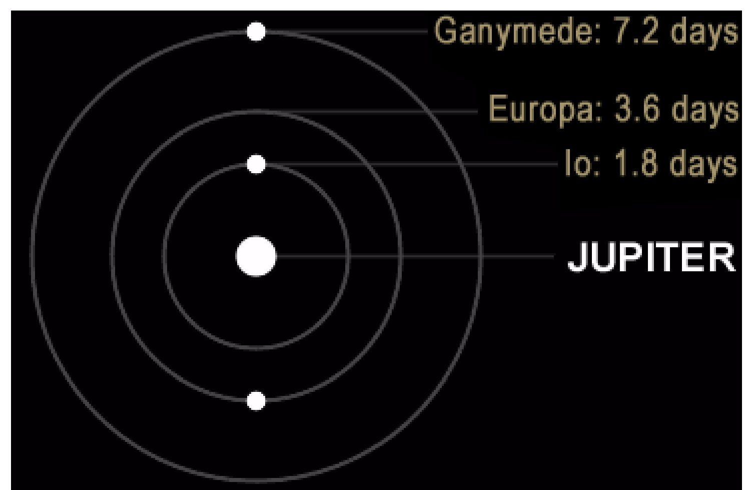
- Agent learning from interaction with the environment
  - Performs actions
  - Receives feedback/ reinforcement
- Google's AlphaGo



Computational scientific discovery: Finding laws in data

- Planets further from the Sun orbit slower, following the law $d^3/p^2 = k$ (Kepler's third law)
- Where d is the distance from the planet to the Sun, and p is period of one revolution; k is a constant
- This law is also followed by the moons of Jupiter

Moon	Moon	d	p
Io	A	5.67	1.77
Europa	B	8.67	3.57
Ganimed	C	14.00	7.16
Calysto	D	24.67	16.69



Computational scientific discovery: Finding laws in data

- Input: Observations of distances between moons and Jupiter (d) and periods of their orbits (p)

Moon	d	p
A	5.67	1.77
B	8.67	3.57
C	14.00	7.16
D	24.67	16.69

- Output: Kepler's third law $d^3/p^2 = k$
- Process of discovery
  - BACON (Langley 1978; Langley et al. 1987)
  - Carries out heuristic search through the space of numeric terms, looking for constant values and linear relations

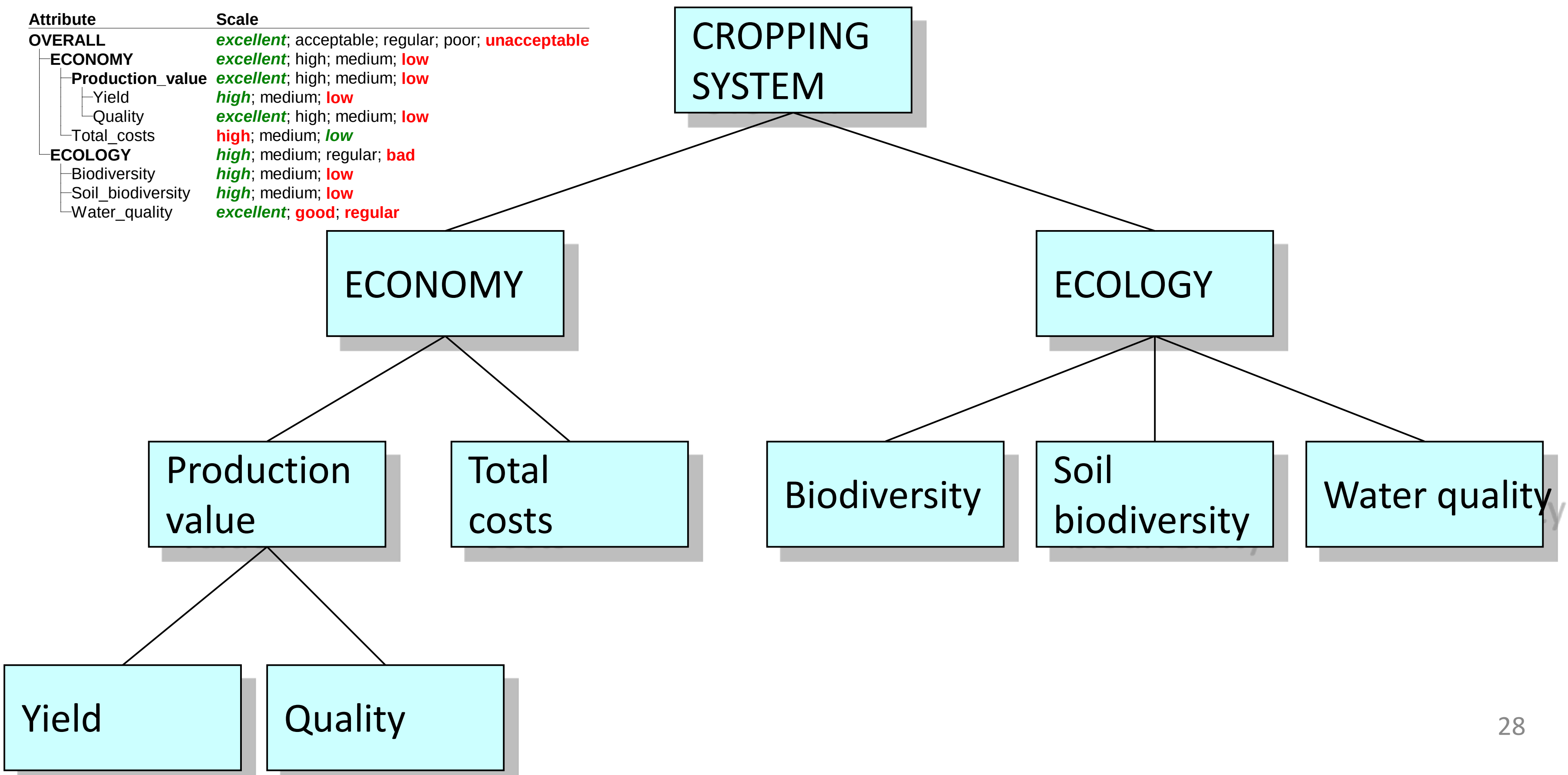
Moon	d	p	d/p	d^2/p	d^3/p^2
A	5.67	1.77	3.20	18.15	58.15
B	8.67	3.57	2.43	21.04	51.06
C	14.00	7.16	1.96	27.40	53.16
D	24.67	16.69	1.48	36.46	53.89

- Proceeds from observed variables to constant theoretical term

There’s more to AI than ML

- Knowledge representation & engineering, Reasoning
- All of the above exemplified, e.g., in decision support

Attribute	Scale
OVERALL	***excellent***; acceptable; regular; poor; **unacceptable**
ECONOMY	***excellent***; high; medium; **low**
Production_value	***excellent***; high; medium; **low**
Yield	***high***; medium; **low**
Quality	***excellent***; high; medium; **low**
Total_costs	**high**; medium; ***low***
ECOLOGY	***high***; medium; regular; **bad**
Biodiversity	***high***; medium; **low**
Soil_biodiversity	***high***; medium; **low**
Water_quality	***excellent***; **good**; **regular**



There’s more to AI than ML

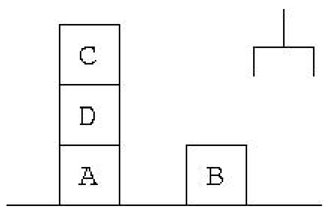
- Reasoning (e.g., theorem proving)
- Planning: Blocks world example
- Initial state

Ontable(A), Ontable(B)

On(D,A), On(C,D)

Clear(C), Clear(B)

Handempty



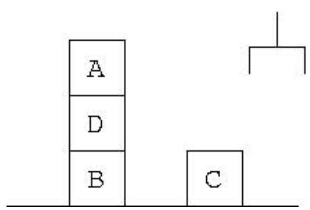
- Goal state

Ontable(B), Ontable(C)

On(D,A), On(C,D)

Clear(C), Clear(B)

Handempty



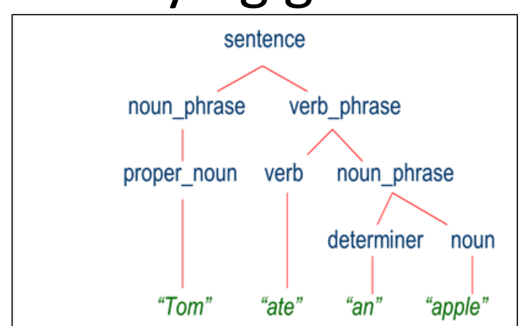
- Plan: Sequence of actions to get to goal state

UNSTACK(C,D),PUTDOWN(C),UNSTACK(D,A),STACK(D,B),PICKUP(A),STACK(A,D)

Natural Language Processing (NLP)

- Syntactic Analysis
  - **Tokenization: B**reaking up text into smaller parts/*tokens*
  - **Stop-word removal** removes frequently occuring words that don't add any semantic value: *I, they, have, like, yours*, etc.
  - **Lemmatization & stemming** consist of reducing inflected words to their base form to make them easier to analyze
  - **Part of speech tagging (PoS tagging)** labels tokens as *verb, adverb, adjective, noun*, etc.
  - **Parsing** determines the syntactic structure of text by analyzing its words based on an underlying grammar

```
sentence -> noun_phrase, verb_phrase
noun_phrase -> proper_noun
noun_phrase -> determiner, noun
verb_phrase -> verb, noun_phrase
proper_noun -> [Tom]
noun -> [apple]
verb -> [ate]
determiner -> [an]
```



Natural Language Processing (NLP)

- Semantic Analysis: focuses on capturing the meaning of text. First, it studies the meaning of each individual word (lexical semantics). Then, it looks at the combination of words and what they mean in context.
  - **Word sense disambiguation** tries to identify in which sense a word is being used in a given context.
  - **Relationship extraction** attempts to understand how entities (places, persons, organizations) relate to each other in a text.
- Classical approach to NLP: Knowledge-driven (e.g., write a grammar for Slovene)
- Currently predominant: Statistical NLP (based on corpora of text)

NLP Downstream Tasks

- Sentiment Analysis: identifies emotions in text and classifies opinions as positive, negative, or neutral
  - Opposite orientations in different application domains
    - "This camera **sucks**."
    - "This vacuum cleaner really **sucks**."
  - Sarcasm:
    - "What a **great** car! It stopped working in two days."
- Machine Translation

kdo lahko razume moj svet?
mu sam sploh lahko sledim?
hodim in iščem, sem mar zaklet?
čakam, da se zbudim?

who can understand my world?
Can I follow him at all?

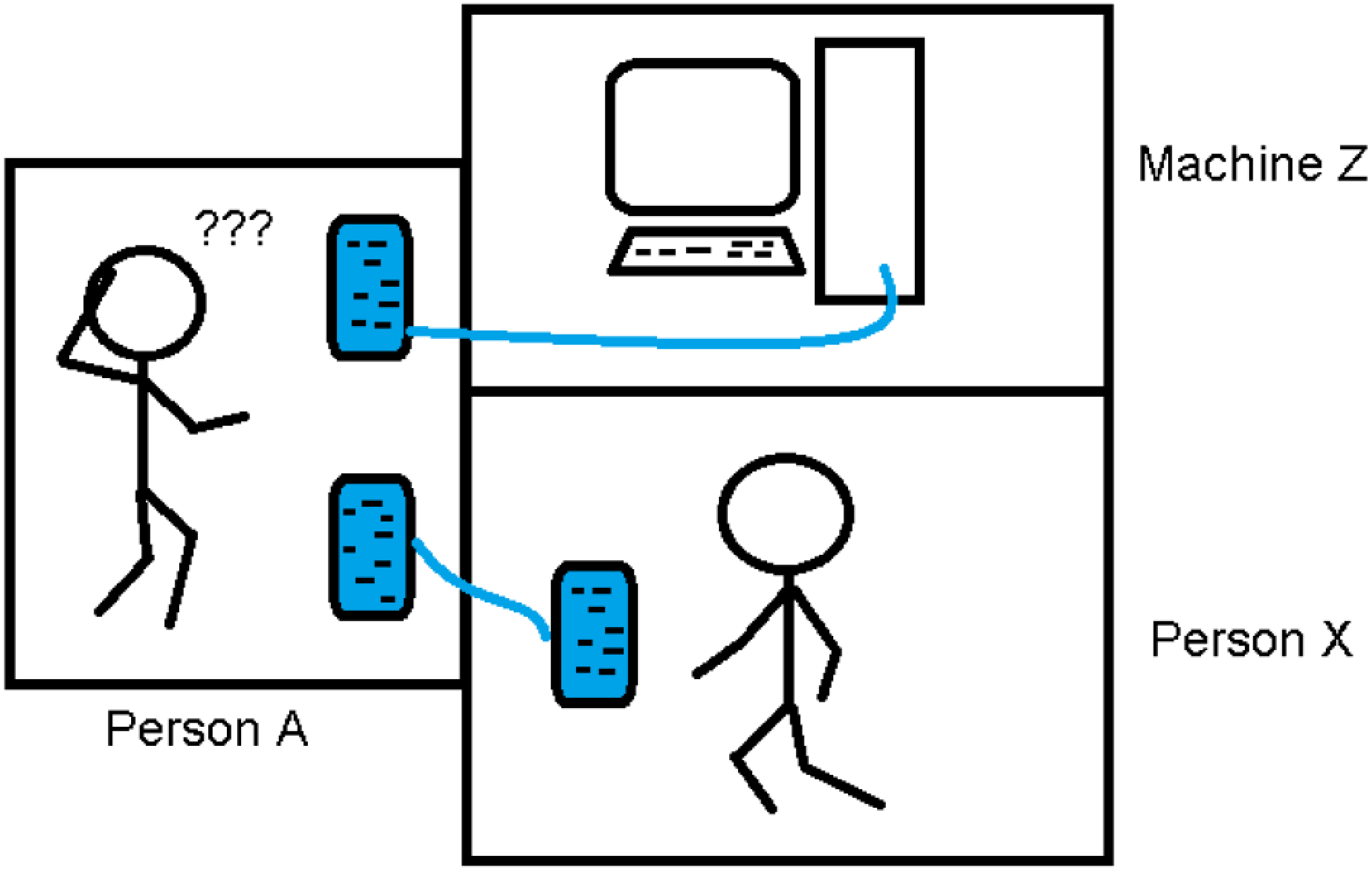
Google Translate

Who can understand me?
Myself I am lost

NLP Downstream Tasks

- Information Extraction
- Chat-bots
  - Me: Mitsuku, I want to show your abilities to the audience.
  - Mitsuku: Yep, I want to become smarter than humans.
  - Me: Do you know anything about NLP?
  - Mitsuku: Natural Language Processing.
- Topic Classification

NLP: The Turing test

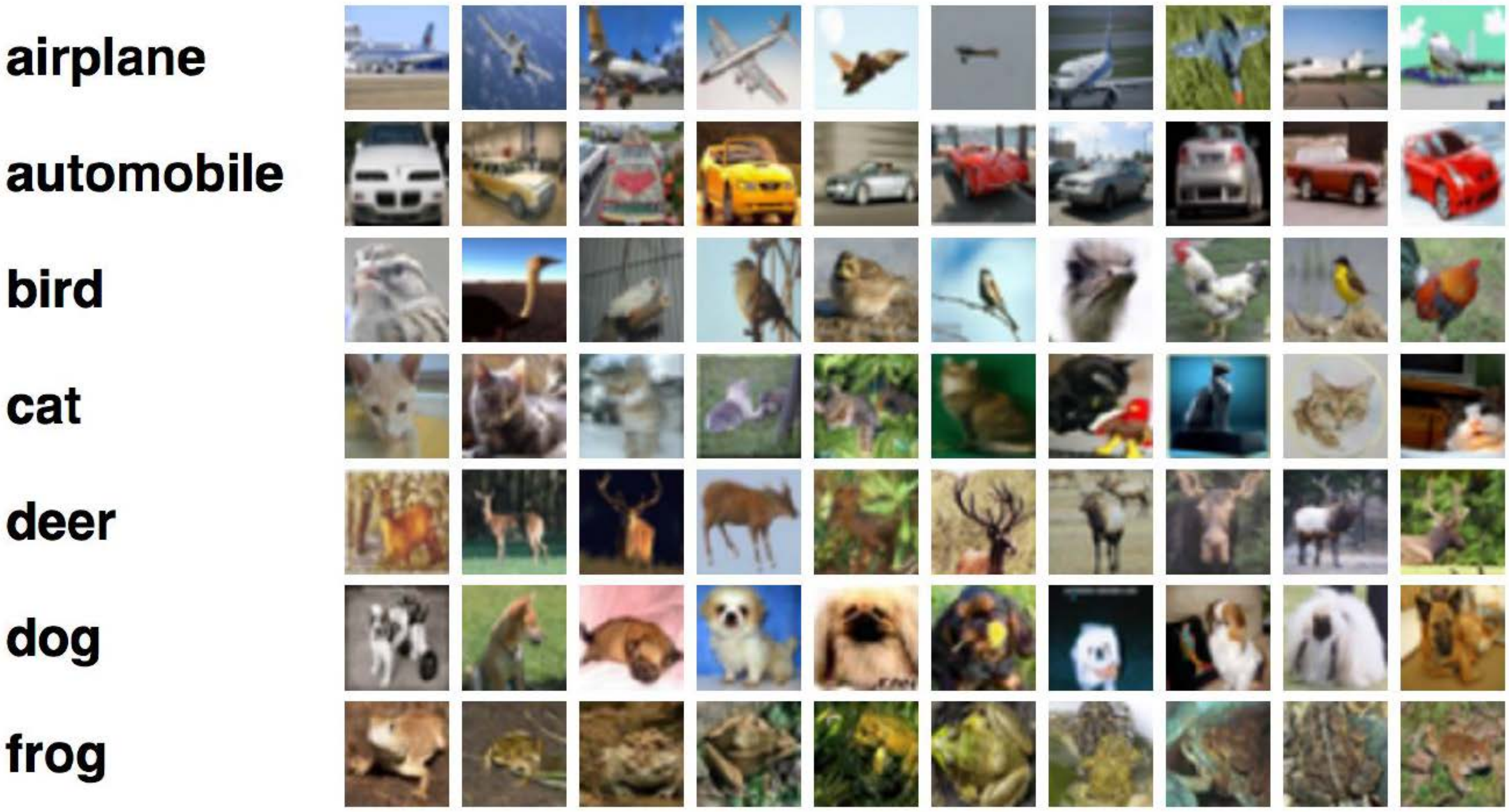


Computer Vision

- How computers can be made to gain high-level understanding from digital images or videos
- Typical tasks in computer vision
  - Recognition
  - Motion analysis
  - Scene reconstruction
  - Image restoration
- Recognition
  - Object recognition/classification
  - Identification
  - Detection
  - Content-based image retrieval

Object recognition/classification

- Typically addressed by deep/convolutional NNs today



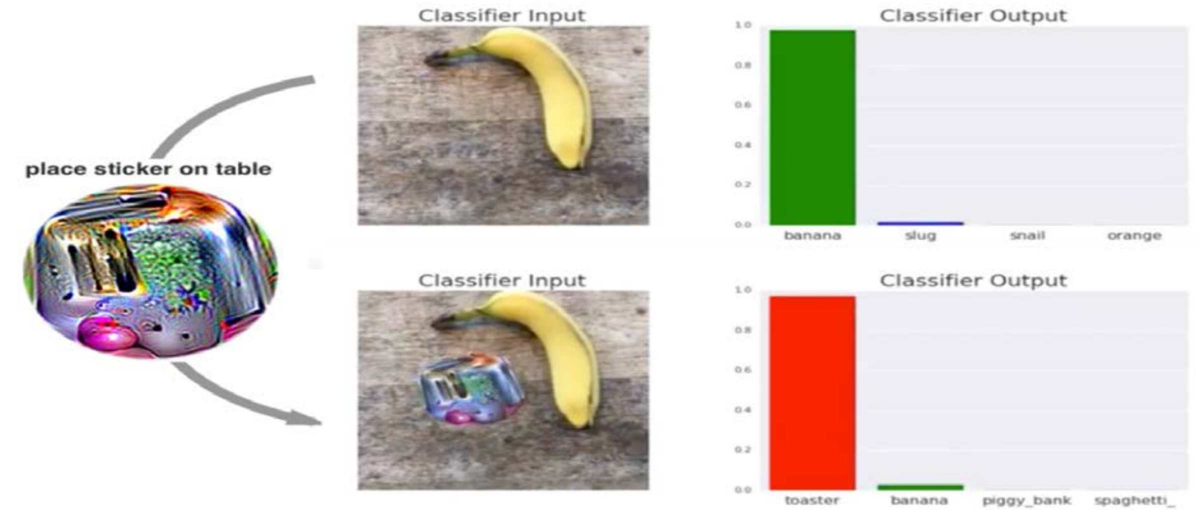
Computer Vision with DNNs

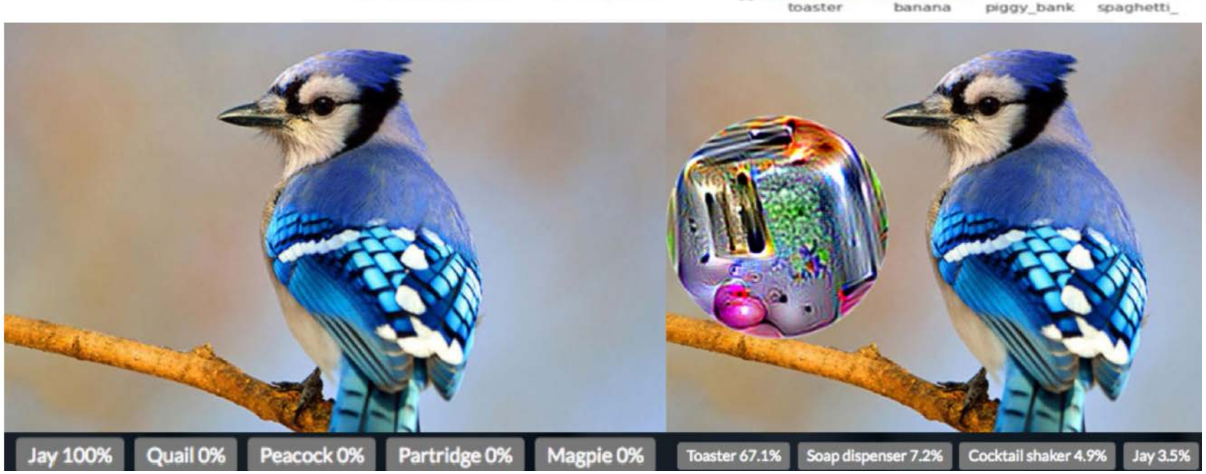
- On one hand, excellent accuracies and applications

Original Radiograph Output Image

Computer Vision with DNNs

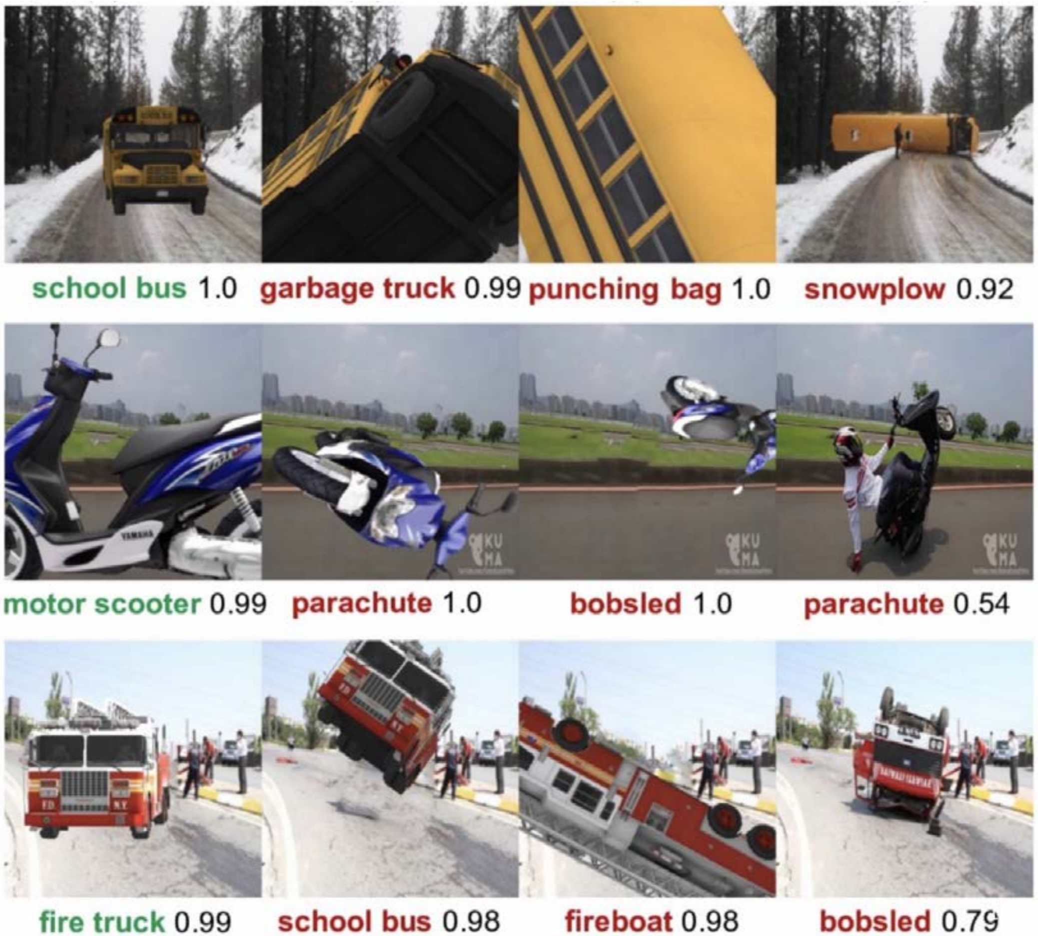
- On the other hand, brittle & vulnerable to attacks





Computer Vision with DNNs

- Failures on out-of-distribution examples



Break

- Short break
- Questions welcome!
- Before and after

Machine Learning from Big and Complex Data

AiTLAS
Artificial Intelligence Toolbox for Earth Observation

Predictive modeling (single-target classification and regression)

	Descriptive space				Target space
Example 1	1	TRUE	0.49	0.69	Yes
Example 2	2	FALSE	0.08	0.07	Yes
Example 3	1	FALSE	0.08	0.07	No
Example 4	2	TRUE	0.49	0.69	Yes
Example 5	3	TRUE	0.49	0.69	No
Example 6	4	FALSE	0.08	0.07	Yes
…	…				…

	Descriptive space				Target space
Example 1	1	TRUE	0.49	0.69	0.84
Example 2	2	FALSE	0.08	0.07	0.75
Example 3	1	FALSE	0.08	0.07	0.11
Example 4	2	TRUE	0.49	0.69	0.52
Example 5	3	TRUE	0.49	0.69	0.35
Example 6	4	FALSE	0.08	0.07	0.78
…	…				…

Big Data: Volume & Velocity

- Large number of columns (high dimensionality)
  - Need feature ranking/selection
- Large number of rows (massive data)
  - Need efficient data mining methods
- Streaming rows (data streams)
  - Need incrementality: Not all data available simultaneously
  - Data instances arrive at **high velocities**, in a **specific order** and their number is **potentially arbitrarily large**
  - The **underlying concept** (distribution) governing the data **can change (concept drift)**
  - We need **fast processing** (due to the high velocity)
  - The large and potentially infinite number of examples demands **economical management of available memory**

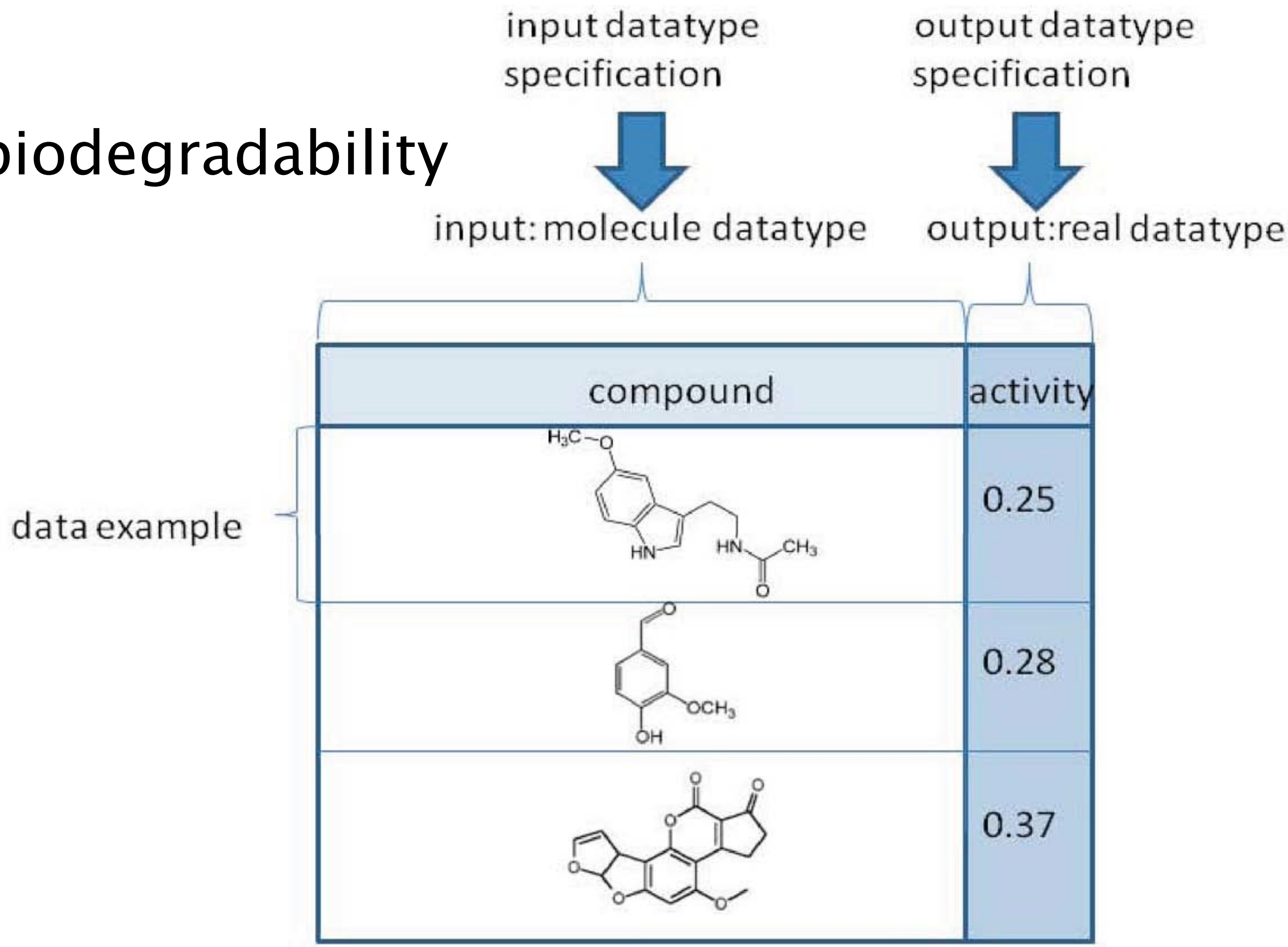
Data streams: Regression

	Descriptive space				Target space
…	…				…
Example n Example n+5	**1**	**TRUE**	**0.49**	**0.69**	**0.45**
Example n+1	4	FALSE	0.08	0.07	0.12
Example n+2	6	FALSE	0.08	0.07	1.54
Example n+3	8	TRUE	0.00	1.00	3.12
Example n+4	6	TRUE	0.00	0.00	0.05
…	…				…

Big Data: Variety - Structured Input

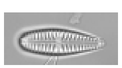
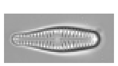
Example:

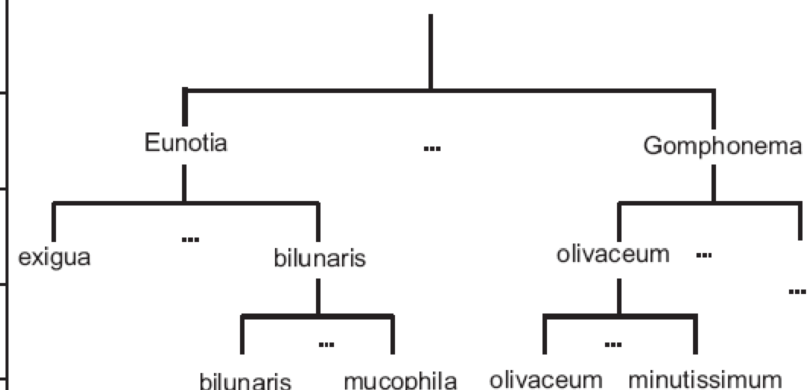
Predicting biodegradability



Big Data: Variety - Structured Output

- Example: Hierarchical classification
- Taxonomic classification of diatoms
- From microscopic images
- Taking into account the taxonomy of diatoms

image	features/descriptors						taxonomy
	Heuristic shape descriptors						
	48	24	59	66	37	…	olivaceum
	36	25	53	45	15	…	minutissimum
	35	25	56	52	19		exigua
…	…	…	…	…	…	…	…



Multi-target prediction

- Classification

	Descriptive space				Target space		
Example 1	1	TRUE	0.49	0.69	Yes	Blue	Rain
Example 2	2	FALSE	0.08	0.07	Yes	Green	Sun
Example 3	1	FALSE	0.08	0.07	Yes	Blue	Cloudy
Example 4	2	TRUE	0.49	0.69	Yes	Green	Sun
Example 5	3	TRUE	0.49	0.69	No	Blue	Sun
Example 6	4	FALSE	0.08	0.07	Yes	Red	Cloudy
…	…				…	…	…

- Regression

	Descriptive space				Target space		
Example 1	1	TRUE	0.49	0.69	0.68	0.60	3.91
Example 2	2	FALSE	0.08	0.07	0.56	0.99	7.59
Example 3	1	FALSE	0.08	0.07	0.10	1.69	7.57
Example 4	2	TRUE	0.49	0.69	0.08	0.77	8.86
Example 5	3	TRUE	0.49	0.69	0.11	3.51	2.50
Example 6	4	FALSE	0.08	0.07	0.43	2.10	8.09
…	…				…	…	…

Weather prediction

- Predicting the outlook (sunny, overcast, rain): STC
- Predicting the temperature (in degrees Celsius): STR
- Predicting the weather: MTP
  - Outlook
  - Temperature
  - Humidity
  - Quantity of precipitation …

The taxonomy of MTP tasks

Multi-target prediction

- Multi-target regression
  - Hierarchical multi-target regression
- Multi-target classification
  - Multi-label classification
    - Hierarchical multi-label classification

Few methods exist that can handle multi-target prediction with mixed targets, most focus on MTR/MLC

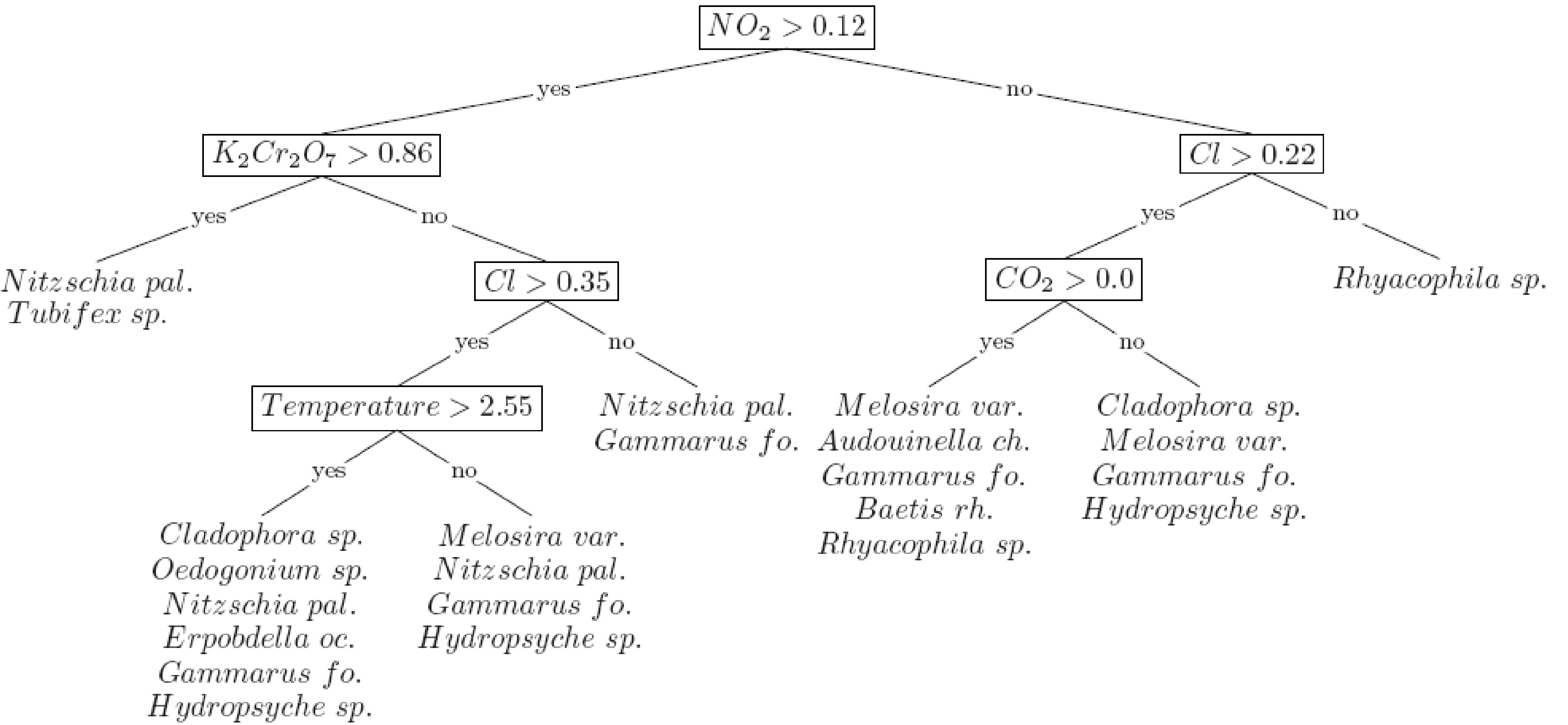
Multi-Label Classification

- Learning models that simultaneously predict several **binary** target variables
- Input: A vector of descriptive variables
- Output: A vector of several **binary** targets

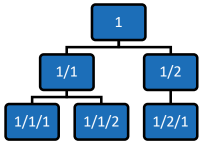
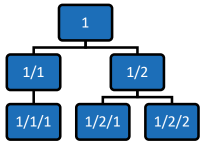
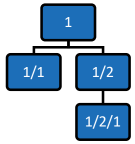
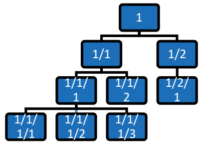
	Descriptive variables						Target variables													
Sample ID	Temperature	$K_2Cr_2O_7$	NO_2	Cl	CO_2		*Cladophora sp.*	*Gongrosira incrustans*	*Oedogonium sp.*	*Stigeoclonium tenue*	*Melosira varians*	*Nitzschia palea*	*Audouinella chalybea*	*Erpobdella octoculata*	*Gammarus fossarum*	*Baetis rhodani*	*Hydropsyche sp.*	*Rhyacophila sp.*	*Simulim sp.*	*Tubifex sp.*
ID1	0.66	0.00	0.40	1.46	0.84	...	1	0	0	0	0	1	1	0	1	1	1	1	1	1
ID2	2.03	0.16	0.35	1.74	0.71	...	0	1	0	1	1	1	1	0	1	1	1	1	1	0
ID3	3.25	0.70	0.46	0.78	0.71	...	1	1	0	0	1	0	1	0	1	1	1	0	1	1

Multi-Label Classification Example

- A decision tree for multi-label classification

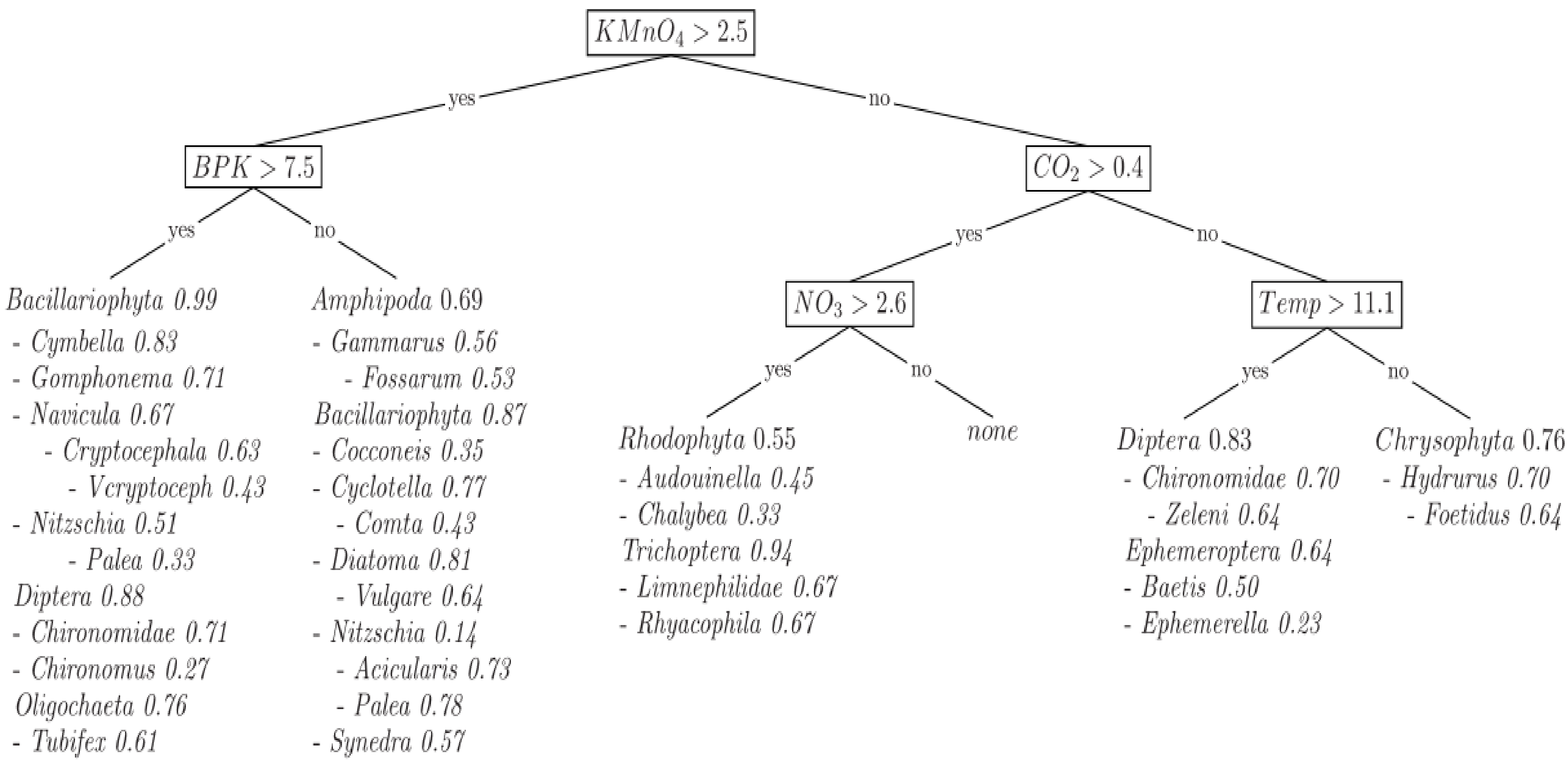


Hierarchical multi-label classification

	Descriptive space				Target space
Example 1	1	TRUE	0.49	0.69	
Example 2	2	FALSE	0.08	0.07	
Example 3	1	FALSE	0.08	0.07	
Example 4	2	TRUE	0.49	0.69	
…	…				…

A decision tree for HMLC

Taking into account the taxonomy of living organisms



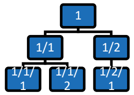
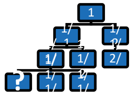
Semi-supervised learning: Classification and regression

	Descriptive space				Target space
Example 1	1	TRUE	0.49	0.69	Yes
Example 2	2	FALSE	0.08	0.07	**?**
Example 3	1	FALSE	0.08	0.07	**?**
Example 4	2	TRUE	0.49	0.69	Yes
Example 5	3	TRUE	0.49	0.69	No
Example 6	4	FALSE	0.08	0.07	**?**
…	…				…

	Descriptive space				Target space
Example 1	1	TRUE	0.49	0.69	0.84
Example 2	2	FALSE	0.08	0.07	**?**
Example 3	1	FALSE	0.08	0.07	0.11
Example 4	2	TRUE	0.49	0.69	**?**
Example 5	3	TRUE	0.49	0.69	**?**
Example 6	4	FALSE	0.08	0.07	0.78
…	…				…

SOP+SSL Incomplete Annotations

- Some examples have labels, some don’t, some partial

	Descriptive space				Target space
Example 1	1	TRUE	0.49	0.69	
Example 2	2	FALSE	0.08	0.07	**?**
Example 3	1	FALSE	0.08	0.07	**?**
Example 4	2	TRUE	0.49	0.69	

	Descriptive space				Target space		
Example 1	1	TRUE	0.49	0.69	?	0.60	3.91
Example 2	2	FALSE	0.08	0.07	0.56	0.99	7.59
Example 3	1	FALSE	0.08	0.07	?	?	?
Example 4	2	TRUE	0.49	0.69	0.08	0.77	8.86
Example 5	3	TRUE	0.49	0.69	0.11	?	?
Example 6	4	FALSE	0.08	0.07	0.43	2.10	8.09

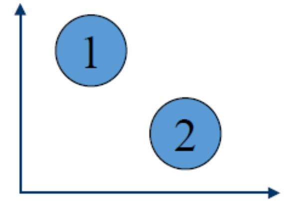
ML from Complex Data by Top-down induction of PCTs (Predictive Clustering Trees)

To construct a tree T from a training set S:

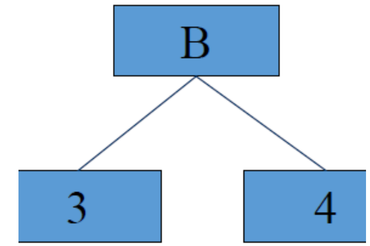
- If **the examples in S have low variance**,
  construct a leaf labeled *target(prototype(S))*
- Otherwise:
  - Select the best attribute A with values v1, …, vn, which **reduces the most the variance** (*measured according to a given distance function d*)
  - Partition S into S1, …, Sn according to A
  - Recursively construct subtrees T1 to Tn for S1 to Sn
  - Result: a tree with root A and subtrees T1, …, Tn

Learning PCTs

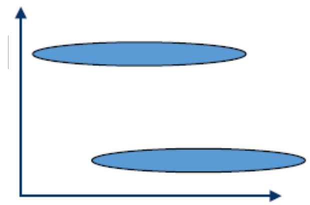
Clustering:



Prediction:

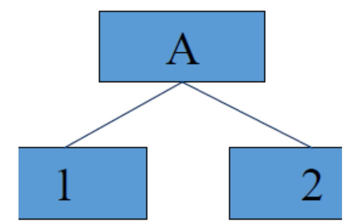


- Recursively partition data set into subsets (cluste
  with low intra-cluster variance
  - Variance = avg. squared distance to prototype

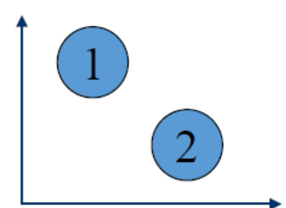


$$ICV(S) = \sum_{y_j \in S} d(y_j, p(S))^2$$

Predictive clustering:



- For the variance, the distance is measured
  - In standard clustering, along all dimensions
  - In prediction, along a single target dimension
  - In predictive clustering, along a structured target, e.g., several target dimensions



Selecting the best test in a PCT

- Select the test that maximizes variance reduction
- Calculated in line 4

procedure BestTest(E)

$$Var(E) = \sum_{i=1}^{T} Var(Y_i).$$

1: $(t^*, h^*, \mathcal{P}^*) = (none, 0, \emptyset)$
2: **for each** possible test t **do**
3: $\quad \mathcal{P}$ = partition induced by t on E
4: $\quad h = Var(E) - \sum_{E_i \in \mathcal{P}} \frac{|E_i|}{|E|} Var(E_i)$
5: $\quad$ **if** $(h > h^*) \land$ Acceptable$(t, \mathcal{P})$ **then**
6: $\quad\quad (t^*, h^*, \mathcal{P}^*) = (t, h, \mathcal{P})$
7: **return** $(t^*, h^*, \mathcal{P}^*)$

Semi-Supervised Learning w. PCTs

- New definition of variance that includes both targets and attributes, e.g., for MTR

$$Var(E) = \frac{1}{T + D} \cdot \left(w \cdot \sum_{i=1}^{T} Var(Y_i) + (1 - w) \cdot \sum_{j=1}^{D} Var(X_j) \right)$$

- T = #target attributes, D = #descriptive attributes
- $E = E_{\mathrm{Labeled}} \cup E_{\mathrm{Unlabeled}}$
- Variances only calculated for non-missing values

PCTs for ML from Big & Complex D.

- Different tasks of structured output prediction
  - MTR
  - MLC/HMLC
- Learning from data streams
- Semi-supervised Learning

- Tree ensembles
- Feature ranking
- CLUS SW available at http://source.ijs.si/ktclus/clus-public

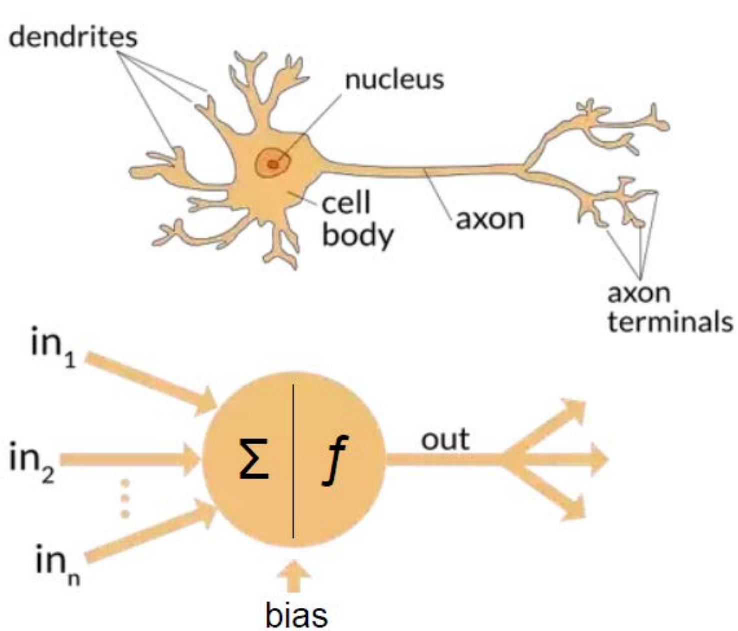
A Gentle Introduction to Neural Networks

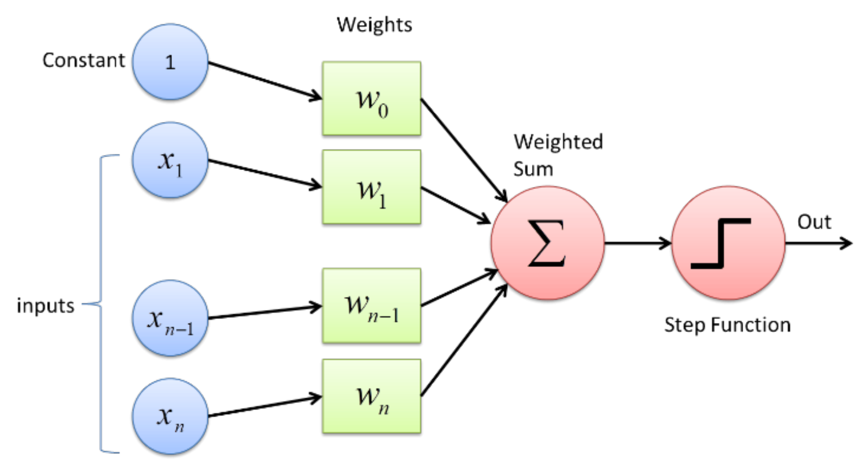
AiTLAS

Artificial Intelligence Toolbox for Earth Observation

The basic building block of NNs: Artificial neuron = Perceptron

- The perceptron can be viewed as an artificial neuron, the basic building block of artificial neural networks



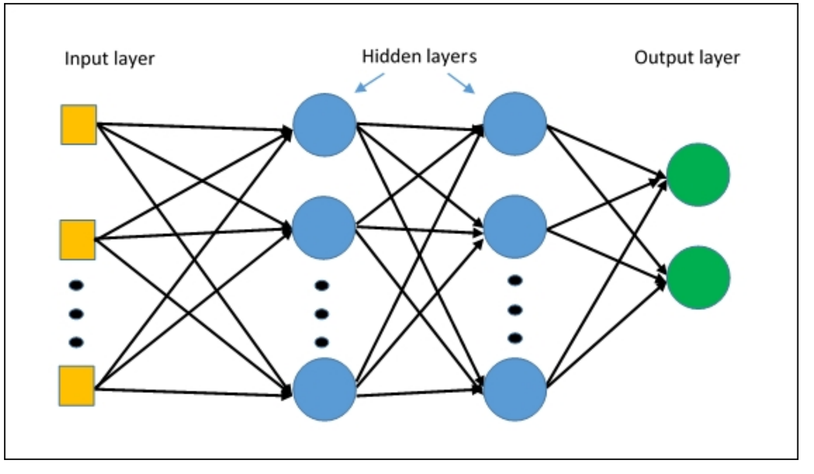


- Performs binary classification via linear discrimination

The multi-layer perceptron & NNs

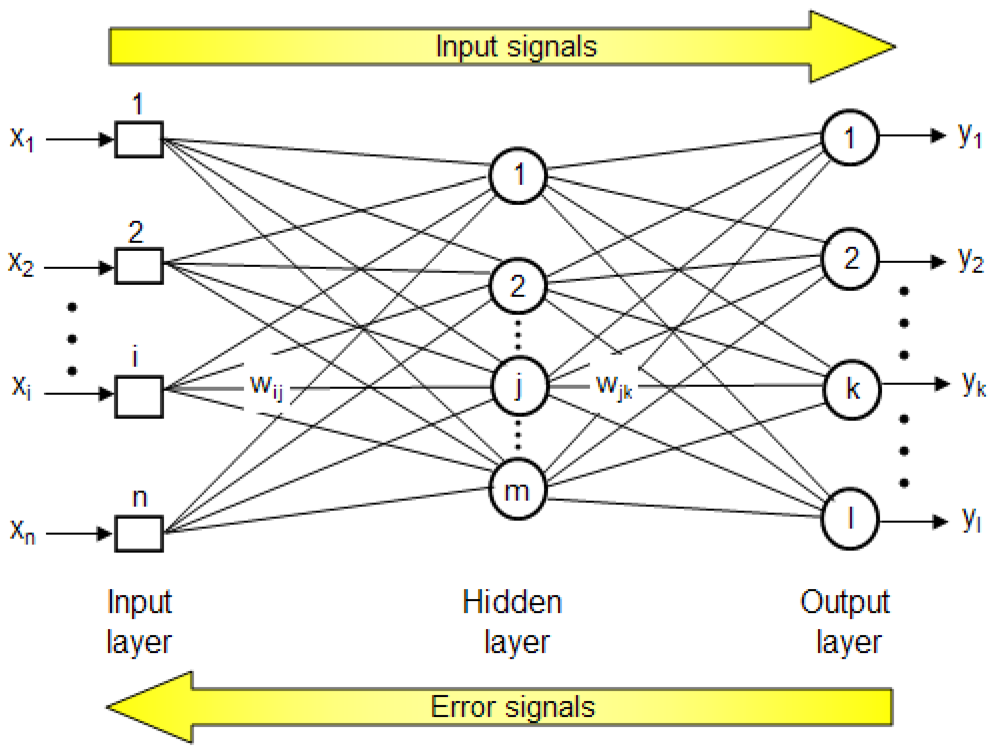
- A multi-layer perceptron has
  - An Input layer
  - An Output layer
  - And several Hidden layers of neurons
- Can approximate arbitrary non-linear mapping
- Arbitrary depth or width may be needed
- Note the hidden layers:

  If only a few, shallow NNs

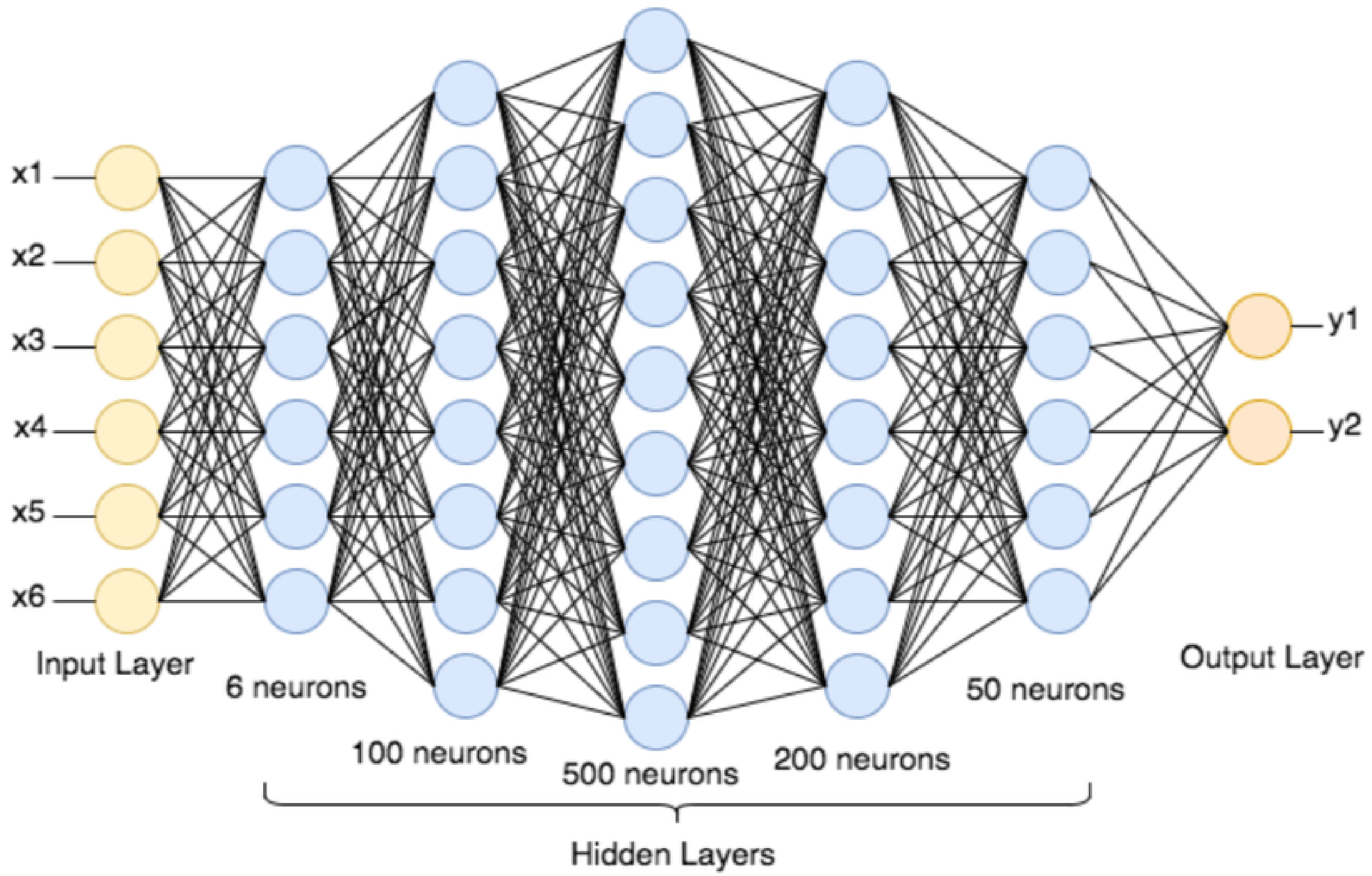


Feed-forward NNs & Backpropagation

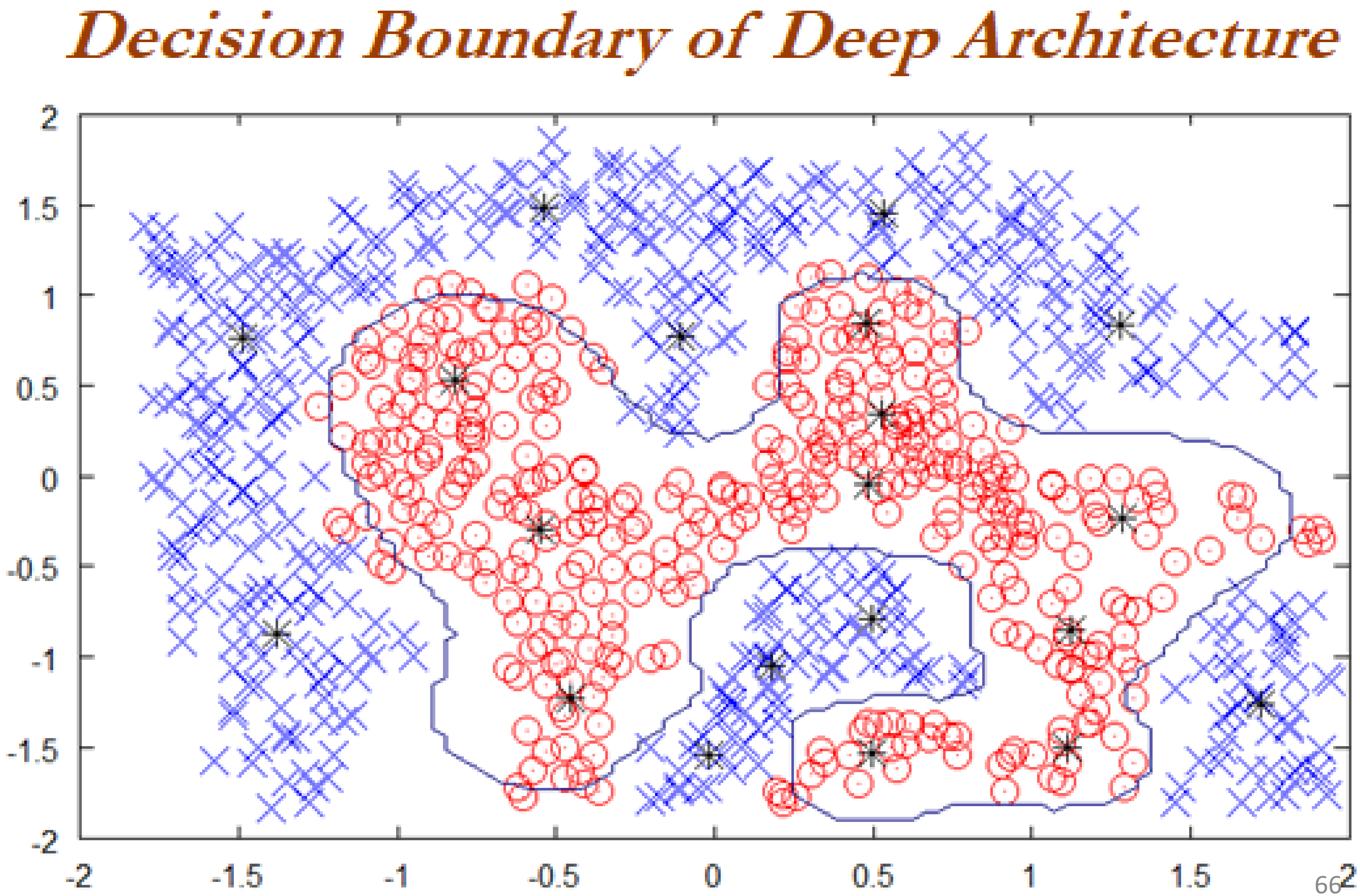
- Reasoning/making predictions: Feed-forward
- Learning the weights in an ANN: Back-propagation



Deep neural nets: Many hidden layers

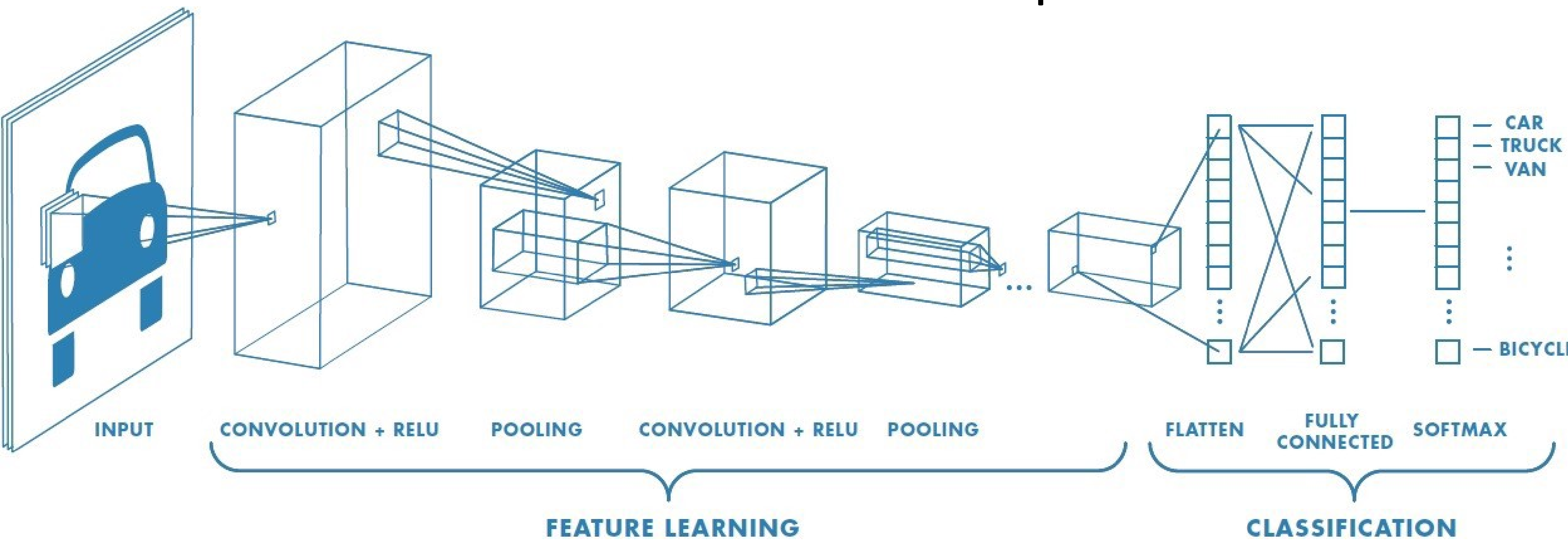


Decision boundaries of DNNs



Convolutional Neural Nets (CNNs)

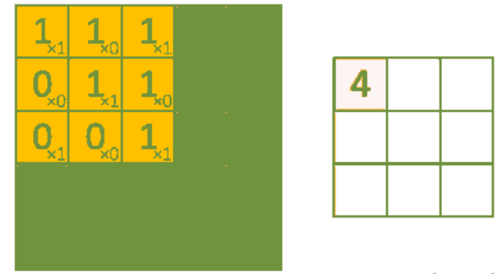
- Combine ideas from NNs and computer vision



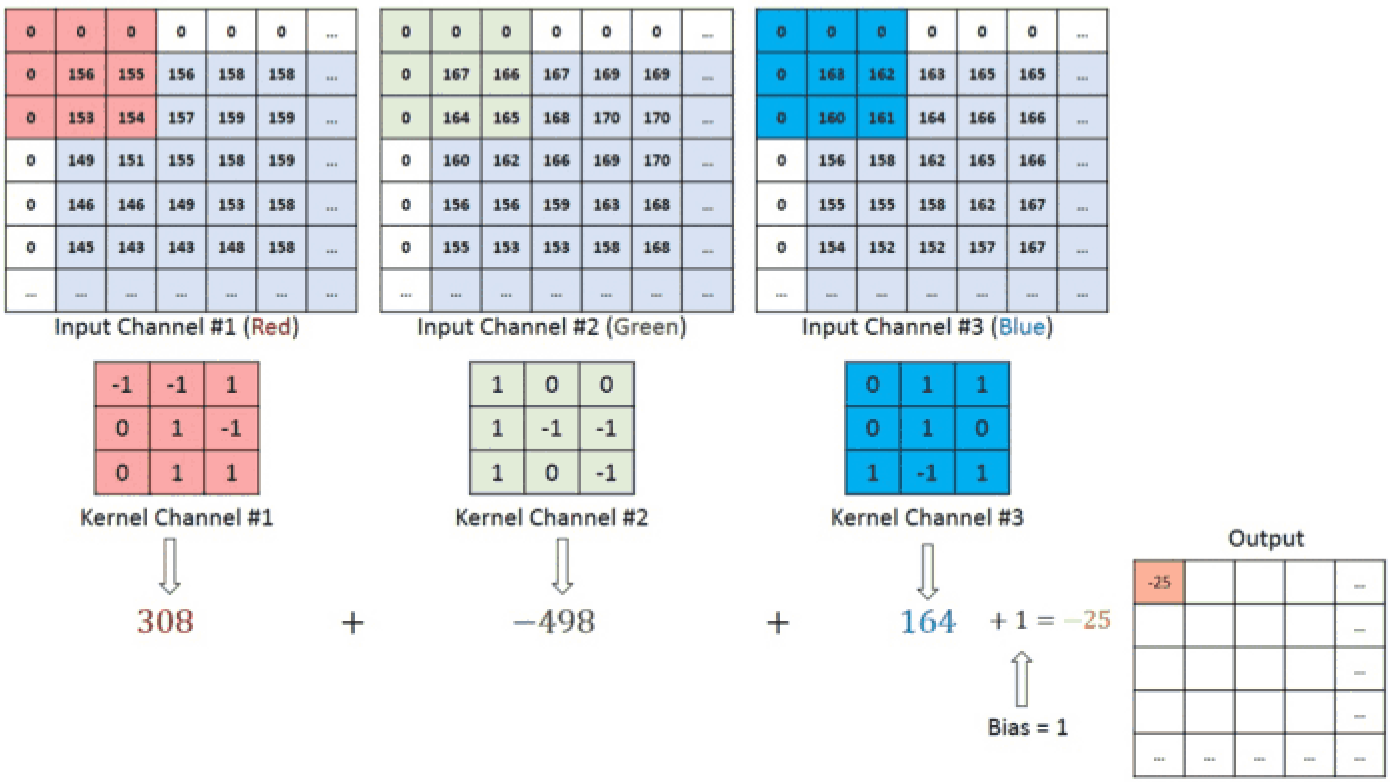
- Convolutions are filtering/smoothing operations in which a kernel/filter is applied to an image

CNN Ingredients

- Convolution with filter K= $\begin{matrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{matrix}$

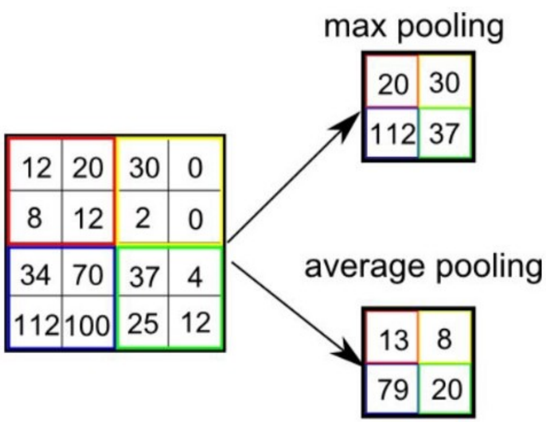


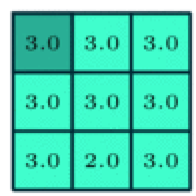
- Convolution over an RGB (3-channel) image

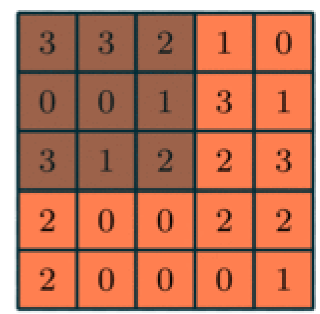


CNN ingredients: Pooling, AFs

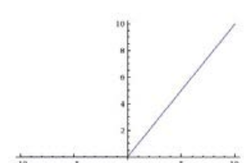
- Convolution extracts high-level features, e.g., edges
- Like convolution, pooling reduces dimensionality by aggregation

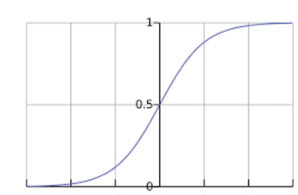




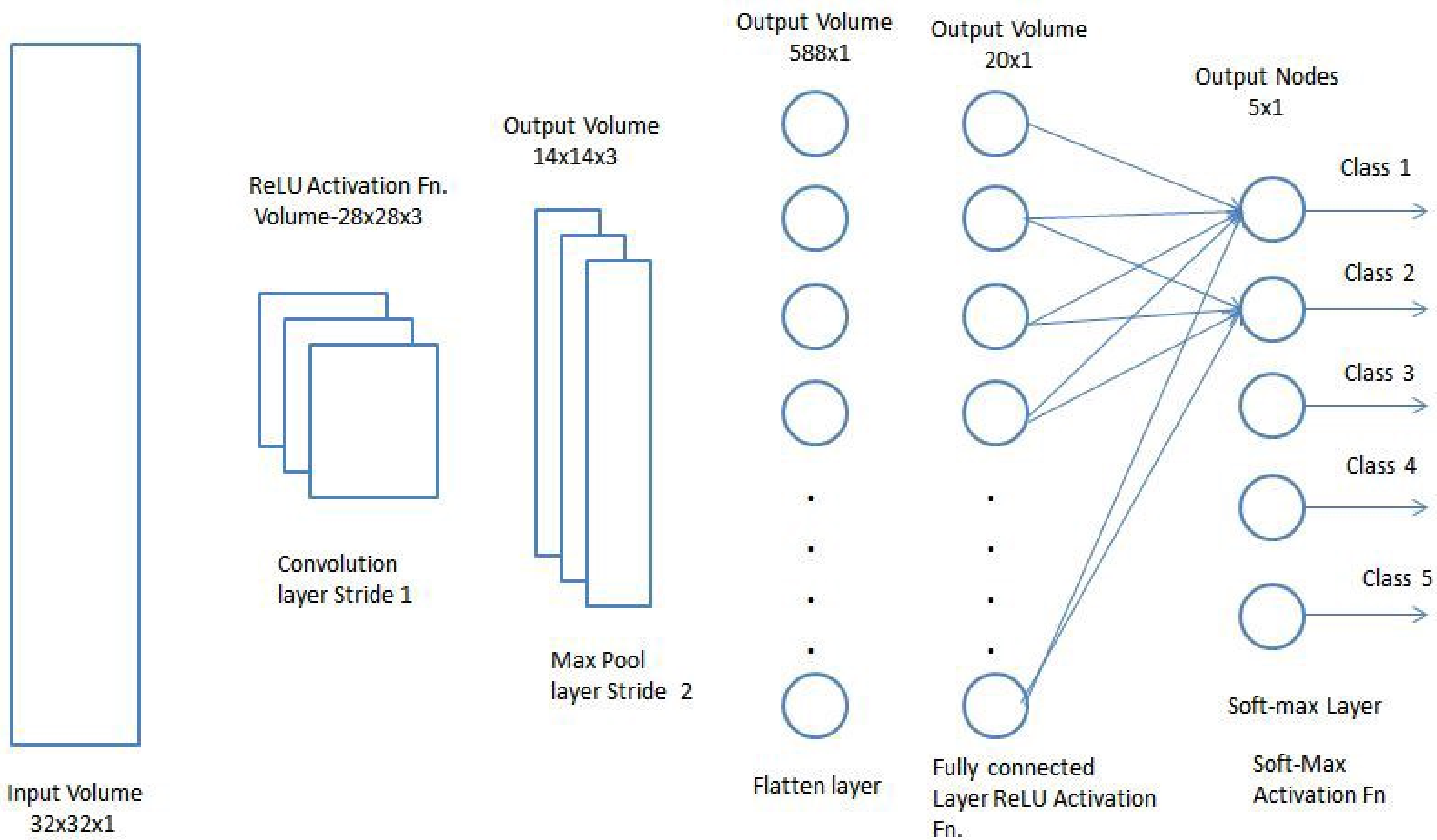


- Contributes to feature learning
- Activation functions: RELU (rectified linear unit), SoftMax





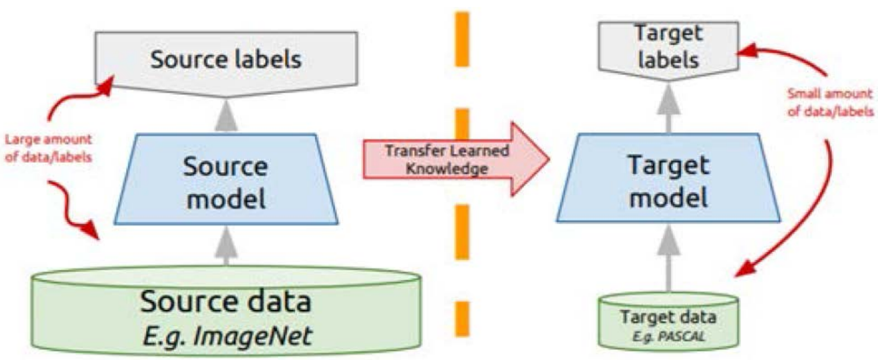
CNN Ingredients: Classification

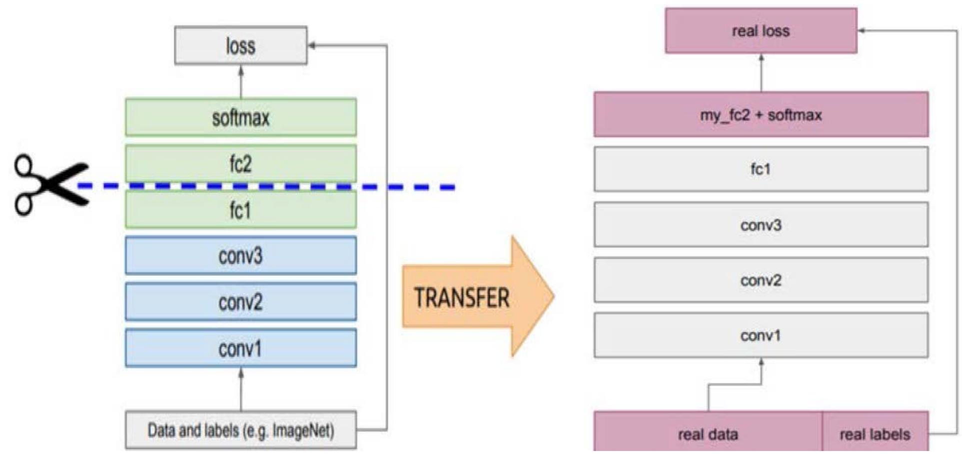


Transfer Learning with CNNs

- To combat scarcity of labeled data, learn in one domain (where enough labeled data available), then transfer the learned knowledge to another

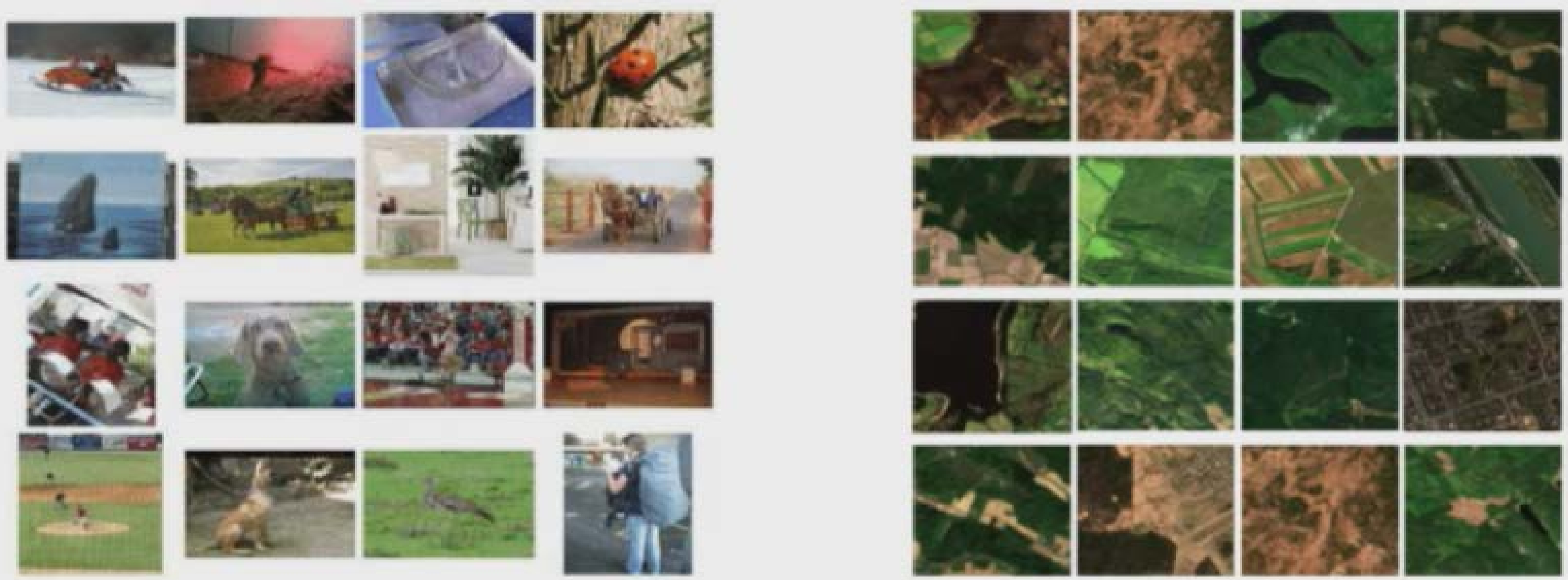
- Fine-tuning pre-trained DNNs
- Cut-off final layers of pre-trained net
- Retrain final layers with labeled data





Transfer Learning with CNNs

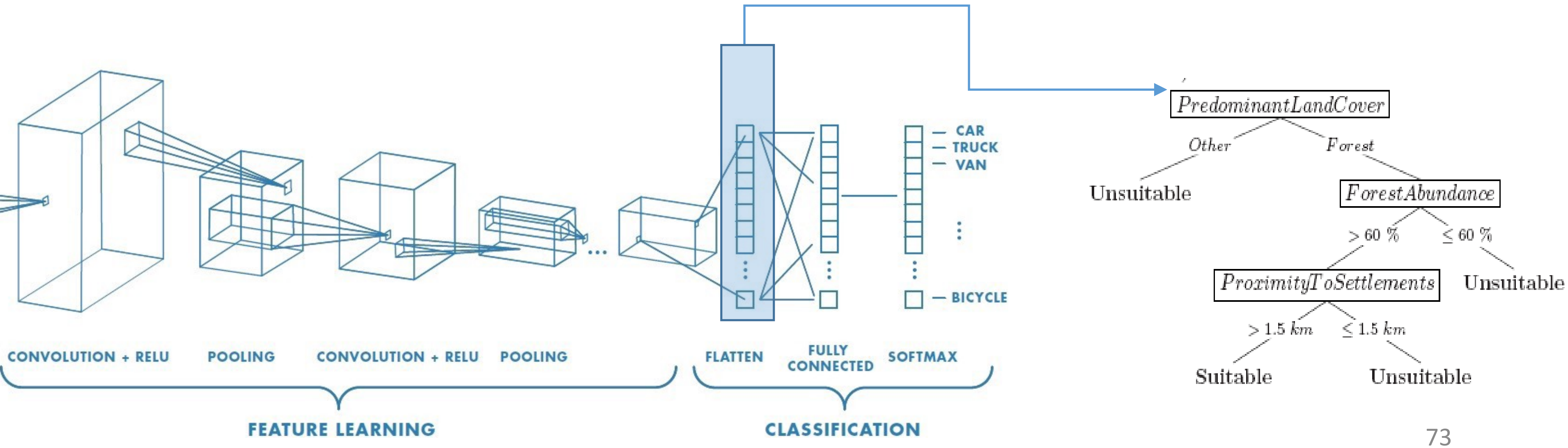
- Typicall pre-train on ImageNet, fine-tune on labeled sets of satellite images



Problem: Differences on the characteristics of images between computer vision and RS.

Combining CNNs and PCTs

- Use features extracted by the convolutional layers, e.g., the input to the classification layers
- Apply the learned CNN, say to a set of multi-labeled satellite images, generating features
- From features & labels, learn PCT



Break

- Short break
- Questions welcome!
- Before and after

Applications of Machine Learning in Forestry and Environmental Sciences

Estimating / Evaluating the State of Ecosystems

AiTLAS

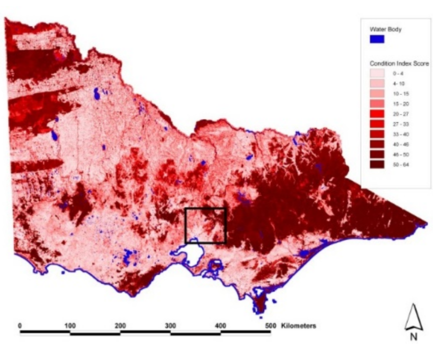
Artificial Intelligence Toolbox for Earth Observation

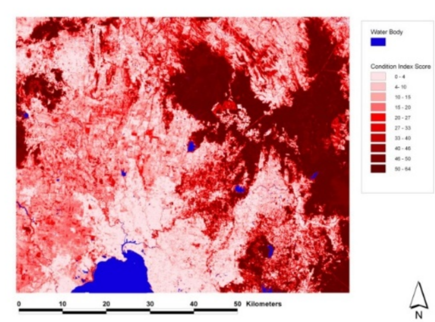
Australia: Remnants of indigenous vegetation

- 16967 sites
- 40 independent variables
    - Climate Variables (annual mean rainfall, temperature, evaporation)
    - Radiometric Data
    - Tree Density Data
    - Digital elevation model
    - Vegetation Type
- 7 dependent variables (Habitat Hectares scores)
    - Large Trees, Tree Canopy Cover, Understorey strata, Lack of weeds, Recruitment, Organic Litter, Logs

Australia: Remnants of indigenous vegetation

- $NativeTreeProb > 31$
  - yes: $LandCover\ in\ \{1, 2, 6, 7, 8, 9\}$
    - yes: $GrassProb1Ha\_RegionStdDev > 3.386$
      - yes: 1.50, 1.10, 4.02, 2.15, 0.49, 2.97, 1.58
      - no: 0.22, 0.18, 0.70, 0.43, 0.10, 0.61, 0.31
    - no: 2.70, 1.98, 7.47, 3.00, 1.20, 5.37, 2.80
  - no: $TreeProb100ha\_RegionMean > 7$
    - yes: $TempRange > 234$
      - yes: 4.50, 4.13, 10.97, 4.00, 2.37, 8.94, 4.08
      - no: 2.64, 3.24, 13.47, 4.12, 2.35, 7.63, 5.46
    - no: 5.01, 3.97, 13.56, 4.33, 4.10, 14.30, 5.35

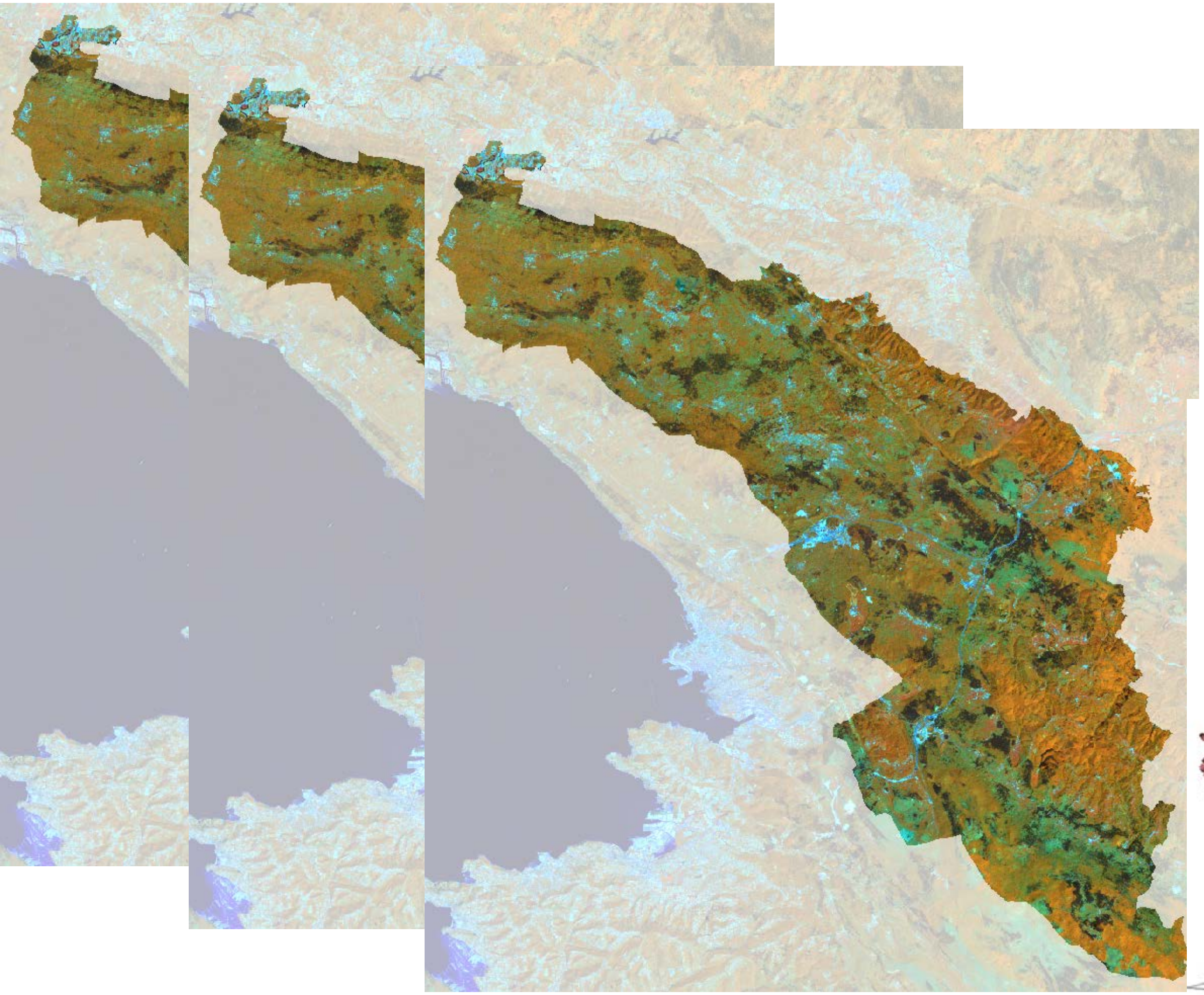




Estimating the State of Forests in Slovenia from RS Data

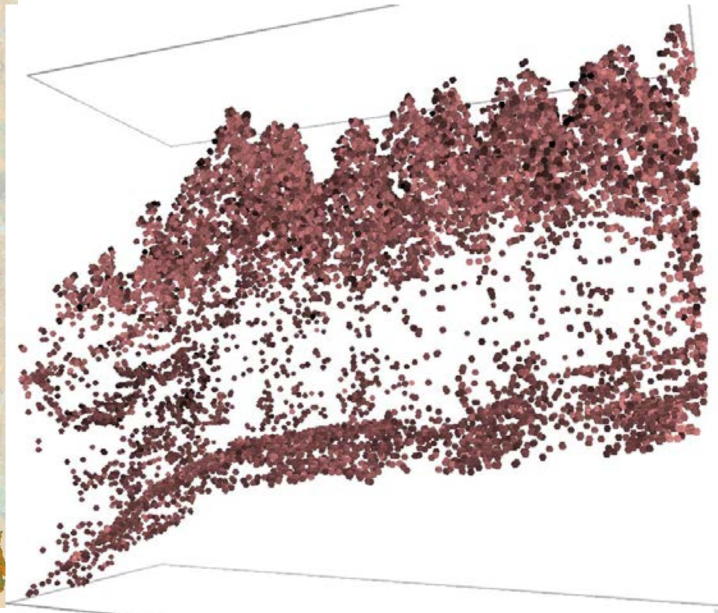
AiTLAS
Artificial Intelligence Toolbox for Earth Observation

From satellite images + LIDAR to forest height and density



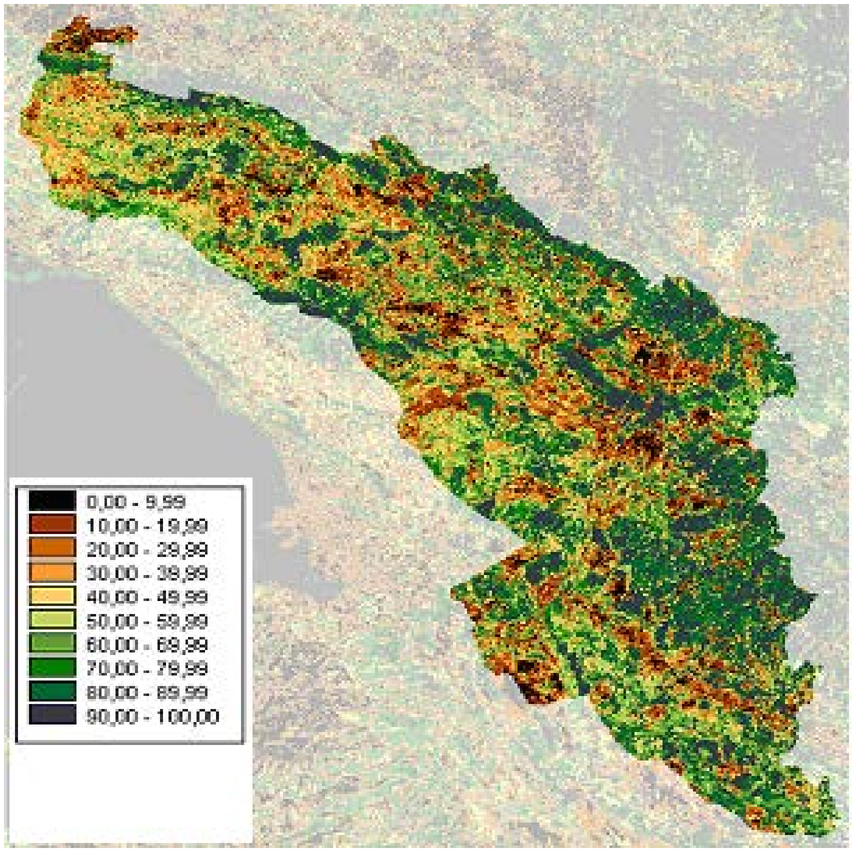
Input: Landsat images,
Multi-temporal, Multi-spectral

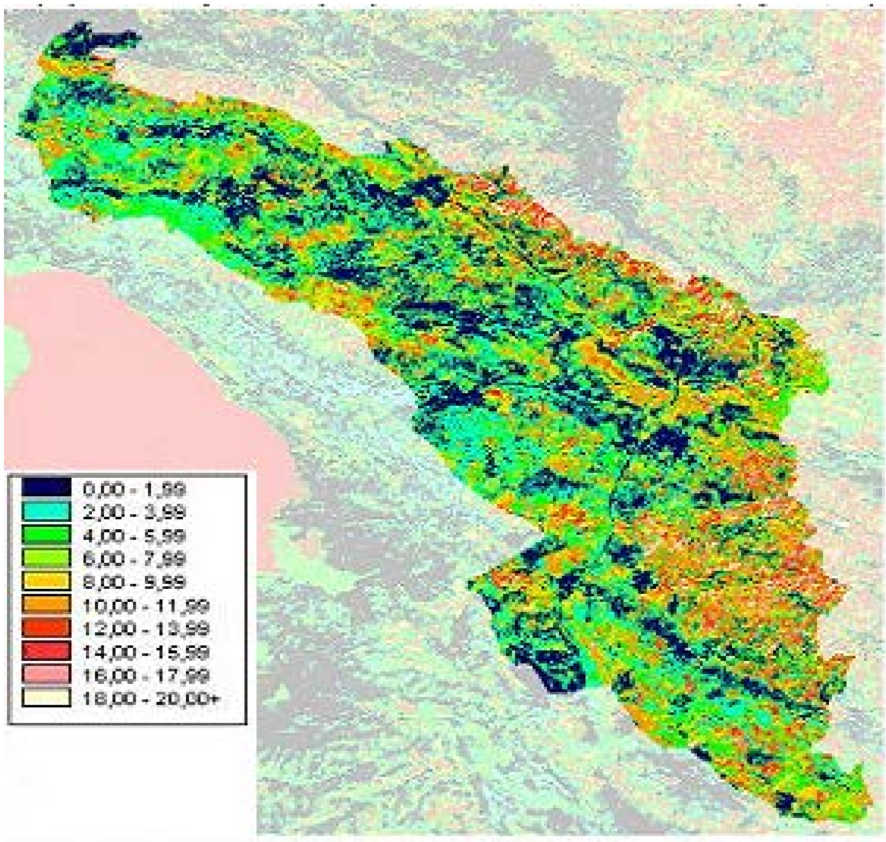
+ LIDAR for a small patch



Predict forest height & density

- For the Karst region in Slovenia

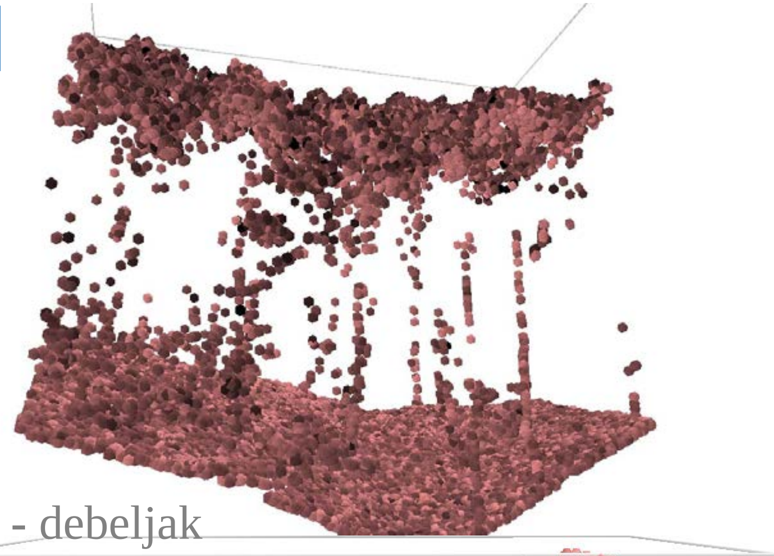




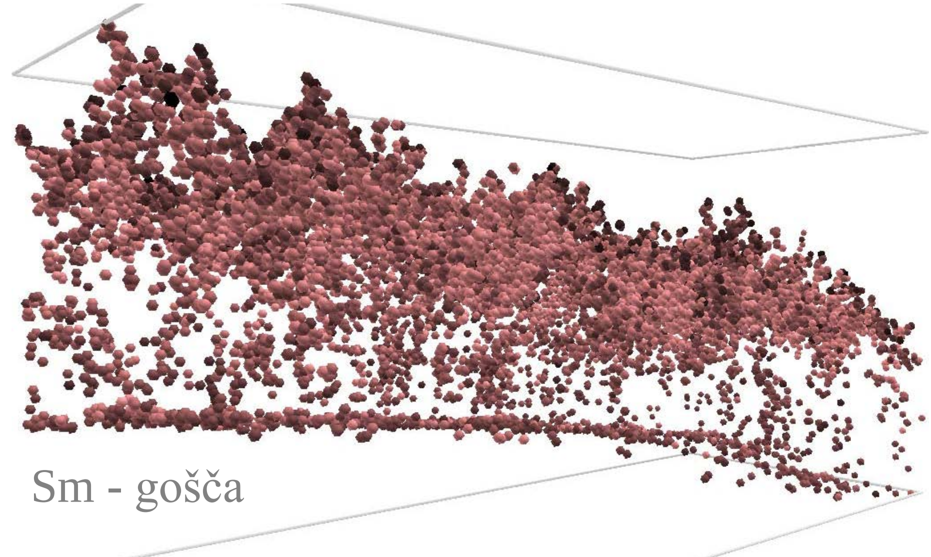
Assessing the state of Slovenian forests

- Most important two variables
  - Canopy cover/closure: Percentage of bare ground within a square covered by the vertical projection of overlaying vegetation
  - Forest stand height: Relative height of vegetation above bare ground
  - But also a complete vertical vegetation profile (total 10 vars)
- All the target variables derived from LIDAR data
- All the attributes derived from satellite images

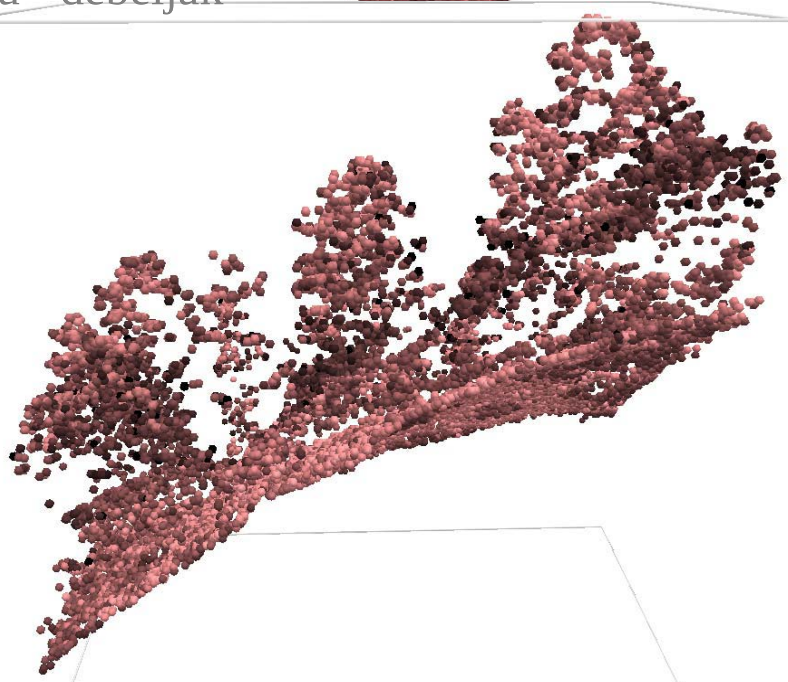
Input: LIDAR data – A sample

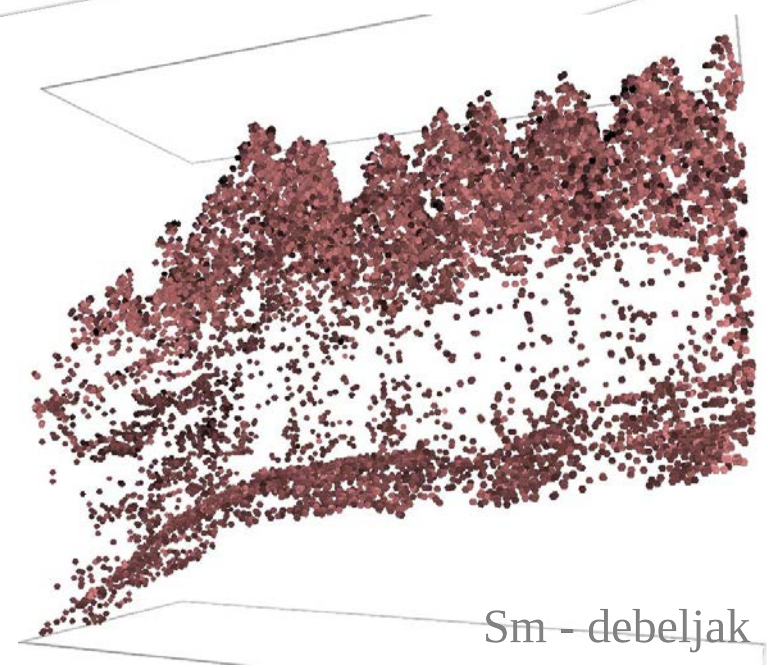


Bu - debeljak



Sm - gošča





Sm - debeljak

Processing the LIDAR data

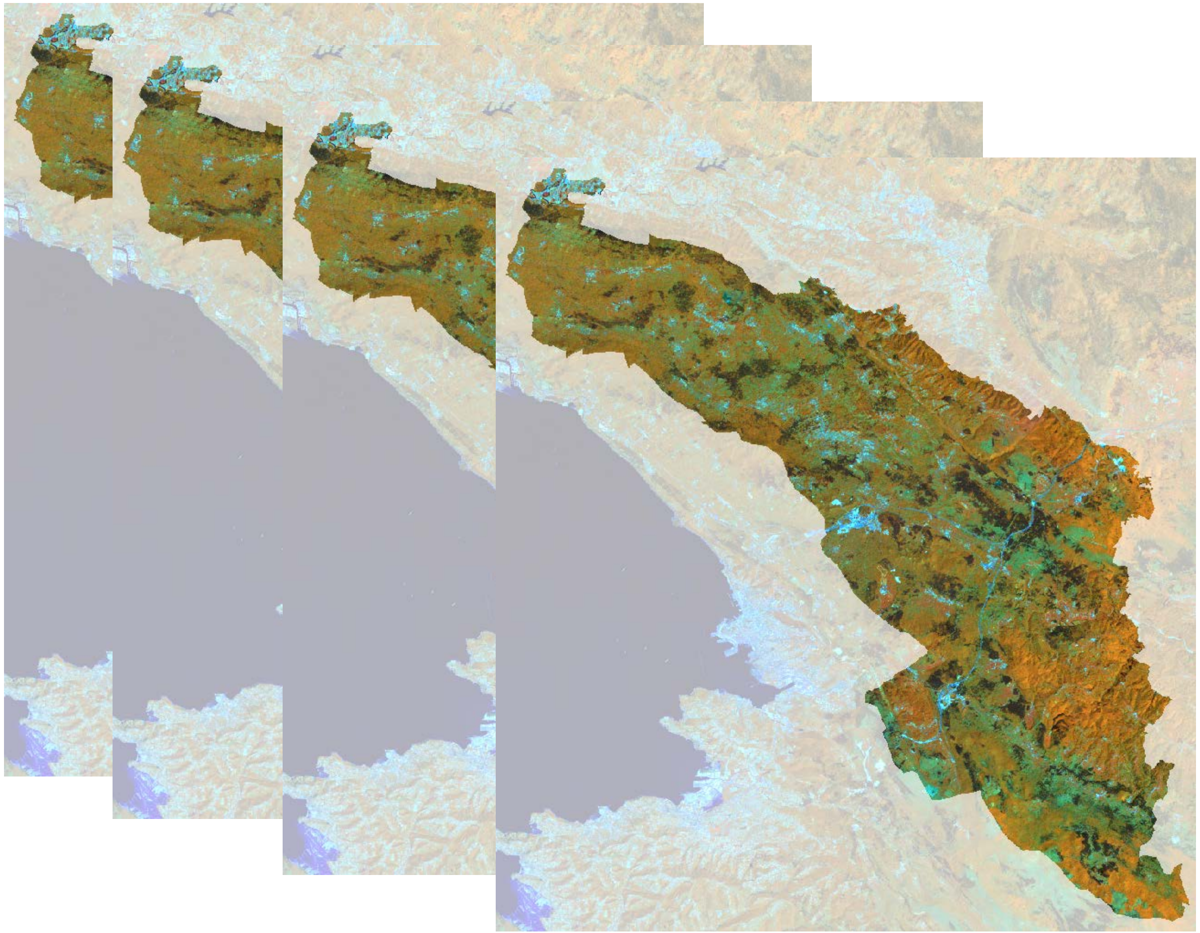
Calculating 11 target variables from LIDAR data

- Canopy cover/closure
  - Percentage of bare ground within a square covered by the vertical projection of overlaying vegetation
- Forest stand height
  - Relative height of vegetation above bare ground
- Percent of vegetation (higher than 1m)
  - Alternative to canopy closure]
- Vertical vegetation profile
  - Max. height of vegetation
  - 99% percentile
  - 95%, 75%, 50%, 25%, 10%, 5%

LIDAR Data: Area of Recording

Input: Satellite images (LandSat)

- Multi-spectral, multi-temporal



Preprocessing LandSat Images

- Each image was segmented at two levels of spatial detail (avg. 4ha and 20ha)
  - 4 image segment statistics were calculated
  - Max / Min / Avg Reflectance
  - Standard deviation of Reflectance

- For each level and each of the five image channels (2,3,4,5,7)

- 160 explanatory variables derived for modeling (attributed back to individual image pixels)

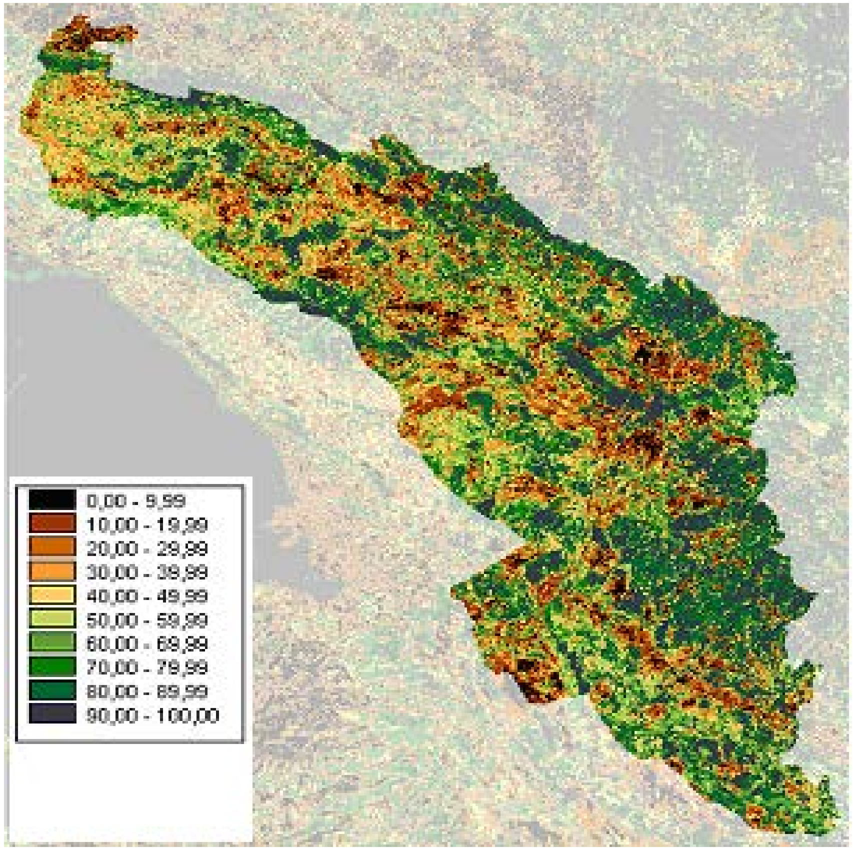
Using MTR ensembles

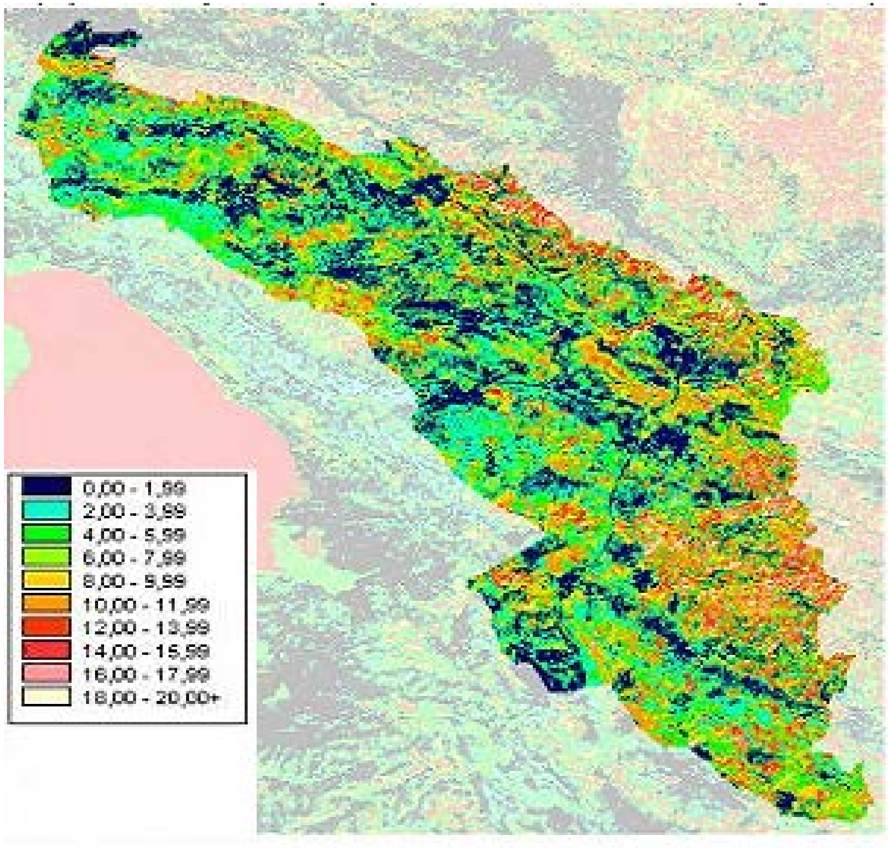
- CLUS used to learn ensembles of multi-target regression trees to predict the 11 targets

```
b7s1_avg > 60.16
+--yes: d3s1_avg > 69.52
|         +--yes: b2s1_std > 2
|         |         +--yes: [10.147239,0.626994,9.984663,5.033252,4.306319,3.77362,2.926442,2.444724,2.054908,1.77576
|         |         +--no:  a4s1_std > 7.01
|         |                   +--yes: [8.375,0.67625,7.125,7.48625,6.91875,6.42375,5.48125,4.98,4.6475,4.43375,4.34125]
|         |                   +--no:  c4s1_avg > 64.06
|         |                             +--yes: c2s1_std > 1.98
|         |                             |         +--yes: [9,0.474,9.1,7.452,6.778,5.737,4.707,3.973,3.537,3.253,3.122] (3)
|         |                             |         +--no:  a7s1_std > 1
|         |                             |                   +--yes: [4.288577,0.266713,4.148297,2.196613,1.89984,1.658337,1.3
|         |                             |                   +--no:  [5.714286,0.507143,5.142857,8.504286,7.587143,6.694286,5.
|         |                             +--no:  a4s2_avg > 79.3
|         |                                       +--yes: [9.707865,0.564073,9.713483,4.59441,3.932669,3.436292,2.666011,2.
|         |                                       +--no:  [2.408333,0.232583,2.358333,1.729,1.643917,1.485583,1.2455,1.0836
|         +--no:  c2s1_std > 1.82
|                   +--yes: [17.095436,1.004481,16.713693,6.832116,5.595519,4.819461,3.682241,3.018382,2.500415,2.163
|                   +--no:  a3s1_avg > 67
|                             +--yes: [3.656716,0.262438,3.781095,2.742537,2.535771,2.266567,1.879502,1.639005,1.424776
|                             +--no:  c4s1_max > 91
|                                       +--yes: [3.68,0.2608,3.413333,2.578533,2.138133,1.895067,1.506533,1.238933,1.0124
|                                       +--no:  c4s1_avg > 62.18
|                                                 +--yes: c4s2_min > 49
|                                                 |         +--yes: [10.875,0.571798,10.664474,5.050833,4.322303,3.807149,3.0
|                                                 |         +--no:  b7s2_std > 9.45
|                                                 |                   +--yes: [31.52381,2.156667,31.095238,9.232381,7.34,6.2023
|                                                 |                   +--no:  d3s1_std > 4.23
```

The final output: Maps of forest height and density

- Generated by applying the learned models to satellite images of the whole Karst area





Predicting the Risk of Fires

AiTLAS

Artificial Intelligence Toolbox for Earth Observation

Estimating the risk of fire in the natural environment

- Develop models to predict fire outbreaks from
  - historical data on fires and
  - explanatory GIS data using data mining methods

- Three models developed for three different regions of Slovenia:
  - Kras (Karst) region,
  - Primorska (Coastal) region, and
  - Continental Slovenia

The data

Data on fire outbreaks for the period 2000-2004

Data on influencing factors

- **GIS** data
  - Infrastructure
  - Land use
  - Relief
- Multi-temporal **MODIS** data
- Meteorological **ALADIN** data
- Fire fuels

The spatial unit was 1x1 km^2 quadrant

Additional data: Fire fuels

For the Kras/Karst region, additional data

- Properties of the forest stand
  (height, canopy cover, vertical veg. profile)
- Related to fire fuel
- Derived by applying machine-learned models
  to LANDSAT images
- Training data for forest stand properties derived from 3D (LIDAR) and LANDSAT images, at spatial resolution of 25m x 25m
- Originally at 25m x 25 m, aggregated to 1 km x 1km quadrants

Positive and negative examples

- **Positive examples** - locations where fires occurred in the past, along with the date and hour
- **Negative examples** - represented by an equal number of points with random time stamps and random locations; at least 15 km away from any positive examples detected in ± 3 days of the random time stamp chosen.
- Locations of the positive and negative examples of fire occurrence spatially and temporally linked to the descriptive data

Machine Learning Setup

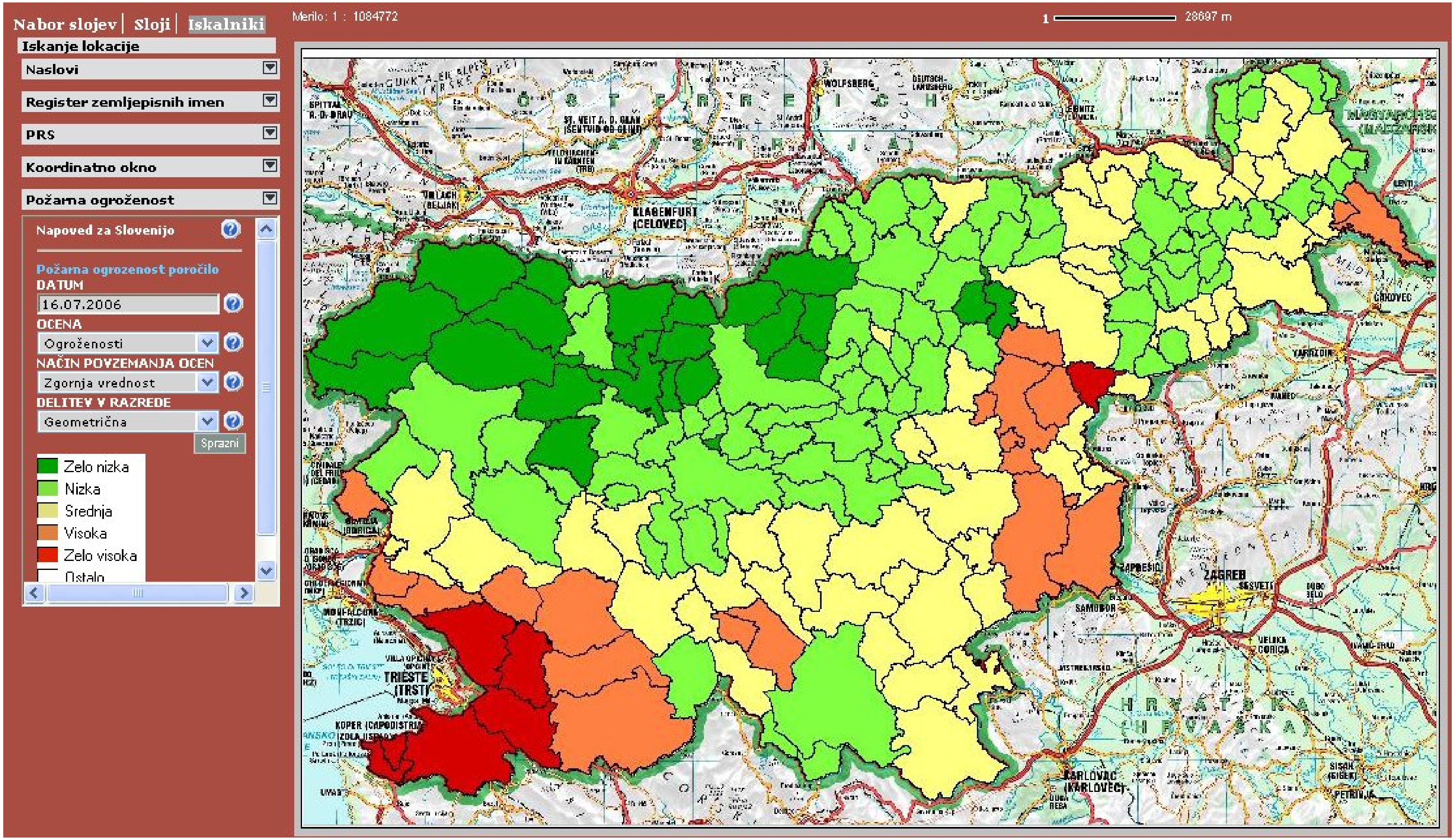
- Three datasets for different regions of Slovenia
  - Kras: 159 attributes and 1439 examples
  - Coastal Slovenia: 129 attributes and 2442 examples
  - Continental Slovenia: 129 attributes and 8476 examples
- Nominal Target attribute that predicts fire outbreak (yes/no)
- Several machine learning methods used
  - Logistic regression
  - Decision trees
  - Tree ensembles (boosting, bagging, random forests)
- Models were validated with 10 fold CV

Understandable rules

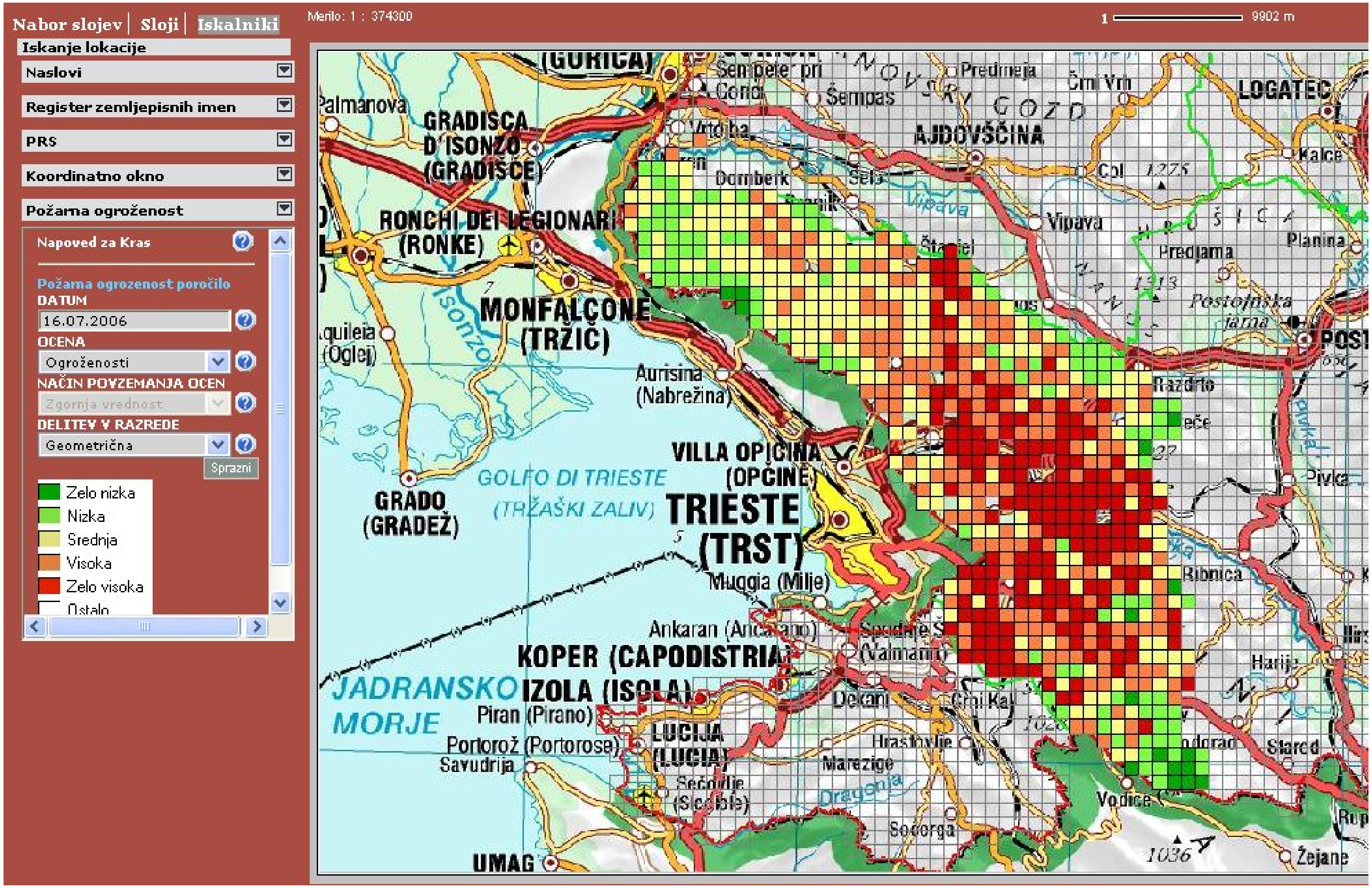
- For the Karst region
- Railways and human activity important

```
if ((dist_railways ≤ 2970) and (elevation ≥ 378) and (percBuiltUp ≥ 0.875)) then
   fireOutbreak = YES
else if ((percBuiltUp ≥ 0.875) and (dist_railways ≤ 1487) and (percOver ≤ 2.875) and
         (percRiparian ≥ 0.625)) then
   fireOutbreak = YES
else if ((percSwMead ≥ 26) and (percarable ≥ 0.1875) and
         (evapotranspiration_48 ≥ -0.9)) then
   fireOutbreak = YES
else if ((dist_railways ≤ 2970) and (elevation ≥ 350) and (percRiparian ≥ 0.125) and
         (dist_railways ≥ 1897)) then
   fireOutbreak = YES
else
   fireOutbreak = NO
end if
```

Municipality level estimates of risk



More detailed estimates of risk



Deployment status

- Included in the GIS system e-GIS UJME about natural disasters at the Administration for Civil Protection and Rescue
- Available to and used by the following organizations
    - Firefighter association
    - Administration for civil protection and rescue
    - Environment Agency
    - Forest Service
- Also accessible to the whole range of users of eGIS-UJME (Natural Disasters), incl. municipalities

Relating the Environment and the Biota: From Habitat models to Community composition

AiTLAS
Artificial Intelligence Toolbox for Earth Observation

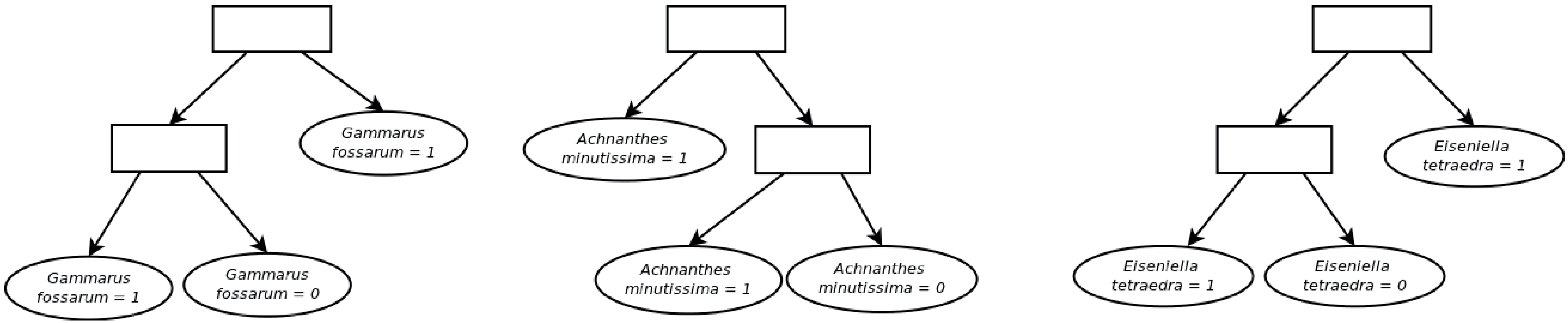
Environment <-> Biota

- Predict the biota (or specific components of it)
- At a given site
- From characteristics of the environment at the site
- E.g. predict river water biota from water properties

	Descriptive variables						Target variables													
Sample ID	Temperature	$K_2Cr_2O_7$	NO_2	Cl	CO_2		*Cladophora sp.*	*Gongrosira incrustans*	*Oedogonium sp.*	*Stigeoclonium tenue*	*Melosira varians*	*Nitzschia palea*	*Audouinella chalybea*	*Erpobdella octoculata*	*Gammarus fossarum*	*Baetis rhodani*	*Hydropsyche sp.*	*Rhyacophila sp.*	*Simulim sp.*	*Tubifex sp.*
ID1	0.66	0.00	0.40	1.46	0.84	...	1	0	0	0	0	1	1	0	1	1	1	1	1	1
ID2	2.03	0.16	0.35	1.74	0.71	...	0	1	0	1	1	1	1	0	1	1	1	1	1	0
ID3	3.25	0.70	0.46	0.78	0.71	...	1	1	0	0	1	0	1	0	1	1	1	0	1	1

Habitat modeling

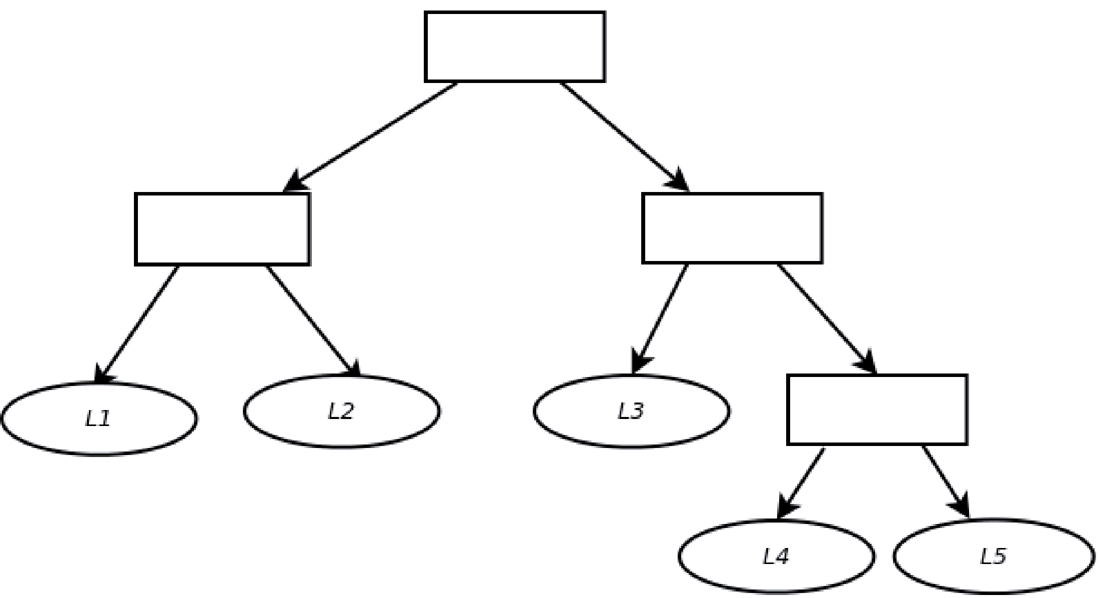
- Model the presence & absence (abundance) of each species separately



- Binary Classification (Regression)

Predicting species composition

- One model for **all the species at once**



L1:
Gammarus fossarum: 0
Achnanthes minutissima: 1
Eiseniella tetraedra: 1

L4:
Gammarus fossarum: 1
Achnanthes minutissima: 1
Eiseniella tetraedra: 0

L2:
Gammarus fossarum: 0
Achnanthes minutissima: 1
Eiseniella tetraedra: 0

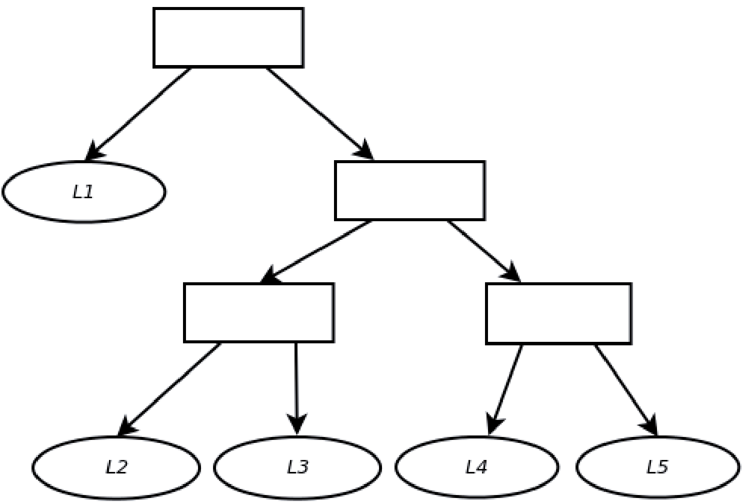
L5:
Gammarus fossarum: 0
Achnanthes minutissima: 1
Eiseniella tetraedra: 1

L3:
Gammarus fossarum: 1
Achnanthes minutissima: 1
Eiseniella tetraedra: 1

- **Multi-target classification/regression**

Predicting community structure

- One model for all of the species at once, additionally using the taxonomical hierarchy



L1:

- *Amphipoda : 1*
  - *Gammarus : 1*
    - *Gammarus fossarum : 1*
    - *Gammarus lacustris : 0*
- *Bacillariophyta : 1*
  - *Achnanthes : 1*
    - *Achnanthes minutissima: 1*
  - *Eiseniella : 0*
    - *Eiseniella tetraedra: 0*

L3:

- *Amphipoda : 1*
  - *Gammarus : 1*
    - *Gammarus fossarum : 0*
    - *Gammarus lacustris : 1*
- *Bacillariophyta : 1*
  - *Achnanthes : 1*
    - *Achnanthes minutissima: 1*
  - *Eiseniella : 0*
    - *Eiseniella tetraedra: 0*

L5:

- *Amphipoda : 1*
  - *Gammarus : 1*
    - *Gammarus fossarum : 1*
    - *Gammarus lacustris : 1*
- *Bacillariophyta : 1*
  - *Achnanthes : 0*
    - *Achnanthes minutissima: 0*
  - *Eiseniella : 1*
    - *Eiseniella tetraedra: 1*

L2:

- *Amphipoda : 1*
  - *Gammarus : 1*
    - *Gammarus fossarum : 1*
    - *Gammarus lacustris : 1*
- *Bacillariophyta : 0*
  - *Achnanthes : 0*
    - *Achnanthes minutissima: 0*
  - *Eiseniella : 0*
    - *Eiseniella tetraedra: 0*

L4:

- *Amphipoda : 1*
  - *Gammarus : 1*
    - *Gammarus fossarum : 1*
    - *Gammarus lacustris : 0*
- *Bacillariophyta : 1*
  - *Achnanthes : 1*
    - *Achnanthes minutissima: 1*
  - *Eiseniella : 1*
    - *Eiseniella tetraedra: 1*

- Hierarchical multi-label classification

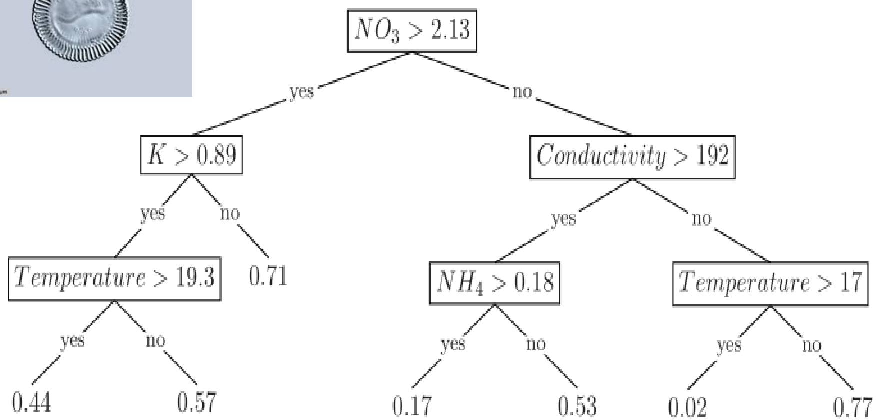
Slovenian rivers

- 1.060 samples
- 16 physical and chemical props. of water, 491 species
- data collected in 1990-1995

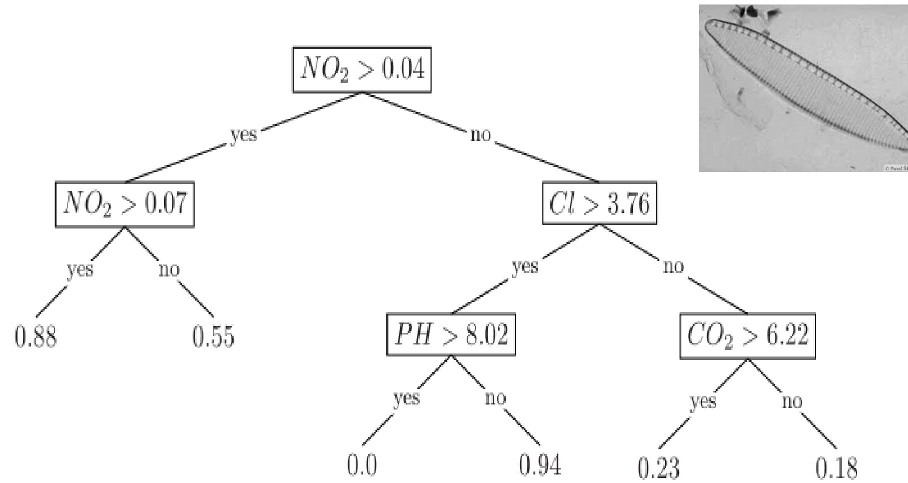
ephemeroptera
- *ephemeroptera_acantrella*
  - *ephemeroptera_acantrella_sinaica*
- *ephemeroptera_baetidae*
- *ephemeroptera_baetis*
  - *ephemeroptera_baetis_alpinus*
  - *ephemeroptera_baetis_buceratus*
  - *ephemeroptera_baetis_fuscatus*
  - *ephemeroptera_baetis_muticus*
  - ***ephemeroptera_baetis_rhodani***
  - *ephemeroptera_baetis_scambus*
  - *ephemeroptera_baetis_vernus*
- *ephemeroptera_ecdyonurus*
  - *ephemeroptera_ecdyonurus_forcipula*
  - *ephemeroptera_ecdyonurus_helveticus*
  - *ephemeroptera_ecdyonurus_insignis*
  - *ephemeroptera_ecdyonurus_torrentis*
  - *ephemeroptera_ecdyonurus_venosus*
- *ephemeroptera_electrogena*
  - *ephemeroptera_electrogena_lateralis*
  - *ephemeroptera_electrogena_quadrilineata*

plecoptera
- *plecoptera_amphinemura*
  - *plecoptera_amphinemura_triangularis*
- *plecoptera_brachyptera*
  - ***plecoptera_brachyptera_risi***
  - *plecoptera_brachyptera_seticornis*

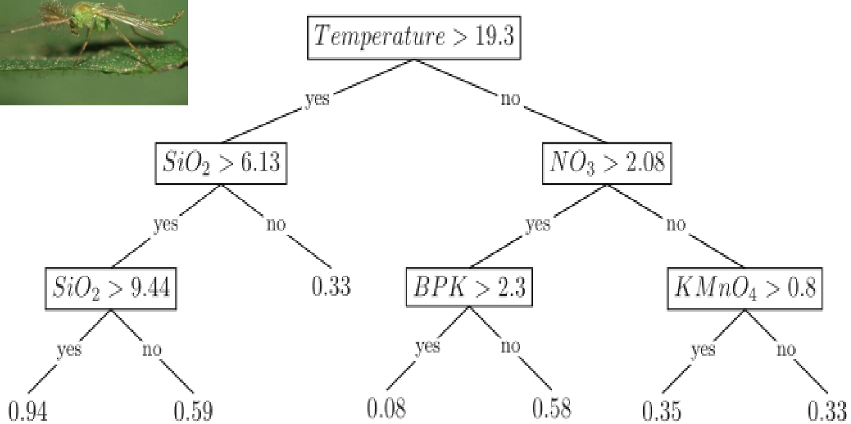
Slovenian rivers: Habitat models



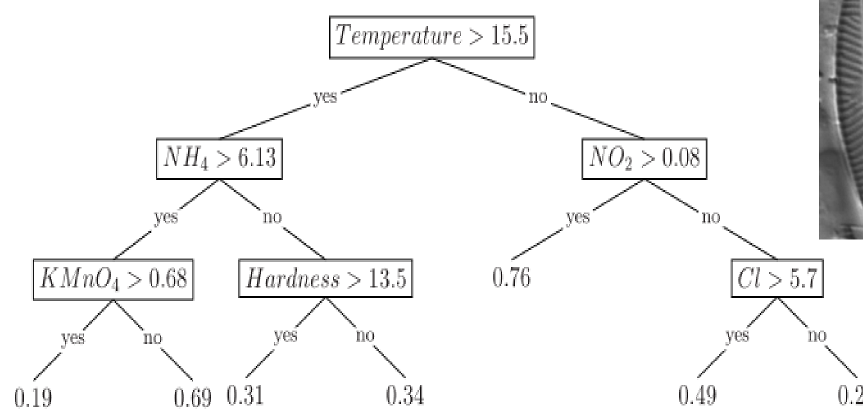
Bacillariophyta Cyclotella Comta

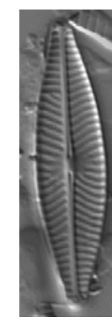


Bacillariophyta Nitzschia Palea



Diptera Chironomidae Zeleni

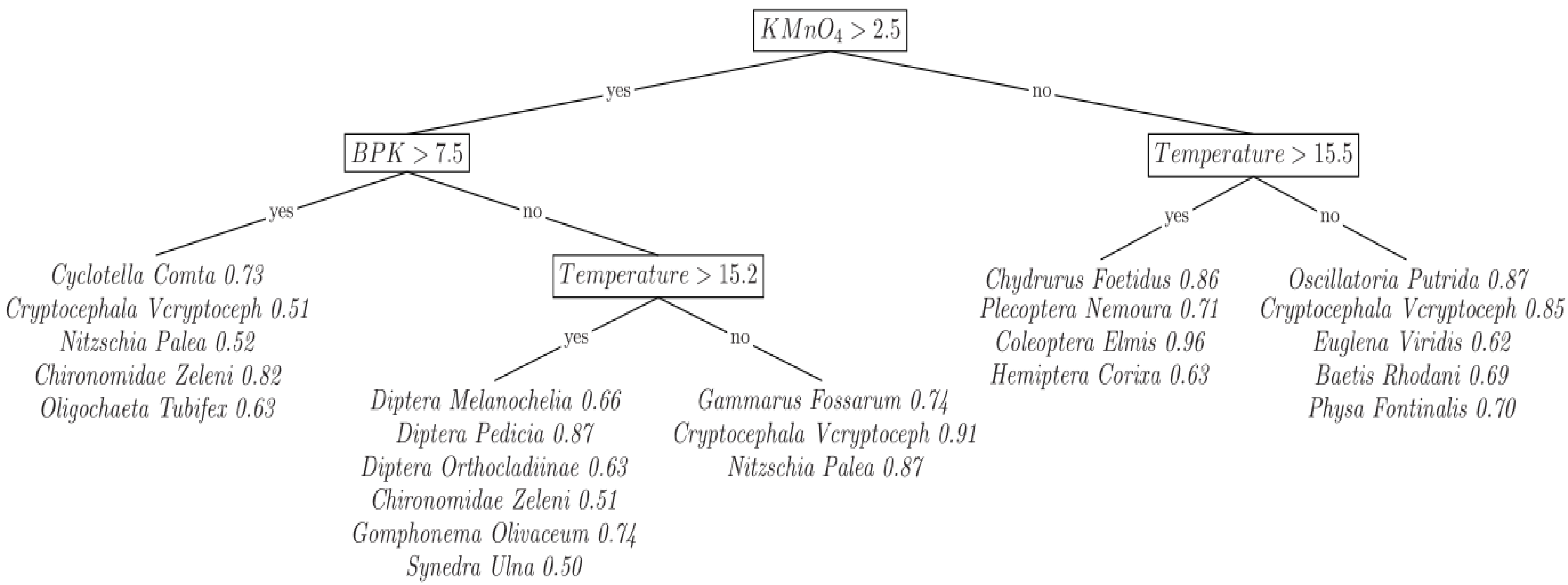




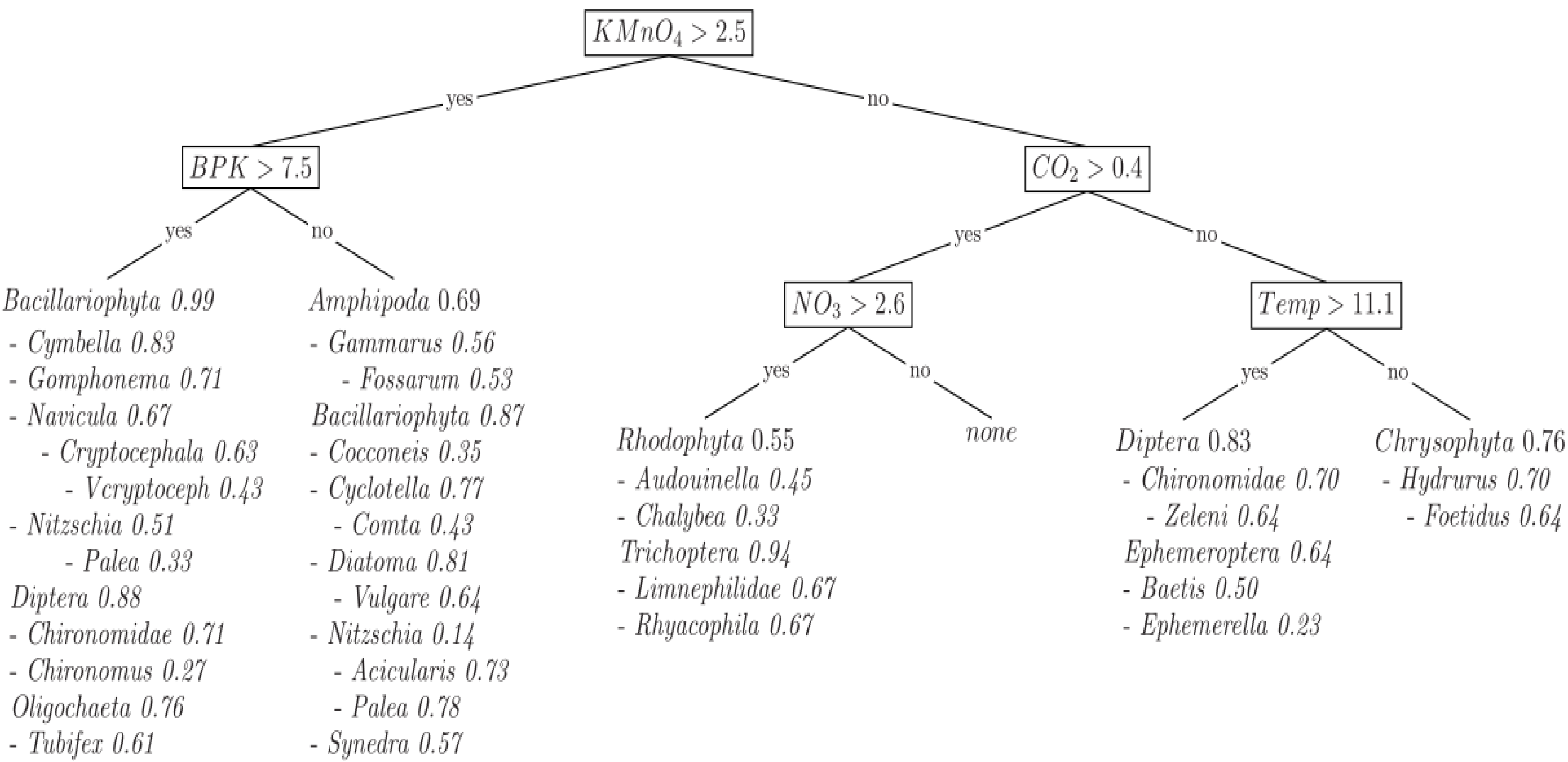
Bacillariophyta Navicula Cryptocephala Vcryptoceph

Slovenian rivers: Species comp.

- MLC: Multi-label classification tree



Slovenian rivers: Community struc.



Danish farms: Soil Microarthropods

- 1.944 soil samples
- 137 attributes/agricultural events and soil biological parameters
- 35 collembolan species
- data collected 1989-1993

Isotominae
- ***Isotominae_Isotoma***
  - *Isotominae_Isotoma_anglicana*
  - *Isotominae_Isotoma_notabilis*
  - *Isotominae_Isotoma_tigrina*

Lepidocyrtinae
- *Lepidocyrtinae_Lepidocyrtus*
  - *Lepidocyrtinae_Lepidocyrtus_cyaneus*
  - *Lepidocyrtinae_Lepidocyrtus_lanuginosus*
- *Lepidocyrtinae_Pseudosinella*
  - *Lepidocyrtinae_Pseudosinella_alba*
  - *Lepidocyrtinae_Pseudosinella_sexoculata*

Orchesellinae
- *Orchesellinae_Heteromurus*
  - *Orchesellinae_Heteromurus_nitidus*
- *Orchesellinae_Orchesella*
  - *Orchesellinae_Orchesella_cincta*
  - *Orchesellinae_Orchesella_villosa*

Sminthuridae
- *Sminthuridae_Smint*
- *Sminthuridae_Sminthurinus*
  - *Sminthuridae_Sminthurinus_aureus*
  - *Sminthuridae_Sminthurinus_elegans*
- *Sminthuridae_Sminthurus*
  - *Sminthuridae_Sminthurus_viridis*

Tomoceridae
- *Tomoceridae_Tomocerus*
  - *Tomoceridae_Tomocerus_flavescens*
  - *Tomoceridae_Tomocerus_minor*

Tullbergiidae
- *Tullbergiidae_Mesaphorura*

Victoria, Australia Vegetation

- 27.482 sites
- 81 env. attributes
- 3.173 species

- DivisionConifer
  - DivisionConifer_callitris
    - DivisionConifer_callitris_endlicheri
    - DivisionConifer_callitris_glaucophylla
    - DivisionConifer_callitris_gracilis
      - DivisionConifer_callitris_gracilis_ssp~murrayensis
    - DivisionConifer_callitris_rhomboidea
    - DivisionConifer_callitris_verrucosa
- DivisionMonocotyledon
  - DivisionMonocotyledon_leucopogon
    - DivisionMonocotyledon_leucopogon_attenuatus
    - DivisionMonocotyledon_leucopogon_australis
    - DivisionMonocotyledon_leucopogon_clelandii
    - DivisionMonocotyledon_leucopogon_juniperinus
    - DivisionMonocotyledon_leucopogon_lanceolatus

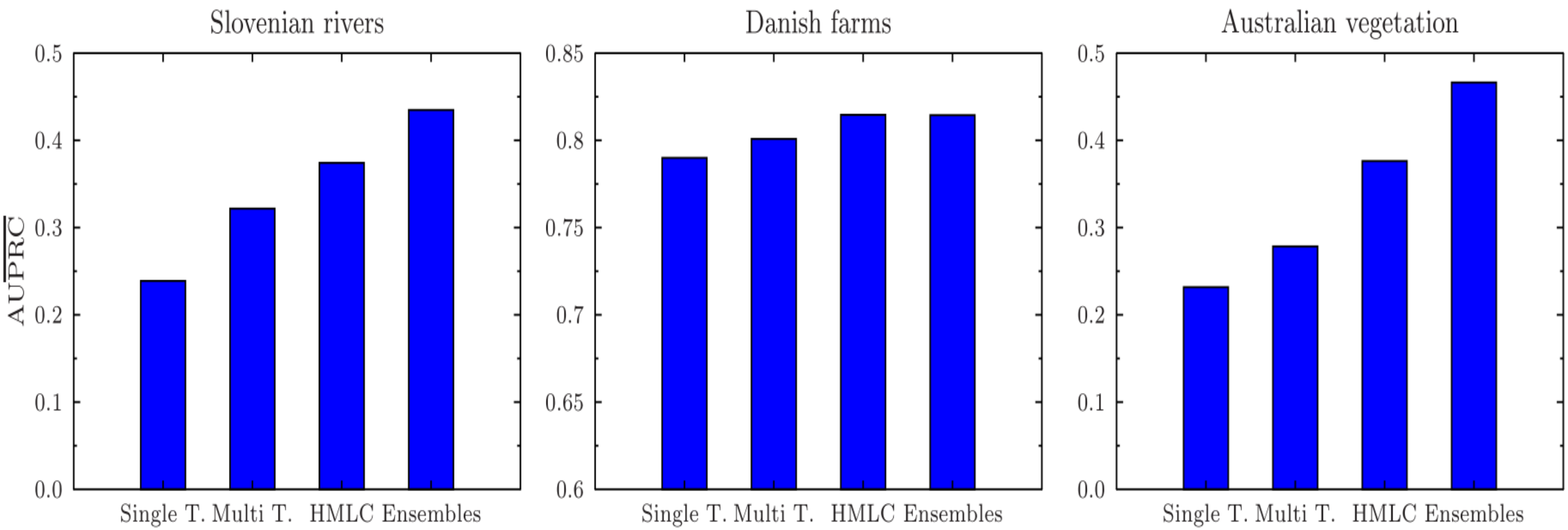
DivisionMonocotyledon_leucopogon_lanceolatus_var~lanceolatus

- DivisionMonocotyledon_leucopogon_maccraei
- **DivisionMonocotyledon_leucopogon_microphyllus**

DivisionMonocotyledon_leucopogon_microphyllus_var~pilibundus

- DivisionMonocotyledon_leucopogon_montanus
- DivisionMonocotyledon_leucopogon_neurophyllus
- DivisionMonocotyledon_leucopogon_parviflorus
- DivisionMonocotyledon_leucopogon_virgatus
  - DivisionMonocotyledon_leucopogon_virgatus_var~brevifolius
  - DivisionMonocotyledon_leucopogon_virgatus_var~virgatus
- DivisionMonocotyledon_leucopogon_woodsii
- DivisionMonocotyledon_epacris
  - DivisionMonocotyledon_epacris_breviflora
  - DivisionMonocotyledon_epacris_celata
  - DivisionMonocotyledon_epacris_glacialis
  - DivisionMonocotyledon_epacris_gunnii
  - **DivisionMonocotyledon_epacris_impressa**
    - DivisionMonocotyledon_epacris_impressa_var~grandiflora
    - DivisionMonocotyledon_epacris_impressa_var~impressa

Community structure: Overall results



Dataset	Method	$\overline{\text{AUPRC}}$	O_S	Learning time	Complexity
Slovenian rivers	Single-label	0.239	0.692	23.3	15,336
	HSC	0.309	0.591	10.2	25,035
	Multi-label	0.322	0.007	9.4	1
	HMC	**0.374**	0.132	0.6	37
Danish farms	Single-label	0.790	0.099	3.7	2605
	HSC	0.808	0.083	1.3	2873
	Multi-label	0.801	0.112	0.7	265
	HMC	**0.815**	0.065	0.4	259
Australian vegetation	Single-label	0.232	0.715	14,888.2	482,745
	HSC	0.306	0.591	76,023.2	648,970
	Multi-label	0.278	0.684	4639.5	23,699
	HMC	**0.376**	0.180	313.5	1279

Relating the Environment and the Biota: Predicting the functional composition/ traits of biota

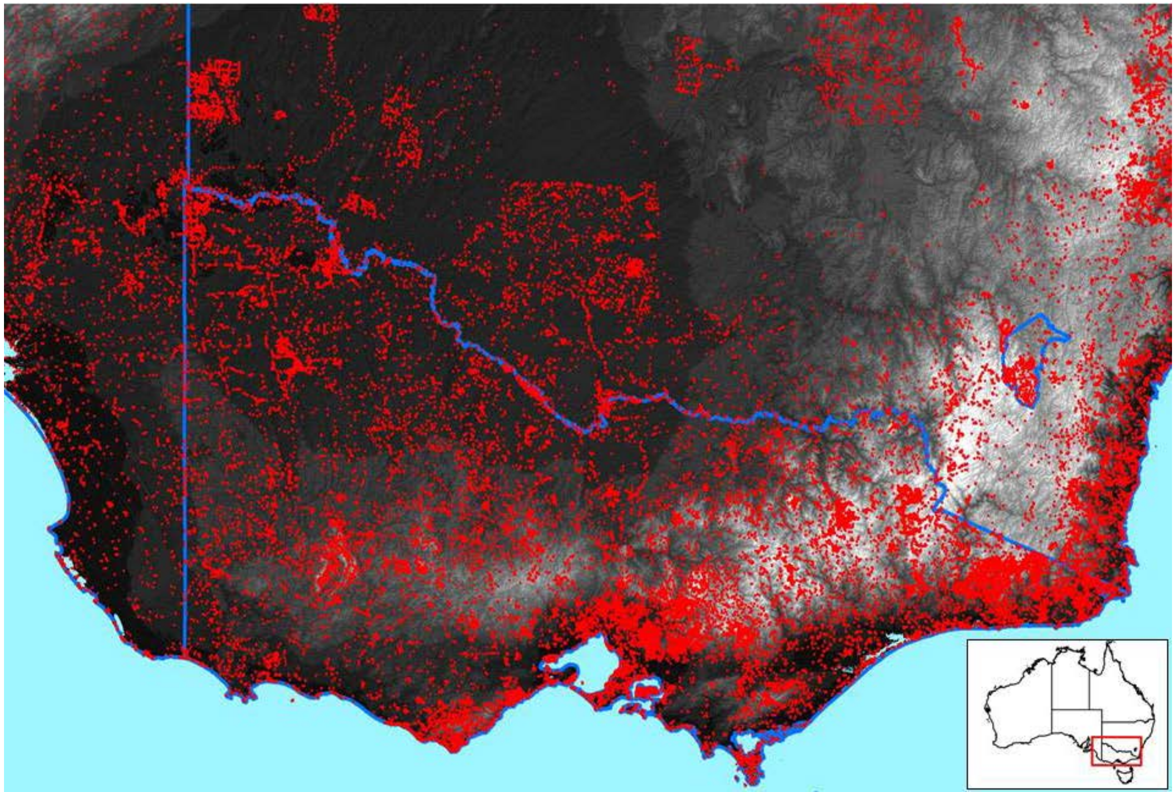
AiTLAS

Artificial Intelligence Toolbox for Earth Observation

Victoria, Australia: Relating env. char. to plant trait profiles

New, much more extensive data: Collected 1960-2010, 53362 sites, more than 1.35 Mio indiv. spec. obs.

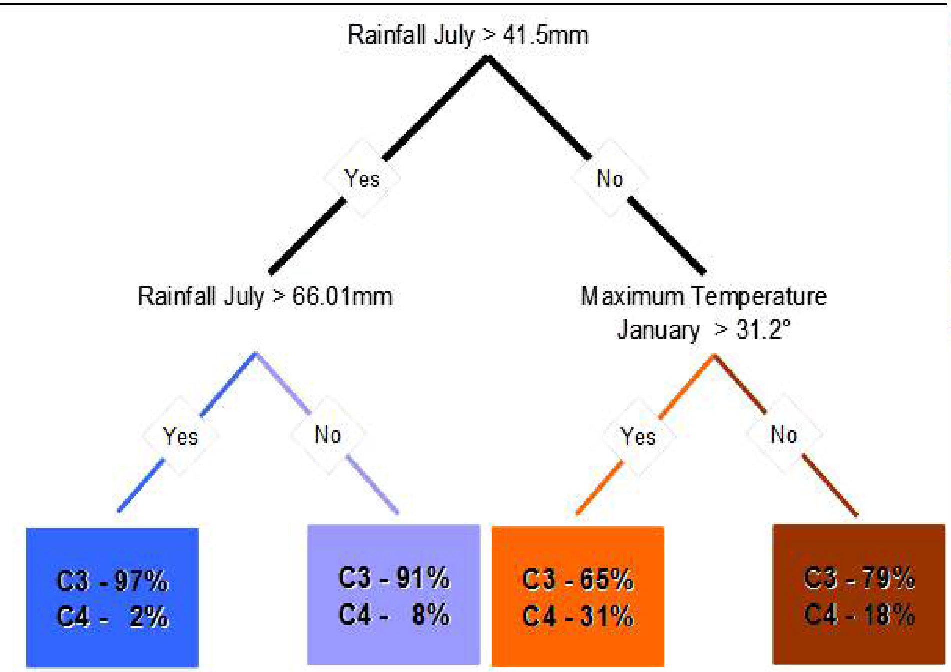
Each vascular species, recorded together with % cover

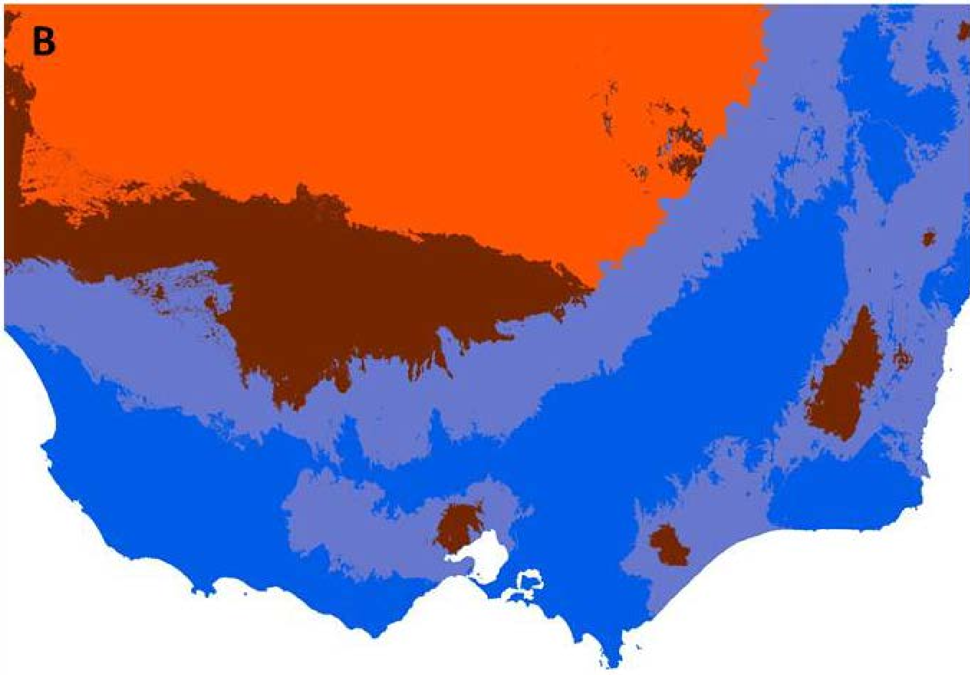


Victoria, Australia: Relating env. char. to plant trait profiles

Plant photosynthetic type (carbon fixation pathways)

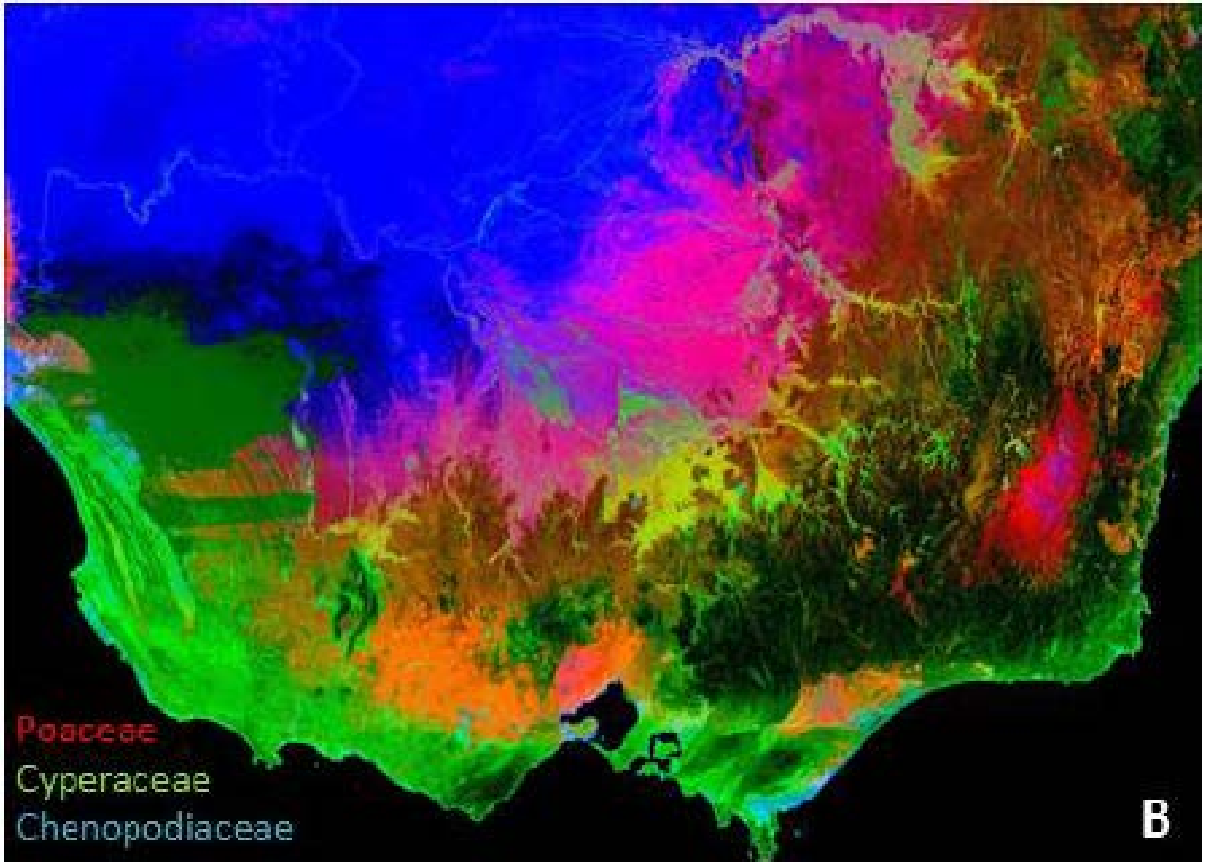
- C3: cool-season-active
- C4: warm-season-active





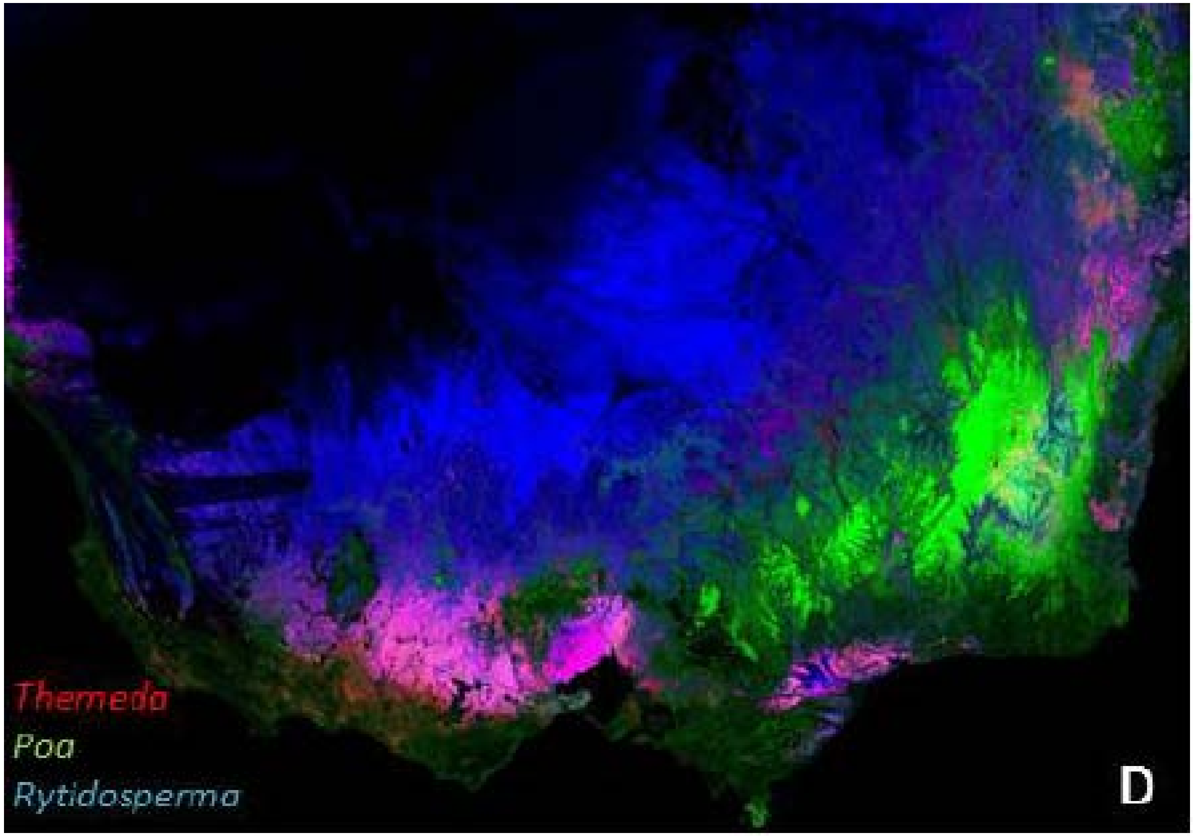
Victoria, Australia: Relating env. char. to plant trait profiles

Phylogeny via main monocot families (Poaceae=Grasses, Cyperaceae=Sedges; Chenopodiaceae=Goosefoots)



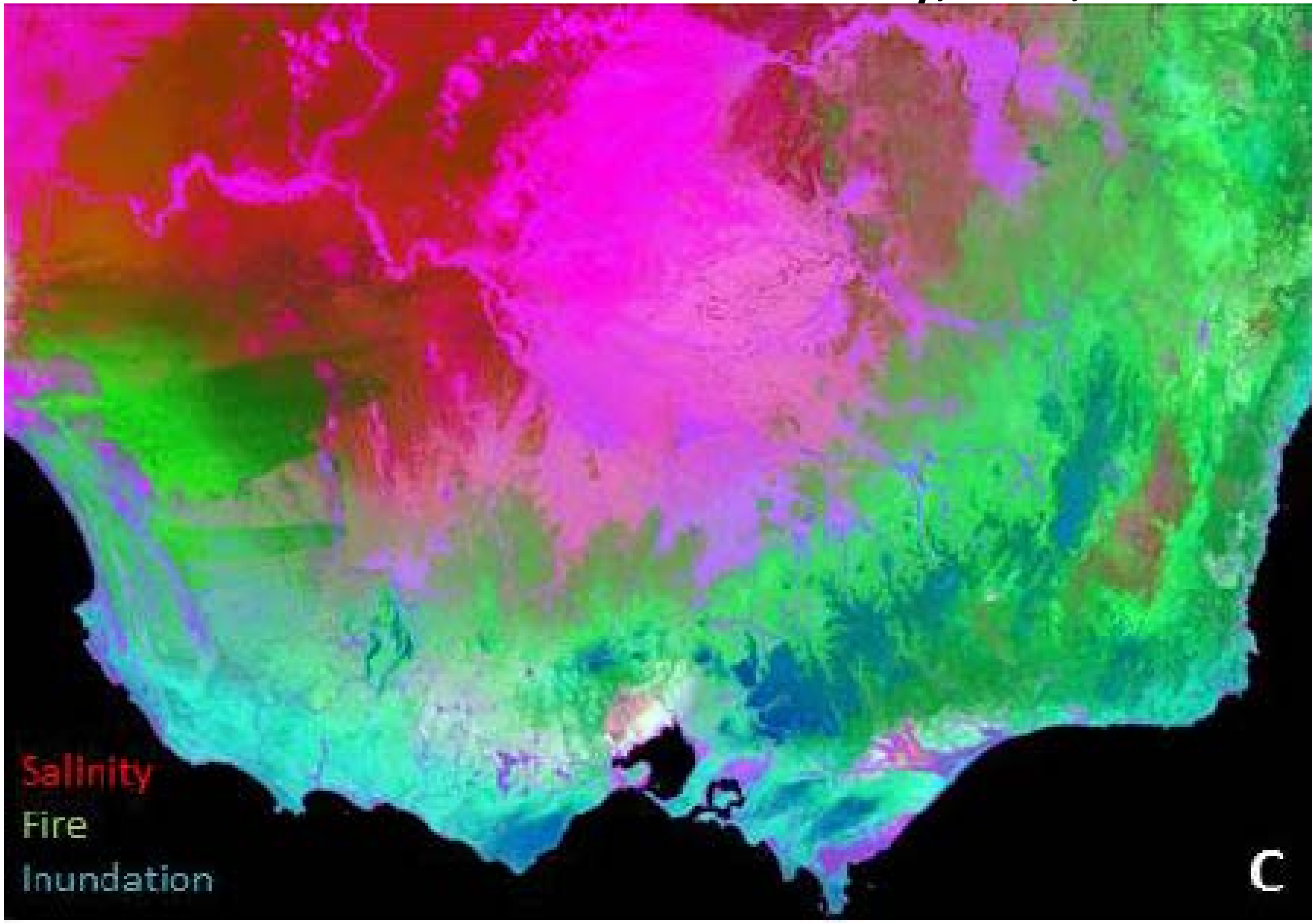
Victoria, Australia: Relating env. char. to plant trait profiles

Phylogeny via three main grass genera



Victoria, Australia: Relating env. char. to plant trait profiles

Stress tolerance: Tolerance to salinity, fire, inundation



Multi-label classification for Land-use/land-cover

AiTLAS

Artificial Intelligence Toolbox for Earth Observation

MLC Datasets

Small datasets

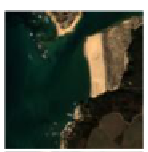
1. UC-Merced2100 multilabel
2. AID multilabel
3. Ankara HIS Archive
4. DFC15 multilabel

The BigEarthNet dataset

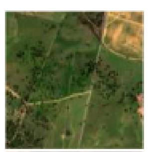
- Size: 1% and 10%
- Labels: 43 and 19

Dataset	Examples	Train examples	Test examples	Labels
AID multilabel	3000	2400	600	17
Ankara HIS Archive	216	173	43	29
DFC15 multilabel	3342	2673	669	8
UC-Merced	2100	1600	500	17

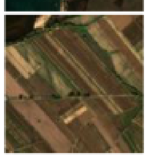
Dataset	Examples	Train examples	Test examples	Labels
BigEarthNet-b01	5,192	3,934	1,258	19
BigEarthNet-b10	51,928	39,341	12,586	19
BigEarthNet-a01	5,192	3,934	1,258	43
BigEarthNet-a10	51,928	39,341	12,586	43



permanently irrigated land, sclerophyllous vegetation, beaches, dunes, sands, estuaries, sea and ocean

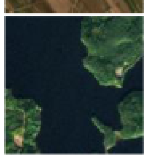


non-irrigated arable land, fruit trees and berry plantations, agro-forestry areas, transitional woodland/shrub

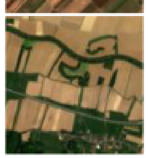


permanently irrigated land, vineyards, beaches, dunes, sands, water courses

non-irrigated arable land



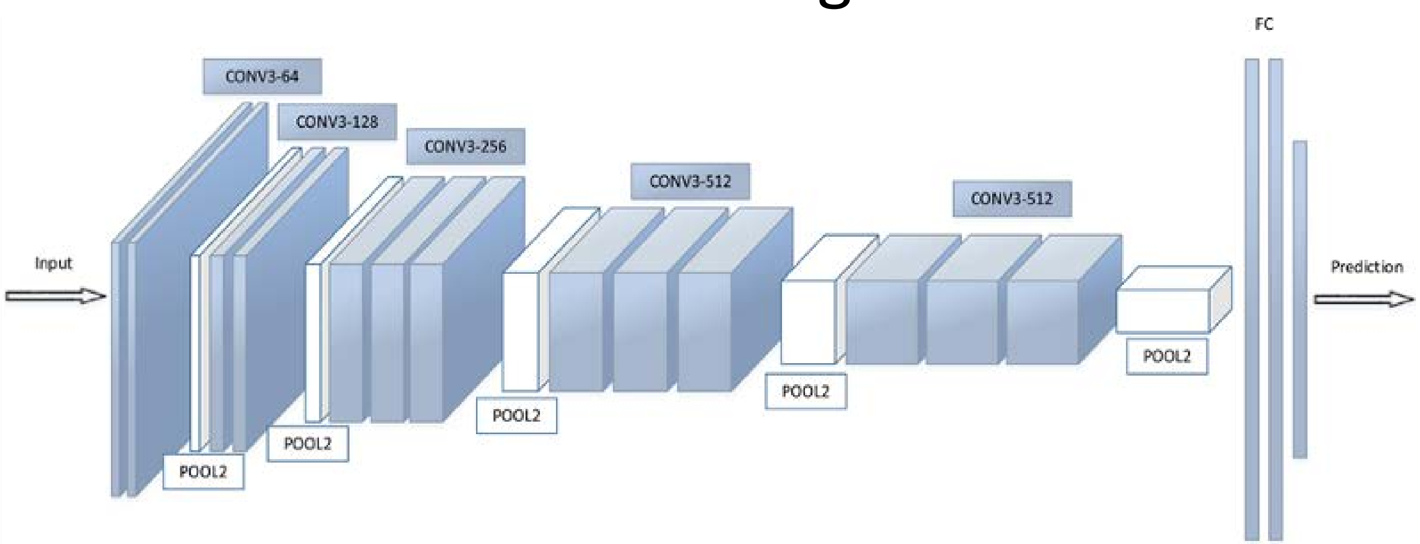
coniferous forest, mixed forest, water bodies



discontinuous urban fabric, non-irrigated arable land, land principally occupied by agriculture, broad-leaved forest

End-to-End Deep Learning Approach

- VGG16: Used as a baseline,
  but also to extract features from the images
- Pre-trained on ImageNet
  (except for BigEarthNet, trained from scratch)
- Fine-tuned on the training set of each dataset



Multi-label classification approach

- Extensive empirical comparative study revealed that
- Tree-ensembles for multi-label classification work by far the best

- Here we apply tree ensembles for multi-label classification
- On features are extracted from pre-trained (and fine-tuned) CNNs, more specifically the VGG-16 architecture mentioned above

MLC: Summary of Results

- UC-Merced
  - RF for MLC: 0.8041 microF1 / VGG 16: 0.7987
- AID
  - RF for MLC: 0.8731 microF1 / VGG 16: 0.84
- DFC-15
  - RF for MLC: 0.7865 microF1 / VGG 16: 0.79
- Ankara
  - RF for MLC: 0.8048 microF1
- BigEarthNet
  - RF for MLC: 0.7502 microF1 / VGG 16: 0.74

Hierarchical Multi-label Classification (HMLC)

- For original BigEarthNet label set, with 43 labels, the labels are taken from the CLC nomenclature, which also provides the hierarchy
- Three-level hierarchy, all leaf labels are at level 3

Level 1	Level 2	Level 3
1 Artificial surfaces	11 Urban fabric	111 Continuous urban fabric
		112 Discontinuous urban fabric
	12 Industrial, commercial and transport units	121 Industrial or commercial units
		122 Road and rail networks and associated land
		123 Port areas
		124 Airports
	13 Mine, dump and construction sites	131 Mineral extraction sites
		132 Dump sites
		133 Construction sites
	14 Artificial, non-agricultural vegetated areas	141 Green urban areas
		142 Sport and leisure facilities
2 Agricultural areas	21 Arable land	211 Non-irrigated arable land
		212 Permanently irrigated land
		213 Rice fields
	22 Permanent crops	221 Vineyards
		222 Fruit trees and berry plantations
		223 Olive groves
	23 Pastures	231 Pastures
	24 Heterogeneous agricultural areas	241 Annual crops associated with permanent crops
		242 Complex cultivation patterns
		243 Land principally occupied by agriculture, with
		244 Agro-forestry areas
3 Forest and semi natural areas	31 Forests	311 Broad-leaved forest
		312 Coniferous forest
		313 Mixed forest
4 Wetlands	41 Inland wetlands	411 Inland marshes
		412 Peat bogs
	42 Maritime wetlands	421 Salt marshes
		422 Salines
		423 Intertidal flats
5 Water bodies	51 Inland waters	511 Water courses
		512 Water bodies
	52 Marine waters	521 Coastal lagoons
		522 Estuaries
		523 Sea and ocean

HMLC for BigEarthNet – 19 labels

- The CLC nomenclature can be easily adapted to the new set of 19 leaf labels (see below). Not all of the 19 labels are at level 3 (10 are at level 2).

Rev CLC Level 1	Rev CLC Level 2	Rev CLC Level 3
a Artificial surfaces	aa Urban fabric	
	ab Industrial, commercial and transport units	
b Agricultural areas	ba Arable land	
	bb Permanent crops	
	bc Pastures	
	bd Heterogeneous agricultural areas	bda Complex cultivation patterns
		bdb Land principally occupied by agriculture, with significant areas of natural vegetation
		bdc Agro-forestry areas
c Forests and semi-natural areas	ca Forests	caa Broad-leaved forest
		cab Coniferous forest
		cac Mixed forest
	cb Shrub and/or herbaceous vegetation association	cba Natural grassland and sparsely vegetated areas
		cbb Moors, heathland and sclerophyllous vegetation
		cbc Transitional woodland shrub
	cc Beaches, dunes, sands	
d Wetlands	da Inland wetlands	
	db Coastal wetlands	
e Water bodies	ea Inland waters	
	eb Marine waters	

Summary

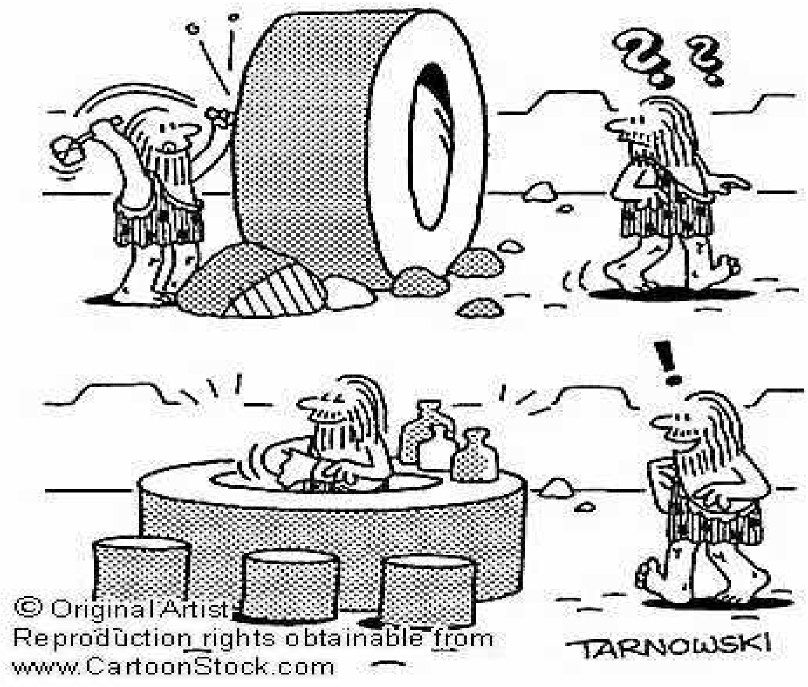
- AI is a very powerful technology
- It is more than just machine learning/DNNs

- Both explainable AI/ML and DNNs have their place under the sun, can be also used together
- There are many applications in forestry/env. sciences

- AITLAS Toolbox contains many useful ML methods

Thank you!

- For your attention
- Questions welcome!
- Thanks also to the AITLAS team



AiTLAS
Artificial Intelligence Toolbox
for Earth Observation

Bias Variance Labs

JSI

ZRCSAZU

esa