Εισαγωγή στην Εκμάθηση Μηχανής

Μάθημα 1 - Εισαγωγή

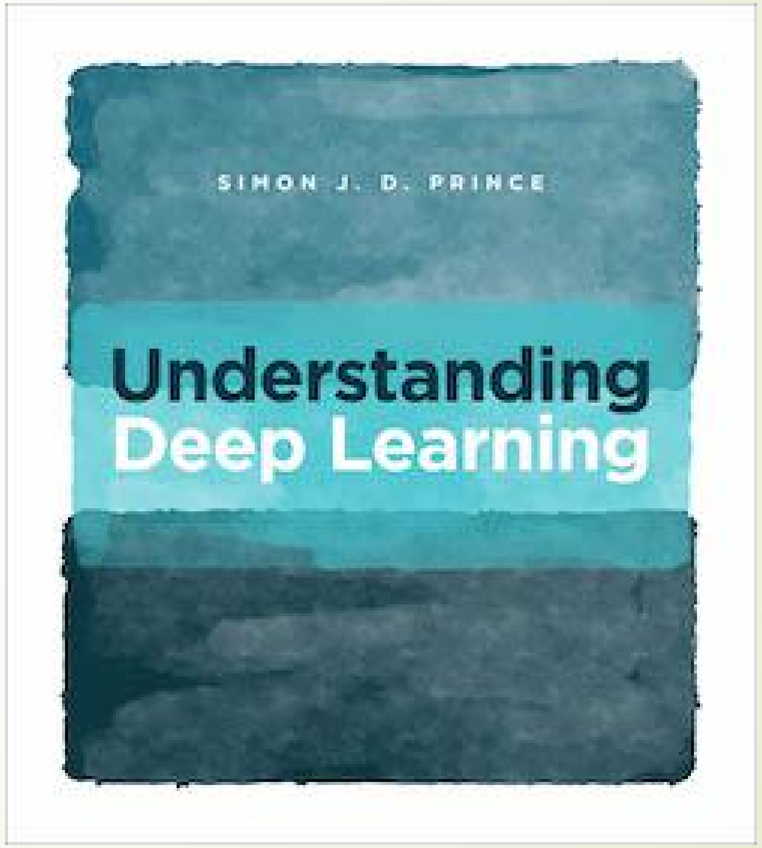
Γιώργος Σφήκας

Γενικά για τον τρόπο διεξαγωγής του μαθήματος

- Χειμερινό εξάμηνο 2025-2026
- 2 ώρες θεωρία (Τετάρτη 1-3)
- 2 ώρες εργαστήριο και ασκήσεις (Τετάρτη 3-5)
- Παράδοση μιας **απλής** άσκησης σε τακτά διαστήματα (“θεωρητική” / Python)
- 50% ασκήσεις + 70% θεωρία

Βιβλίο μαθήματος

- S.J.D. Prince, «Understanding Deep Learning» (MIT Press, 2023)



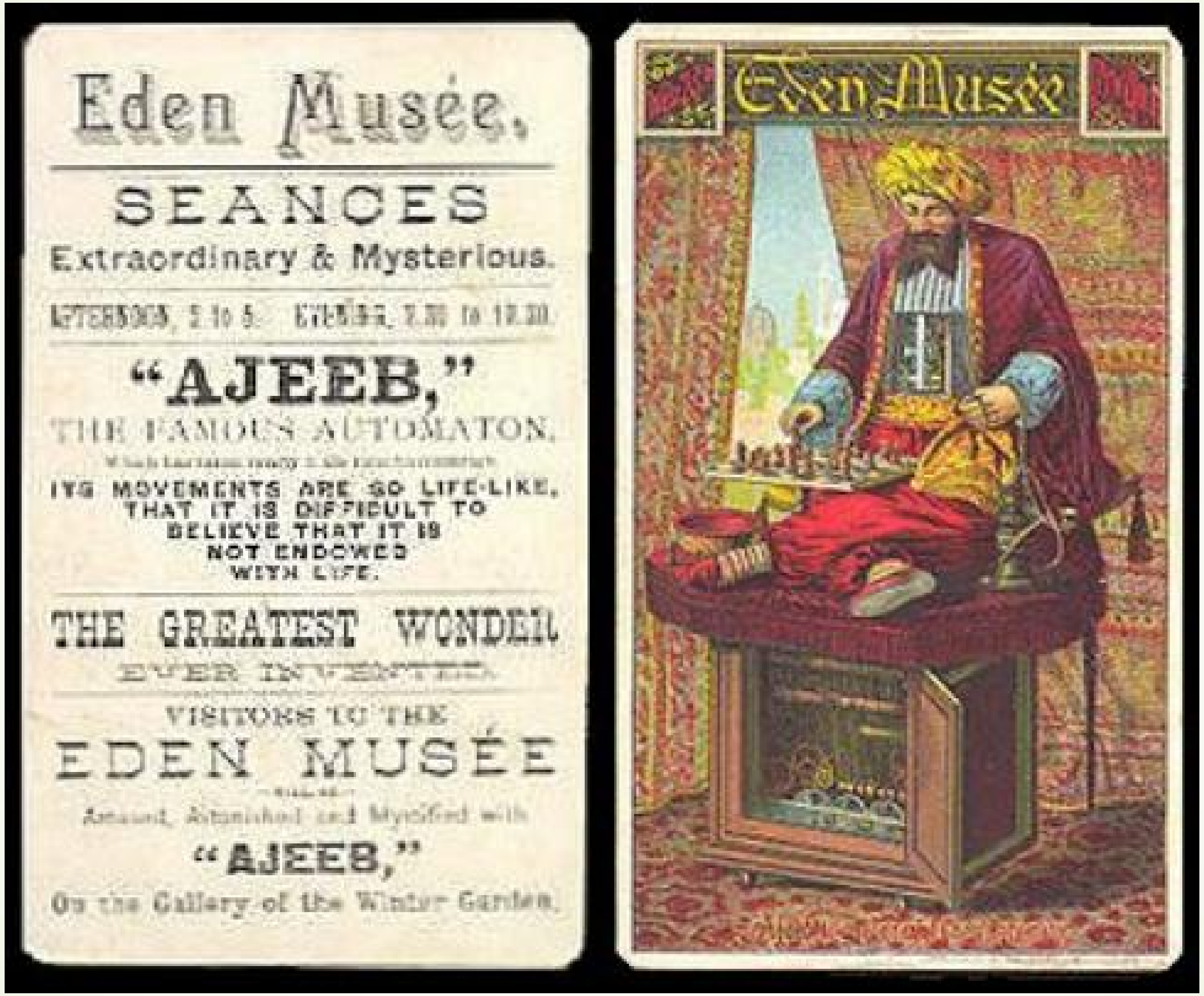
Άλλα εξαιρετικά βιβλία

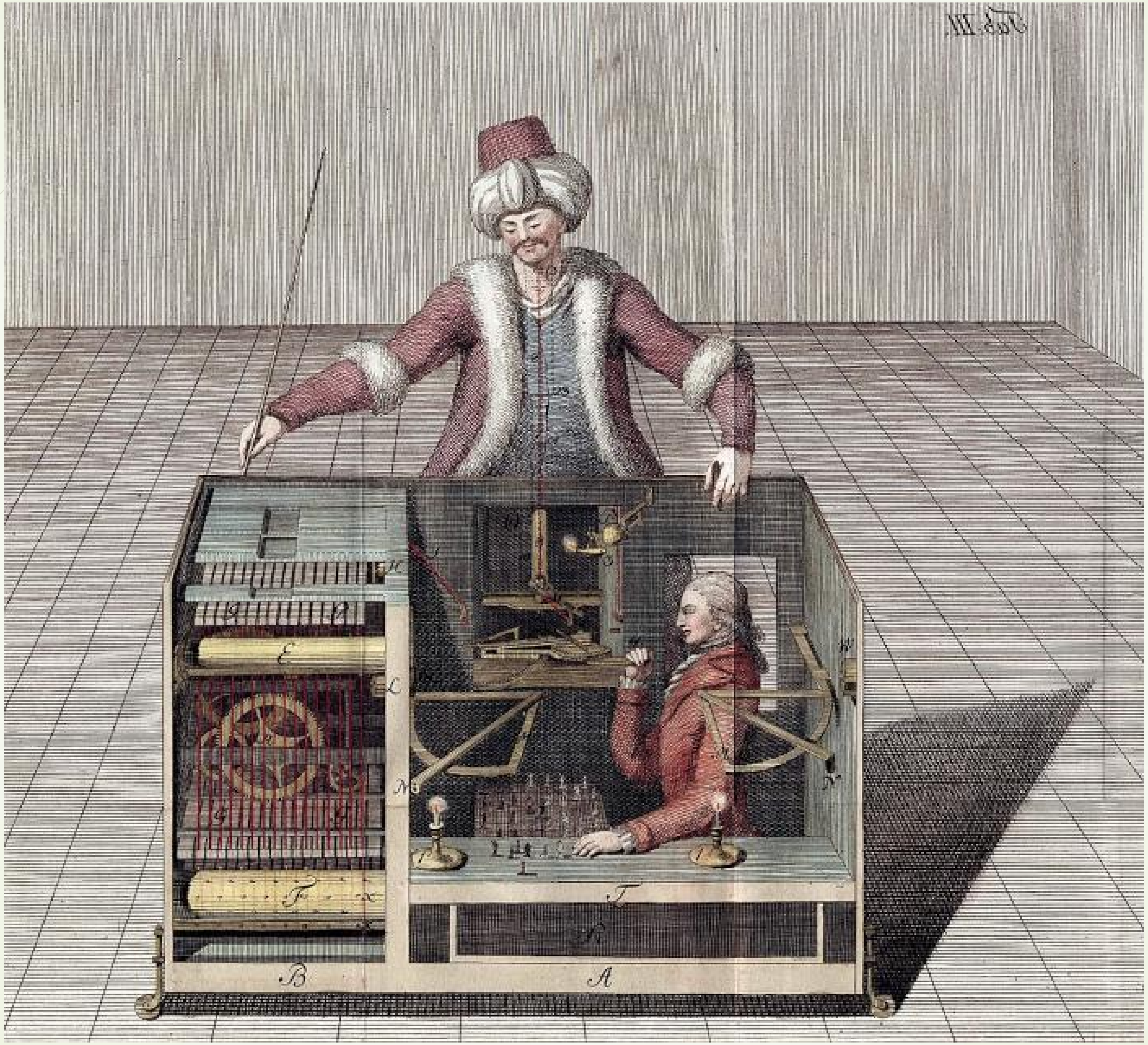
- Σέργιος Θεοδωρίδης, «Machine Learning: A Bayesian and Optimization perspective» (2η έκδοση, 2020)
- Goodfellow, Bengio, Courville, «Deep Learning»
- C.M.Bishop, «Pattern Recognition and Machine Learning» (2006)
- S.J.D. Prince, «Computer Vision: Models, Learning and Inference» (2011)
- G. Strang, «Linear Algebra and its Applications»
- G. Strang, «Linear Algebra and Learning from Data»

Προσοχή στην ορολογία

- Οι μεταφράσεις των όρων σε αυτό το πεδίο είναι ακανθώδες ζήτημα..
- Οι περισσότεροι ελληνικοί όροι που έχουν προταθεί είναι σε αχρηστία
- Regression = Παλινδρόμηση
- Class = Κλάση, τάξη
- Machine Learning = Εκμάθηση μηχανής (ή μηχανική μάθηση..)
- ...

Artificial intelligence


Eden Musée.
SEANCES
Extraordinary & Mysterious.
"AJEEB,"
THE FAMOUS AUTOMATON,
ITS MOVEMENTS ARE SO LIFE-LIKE, THAT IT IS DIFFICULT TO BELIEVE THAT IT IS NOT ENDOWED WITH LIFE.
THE GREATEST WONDER
VISITORS TO THE
EDEN MUSÉE
Amused, Astonished and Mystified with
"AJEEB,"
On the Gallery of the Winter Garden.
Eden Musée


Tab. III
E
L
F
T
R
B
A

Το τεστ του Turing (1950)

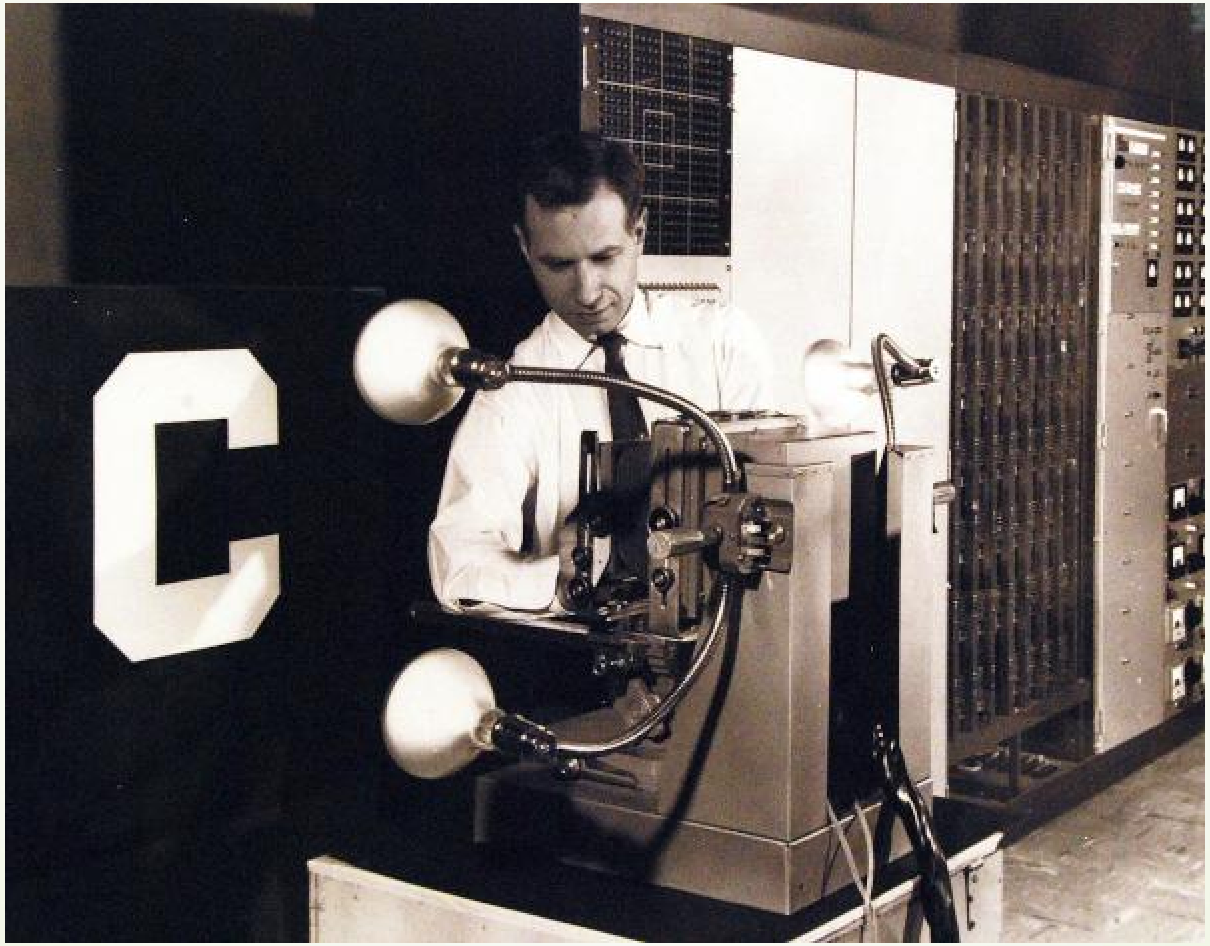
I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

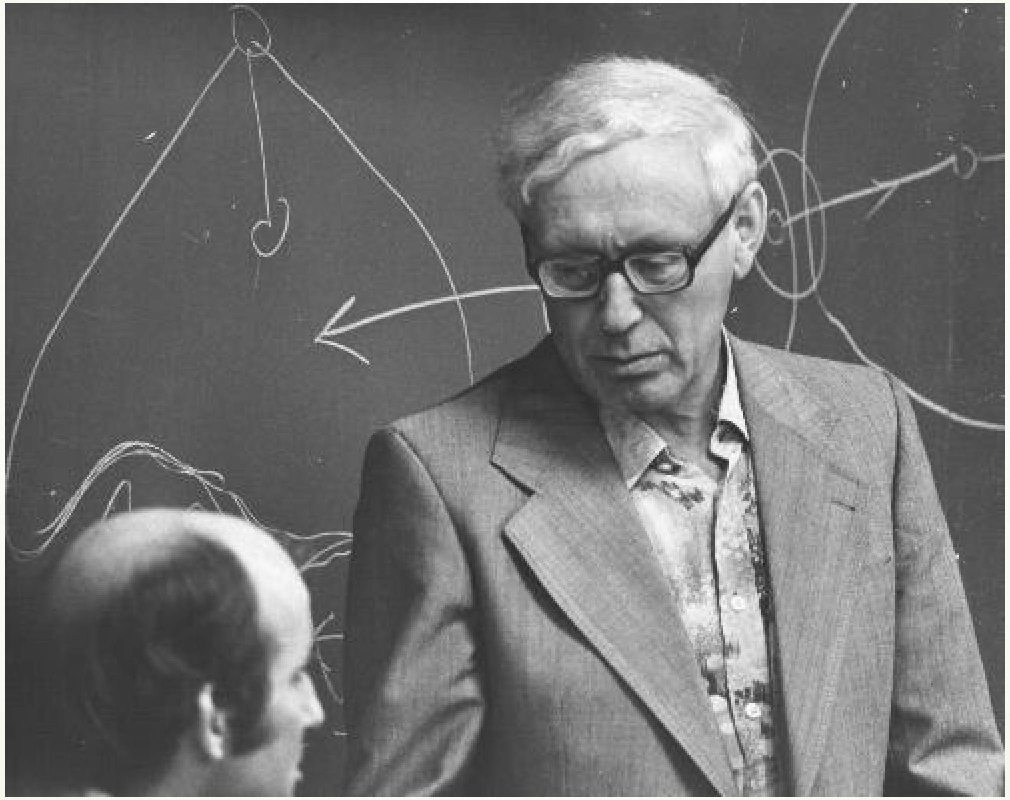
1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as

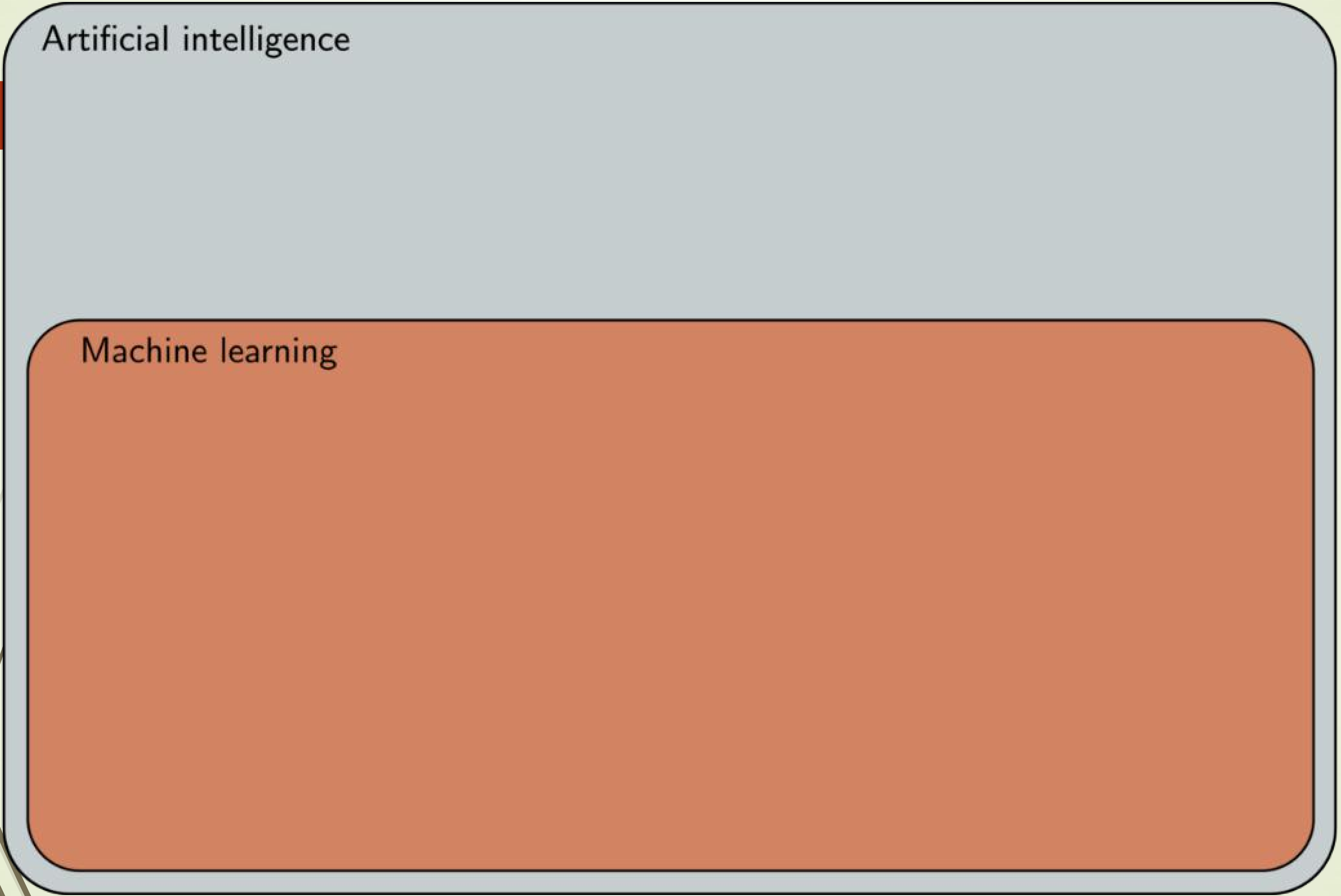
Ο προσλήπτης (perceptron) του Rosenblatt (1958)

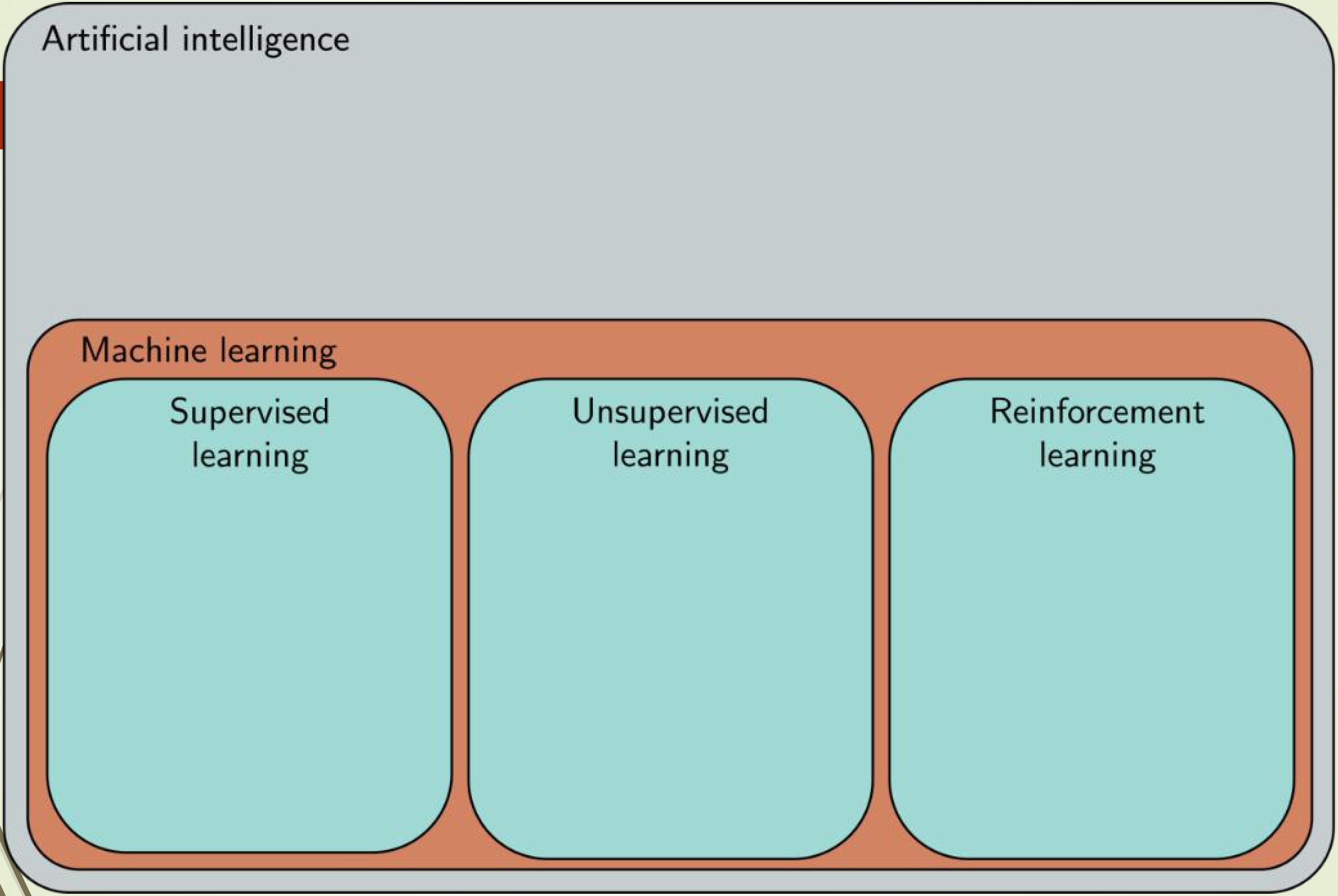


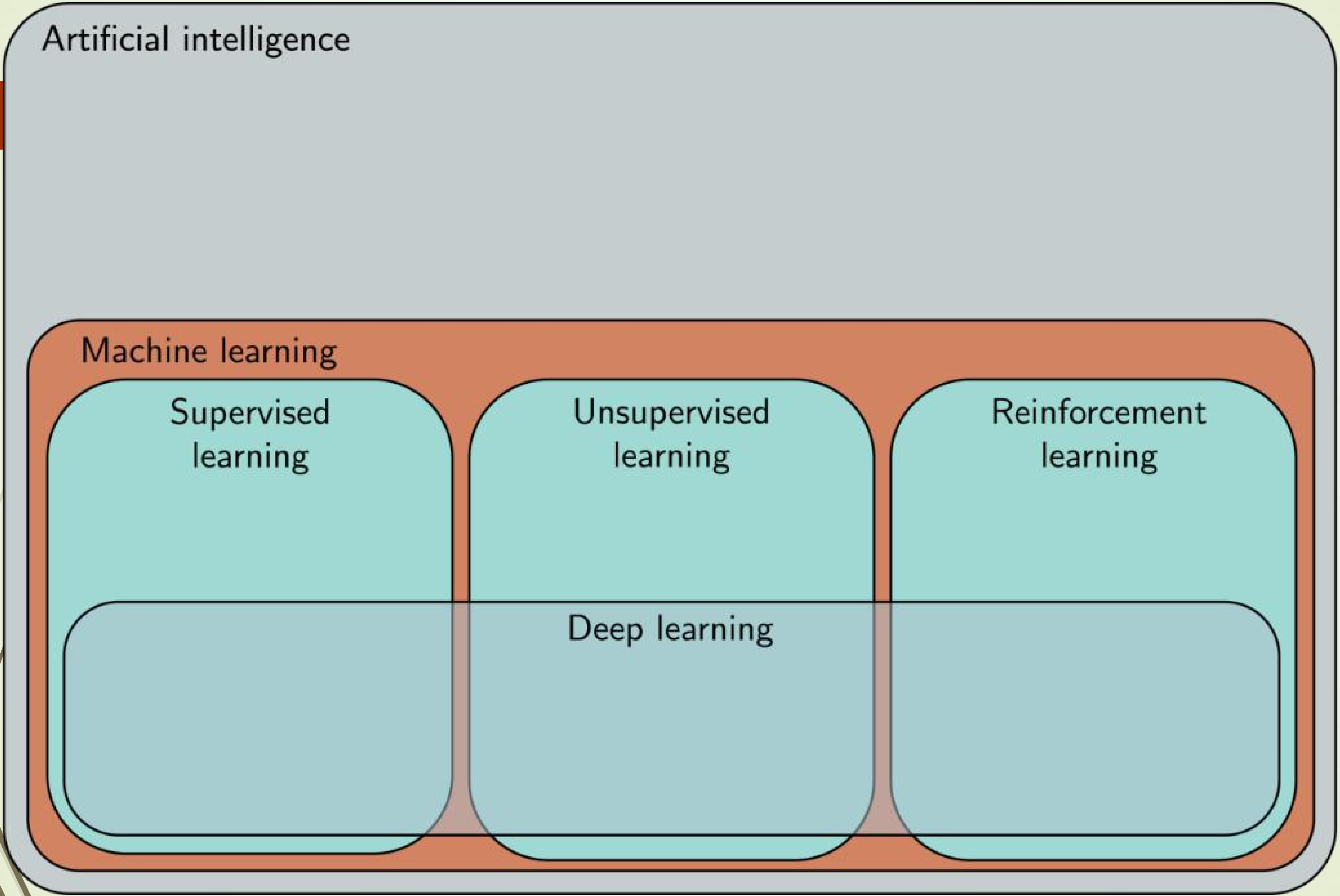
Ο “Πιονιέρος” (Пионер, ~1950 ως ~1980)



Artificial intelligence


Artificial intelligence
Machine learning

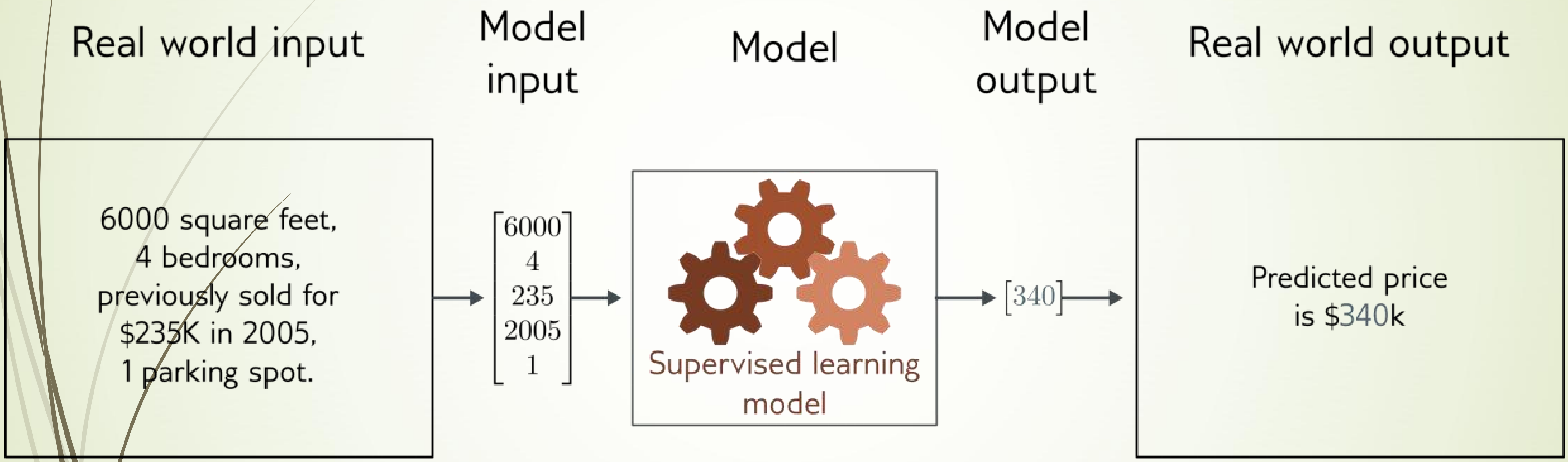

Artificial intelligence
Machine learning
Supervised learning
Unsupervised learning
Reinforcement learning


Artificial intelligence
Machine learning
Supervised learning
Unsupervised learning
Reinforcement learning
Deep learning

Μάθηση με επίβλεψη (Supervised learning)

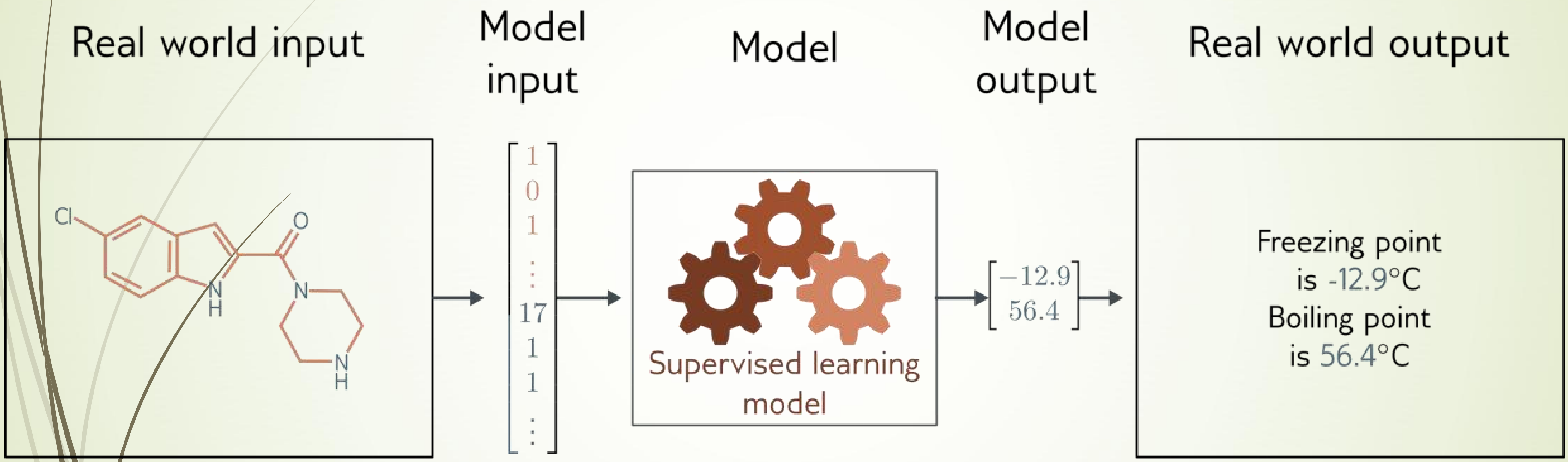
- Ορισμός μιας **απεικόνισης** της “**εισόδου**” στην “**έξοδο**”
- **Εκμάθηση** αυτής της απεικόνισης από **ζεύγη παραδειγμάτων** εισόδου-εξόδου

Regression



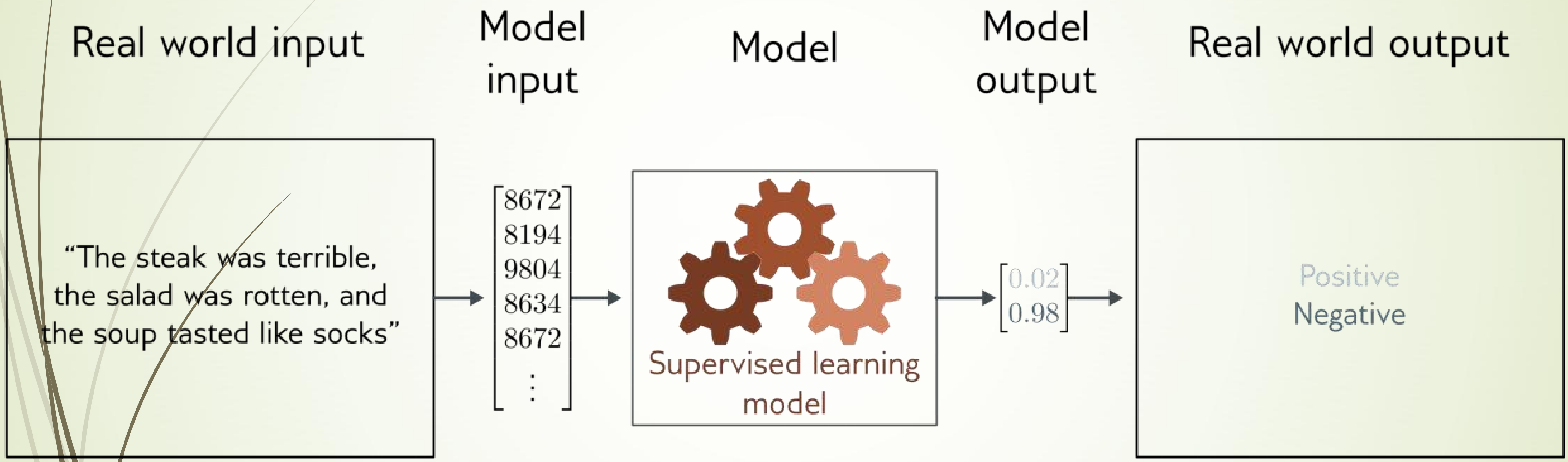
Real world input
Model input
Model
Model output
Real world output
6000 square feet, 4 bedrooms, previously sold for $235K in 2005, 1 parking spot.
6000
4
235
2005
1
Supervised learning model
[340]
Predicted price is $340k

Regression

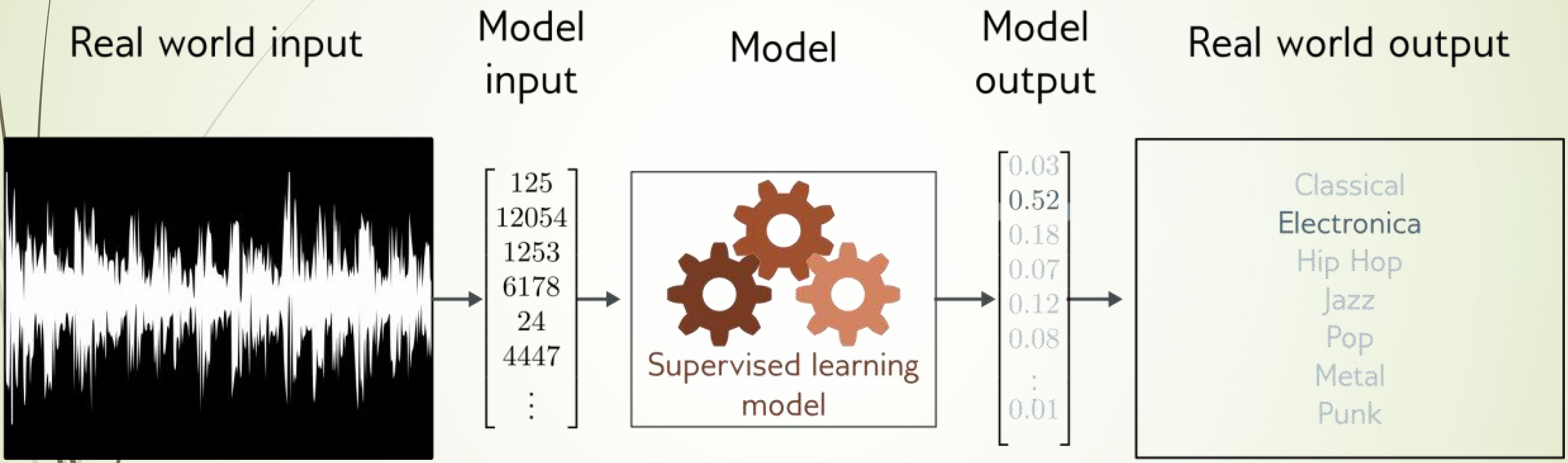


Real world input
Model input
Model
Model output
Real world output
Cl
O
N
NH
1
0
1
⋮
17
1
1
⋮
Supervised learning model
−12.9
56.4
Freezing point is -12.9°C
Boiling point is 56.4°C

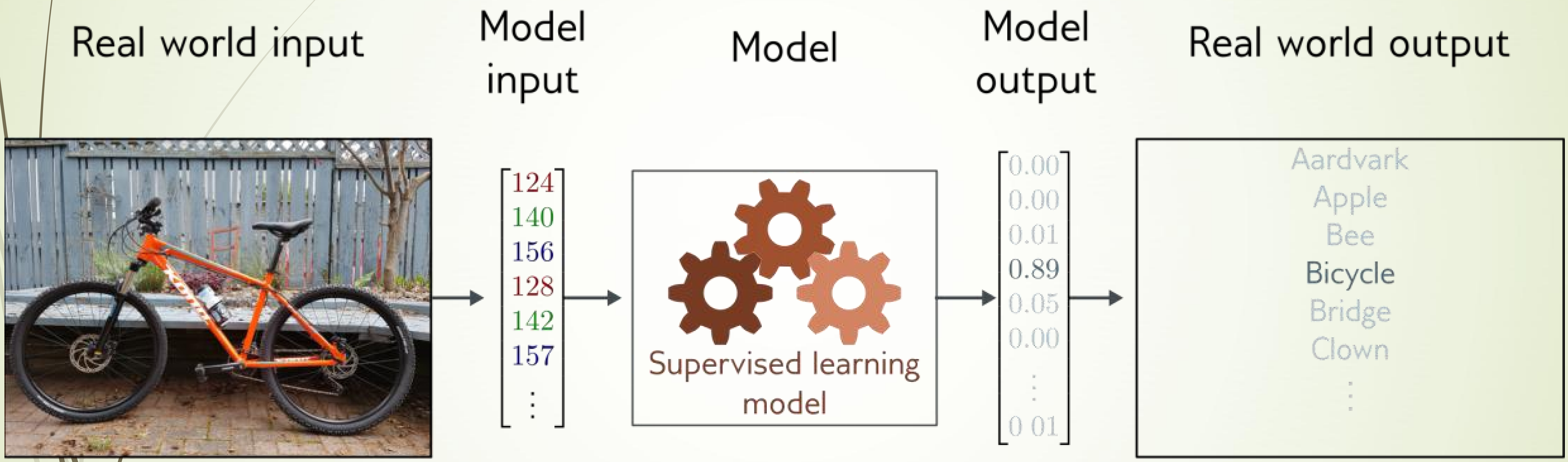
Classification



Classification



Classification



Αλλά *τι ακριβώς είναι* ένα μοντέλο επιβλεπόμενης μάθησης;

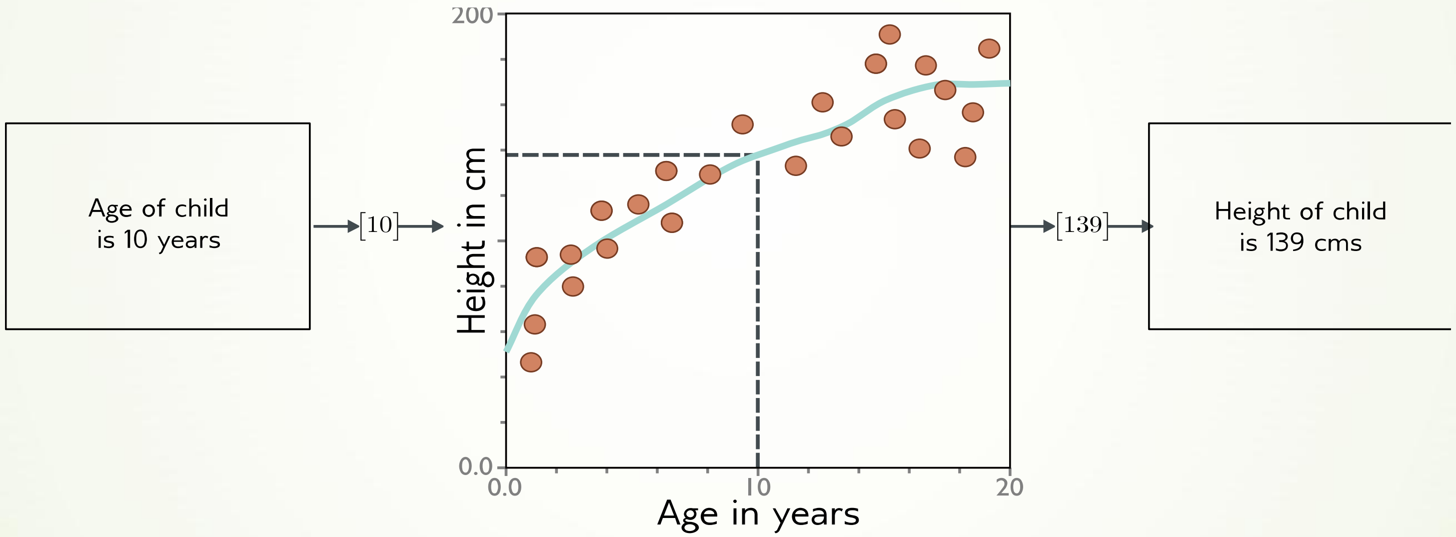
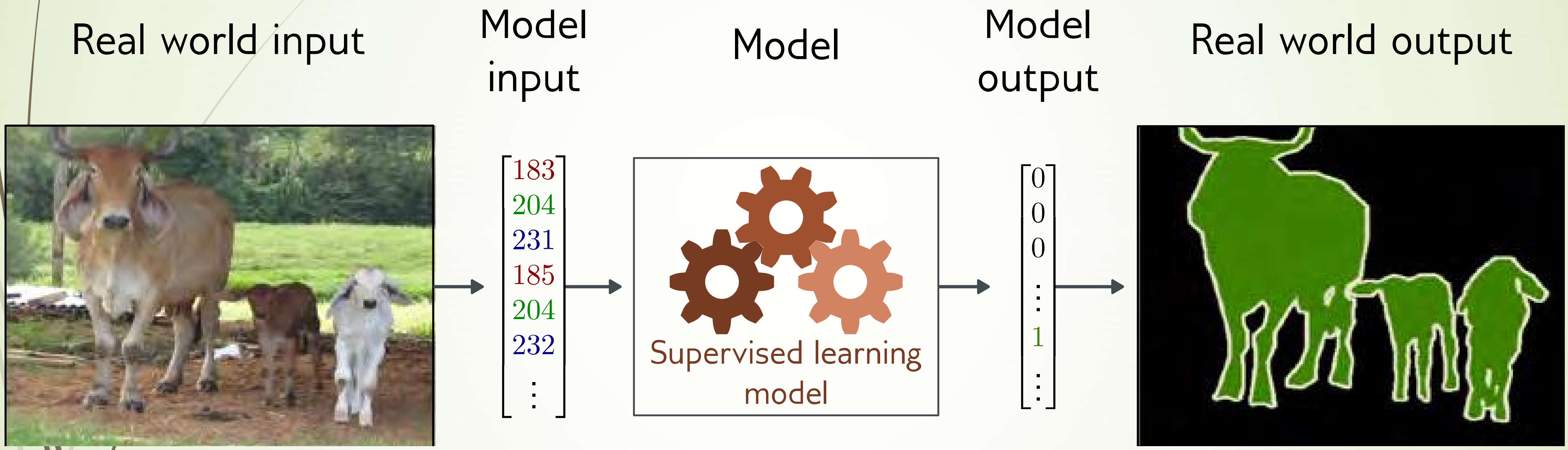
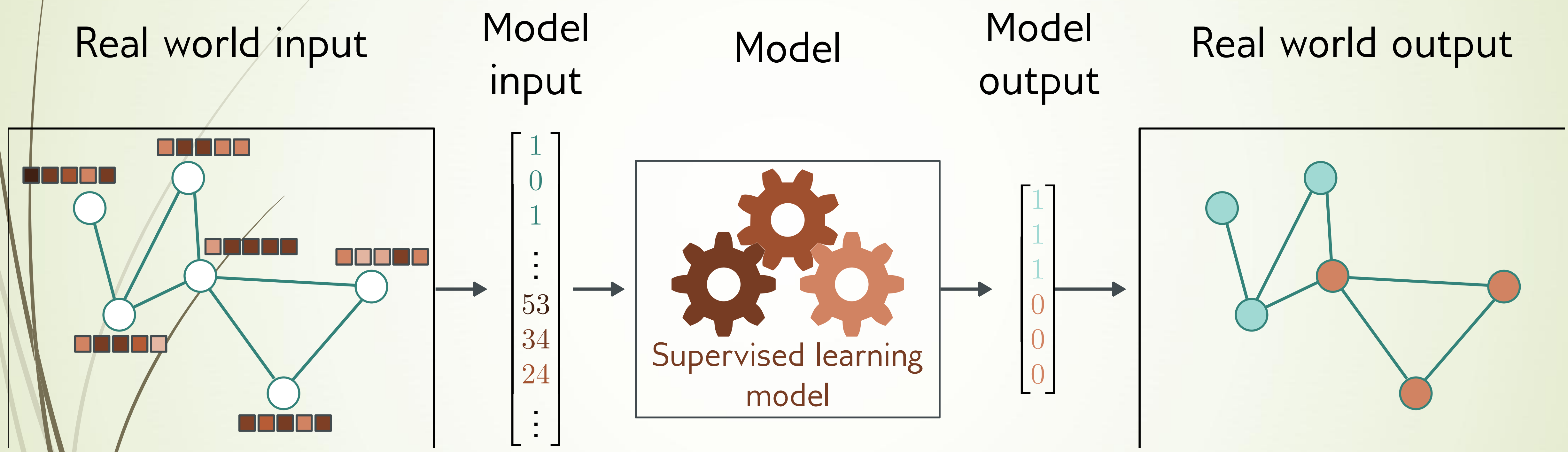


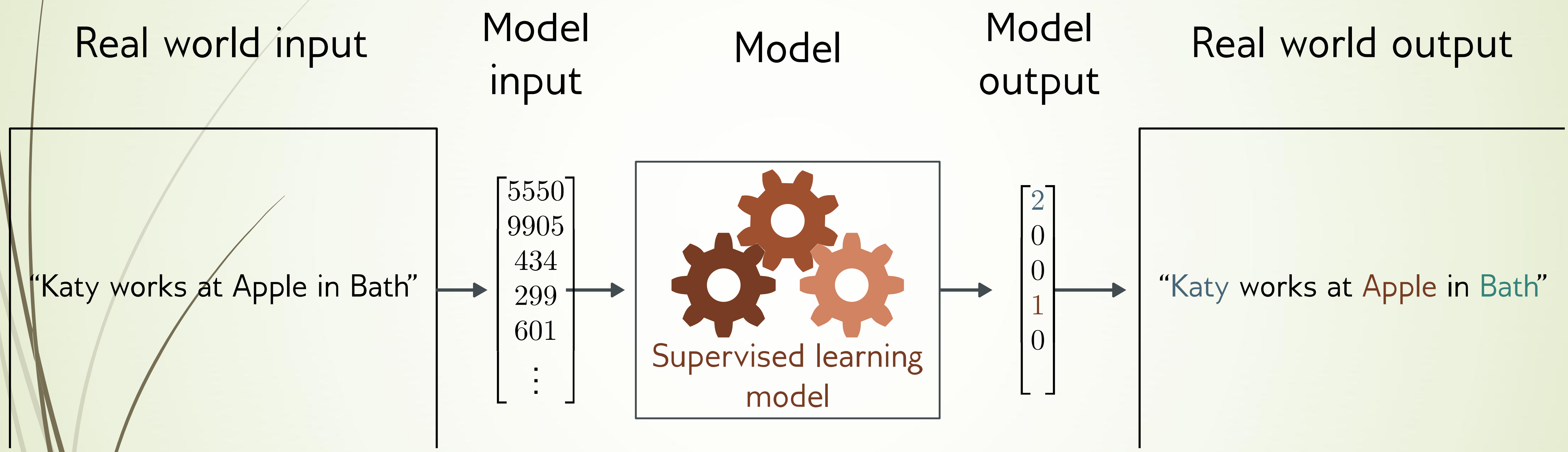
Image Segmentation



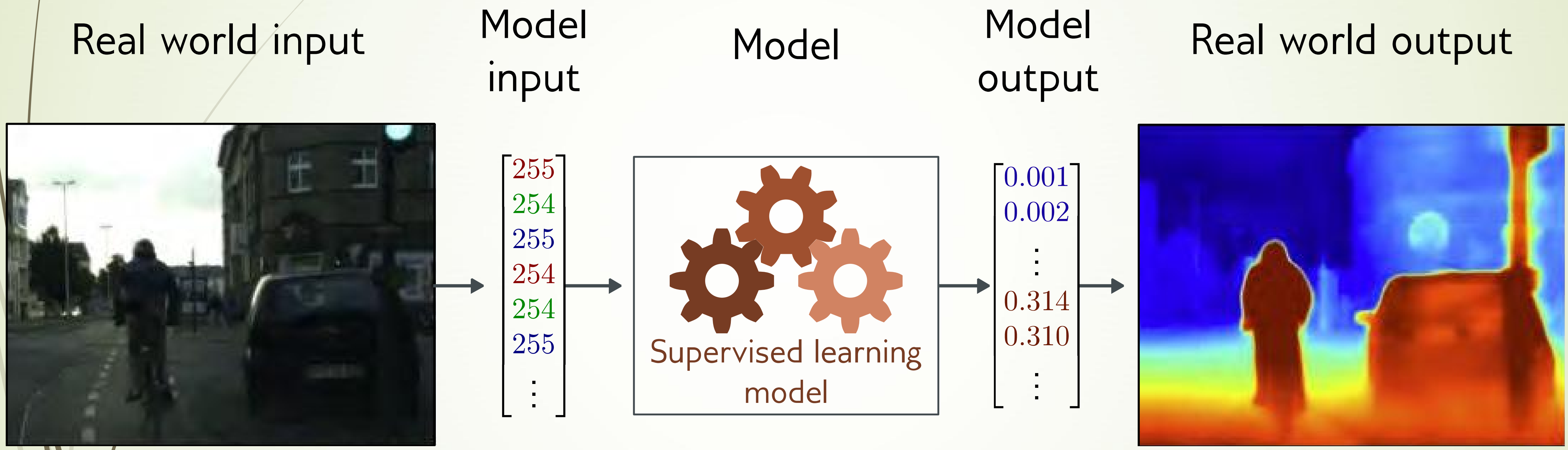
Ταξινόμηση κόμβων γράφου



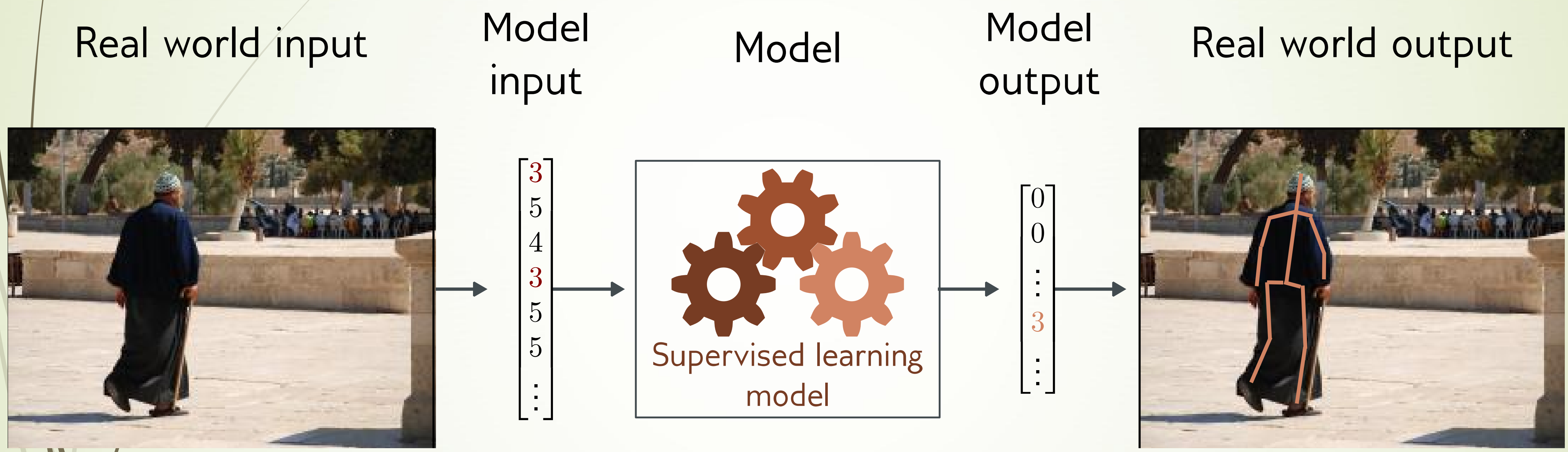
Ταξινόμηση λέξεων



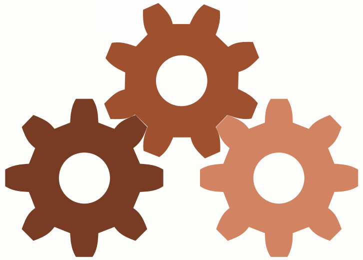
Εκτίμηση βάθους



Εκτίμηση πόζας



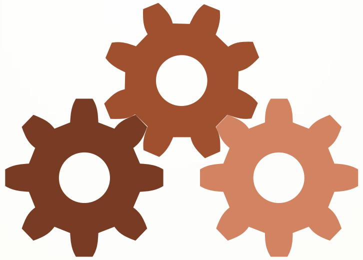
Μετάφραση

Real world input	Model input	Model	Model output	Real world output
"Skill without imagination is craftsmanship and gives us many useful objects such as wickerwork picnic baskets. Imagination without skill gives us modern art."	$\begin{bmatrix} 7800 \\ 9853 \\ 4520 \\ 4596 \\ 987 \\ \vdots \end{bmatrix}$	 Supervised learning model	$\begin{bmatrix} 6003 \\ 3689 \\ 4432 \\ 6003 \\ 2149 \\ \vdots \end{bmatrix}$	"L'habileté sans l'imagination est de l'artisanat et nous donne de nombreux objets utiles tels que des paniers de pique-nique en osier. L'imagination sans habileté nous donne l'art moderne."

Παραγωγή γραπτής περιγραφής για εικόνα

Real world input | Model input | Model | Model output | Real world output

$$\begin{bmatrix} \color{darkred}{183} \\ \color{green}{204} \\ \color{blue}{231} \\ \color{darkred}{185} \\ \color{green}{204} \\ \color{blue}{232} \\ \vdots \end{bmatrix}$$



Supervised learning model

$$\begin{bmatrix} 1 \\ 5593 \\ 7532 \\ 7924 \\ 1 \\ \vdots \end{bmatrix}$$

“A Kazakh man on a horse holding a bird of prey”

Παραγωγή εικόνας από γραπτή περιγραφή

Real world input → Model input → Model → Model output → Real world output

"Teddy bears mixing sparkling chemicals as mad scientists, in a steampunk style."

$$\begin{bmatrix} 8300 \\ 532 \\ 7676 \\ 7898 \\ 883 \\ \vdots \end{bmatrix}$$

Supervised learning model

$$\begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 1 \\ \vdots \end{bmatrix}$$

Τι κοινό έχουν τα προηγούμενα;

- Ιδιαίτερα περίπλοκες σχέσεις και συσχετίσεις μεταξύ εισόδων και εξόδων
- Μερικές φορές έχουμε πάνω από μία δόκιμες απαντήσεις / λύσεις
- Ωστόσο οι έξοδοι "οφείλουν" να ακολουθούν ορισμένους "κανόνες" (όπως και οι είσοδοι)

"A Kazakh man on a horse holding a bird of prey"

Η γλώσσα έχει γραμματικούς κανόνες

Οι φυσικές εικόνες έχουν επίσης "κανόνες"

Semantic Segmentation

Girisha, S., et al. "Semantic segmentation of UAV aerial videos using convolutional neural networks." 2019 IEEE International Conference on Artificial Intelligence and Knowledge Engineering (AIKE). IEEE, 2019.

Semantic Segmentation

Fig. 1. Example images and labels from UAVid dataset. First row shows the images captured by UAV. Second row shows the corresponding ground truth labels. Third row shows the prediction results of MS-Dilation net + PRT + FSO model as in Table 1.

Lyu, Y., Vosselman, G., Xia, G. S., Yilmaz, A., & Yang, M. Y. (2020). UAVid: A semantic segmentation dataset for UAV imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 165, 108-119.

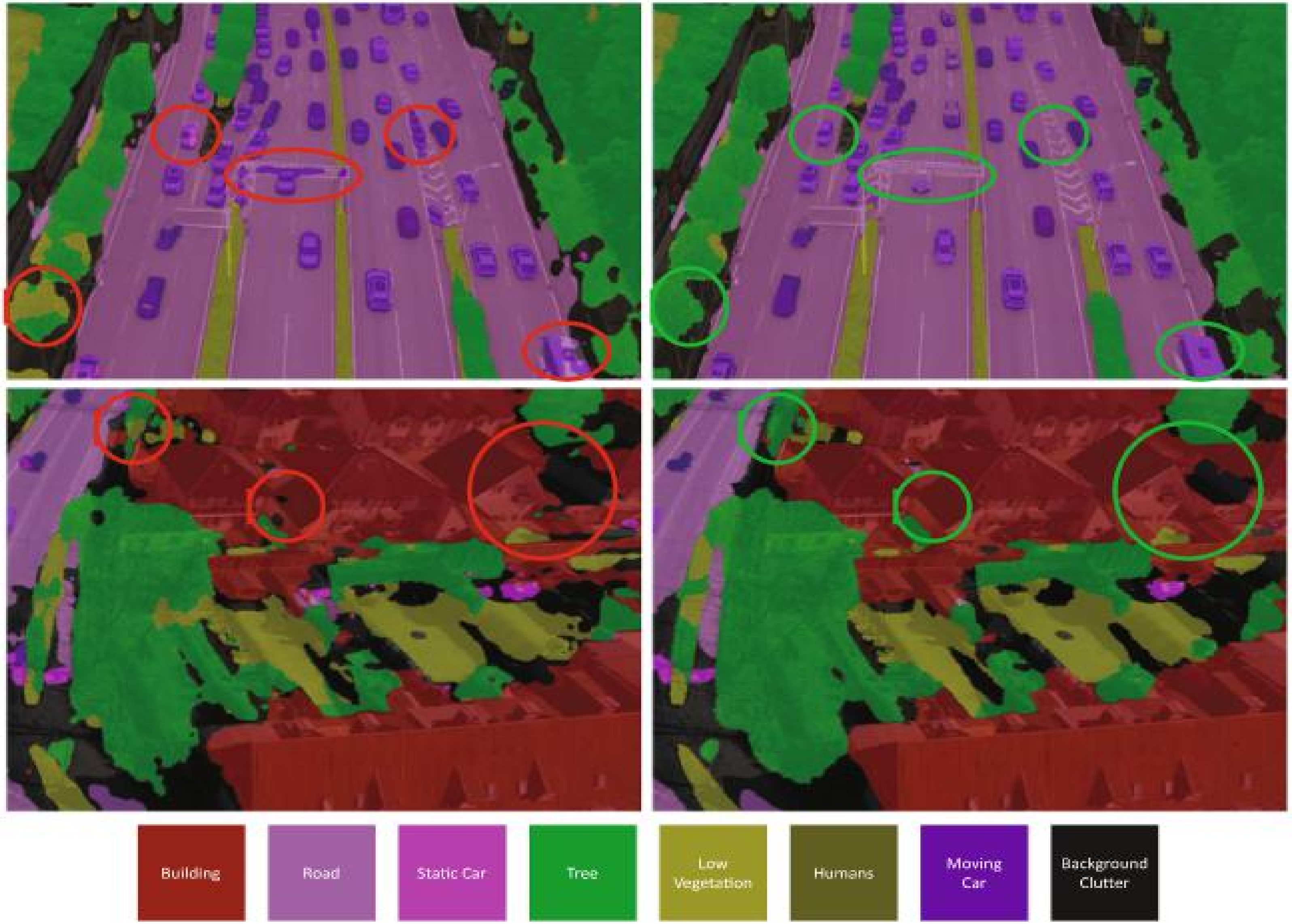
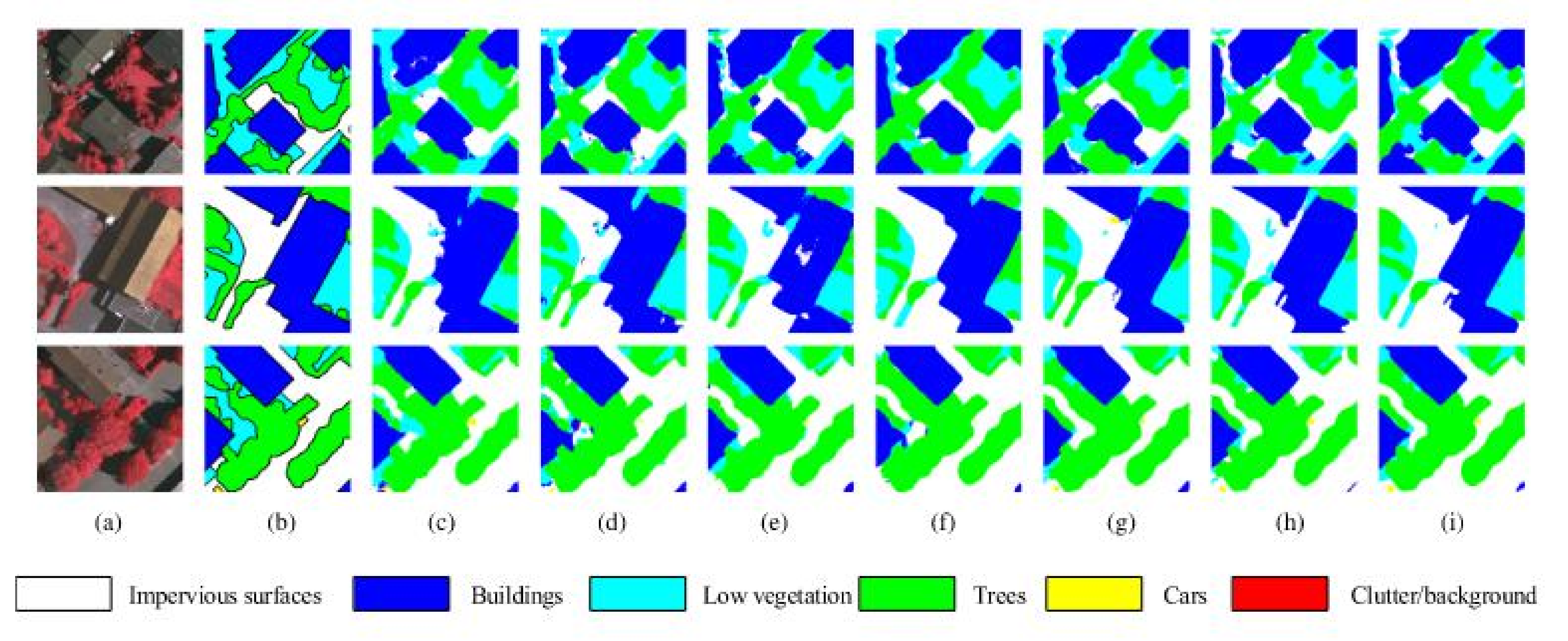


Fig. 12. Examples of spatial-temporal regularization for UAVid image semantic segmentation. The left column shows the prediction without FSO plus 3D CRF refinement. The right column shows the corresponding refined prediction with FSO plus 3D CRF refinement. The most obvious improvements are high-lighted with circles. The spatial-temporal regularization achieves a more coherent prediction for different objects.

Lyu, Y., Vosselman, G., Xia, G. S., Yilmaz, A., & Yang, M. Y. (2020). UAVid: A semantic segmentation dataset for UAV imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 165, 108-119.

Semantic Segmentation



Liu, Siyu, et al. "Light-Weight Semantic Segmentation Network for UAV Remote Sensing Images." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14 (2021): 8287-8296.

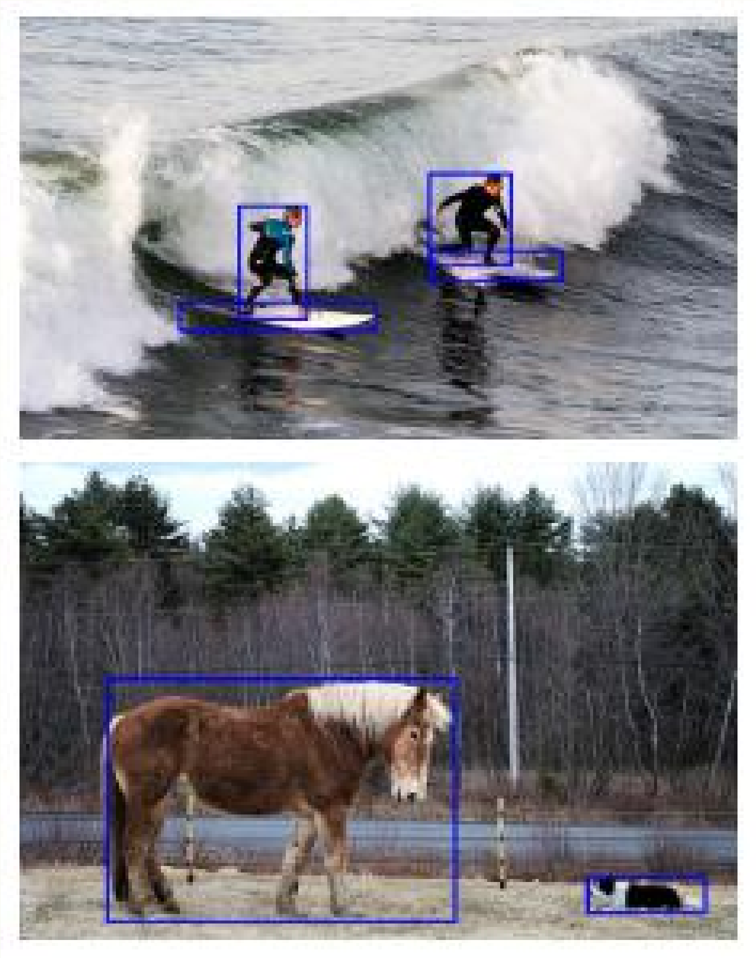
Object detection

(b) SlimYOLOv3-SPP3-95

Figure 8. Visualized detection results of SlimYOLOv3-SPP3-95 and YOLOv3-SPP3 on a challenging frame captured by our drone.

Zhang, Pengyi, Yunxin Zhong, and Xiaoqiong Li. "SlimYOLOv3: Narrower, faster and better for real-time UAV applications." Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops. 2019.

Object detection



Dimitrakopoulos, Panagiotis, Giorgos Sfikas, and Christophoros Nikou. "Variational Feature Pyramid Networks." International Conference on Machine Learning. PMLR, 2022.

Salient object detection

Megir, Victor, et al. "Salient Object Detection with Pretrained Deeplab and k-Means: Application to UAV-Captured Building Imagery." International Conference on Pattern Recognition Workshops. Springer, Cham, 2021.

Μάθηση χωρίς επίβλεψη (Unsupervised learning)

Figure 1.5 Generative models for images. Left: two images were generated from a model trained on pictures of cats. These are not real cats, but samples from a probability model. Right: two images generated from a model trained on images of buildings. Adapted from Karras et al. (2020b).

Μάθηση χωρίς επίβλεψη (Unsupervised learning)

The moon had risen by the time I reached the edge of the forest, and the light that filtered through the trees was silver and cold. I shivered, though I was not cold, and quickened my pace. I had never been so far from the village before, and I was not sure what to expect. I had been walking for hours, and I was tired and hungry. I had left in such a hurry that I had not thought to pack any food, and I had not thought to bring a weapon. I was unarmed and alone in a strange place, and I did not know what I was doing.

I had been walking for so long that I had lost all sense of time, and I had no idea how far I had come. I only knew that I had to keep going. I had to find her. I was getting close. I could feel it. She was nearby, and she was in trouble. I had to find her and help her, before it was too late.

Figure 1.6 Short story synthesized from a generative model of text data. The model describes a probability distribution that assigns a probability to every output string. Sampling from the model creates strings that follow the statistics of the training data (here, short stories) but have never been seen before.

Μάθηση χωρίς επίβλεψη (Unsupervised learning)

Figure 1.7 Inpainting. In the original image (left), the boy is obscured by metal cables. These undesirable regions (center) are removed and the generative model synthesizes a new image (right) under the constraint that the remaining pixels must stay the same. Adapted from Saharia et al. (2022a).

Μάθηση χωρίς επίβλεψη (Unsupervised learning)

I was a little nervous before my first lecture at the University of Bath. It seemed like there were hundreds of students and they looked intimidating. I stepped up to the lectern and was about to speak when something bizarre happened.

Suddenly, the room was filled with a deafening noise, like a giant roar. It was so loud that I couldn't hear anything else and I had to cover my ears. I could see the students looking around, confused and frightened. Then, as quickly as it had started, the noise stopped and the room was silent again.

I stood there for a few moments, trying to make sense of what had just happened. Then I realized that the students were all staring at me, waiting for me to say something. I tried to think of something witty or clever to say, but my mind was blank. So I just said, "Well, that was strange,' and then I started my lecture.

Figure 1.8 Conditional text synthesis. Given an initial body of text (in black), generative models of text can continue the string plausibly by synthesizing the "missing" remaining part of the string. Generated by GPT3 (Brown et al., 2020).

Μάθηση χωρίς επίβλεψη (Unsupervised learning)

Figure 1.9 Variation of the human face. The human face contains roughly 42 muscles, so it's possible to describe most of the variation in images of the same person in the same lighting with just 42 numbers. In general, datasets of images, music, and text can be described by a relatively small number of underlying variables although it is typically more difficult to tie these to particular physical mechanisms. Images from Dynamic FACES database (Holland et al., 2019).

Μάθηση χωρίς επίβλεψη (Unsupervised learning)

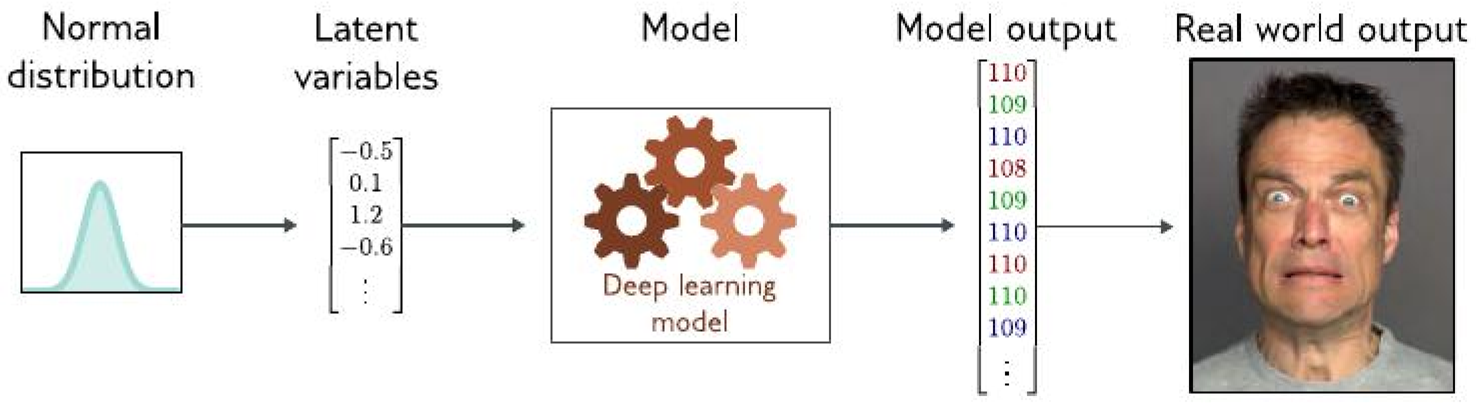


Figure 1.10 Latent variables. Many generative models use a deep learning model to describe the relationship between a low-dimensional "latent" variable and the observed high-dimensional data. The latent variables have a simple probability distribution by design. Hence, new examples can be generated by sampling from the simple distribution over the latent variables and then using the deep learning model to map the sample to the observed data space.

Μάθηση χωρίς επίβλεψη (Unsupervised learning)

Figure 1.11 Image interpolation. In each row the left and right images are real and the three images in between represent a sequence of interpolations created by a generative model. The generative models that underpin these interpolations have learned that all images can be created by a set of underlying latent variables. By finding these variables for the two real images, interpolating their values, and then using these intermediate variables to create new images, we can generate intermediate results that are both visually plausible and mix the characteristics of the two original images. Top row adapted from Sauer et al. (2022). Bottom row adapted from Ramesh et al. (2022).

Ενισχυτική μάθηση (Reinforcement learning)

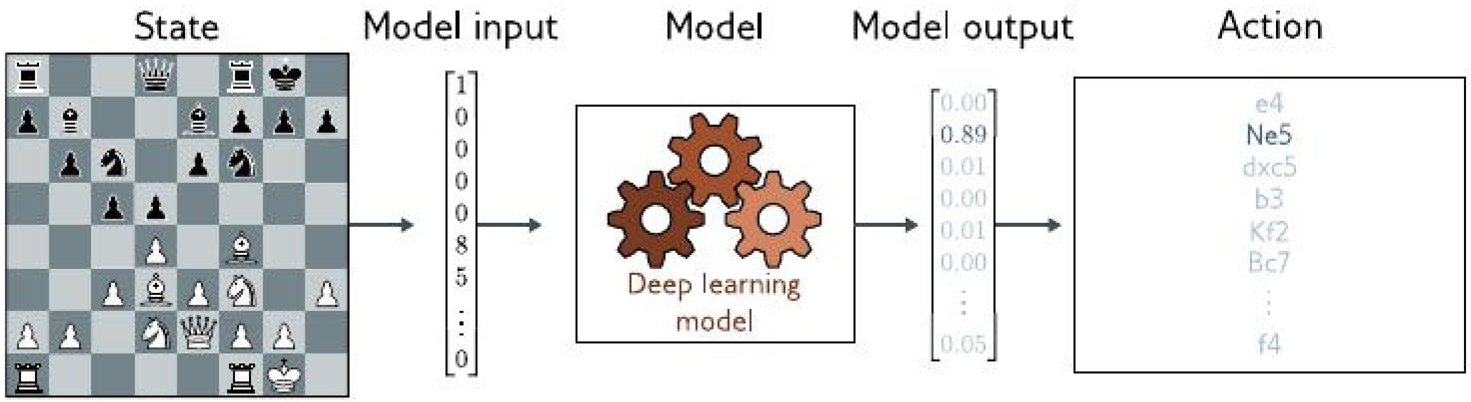


Figure 1.13 Policy networks for reinforcement learning. One way to incorporate deep neural networks into reinforcement learning is to use them to define a mapping from the state (here position on chessboard) to the actions (possible moves). This mapping is known as a *policy*. Adapted from Pablok (2017).

Επίμετρο:
Τεχνητή Νοημοσύνη και κοινωνία

- Bias and Fairness
- Explainability
- Weaponizing AI
- Concentrating power
- Existential risk
- Problem of «alignment»
- https://ethics-of-ai.mooc.fi/

Ανακεφαλαίωση

Με τη μορφή κάποιων χρήσιμων κατηγοριών

- AI vs Machine Learning vs Deep Learning
- Supervised vs Unsupervised vs Reinforcement learning
- Μοντέλο vs Μέθοδος
- Classification vs Regression
- Discrete vs Continuous

Μελέτη από το βιβλίο

- Chapter 1, «Introduction»

Στο εργαστήριο..

- Βασικά χρήσης git και github
- Βασικά μαθηματικά και Python

Ερωτήσεις ... ;;;