Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Introduction To Machine Learning

Krisztian Buza

Department of Artificial Intelligence
Eötvös Loránd University
Budapest, Hungary

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Outline

- Examples from the Everyday-Life
- Some Spectacular Success Stories
- Some of our Projects
- An Attempt to Demystify Artificial Intelligence
- Machine Learning as a Programming Paradigm
- Artificial Narrow Intelligence vs.
  Artificial General Intelligence
- What to Expect in this Course
  - Tentative Schedule, Homeworks, Exams...
- Quick Test (counts towards the final grade!)

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Examples from the Everyday-Life

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Examples from the Everyday-Life

- Recommender Systems
  - youtube, Netflix, amazon...
- ZIP codes are read by machines
- Speech recognition in cell phones, smart devices
- Self-driving vehicles

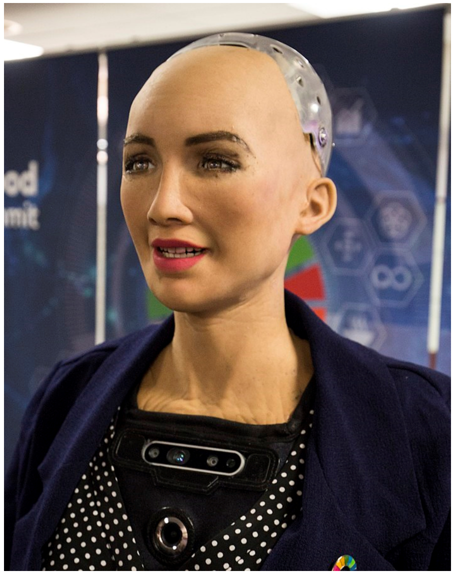
Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Some Spectacular Success Stories

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Some Spectacular Success Stories

- „Mastering the game of Go with deep neural networks and tree search“
- „Dermatologist-level classification of skin cancer with deep neural networks“
- „Clinically applicable deep learning for diagnosis and referral in retinal disease“
- „Blue Jeans and Bloody Tears“ Eurovision song
- Photo → classic painting
- „Towards reconstructing intelligible speech from the human auditory cortex“
- Sofia



ITU Pictures from Geneva, Switzerland

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Some of our own Projects

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Matrix Completion for Drug-Target Interaction Prediction

drugs

targets

-	-	1	-
-	-	-	1
-	1	-	-
-	-	-	-
-	-	-	1

drugs

targets

0.11	0.47	1.00	0.42
0.16	1.53	0.85	1.82
0.04	1.06	0.70	1.24
0.02	0.13	0.15	0.14
0.10	0.86	0.60	1.00

L. Peska, K. Buza, J. Koller (2017): Drug-target interaction prediction: A Bayesian ranking approach, Computer Methods and Programs in Biomedicine, Vol. 152, pp. 15-21

K. Buza, L. Peska (2017): ALADIN: A New Approach for Drug-Target Interaction Prediction, ECML-PKDD, Springer.

K. Buza, L. Peska (2017): Drug-target interaction prediction with Bipartite Local Models and hubness-aware regression, Neurocomputing, Volume 260, pp. 284-293

Predicted Drug-Target i

www.ksi.mff.cuni.cz/~peska/BRDTI/newDTI/predictions_ic.php

D00528 : Caffeine (Found 1 out of 36 validated targets)

Top-10 predicted interaction for: D00528 (Caffeine)

Rank	Target	Score	Validated	KEGG	DrugBank	Matador
1	hsa1080 (cystic fibrosis transmembrane conductance regulator)	0.955604	0	0	0	0
2	hsa3783 (potassium calcium-activated channel subfamily N member 4)	0.921328	0	0	0	0
3	hsa3359 (5-hydroxytryptamine receptor 3A)	0.776884	0	0	0	0
4	hsa2554 (gamma-aminobutyric acid type A receptor alpha1 subunit)	0.769335	0	0	0	0
5	hsa1141 (cholinergic receptor nicotinic beta 2 subunit)	0.749916	1	0	0	1
6	hsa3758 (potassium voltage-gated channel sub					
7	hsa3782 (potassium calcium-activated channel					
8	hsa6833 (ATP binding cassette subfamily C me					
9	hsa285242 (5-hydroxytryptamine receptor 3E)					
10	hsa9177 (5-hydroxytryptamine receptor 3B)					

Journal List › Clin Pharmacol › v.8; 2016 › PMC5036583

Dovepress
open access to scientific and medical research
Clinical Pharmacology: Advances and Applications
Dove Medical Press | This Article | Suscribe | Submit a Manuscript | Search | Follow

Clin Pharmacol

Clin Pharmacol. 2016; 8: 127–140.
Published online 2016 Sep 21. doi: [10.2147/CPAA.S100759]

PMCID: PMC5036583
PMID: 27703398

Cystic fibrosis transmembrane conductance regulator modulators in cystic fibrosis: current perspectives

Béla Z Schmidt,[1] Jérémy B Haaf,[2] Teresinha Leal,[2] and Sabrina Noel[2]

Author information ▶ Copyright and License information ▶ Disclaimer

This article has been cited by other articles in PMC.

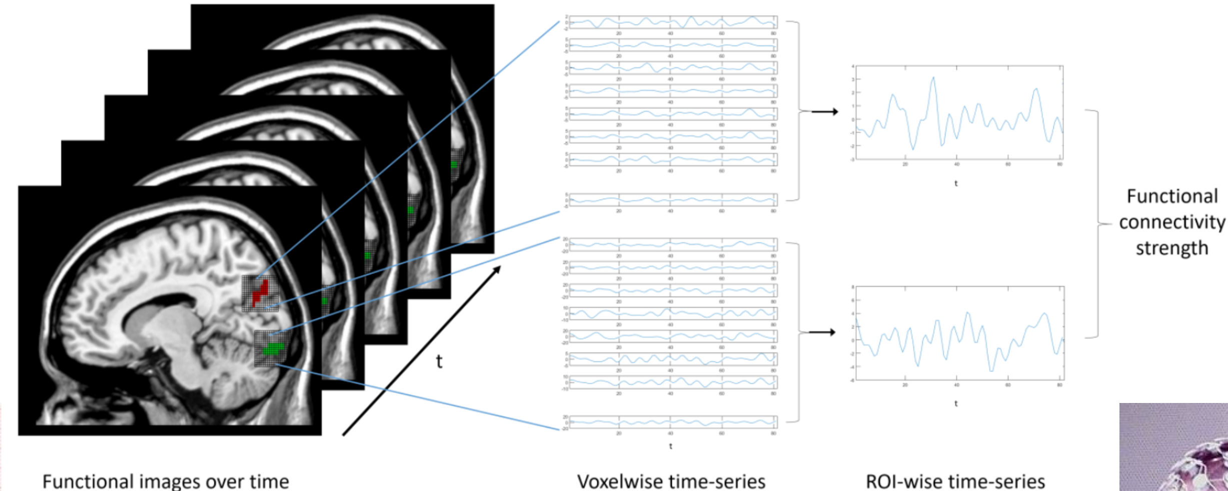
Abstract

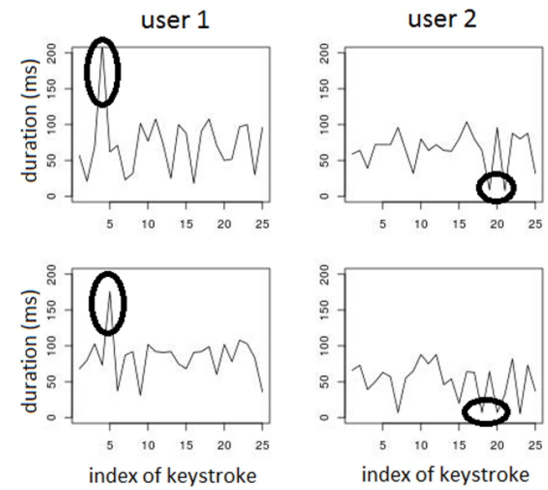
Go to:

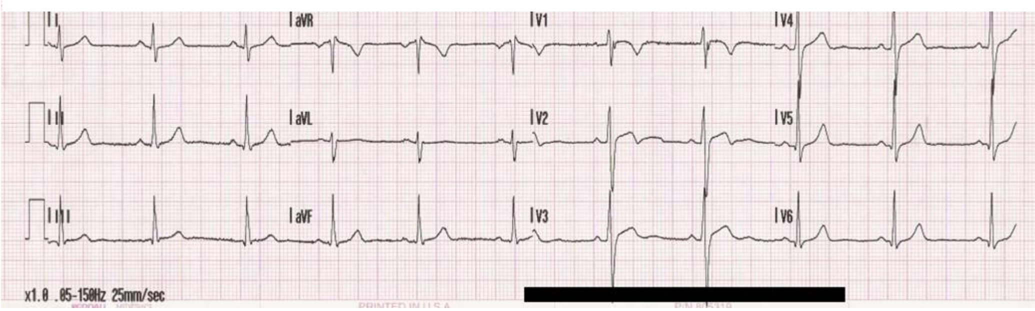
Mutations of the *CFTR* gene cause cystic fibrosis (CF), the most common recessive monogenic disease worldwide. These mutations alter the synthesis, processing, function, or half-life of CFTR, the main chloride channel expressed in the apical membrane of epithelial cells in the airway, intestine, pancreas, and reproductive tract. Lung disease is the most critical manifestation of CF. It is characterized by airway

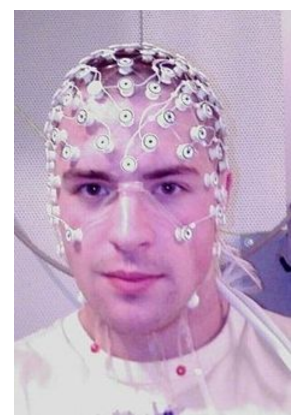
Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Time Series Classification









Images in the bottom, from left to right:
By MoodyGroove - 2007-01-24 (original upload date) Original uploader was MoodyGroove at en.wikipedia, Public Domain, https://commons.wikimedia.org/w/index.php?curid=5266589
By Thuglas at English Wikipedia - Transferred from en.wikipedia to Commons by Sreejithk2000 using CommonsHelper., Public Domain, https://commons.wikimedia.org/w/index.php?curid=10827060
By JSquish - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=16181727

K. Buza, L. Peska (2019): Individualized Warping Window Size for Dynamic Time Warping, Time Series Workshop at the 36th International Conference on Machine Learning

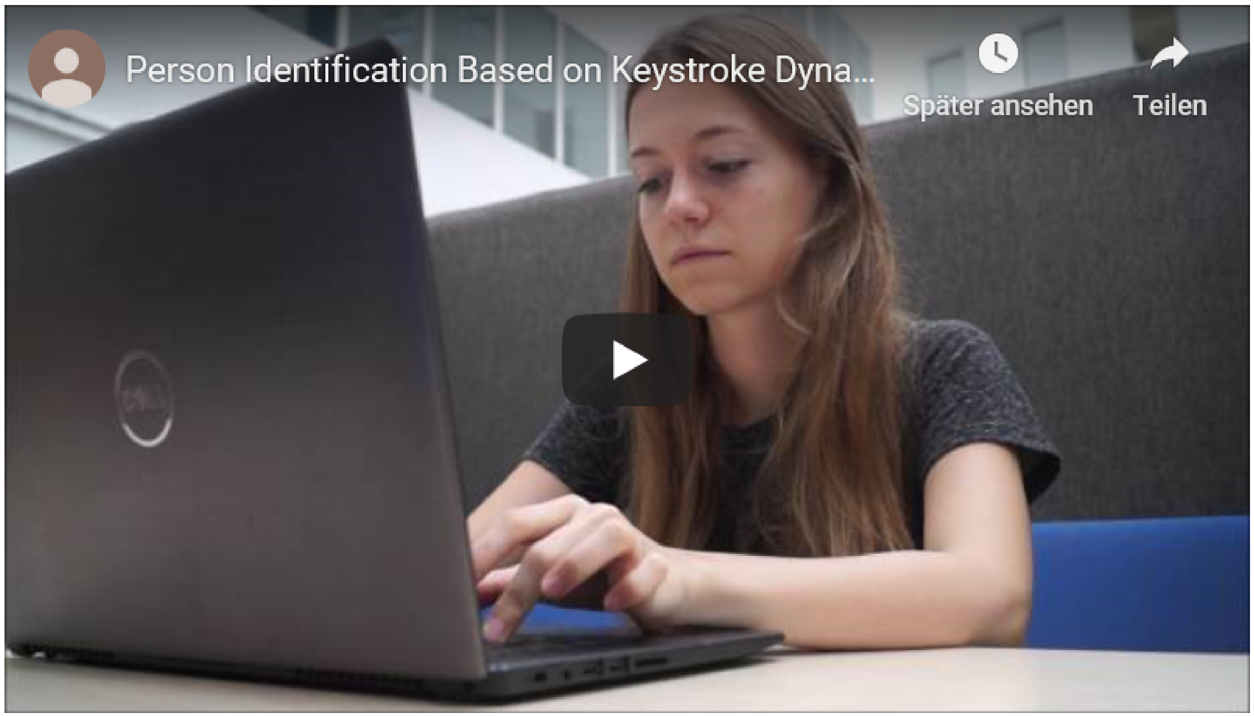
K. Buza (2018): Time Series Classification and its Applications, tutorial at the 8th International Conference on Web Intelligence, Mining and Semantics.

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

www.biointelligence.hu/typing.html

Resources Related to Keystroke Dynamics and Person Identification

Short video on person identification based on keystroke dynamics



M. Antal, L.Z. Szabó, I. László (2015): Keystroke dynamics on android platform, Procedia Technology, 19, pp. 820–826.

Z. Farou, K. Buza (2019): The Warping Window Size Effects the Accuracy of Person Identification based on Keystroke Dynamics, Computational Intelligence and Data Mining workshop at the „Information technologies - Applications and Theory“ Conference

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

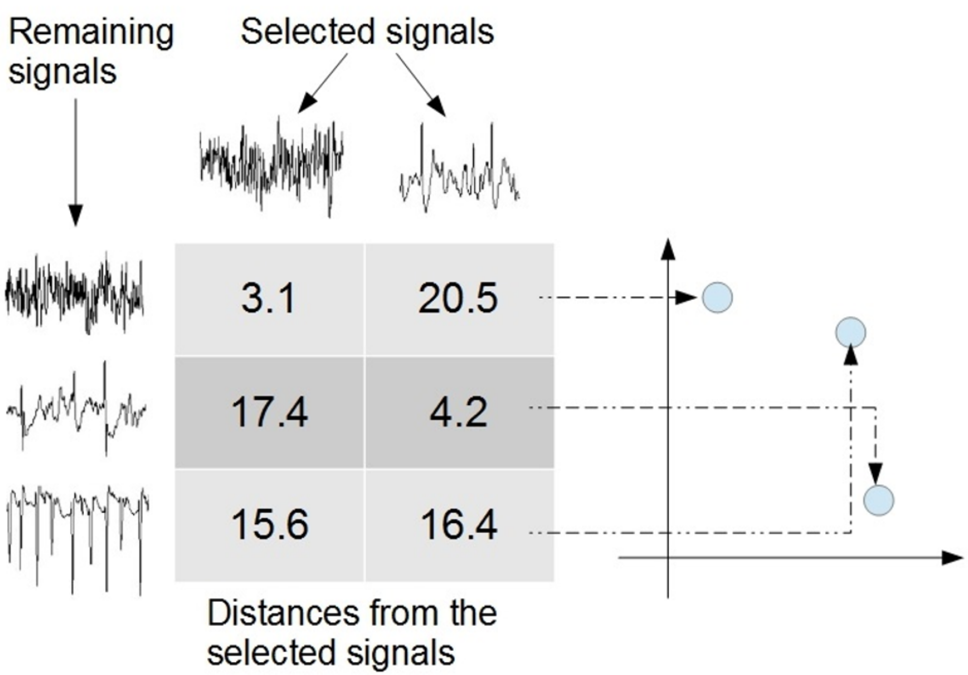
Estimation of UPDRS score based on voice recordings

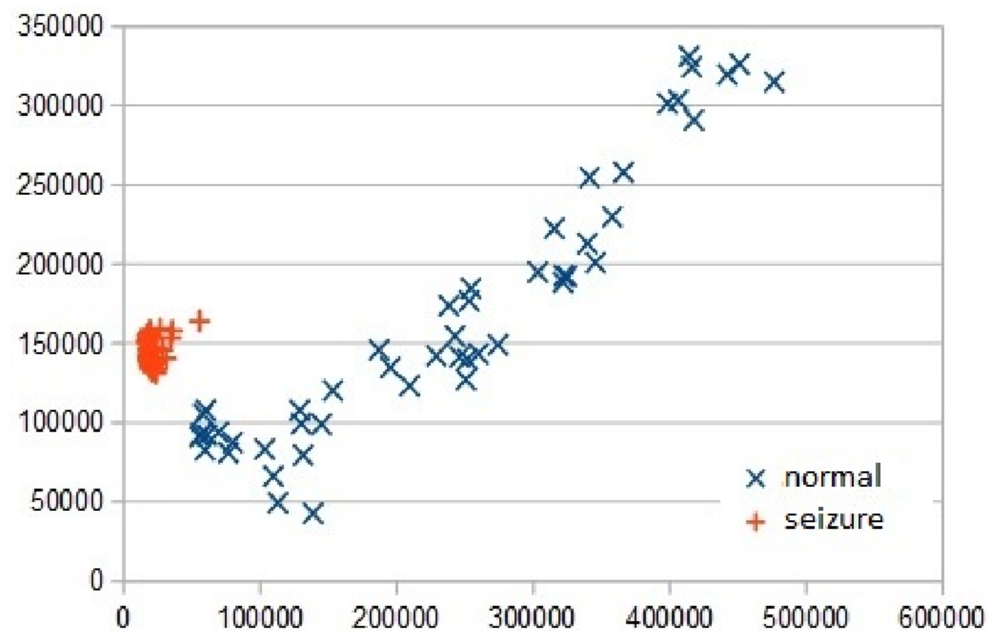
K. Buza, N.Á. Varga (2016):
ParkinsoNET: Estimation of UPDRS Score using Hubness-aware Feed-Forward Neural Networks, Applied Artificial Intelligence, special issue on Intelligent methods applied to healthcare information systems

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Classification of Brain Activity Data

- Electroencephalograph (EEG) data
- Logistic regression using DTW-distance from randomly selected time series as features

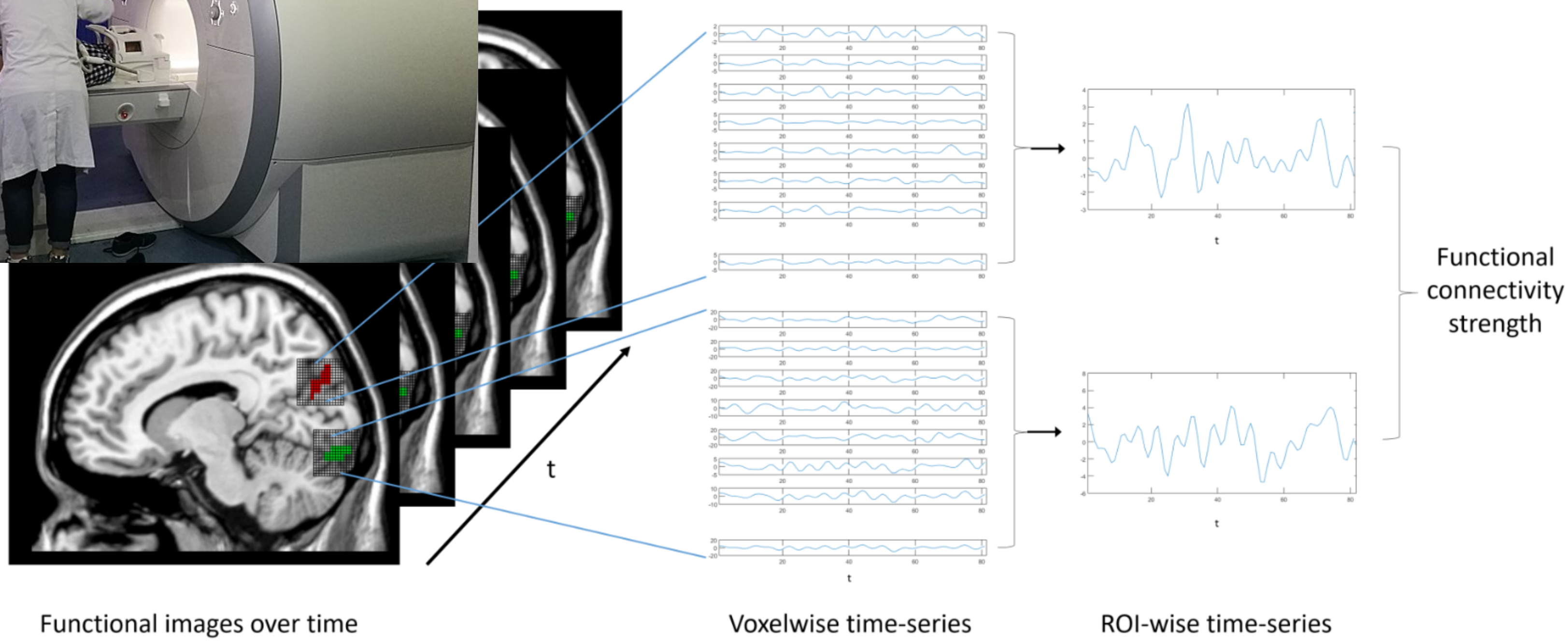




K. Buza, J. Koller, K. Marussy (2015): PROCESS: Projection-Based Classification of Electroencephalograph Signals, ICAISC, LNCS Vol. 9120, pp. 91-100, Springer.

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Classification of Brain Imaging Data



R. J. Meszlényi, K. Buza, Z. Vidnyánszky (2017): Resting State fMRI Functional Connectivity-Based Classification Using a Convolutional Neural Network Architecture, Frontiers in Neuroinformatics

A. Szenkovits, R. Meszlényi, K. Buza, N. Gaskó, R.I. Lung, M. Suciu (2018): Feature Selection with a Genetic Algorithm for Classification of Brain Imaging Data, in: Advances in Feature Selection for Data and Pattern Recognition, Springer

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

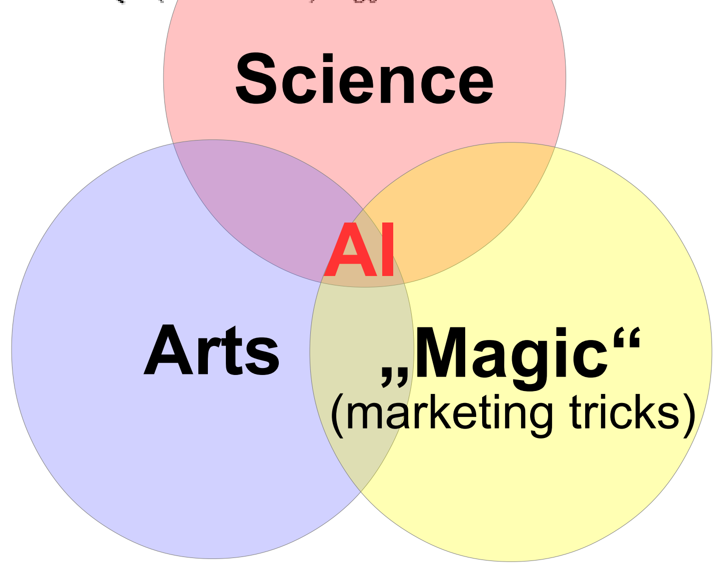
An Attempt to Demystify Artificial Intelligence

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How Artificial Intelligence Appears in Popular News

(personal opinion)

$$\frac{\partial \mathrm{err}(f_\theta, x, y)}{\partial \theta} = \begin{cases} 0 & \text{if } f_\theta(x) > +1 \text{ and } y = +1 \\ 0 & \text{if } f_\theta(x) < -1 \text{ and } y = -1 \\ 2\big(f_\theta(x) - y\big)\frac{\partial f_\theta(x)}{\partial \theta} & \text{otherwise.} \end{cases}$$

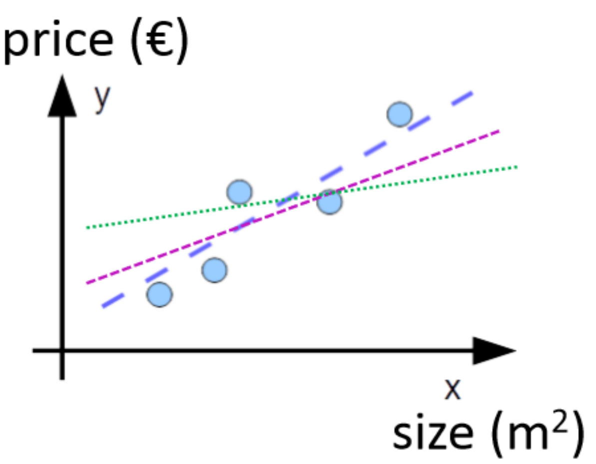


Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

What is behind „Intelligent“ Behavior in Engineering Applications?

high <u>computational power,</u> often coupled with a <u>large database</u>

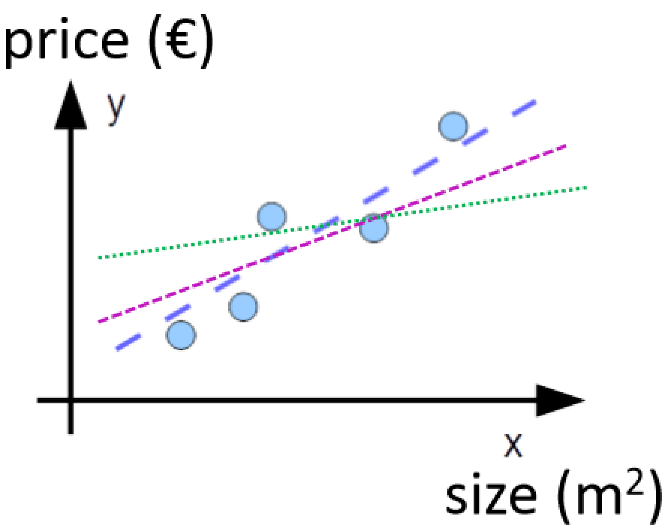
<u>Mathematical model</u> with many parameters

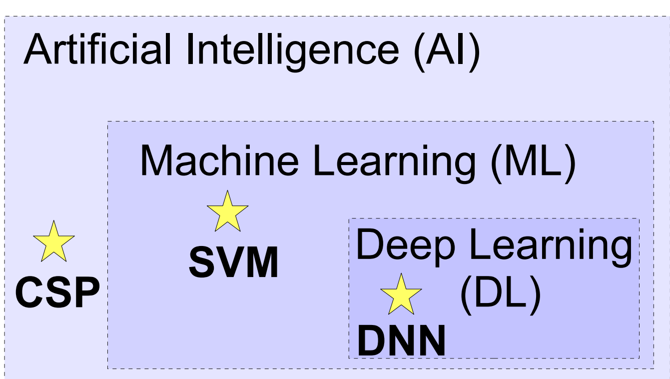


Rubic cube in the top left: from Wikipedia, author: „Alvaro qc“
[CC BY 3.0 (https://creativecommons.org/licenses/by/3.0)]
Server in the bottom right: from Wikipedia, Jfreyre – own work,
CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=2817411

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Machine Learning: an Extremely Simple Example

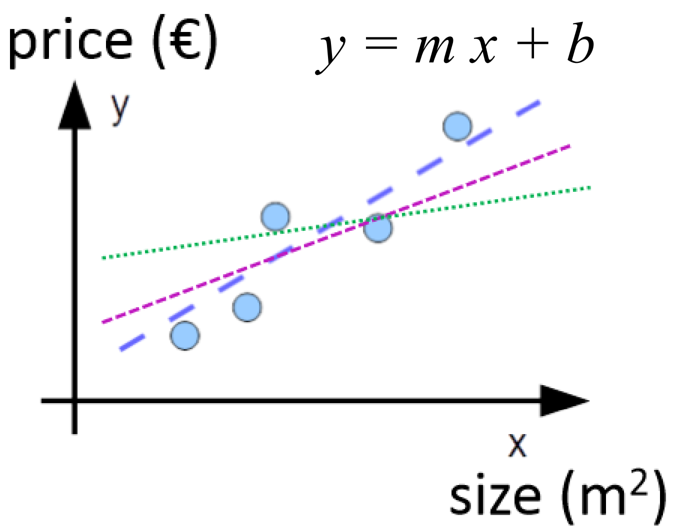




- Mathematical model:
  $y = m\,x + b$
- Parameters of the model: m, b
- *Learning* refers to finding the model parameters in an automated way
- This is true for more complex models as well (e.g. neural networks)
- In realistic applications, the number of model parametes is high (up to millions)
- Recent success stories: recognition of skin cancer, Go, self-driving cars...

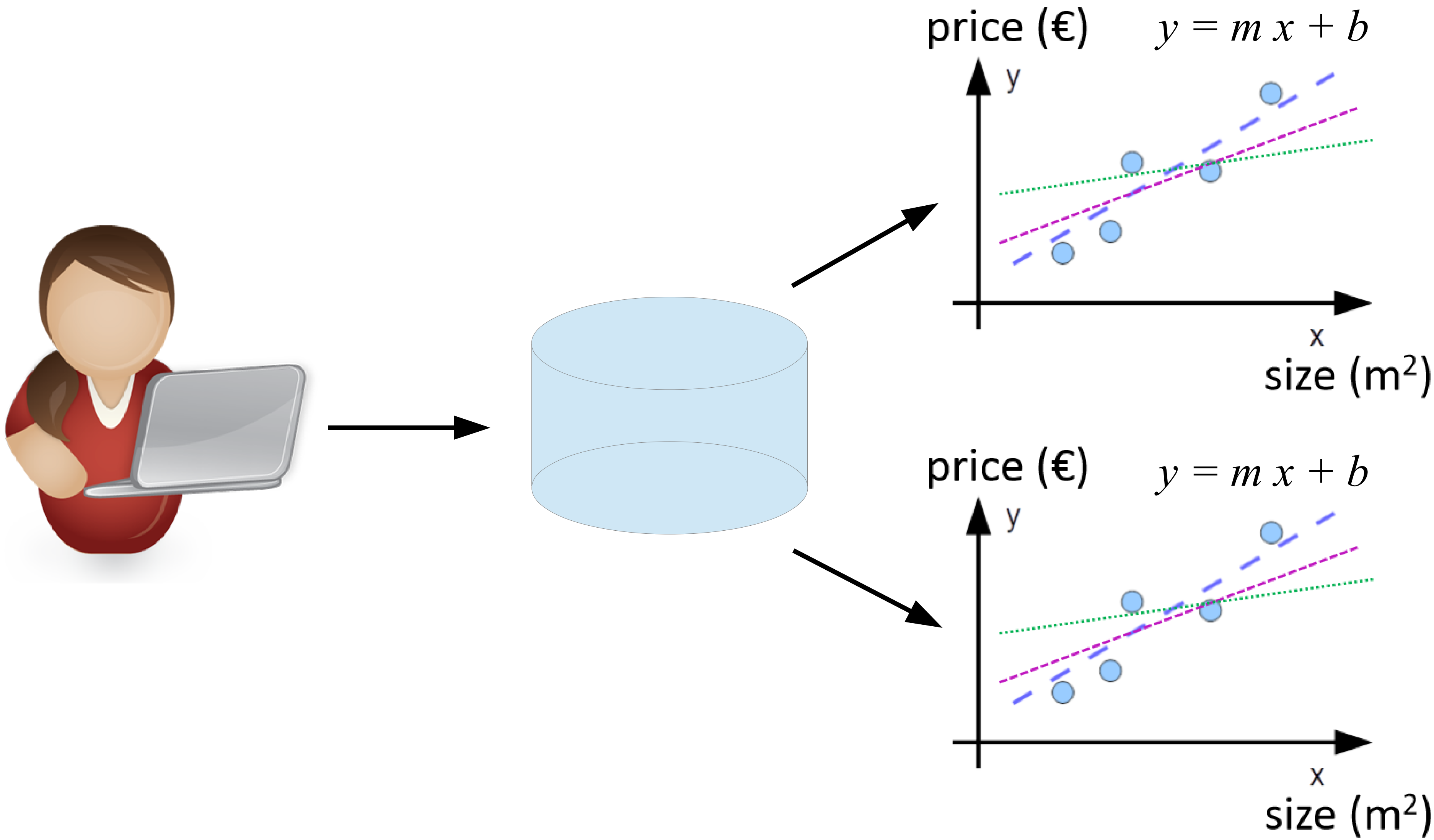
Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How do we „teach“ the machine?



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

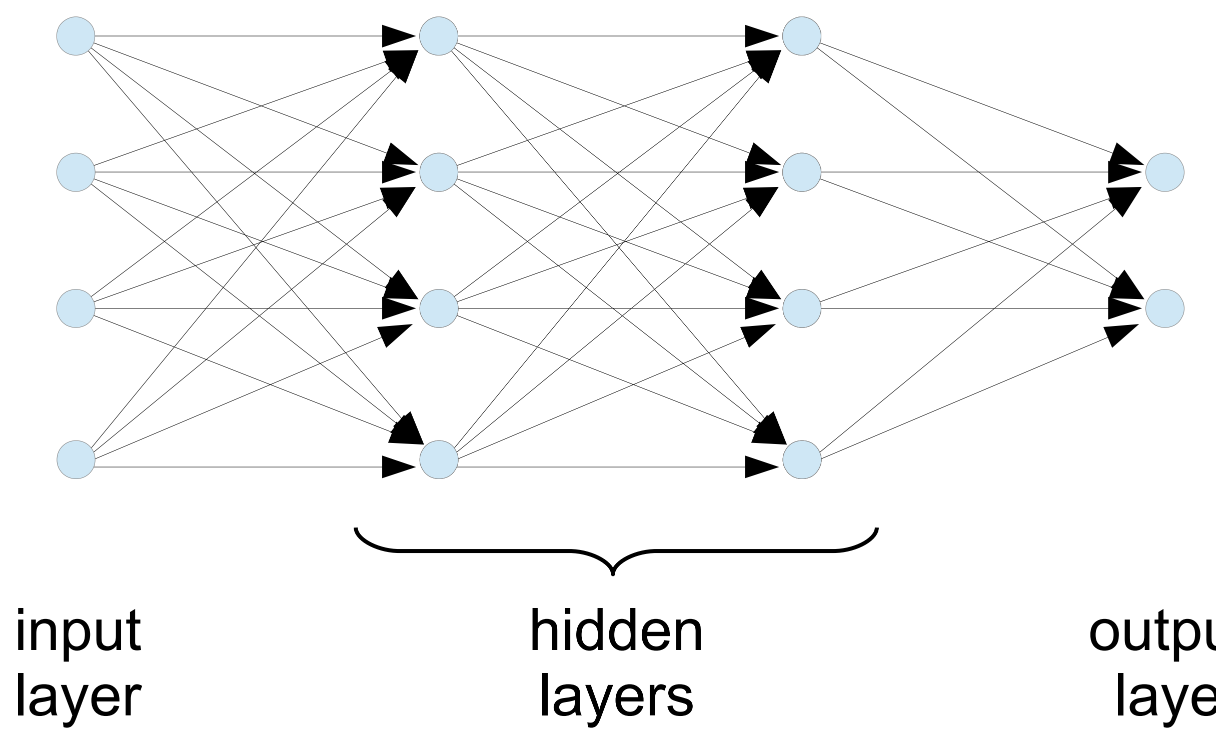
How do we „teach“ the machine?



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

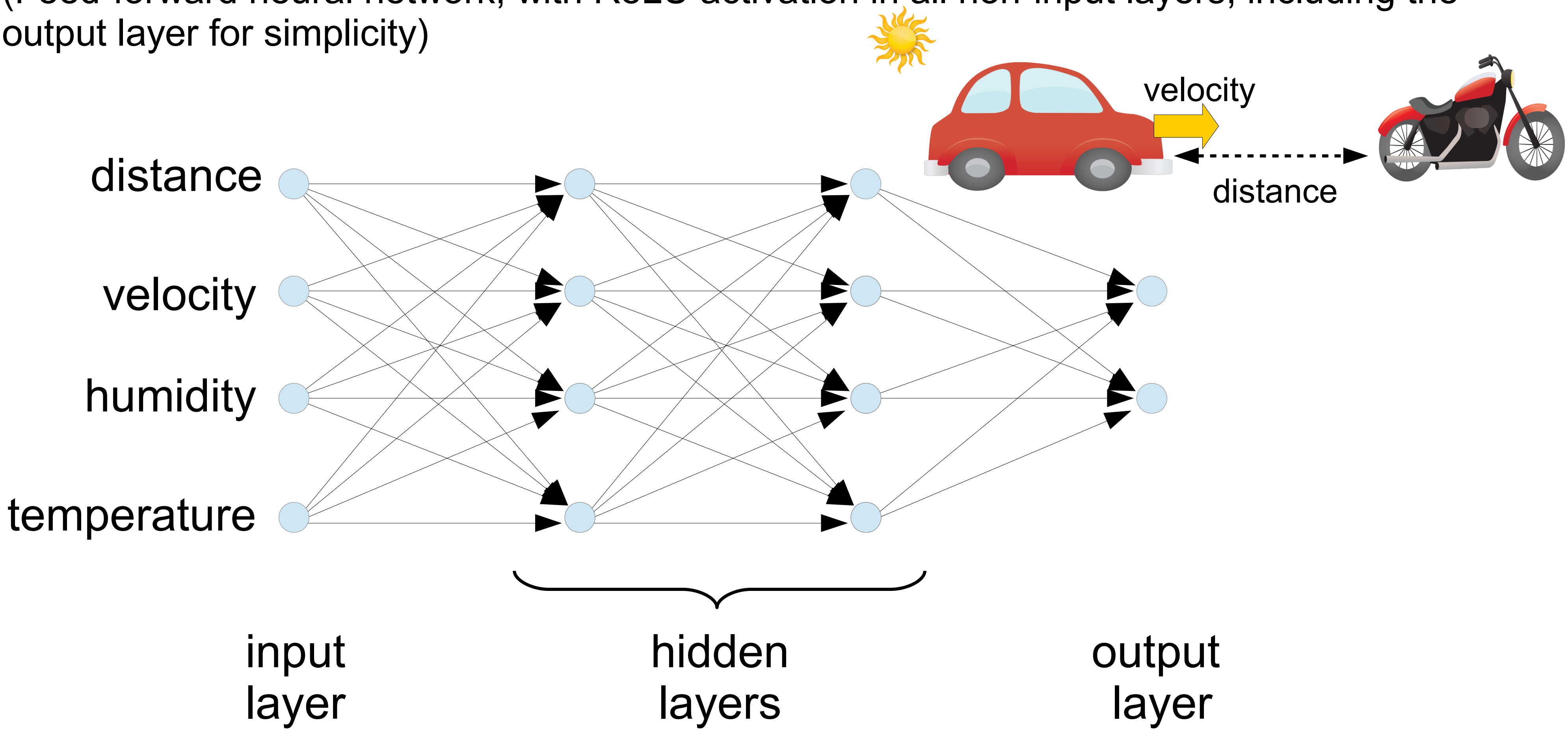
(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

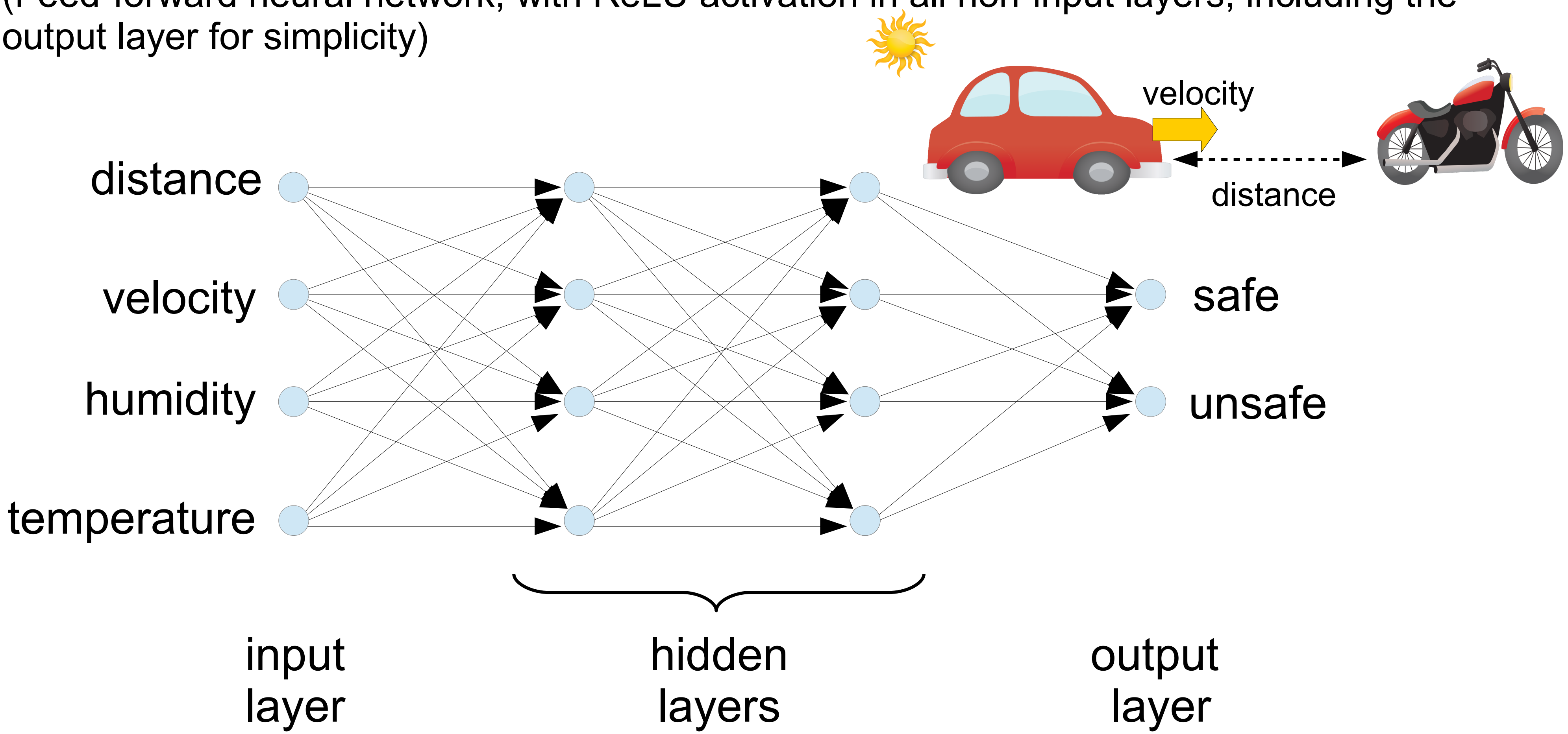
(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

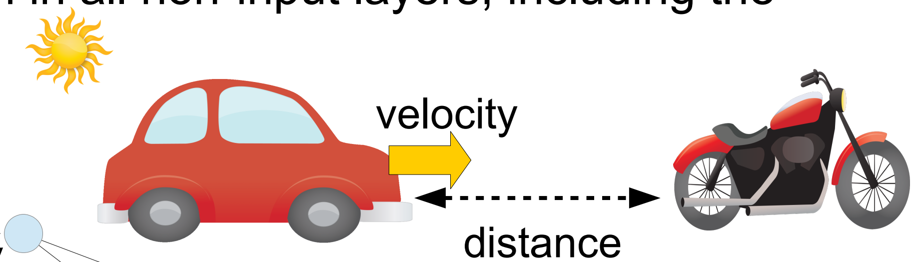
(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)

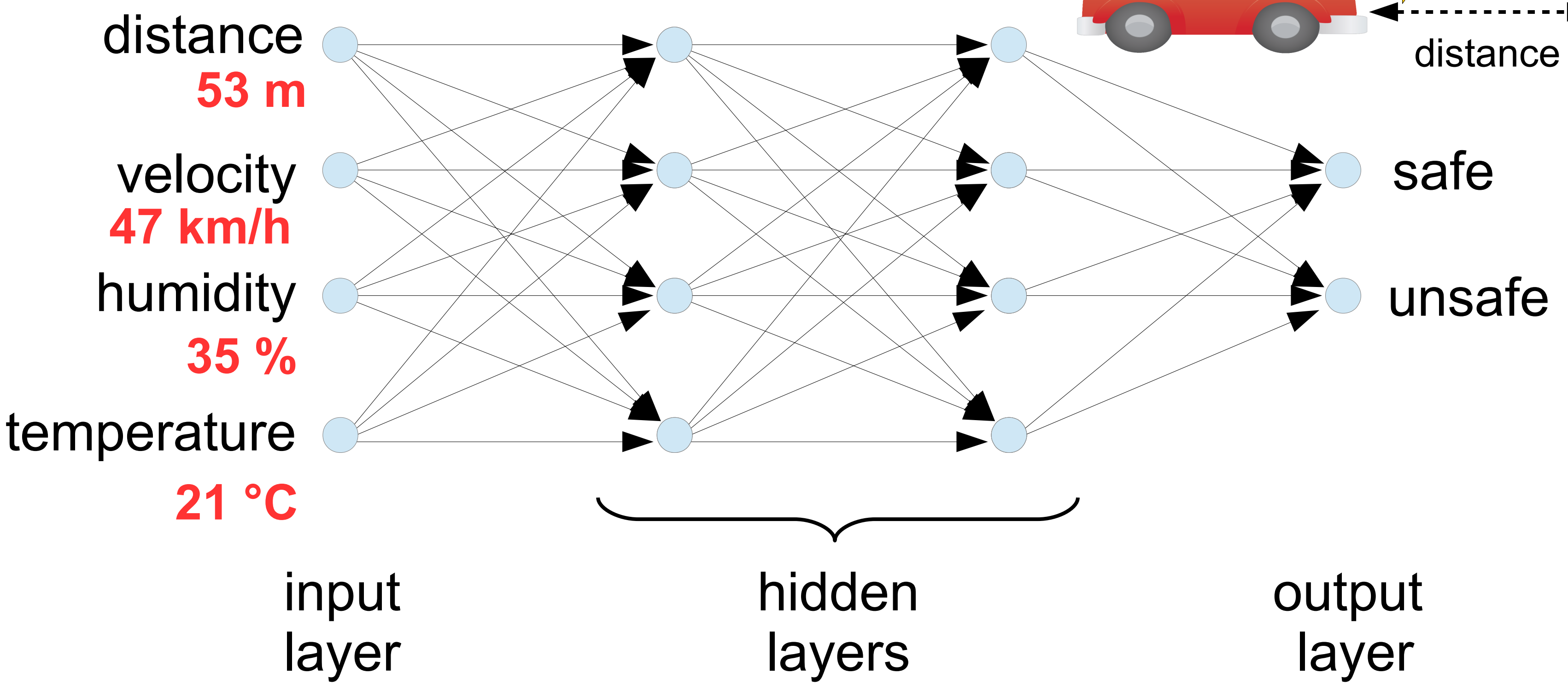


Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)

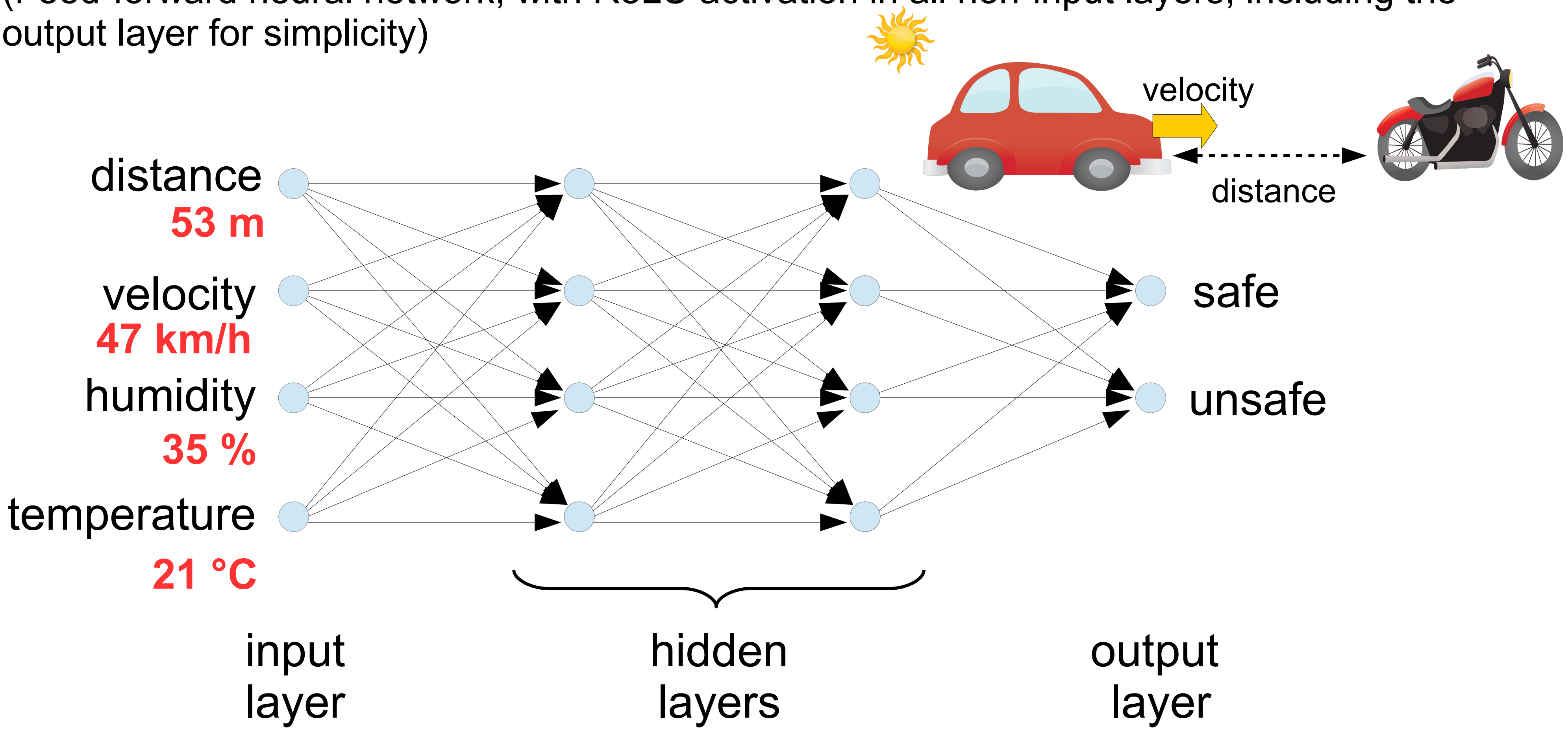




Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)

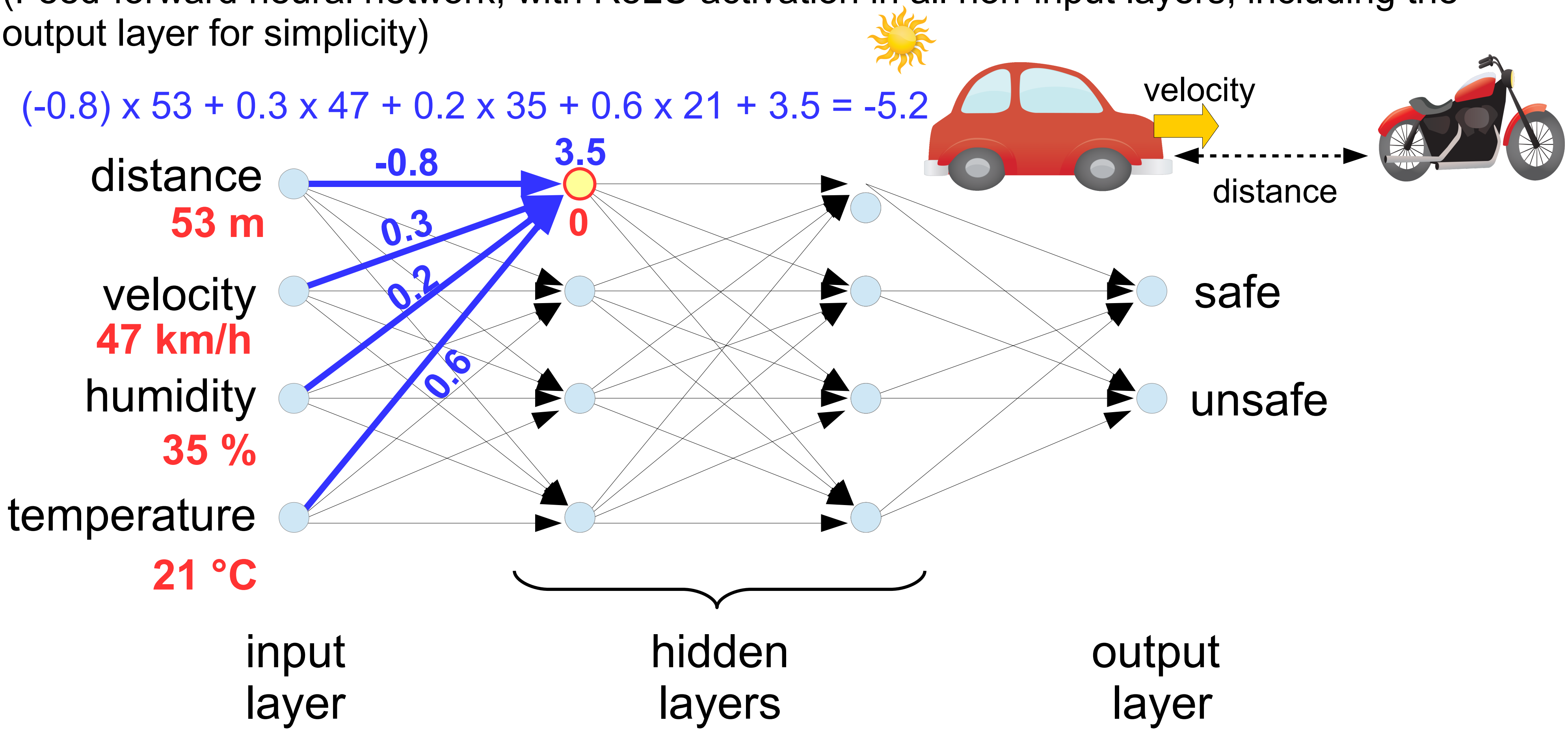


Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)

(-0.8) x 53 + 0.3 x 47 + 0.2 x 35 + 0.6 x 21 + 3.5 = -5.2

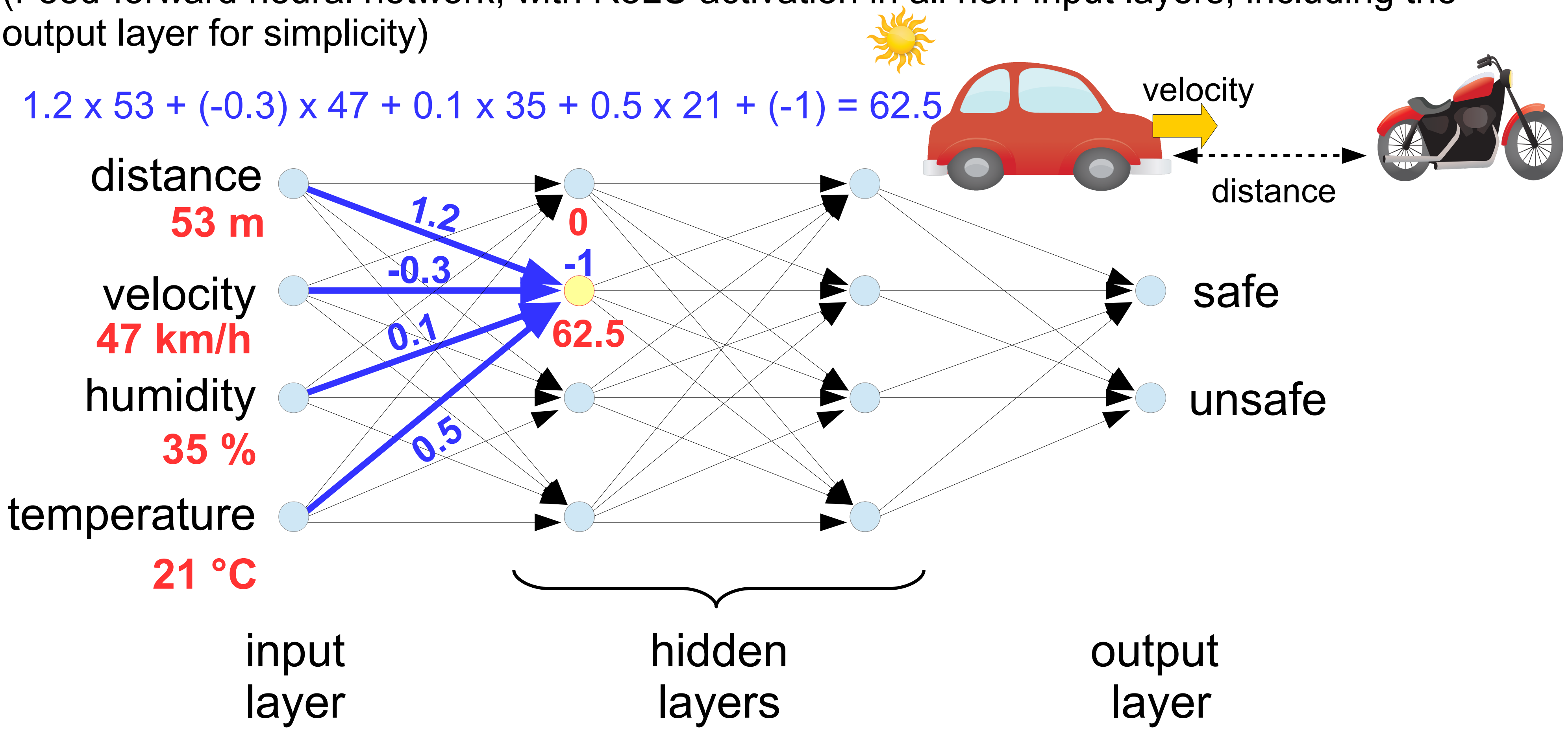


Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)

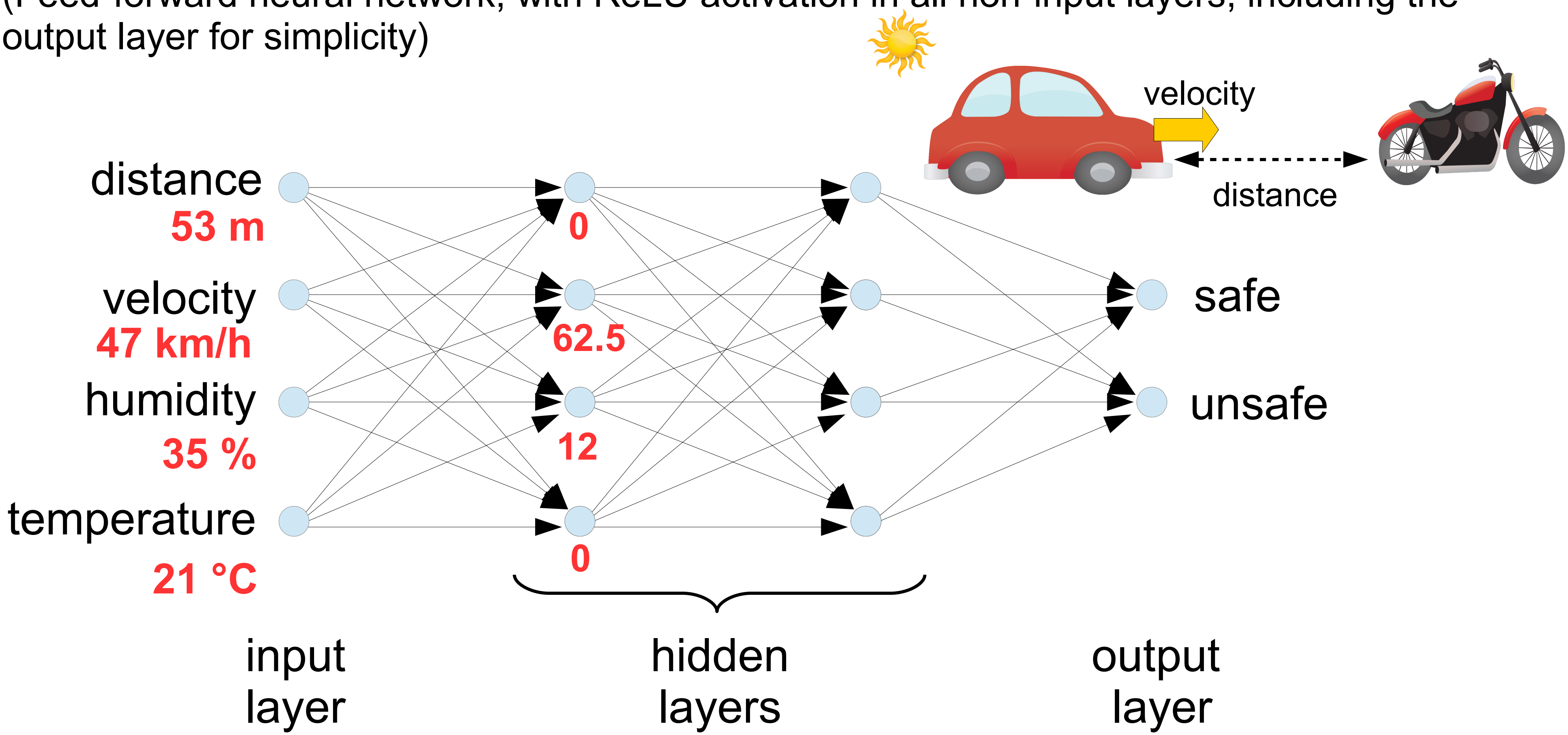
$1.2 \times 53 + (-0.3) \times 47 + 0.1 \times 35 + 0.5 \times 21 + (-1) = 62.5$



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

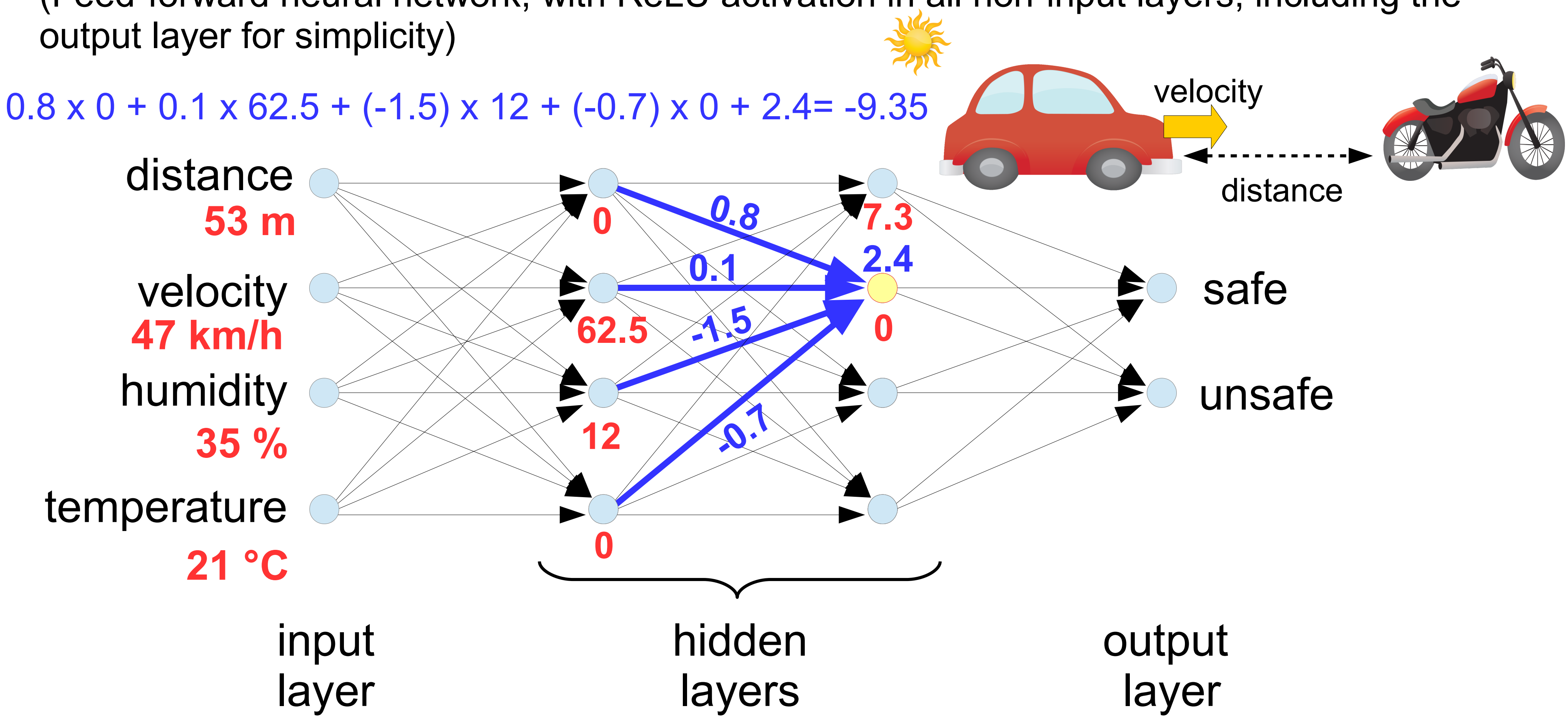
(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

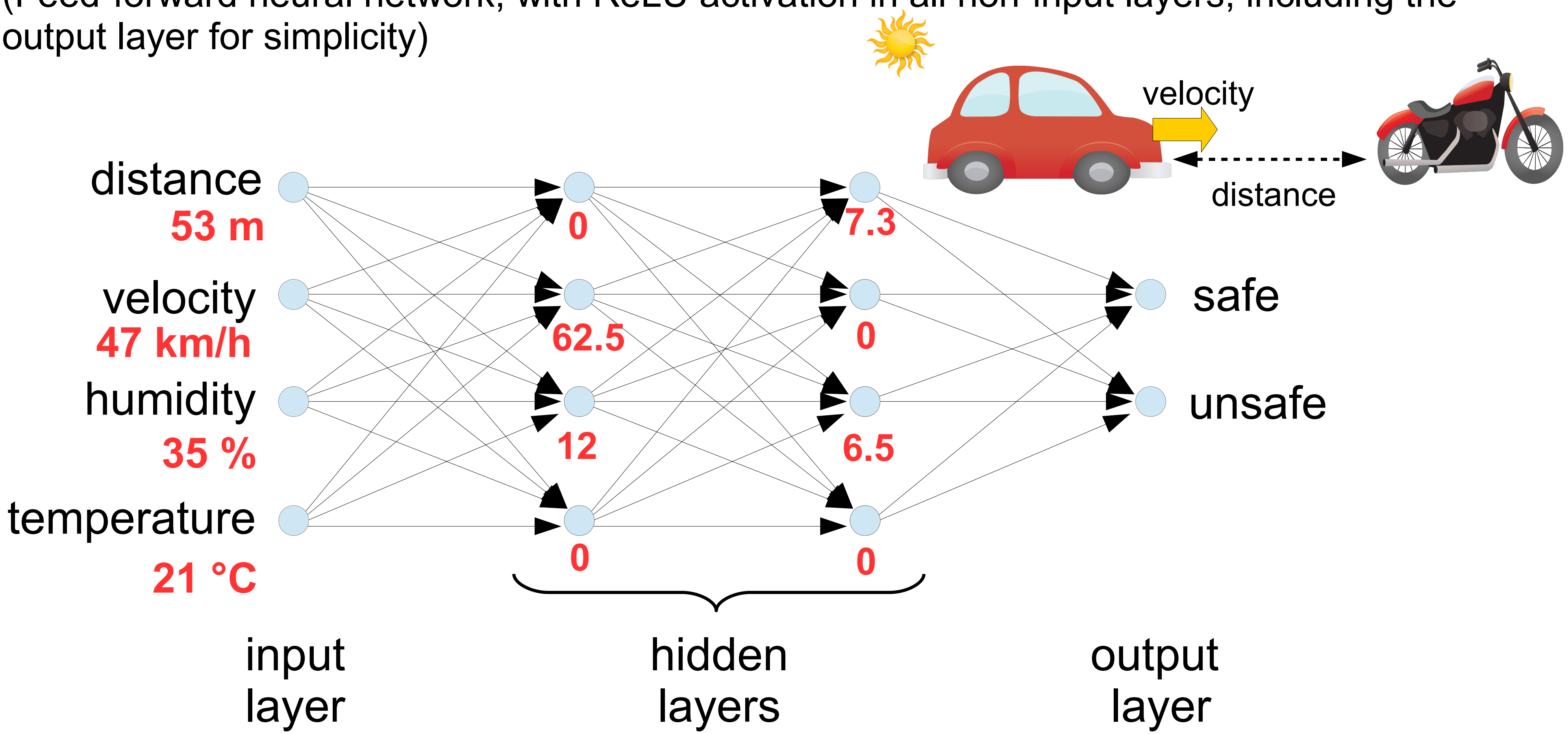
(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

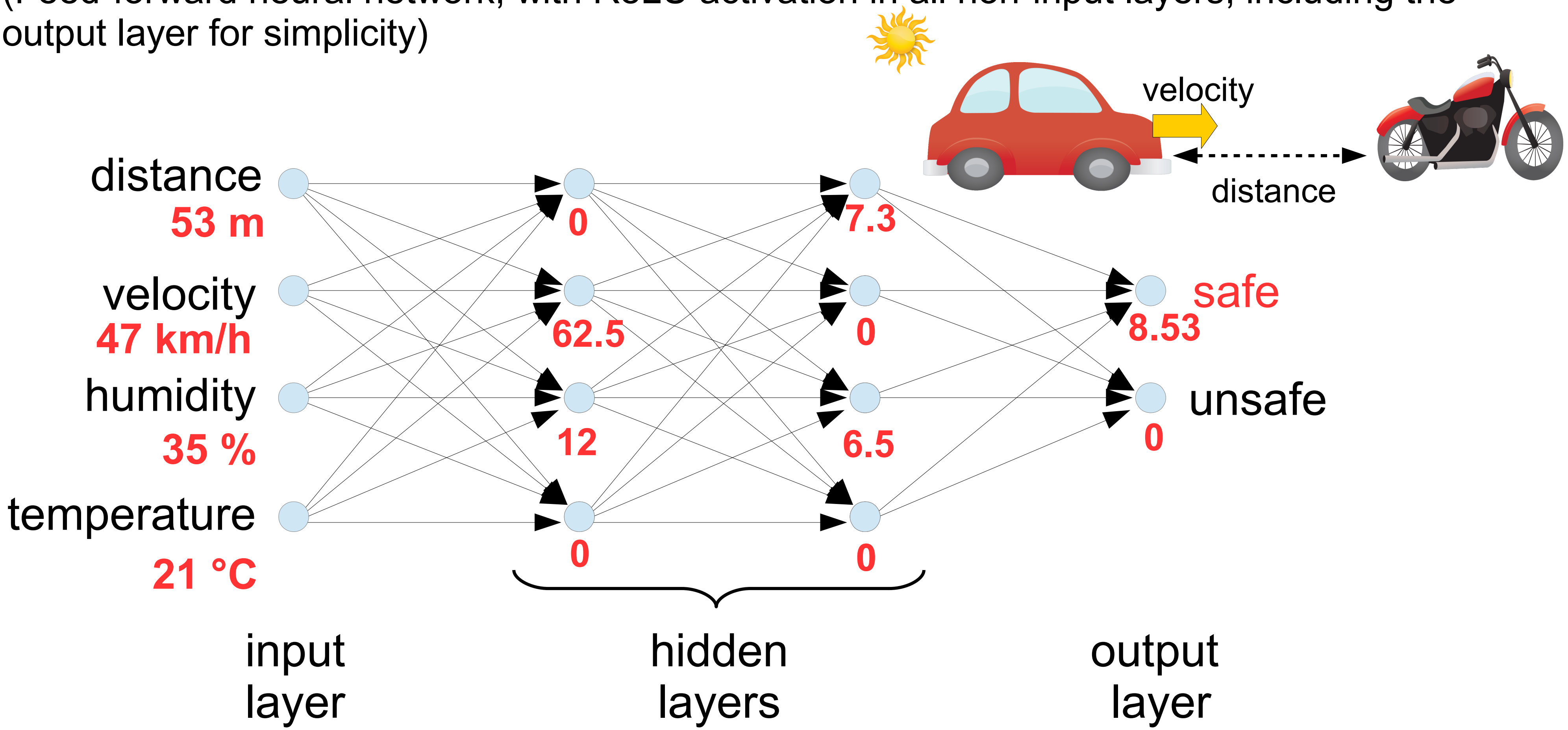
(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

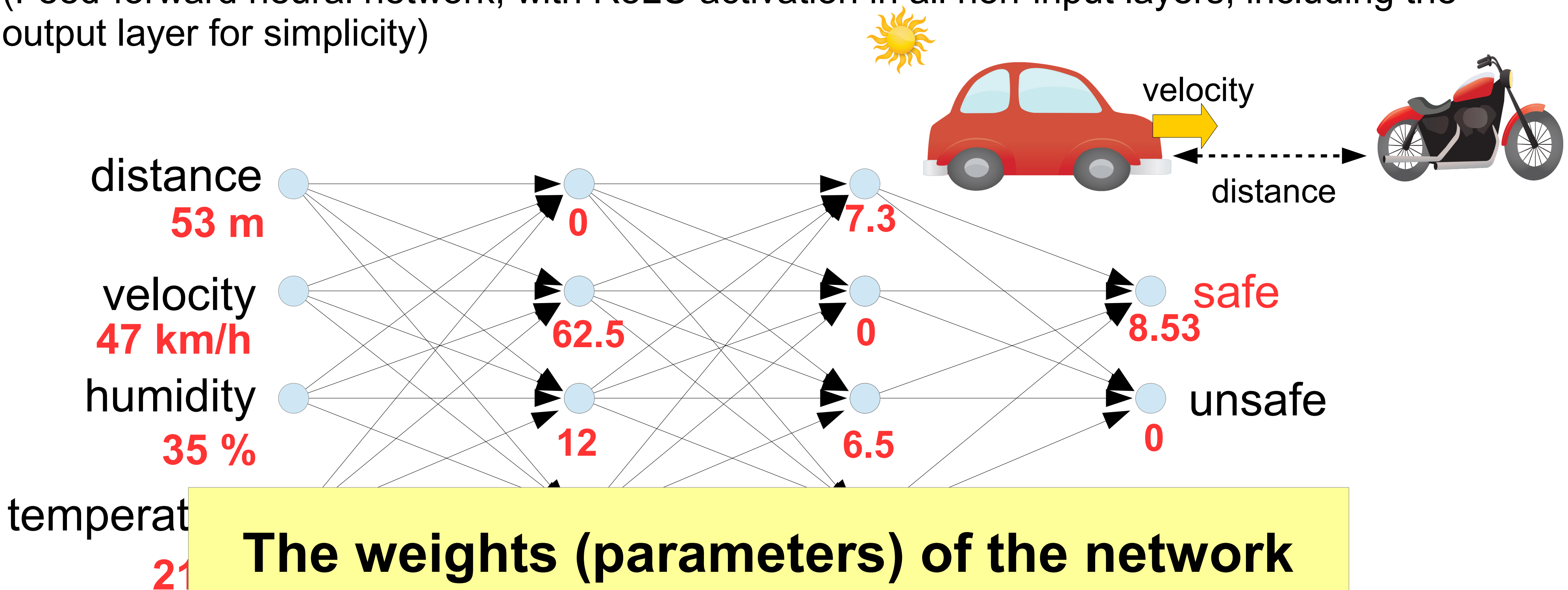
(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How does a Neural Network Work?

(Feed-forward neural network, with ReLU activation in all non-input layers, including the output layer for simplicity)

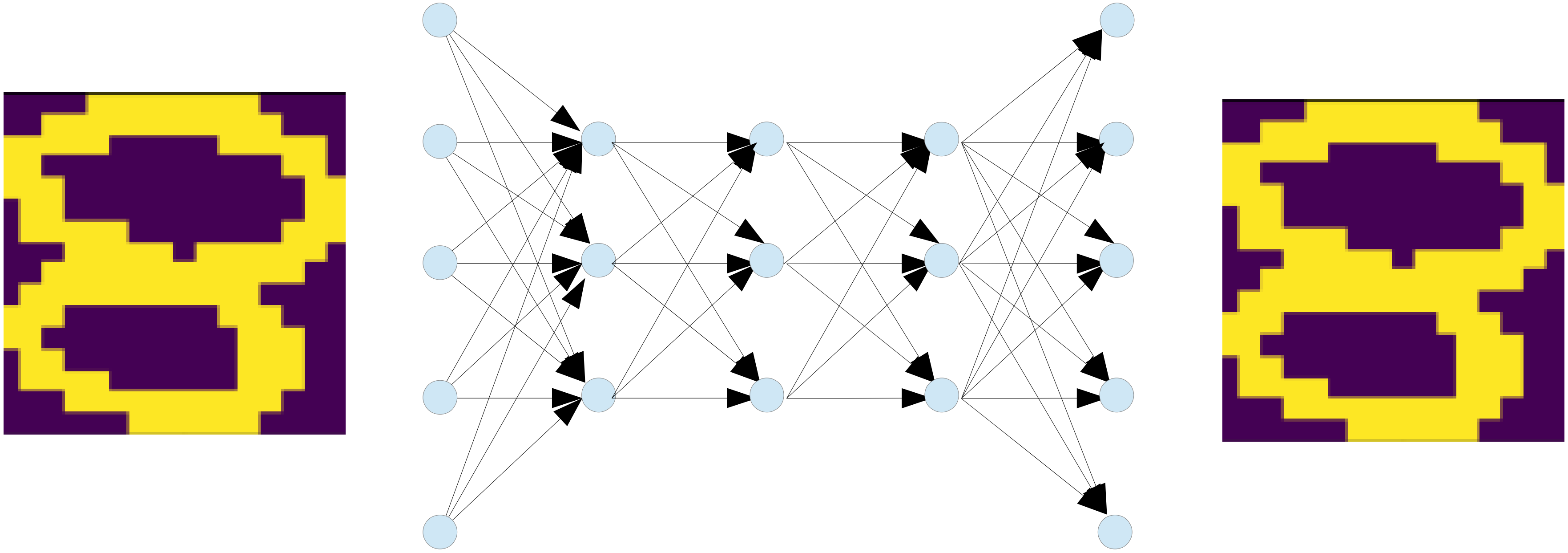


The weights (parameters) of the network are determined by the learning algorithm, NOT by a human expert!

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Autoencoders

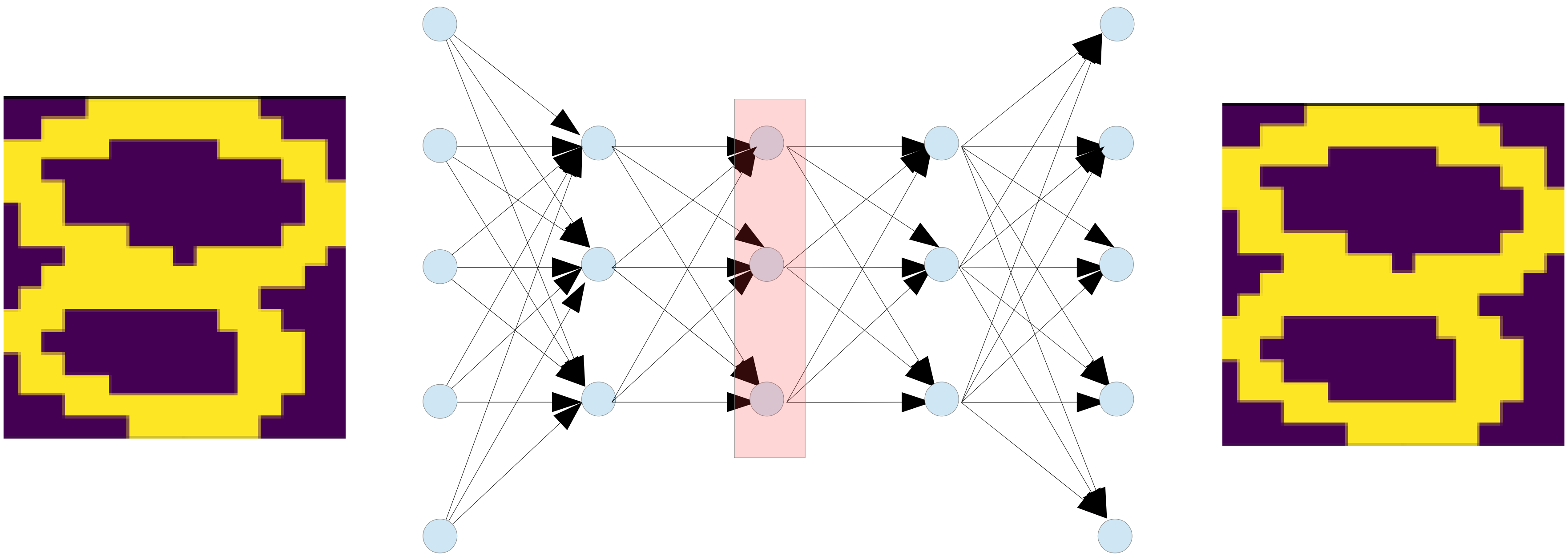
(Note: in real autoencoders, the number of units as well as the number of layers is much higher.)



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Autoencoders – image compression

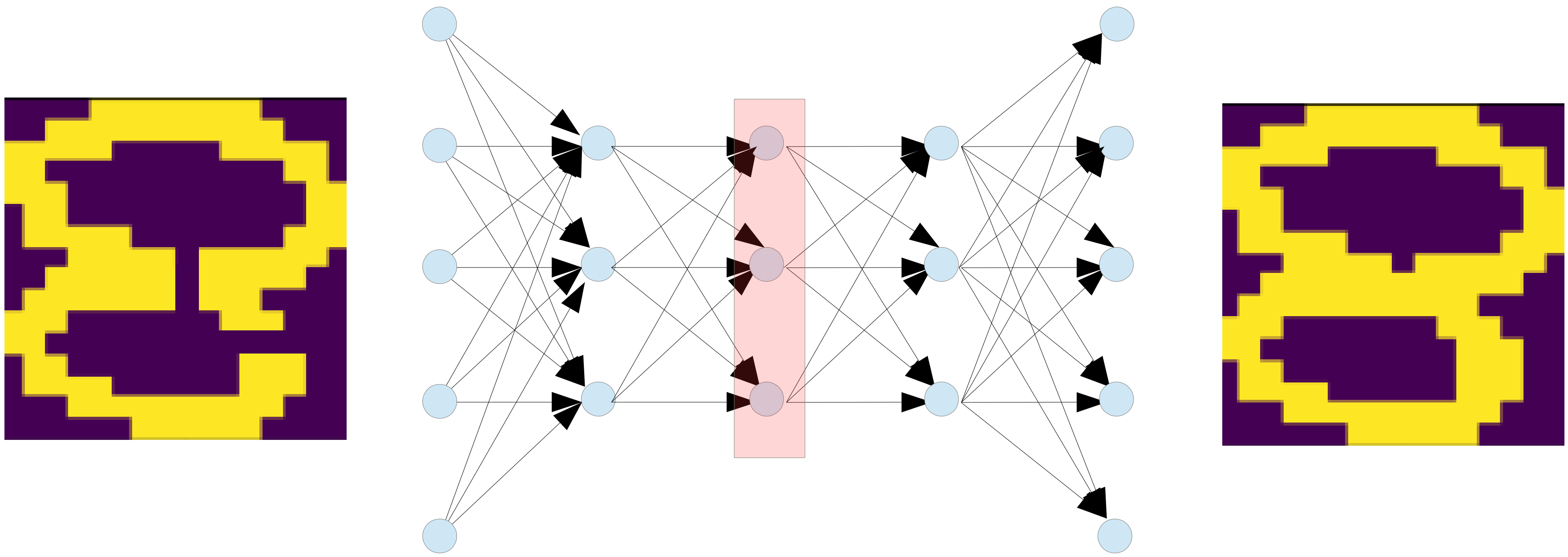
(Note: in real autoencoders, the number of units as well as the number of layers is much higher.)



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Autoencoders – image correction

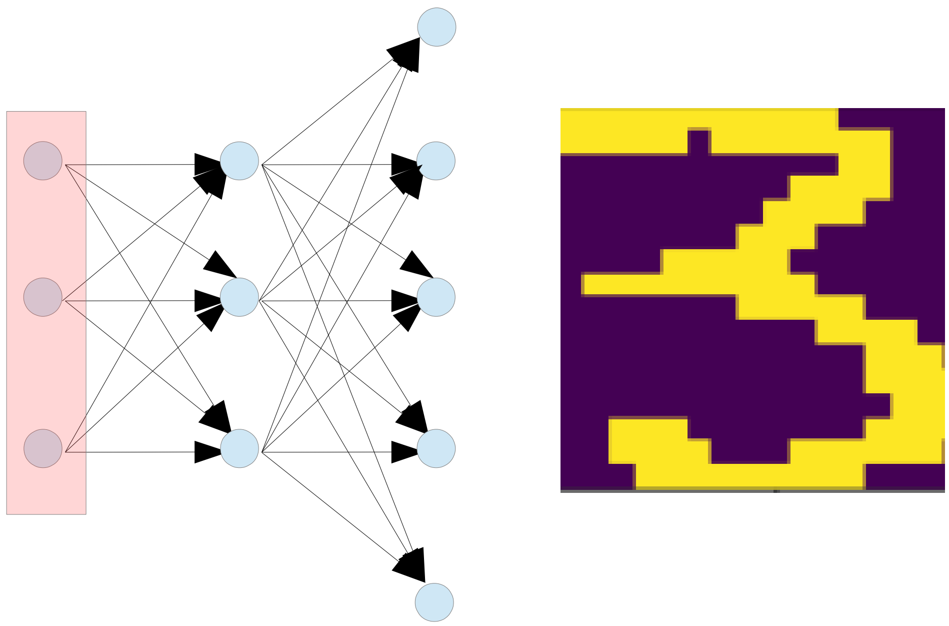
(Note: in real autoencoders, the number of units as well as the number of layers is much higher.)



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Autoencoders – generation of a new image

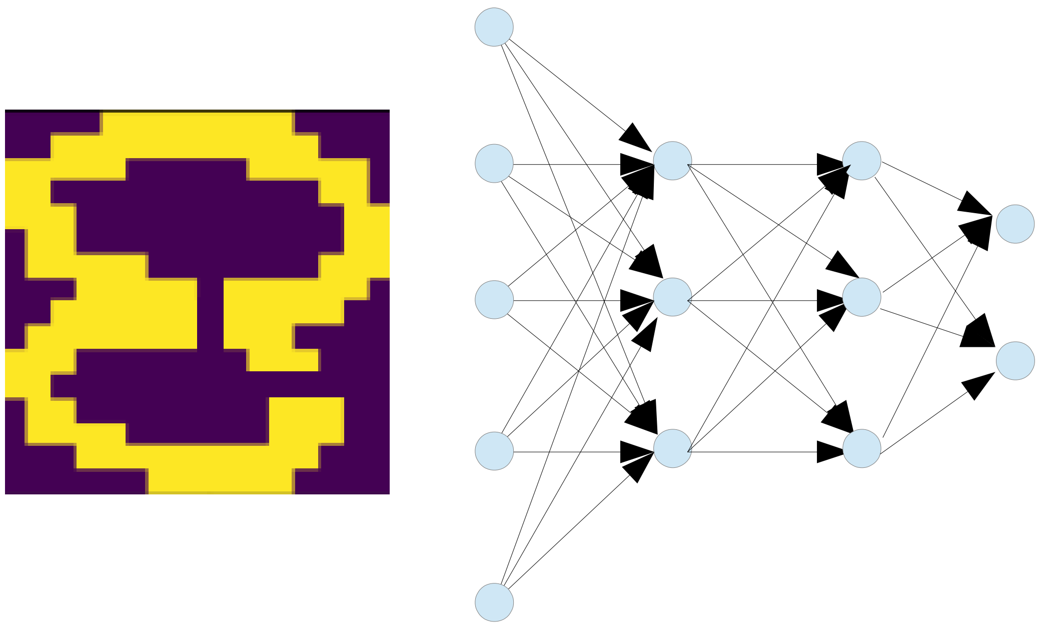
(Note: in real autoencoders, the number of units as well as the number of layers is much higher.)



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Image classification

(Note: in real autoencoders, the number of units as well as the number of layers is much higher.)



„8“

„not 8“

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Machine Learning as a Programming Paradigm

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Machine Learning as a Programming Paradigm


```
model = tf.keras.models.Sequential([
  tf.keras.layers.Conv2D(16, (3,3),
     activation='relu',
     input_shape=(28, 28, 1)),
  tf.keras.layers.MaxPooling2D(2, 2),
  tf.keras.layers.Conv2D(32, (3,3),
     activation='relu'),
  tf.keras.layers.MaxPooling2D(2,2),
  tf.keras.layers.Flatten(),
  tf.keras.layers.Dense(128,
     activation=tf.nn.relu),
  tf.keras.layers.Dense(10,
     activation=tf.nn.softmax)
])

model.compile(optimizer='adam',
  loss='sparse_categorical_crossentropy',
  metrics=['accuracy'])

model.fit(x_train, y_train, epochs=10)
```

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

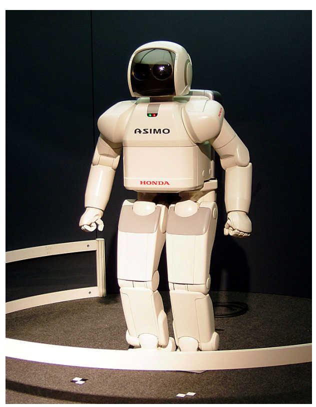
Artificial Narrow Intelligence versus Artificial General Intelligence

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

ANI versus AGI

Artificial Intelligence

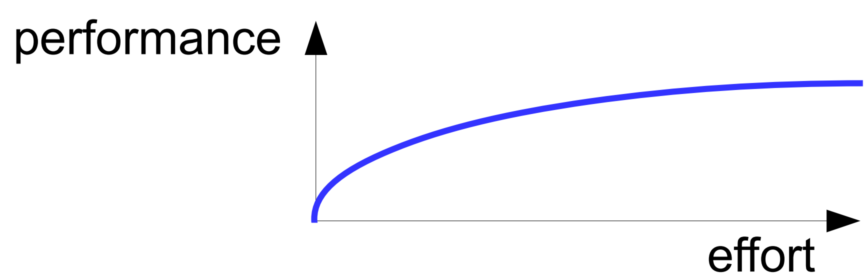
Artificial General Intelligence (AGI)



https://commons.wikimedia.org/wiki/File:HONDA_ASIMO.jpg

Artificial Narrow Intelligence (ANI)

- recognition of skin cancer
- self-driving cars
- playing chess, Go,...
- solving puzzles like the Rubic's cube
- detection of spam
- ...



Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Summary

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Take-home-message

- Behavior that seems to be intelligent is based on a mathematical model and high computational power.
- Machine learning is just a fancy phrase for finding appropriate parameters of the model.
- Most of the breakthroughs of the recent years belong to the category of „artificial narrow intelligence“.
- Medical doctors will not lose their jobs, but our opinion about creativity may change.

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

What to Expect in This Course

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Goals of this Course

- By the end of the course, students should
  - understand essential concepts of machine learning (ML), such as „model“, „training“, „overfitting“...
  - be familiar with some of the most popular ML methods (including deep neural networks)
  - improve their mathematical skills in order to be fit for more advanced machine learning courses
  - improve their skills to work in a team
  - have basic knowledge about some ML software libraries (scikit learn, tensorflow), and issues related to implementation of ML algorithms (vectorized implementation, paralellization, importance of low-level optimizations)
  - have a realistic view about social and economical impact of AI

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

Tentative Schedule

- **Week 1** Introduction (this lecture)
- **Week 2–6** Most important machine learning methods: From linear regression to deep neural networks
- **Week 7** No class (23rd Oct)

Autumn break

- **Week 8** Presentation of Programming Homeworks (6th Nov)
- **Week 9** Outlook: applications...
- **Week 10–11** Presentations related to social, economical and ethical aspects of AI (20th Nov, 27th Nov)
- **Week 12** Summary of take-home-messages of the course, preparation for the test
- **Week 13** Written Test (11th Dec)

Krisztian Buza, buza@inf.elte.hu
http://biointelligence.hu

How to Pass this Course?

- Quick tests (1 or 2 questions) both at the **beginning** as well as at the **end** of the lectures
  - Week 2 – 6, Week 9
  - Week 1: only at the end of the lecture
  - In total: 13 short test, 2 points each → 26 points
- Programming homework
  (week 8, team work, 1 team = 2 or 3 students) → 24 points
- Presentation about a topic related to social, economical or moral aspects of AI (week 10 or week 11, team work,
  1 team = 2 or 3 students, ≈10 minutes per team) → 20 points
- Test (week 13)
  - most essential concepts and algorithms related to the mathematical background of machine learning → 30 points