DRIVES

Development and Research on Innovative Vocational Education Skills

U3 DEEP LEARNING AND NEURAL NETWORKS

U3.E2 DISCLOSURE OF THE LEARNING PROBLEMS

Artificial Intelligence Technician

May 2021, Version 1

Co-funded by the Erasmus+ Programme of the European Union

The Development and Research on Innovative Vocational Educational Skills project (DRIVES) is co-funded by the Erasmus+ Programme of the European Union under the agreement 591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B. The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

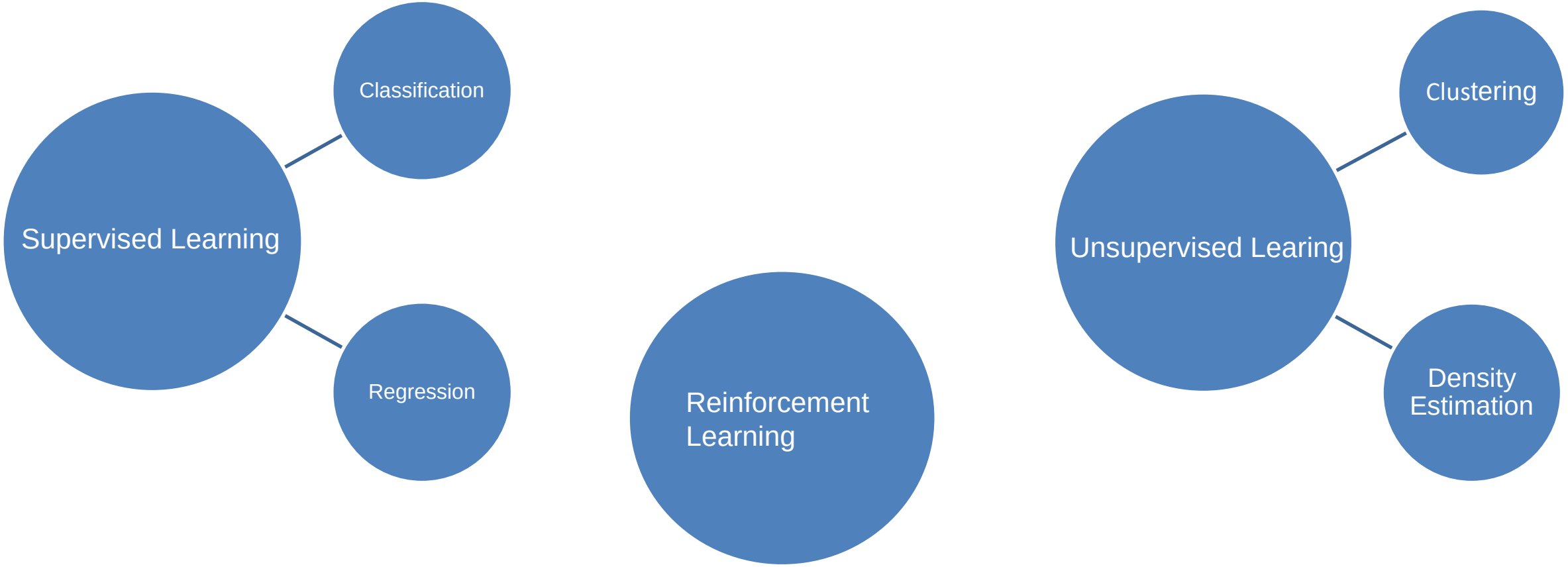
LEARNING OBJECTIVES

The student is able to

AIT.U3.E2.PC1	Identify several learning problems.
AIT.U3.E2.PC2	Understand how AI can help solve those problems.

LEARNING TECHNIQUES

To alleviate problems is used a deep architecture that is capable of learning feature representations from both the labeled and unlabeled data incorporating both unsupervised pre-training and supervised fine-tuning strategies to construct the models.



SUPERVISED LEARNING

Goal: construct a predictor, to learn a mapping from learning *x* to *y*, given a training set made of pairs (x_i, y_i). The **objective** is to define the mapping function accurately that the output variables for that system predict when a new input data occurs.

Estimate the function of the relationship between *x* and *y* in order to minimize the function of the (joint) data distribution, like the probability of classification error.

The training data set is analyzed by experts to identify features that are relevant to the desired output of the system, and to label each sample of the training data set as one of the possible target predictions of the system.

SUPERVISED LEARNING

Supervised learning can be grouped into **Regression** and **Classification**.

Classification methods aim to stimulate distinct classes, like genotypes while regression methods forecast real values. Some of the common methods are KNN, DT, NN, GA and SVM.

This simple method for machine learning is calculated through the use of programs like R or Python. The models can be **time-consuming** to train, and the labels for input and output variables demand expertise.

SOLVING PROBLEMS

To solve Supervised learning problems, we must follow the six steps:

1. Determine the type of training examples;
2. Gather a training set;
3. Determine the input feature representation;
4. Determine the structure of the learning function as the algorithm;
5. Complete the design;
6. Evaluate the accuracy of the learned function;

UNSUPERVISED LEARNING

Let X = $(x_1, \ldots, x_n)$ be *n* examples or points. The goal is to find interesting structure in the data X.

It is where there are only input data with no associated output variables.

The machine learning algorithm identifies features of input data that enable segregating of the input samples into unlabeled output groups rather provide features and labels for the input test data in advance.

The objective is to stimulate the distribution of data to better understand the content. This type of learning is not so popular because of lesser applications in day-to-day life.

There are three main tasks: **association, clustering,** and **dimensionality reduction**.

Association discovers the rules that define meaningful portions of data, while **Clustering** is where underlying groups are revealed in the data. **Dimensionality reduction** is used when the number of data dimensions is too high, normally is used in the preprocessing data stage.

UNSUPERVISED LEARNING

The **problem** identified with this learning is in estimating a density that is likely to have generated X. However, there are other **weaker forms**, such as quantile estimation, clustering, outlier detection, and dimensionality reduction.

This type of learning needs powerful tools for working with large amounts of unclassified data. The models are computationally complex due to the size of the training set to produce intended outcomes. Besides this, the unsupervised learning methods need human intervation to validate the output variables to have accurate results.

SEMI-SUPERVISED LEARNING

Semi-Supervised learning is between supervised and unsupervised learning. It is a learning paradigm concerned with the study of how computers and natural systems (humans) learn in the presence of labeled and unlabeled data.

The **goal** is to learn a better prediction rule using all of the available data effectively, not just the labelled data like in supervised learning.

In machine learning and data mining, this type of learning is of great interest because it can use readily available unlabeled data to improve supervised learning tasks when the labeled data are scarce or expensive.

SEMI-SUPERVISED LEARNING

The problems occurs when have a large amount of input data (X) and only some of the data is labeled (Y).

For example, in a photo archive where some photos are labeled (cat, dog, etc) and the majority are unlabeled.

It is ideal for medical images, where a small amount of training data can lead to a significant improvement in accuracy.

Some popular semi-supervised learning models, are self-training, mixture models, co-training and multiview learning, graph-based methods, and semi-supervised support vector machines.

REINFORCEMENT LEARNING (RL)

Reinforcement learning represent a class of machine learning problems where an agent operates in an environment with no fixed training dataset, learning what to do, how to map situations to actions.
RL studies how artificial systems can solve instrumental conditioning problems.

Reinforcement learning is acquainted to be unstable or to be divergent when a nonlinear function like a neural network is used to represent the action-value function.

This type of learning can be applied to autonomous devices due to their flexibility in their environment.

AI CAN HELP...

Supervised learning to achieve effectiveness in marketing campaigns, predict how often should re-target customers, while **unsupervised learning** will tell what is the marketing strategy more effective for each type of customer. The result will be the ability to adapt the marketing and advertising efforts to each customer.

AI isn't yet useful in all areas but as becomes better at recognizing patterns, the use cases for unsupervised learning will continue growing. The areas of marketing and advertising, transport and logistics, supply chain, dynamic pricing, demand planning, and accounting/finance can benefit from custom solutions based on machine learning.

- **Supervised Learning**: the models can be **time-consuming** to train, and the labels for input and output variables demand expertise.
- **Unsupervised Learning**: problems in estimating a density that is likely to have generated X. Needs powerful tools for working with large amounts of unclassified data and need human intervation to v alidate the output variables to have accurate results.
- **Semi-supervised Learning**: the problems occurs when have a large amount of input data (X) and only some of the data is labeled (Y);
- To solve Supervised learning problems, we must follow six steps
- Both types have wide applications in business analytics, business intelligence, bioinformatics, spam detection, image, object and speech recognition and segmentation, genetic clustering, pattern and sequence mining. Unsupervised learning is greater to abstract proposes like detecting fraudulent transactions and climate change abnormalities.

REFERENCES

14 Different Types of Learning in Machine Learning. (n.d.). Retrieved June 2, 2021, from https://machinelearningmastery.com/types-of-learning-in-machine-learning/

Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine Learning and Data Mining Methods in Diabetes Research. *Computational and Structural Biotechnology Journal*, *15*, 104–116. https://doi.org/10.1016/j.csbj.2016.12.005

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. https://doi.org/10.1038/nature14236

Supervised vs. Unsupervised Learning: What's the Difference? | *IBM*. (n.d.). Retrieved June 7, 2021, from https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning

The Progression of AI: from Supervised to Unsupervised Learning - Pegus Digital. (n.d.). Retrieved June 7, 2021, from https://pegus.digital/the-progression-of-ai-from-supervised-to-unsupervised-learning/

XiaojinZhu and Andrew B.Goldberg. Synthesis Lectures on Artificial Intelligence and Machine

Learning, 2009, Vol. 3, No. 1 , Pages 1-130. https://doi.org/10.2200/S00196ED1V01Y200906AIM006

REFERENCE TO AUTHORS

Diana Ferreira

- PhD student in Biomedical Engineering
- Research Collaborator of the Algoritmi Research Center

0000-0003-2326-2153

Regina Sousa

- PhD student in Biomedical Engineering
- Research Collaborator of the Algoritmi Research Center

0000-0002-2988-196X

Ana Luísa Sousa

- PhD student in Information System and Tecnologies
- Research Collaborator of the Algoritmi Research Center

0000-0001-5731-3583

REFERENCE TO AUTHORS

António Abelha

- Assistant Professor at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0001-6457-0756

José Machado

- Associate Professor with Habilitation at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0003-4121-6169

Victor Alves

- Assistant Professor at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0003-1819-7051

REFERENCE TO AUTHORS

This Training Material has been certified according to the rules of **ECQA – European Certification and Qualification Association.**

The Training Material was developed within the international job role committee **“Artificial Intelligence Technician”:**

UMINHO – University of Minho (https://www.uminho.pt/PT)

The development of the training material was partly funded by the EU under Blueprint Project DRIVES.

DRIVES

Development and Research on Innovative Vocational Education Skills

Thank you for your attention

DRIVES project is project under **The Blueprint for Sectoral Cooperation on Skills in Automotive Sector**, as part of New Skills Agenda.

The aim of the Blueprint is **to support an overall sectoral strategy and to develop concrete actions to address short and medium term skills needs.**

Follow DRIVES project at:

More information at:

www.project-drives.eu

Co-funded by the Erasmus+ Programme of the European Union

The Development and Research on Innovative Vocational Educational Skills project (DRIVES) is co-funded by the Erasmus+ Programme of the European Union under the agreement 591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B. The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.