MACHINA

O1-T4: Report on Machine Learning (ML) learning outcomes

Funded by the Erasmus+ Programme of the European Union

TABLE OF CONTENTS

ACRONYMS AND ABBREVIATIONS

AI	Artificial Intelligence
AUF	Francophone University Association
BT	Business Training
CEPIS	Council of European Professional Informatics Societies
C-VET	Continuous Vocational Education and Training
ECVET	European Credit system for Vocational Education and Training
EQF	European Qualification Framework
EU	European Union
EUA	European Universities Association
EuroCIO	European CIO Association
FEDERGON	House of Training, Federation of HR Service Providers
GSRT	General Secreteriat for Research and Technology
IAU	International Universities Association
ICT	Information and Communications Technology
IEEE	Institute of Electrical and Electronics Engineers
IFIP	International Federation for Information Processing
IT STAR	Regional ICT Association in Central, Eastern & SouthernEurope
I-VET	Initial Vocational Education and Training
KPI	Key Performance Indicator
LERU	League of European Research Universities
ML	Machine Learning
M.Sc.	Master of Science
MOOC	Massive Open Online Course
NIST	National Institute of Standards and Technology (USA)
OER	Open Educational Resources
VET	Vocational Education and Training

1 INTRODUCTION

Machine Learning (ML) is a subset of Artificial Intelligence (AI) that has gainedsubstantial prominence in diverse economic & social realities, having become the basis for a series of technological developments such as automated translation systems, medical image analysis, and virtual assistants. ML was born from pattern recognition but has evolved to refer to the use of data & learning algorithms to produce models, predict outcomes and make decisions with minimum human intervention.

MACHINA is an Erasmus+ KA2 project, which aims to tackle this ML skill deficit by increasing the relevance of Continuing & Initial VET provision in the sector,to assure that the existing & future ICT workforce will have the ML specific competences & transversal skills required to respond to modern workplace requirements and succeed in a competitive, fast-growing field. The project willalso make available transnational educational materials in the form of OERs, to ensure wide adoption and support VET provision in a cost-effective, flexibleway.

The current report is the outcome of the intellectual output O1-T4-b, which isbased on the results of deliverable O1-T4-a entitled "Analysis of Evidence gathered" 2021-02-7. The O1-T4-a document provides the analysis of the results of the online questionnaires (field research) and the analysis of the results of the skill mismatches (desk research) between MACHINA skilldemand (i.e. through the analysis of job vacancies descriptions) and MACHINAskill supply (i.e. by analysing existing training programs).

In the next section, we summarise the most important findings of each analysis presented in the abovementioned deliverable. Then, section 4 provides the definitions of the learning outcomes of the project MACHINA VETprogram on Machine Learning technology in the form of statements of what ICT professionals should know, understand and be able to do upon the completion of the MACHINA course. The definitions of the learning outcomes rely on the European reference frameworks and standards (EQF and ECVET) to develop the learning outcomes in terms of definitions of knowledge, skills and competences.

Finally, section 5 presents guidelines on how to prepare learning units corresponding to the MACHINA learning outcomes.

2 AN OVERVIEW OF DATA ANALYSIS RESULTS

In this section, we bring out the main findings of the analysis on the online questionnaire responses (field research), as well as the analysis of the resultsof the skill mismatches (desk research) between MACHINA skill demand and MACHINA skill supply, which is detailed in the report entitled *MACHINA_O1_T4-a Analysis evidence gathered 2021-02-07*.

According to the respondents, the five most needed field knowledge for working on ML applications and related services are, in decreasing order:

- *ML algorithms (supervised, unsupervised, semi-supervised,reinforcement)*
- *Probability and Statistics*
- *ML Programming Languages (e.g. Python, Scala, Java, JavaScript, Lisp,R, C++),*
- *Neural network architectures (feed-forward, recurrent, symmetricallyconnected)*
- *Deep Learning (DL) frameworks (e.g. TensorFlow. Keras. PyTorch,Theano)*

ICT Professionals are lowly knowledgeable on the topics of

- *EU legislation on data protection,*
- *Ethical and social implications of Artificial Intelligence,*
- *Project Management Principles*

Regarding the four most important skills to work as ICT ML professional, therespondents valued the most, in decreasing order:

- *Advise on organisational improvements (governance changes) based on(big) data insights and interpretation*,
- *Lead organisational changes to support the integration of ML enabledsolutions in a business context*
- *Communicate the benefits and limitations of ML enabled solutions (inbusiness terms) to potential customers and business stakeholders*
- *Develop proof of concepts for envisioned ML applications.*

As we highlighted in the report *MACHINA_O1_T3-a_Analysis evidence gathered (desk research section)*:

- *Most training offers types are Online Course and Higher educational level*

- *Most of courses are theoretical or a focus on a specific softwaredevelopment aspect*
- *The industrialization aspect is missing*
- *The concepts are defined from the computer scientist's point of viewand not from the decision makers or users' points of view*

The major application areas of practical applications are:

- *Clinical decisions*
- *Fraud detection*
- *Engineering + Agriculture*
- *Translation/Speech*
- *Mobile support*

The analysis of relevant studies and articles show the following knowledge/skills requirements:

- *Knowledge of mathematical and statistical models*
- *Programming fundamental (Python, R, Apache, Spark, MapReduce,Hadoop …)*
- *Machine Learning Algorithms*
- *Knowledge of deep learning*
- *Knowledge of data structures and data modeling*
- *Software engineering*
- *Knowledge about preprocessing and cleaning of data*
- *Ability to work with Cloud based AI and AI as a Service*
- *Availability to develop and increasing IT security*
- *Ability to develop and implement AI solutions*
- *Business analytics*
- *Communication*

3 DEFINITION OF LEARNING OUTCOMES

Aligned with the provisions of the project Application Form and the O1-T1 deliverable,the definition of the MACHINA learning outcomes is based on the European Qualification Framework (EQF) [1], as the latter acts as a translation device to makenational qualifications more readable and comparable across Europe, aiming to promote workers' and learners' mobility between countries and facilitate their lifelong learning. The EQF relates different countries' national qualifications systems and frameworks together around a common European reference – its eight reference levels based on **"learning outcomes"** (defined in terms of knowledge, skills and competences). Learning outcomes do not describe the learning target or the learningpath, but the result following the completion of a learning process.

According to the 2017 CEDEFOP handbook *Defining, writing and applying learning outcomes* [2], learning outcomes are "statements of what a learner knows, understands and is able to do on completion of a learning process, which are definedin terms of knowledge, skills and competences". We recall their definition, as mentioned in the O1-T1 deliverable:

- **Knowledge**: The outcome of the assimilation of information through learning.Knowledge is the body of facts, principles, theories and practices related to a field of work or study. According to the EQF, knowledge is described as theoretical and/or factual.
- **Skill**: The ability to apply knowledge and use know-how to complete tasks andsolve problems. According to the EQF, skills are described as cognitive(involving the use of logical, intuitive and creative thinking) and practical skills(involving manual dexterity and the use of methods, material tools and instruments).
- **Competence:** The proven ability to use knowledge, skills and attitudes, in work in study situations and in professional and personal development.According to the EQF, competence is described in terms of responsibility and autonomy.

While learning outcomes promote overall transparency and help to clarify the intentions of learning processes, the CEDEFOP handbook [2] also points out some criticism to the learning outcomes approach. Among other imperfections, "it can be argued that learning outcomes can inhibit the learning process, for example when indicating (too) restricted a threshold level. Too much specificity and detail, it is argued, also makes it difficult to give room for innovation and exploit the unexpected present in any situation" ([2], page 39). Indeed, the learning outcomes approach isseen, by some constructivist schools of thought, as 'policy hype' and as a threat to high quality education, training and

innovation. To try to improve these flaws, (Biggs,1999; Biggs and Tang, 2007, [2] page 40) stress the importance of aligning learningoutcomes statements to teaching and learning practices as well as to assessment tasks. Aligning learning outcomes to teaching and learning is about connecting the abstract idea of a learning outcome to what teachers actually do to help students learn, and the things that students do to learn.

In fact, learning-outcomes-based approaches have different origins and have been promoted by different schools of thought. While the behaviouristic tradition emphasises learning outcomes as result-oriented, full-ended, clearly observable and (objectively) measurable, the constructivist approach will emphasise the need for learning outcomes to be process-oriented and open-ended, limiting quantified measurability.

The CEDEFOP handbook [2] also emphasises *the writing and articulation of learningoutcomes must be followed by implementation, through teaching, learning and assessment*. Learning outcomes statements form an important part of curricula. They guide teachers in the teaching process, for example supporting the choice ofmethods, and they inform learners about what they are expected to know/do and understand after a given learning activity.

The definitions and descriptions of learning outcomes as used in curricula are statements and expressions of intentions. They are not outcomes of learning, but desired targets. Achieved learning outcomes can only be identified following the learning process, through assessments and demonstration of achieved learning in real life, for example as the result of following the training.

The European e-Competence Framework (e-CF) [3] - version 3.0 claims to be the first sector-specific implementation of the EQF, arguing to provide a reference of 40competences as required and applied at the Information and Communication Technology (ICT) workplace, using a common language for competences, skills and capability levels that can be understood across Europe. The e-CF is used as guidancein the formulation of the MACHINA learning outcomes as presented in the next section.

Formulation of the MACHINA learning outcomes

It is worth to emphasise, in line with the provisions of the project Application Form and the O1-T1 deliverable, the project MACHINA aims to strengthen the key competences of ICT professionals (namely initiative, entrepreneurship andcommunication skills), rather than focus on the technical and

coding skills associatedwith blockchain technology, which risk being outdated by the time the project will make available its results to its target groups. To this end, the content of MACHINA curricula is based on the training needs analysis and organised in modules as follows:

- Module 1: Machine Learning essentials for ICT professionals
- Module 2: Mathematics foundations
- Module 3: Machine Learning algorithms, Programs and Protocols
- Module 4: Deep Learning
- Module 5: Communicating the merits, challenges and implications of Machine Learning technology to customers and within own organisation
- Module 6: Legislation, Ethics, Project Management related to Machine Learning

The learning outcomes of each module are specified hereafter, in terms of knowledge, skills and competences. Thesecorrespond to statements of what ICT professionals should know, understand and beable to do upon the completion of the MACHINA course.

Table 1: MACHINA Module 1 Learning Outcomes

Module 1	**ML essentials for ICT professionals**		
Learning outcomes correspond to EQF 5	Defines the fundamental features of ML applications.		
	Knowledge Knows / Aware of: - ML terminologies - Types of ML - ML Architectures - Existing ML applications, related structures and architectures - Principles and paradigms of ML systems - Essence of Machine Learning. - Main tasks, methods and basic concepts of Machine Learning. - Types of problems that Machine Learning algorithms can solve nowadays. - Data pre- processing methods. - How to split data into training and test datasets	**Skills** Able to: - Identify ML characteristics in a given setting - Analyse existing ML applications according to a given context - Critically evaluate - Collect, formalise and validate functional and non- functional requirements of a ML application for a particular scenario. - Identify Machine Learning methods suitable for the existing problem in the workflow. - Solve the problems using various Machine Learning techniques. - Articulate how ML algorithms are fundamentally different from traditional programming algorithms. - Apply the basic algorithm for working on ML-project.	**Competences** Able to: - Give an account of the advantages and disadvantages of the features of a specific ML application, - Autonomously explain the operation of ML use cases - Act independently to design application using Machine Learning techniques, using different programming languages (e.g. Python, Scala, Java, JavaScript, Lisp, R, C++) and desirable libraries

Table 2: MACHINA Module 2 learning outcomes

Module 2	**Mathematics foundations (Probability, Statistics, Analysis, Linearalgebra, Computing theory)**		
Learning outcomes correspond to EQF 5	Defines the mathematical foundations required for writing programs and algorithms for ML and AI. Breaks down the difficult mathematical concepts into easier one to understand concepts.		
	Knowledge Knows / Aware of: - Mathematical foundations (functions, limits, derivatives, integrals) - Linear algebra (scalars, vectors, vector space, inner and outer products, metrics, linear independence, orthonormal bases, matrices, tensors) - Axiomatic Probability theory (sets, probability as a measure, conditional probabilities, Bayes theorem, random variables with discrete and continuous distribution bernoulli, binomial, exponential, gaussian, etc.) - Statistics (population, sample, frequencies, distributions, mean, variance, percentile, law of large numbers- central limit theorem- multidimensional distributions- covariance- estimators-	**Skills** Able to: - Perform basic functional analysis - Calculate the domain, range, codomain of a function - Estimate the rate of growth of a function - Calculate the basic linear algebra operations on matrices and vectors - Calculate probabilities on a set - Calculate mean and variance of a distribution - Calculate aggregations on a population by standard aggregation functions - Calculate a univariate linear regression on a set of points (X,Y) - Implement a simple algorithm in some programming language - Implementing a simple class with some object oriented programming language - Design a study, collect	**Competences** Able to: - Describe and analyse a problem in mathematical terms - -Representing date in tabular form (vector, matrices, tensors) - Describe qualitatively and quantitatively the statistics of a population - Code formally a problem and its solution with some programming language - Use exponential functions to solve exponential growth and decrease through programming languages

	maximum likelihood- univariate and multivariate linear regression) - Basic computer theory (Von Neumann machine, algorithms, complexity in time and space, programming languages, variables, type of data, procedures, data and procedures abstraction on classes, objects and methods)	data, and/or select the appropriate representation to make conclusions and generalizations - Calculate probabilities of events under different relationships such as inclusion, disjoint, complementary, independent, and dependent, with and without replacement	

Table 3: MACHINA Module 3 learning outcomes

Module 3	ML Algorithms, Programs and Protocols		
Learning outcomes correspond to EQF 5	Makes links between mathematics foundations and algorithms. Specifies, refines, updates and makes available a formal approach to design solutions, necessary to develop and operate a ML application. Selects appropriate technical options for ML design and implementation.		
	Knowledge Knows / Aware of: - Introduction to machine learning (Modeling, pre-processing, issues, evaluation metrics) - Machine learning by linear models (kNN, Adaline, Perceptron, Naive bayes, regression, regularized regression, …) - Supervised learning algorithms like MLP, decision tree, ensemble methods, logistic regression, etc. - Unsupervised learning (Clustering, Feature extraction, …) - Semi-supervised learning (self-training, co- training, …) - Machine learning programming languages for solving problems (Python, Matlab, R,...). - How to apply and compare ML methods on a dataset. - Statistical validation.	**Skills** Able to: - Plan and design the specifications of ML applications for a given scenario. - Understand the strengths and weaknesses of many popular machine learning approaches. - Plan and design the specifications of ML applications for a given scenario. - Understand best-practices for building machine learning applications. - Understand the differences in analyses enabled by clustering, classification, and regression.	**Competences** Able to: - Build an end-to-end a POC that depends on using the correct machine learning algorithm and best-practices. - Evaluate results and compare methods.

Table 4: ML Module 4 learning outcomes

Module 4	Deep Learning (Advanced)		
Learning outcomes correspond to EQF 5	Defines the important details about deep neural networks and get to know of different deep neural network architecture. Identifies the potential of deep learning in different applications such as natural language processing, computer vision, or recommendation systems.		

	Knowledge
	Knows / Aware of:
	- Fundamental concepts of neural networks:
	  ○ Perceptron Activation functions
	  ○ Cost functions
	  ○ Learning
	  ○ Gradient descent
	  ○ Multi-layer perceptron and its universality
	  ○ Parameters and hyper-parameters
	- When to use neural networks:
	  ○ Interpretability of a machine learning model
	  ○ Assumptions on the model underlying the dataset
	- The black box
	- What is our goal? (description,
	- Analysis, forecast, classification, prescription)
	- -Learning a pattern vs learning a function:
		Skills
		Able to:
		- To use a framework to implement , train and validate a machine learning model using existing neural networks development framework (Pytorch, Tensorflow- Keras, Scikit- learn,mxf,...)
			Competences
			Able to:
			- To understand which class of problems could be actually solved with (and only with) a deep learning approach
			- To design and develop a deep learning model to solve those problems
			- To optimize the used technology for best performances (using scalable technologies, fine tuning parameters and hyper-parameters)

	boundaries and function shape - Linearity vs non-linearity - Value the dataset - How much does it cost to develop a deep learning model at home? - ML and AI platforms: avoid uncovering the wheel - Specialization x generality = network complexity - Architectures and performance optimization (Deep learning): - Verticality vs horizontal in a neural network - -Number of neurons, number of layers - What activations to use - Normalize the dataset - Preprocessing - SGD algorithms and optimal solution convergence speed - Overfitting, penalty techniques, drop-out - Convolutional neural networks - Convolution filters - Edge detection - Padding - Stride convolution - Multi-dimensionality - Recurrent neural networks - Long short term memories		

	- Gated Recurrent Unit - Transformers - Auto-encoders - Hopfield and Kohonen networks (SOM) - Simulated annealing - Restricted Boltzmann machines - Transfer learning - Multi task learning - End to end learning - Composite architectures (eg.: image captioning) - Reinforcement learning - Q-learning - Applications: Extraction of features from images - Image segmentation - Image recognition OCR - Natural Language processing - Recommendation systems		

Table 5: MACHINA Module 5 learning outcomes

Module 5	**Communicating the merits, challenges and implications of ML technology to customers and within own organisation**		
Learning outcomes correspond to EQF 5	Addresses the fundamental communication principles and personal communication strengths and weaknesses. Define the various technologies and approaches to accomplish communicative goals.		
	Knowledge Knows / Aware of: - Fundamental communication principles and practices - Fundamental knowledge of the terminology used in describing the discipline which is being communicated - Personal communication strengths and weaknesses	**Skills** Able to: - Able to explain elements of the discipline to different parties - Engage in active listening - Articulate characteristics of mediated and non-mediated messages - Locate and use information relevant to the goals, audiences, purposes and contexts - Present messages in multiple communication modalities and contexts - Identify contexts, situations and barriers that impede communication self-efficacy - Apply ethical communication principles and practices	**Competences** Able to: - Adapt messages to the diverse needs of individuals, groups and contexts - Differentiate between various approaches of communicating issues - Select creative and appropriate modalities and technologies to accomplish communicative goals - Adjust messages while in the process of communicating - Adapt messages to the diverse needs of individuals, groups and contexts)

Table 6: MACHINA Module 6 learning outcomes

Module 6	Legislation, Ethics, Project Management related to ML		
Learning outcomes correspond to EQF 5	Introduces the legislation and EU law and regulations for AI and ML applications. Defines the life cycle, and cost/benefits for a machine learning project.		
	Knowledge Knows / Aware of: - RGPD and legal constraints - Machine learning characteristics - Data value / costs model - Project management patterns	**Skills** Able to: - Analyze terms of service/consents to manage security and ethical constraints - Characterize a system - Identify business-based requirements - Identify data sources characteristics - Define a data collection project thanks to project management patterns - Evaluate project benefits/costs	**Competences** Able to: - Analyze a system - Build a decision support system using data analysis and machine learning - Define Terms of Service allowing further data analytics - Organise a Data analysis / machine learning project

4 GUIDELINES ON HOW TO PREPARE CORRESPONDING LEARNING UNITS

The MACHINA learning outcomes lay the ground for the formulation of MACHINA learning units. The learning units should be specified in intellectual output 2, according to the project application form. As so, this section intends to provide guidelines on how to prepare the MACHINE learning units. 7 presents a template and example of specification of a learning unit and guidelines for learning materials. A learning unit should include the following items: a title, the module of which the learning unit is part of, the list of topics of the learning unit content, prerequisites, learning materials, planned duration of the learning unit and references.

Training materials will be created and if suitable reused to support learning on the technical components and practical applications, such as lecture notes, slide presentations, case studies, FAQs, and audiovisual aids.

Assessment of learning outcomes means methods and processes used to establish the extent to which a learner has in fact attained particular knowledge, skills and competences. In order to determine whether the learner has acquired the proposed knowledge, skills and competences and to provide learners with the opportunity to evaluate the extent to which they have attained the desirable knowledge and skills, assessment materials for each learning unit will be developed.

Table 7: Template and example of specification of a learning unit

Learning unit title: Unsupervised Learning
Part of module 3 – **ML Algorithms, programs and protocols**
Learning unit objectives
- Give a global vision of unsupervised learning and its utility - Provide different algorithms of unsupervised learning, implementation and deployment

Topics / Content
This learning unit is comprised of the following parts: - Part 1: Metrics, distances and inertia-based calculation - Part 2: Partition-based clustering - Part 3: Hierarchical clustering - Part 4: Density-based clustering - Part 5: Spectral clustering - Part 6: Neural network-based topological clustering - Part 7: Incremental clustering - Part 8: Constrained clustering - Part 9: Co-clustering - Part 10: Multi-view clustering - Part 11: Complete practical examples of applying clustering of data with Python.
Prerequisites
- Have followed module 1 of the MACHINA curricula, if no previous knowledge on mathematics foundations. - Knowledge of Python programming language.
Learning materials
Presentation slides and lecture notes: 40-60 Exercises: 2-4 FAQs: 10-20 Case studies: 2-3

For this learning unit, situation case studies can be useful. The situation case requires ananalysis of the information embodied in the case and asks students to delineate the significant relationships existing among the various items of information. The forum discussion is designed to develop an understanding of why things went wrong and how that could have been avoided.

Assessment material:

Multiple choice questions and quizzes, case studies and application scenarios analysis will comprise the bulk of assessment material.

Multiple choice questions: 20-30

Short response questions: 5-10

case studies and application scenarios analysis: 1-2

All learning materials will be available in English and in the partnership languages (FR,GE, IT, GR, RO).

Planned duration of the learning unit

- Average duration for reading the lecture notes and slide presentations: 2h
- Learner personal work: 5h

References

1. Data Clustering, https://www.oreilly.com/library/view/data-clustering/9781466558229/
2. Applied Unsupervised Learning withPython:https://www.amazon.fr/dp/B07KX42TXY?linkCode=gs2&tag=uuid0c9- 21
3. Unsupervised Learning Algorithms, https://www.amazon.fr/dp/B01EZSHMXK?linkCode=gs2&tag=uuid0c9-21
4. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, https://www.amazon.fr/dp/B00475AS2E?linkCode=gs2&tag=uuid0c9-21

5 REFERENCES

[1] The European Qualifications Framework, https://ec.europa.eu/ploteus/search/site?f%5B0%5D=im_field_entity_type%3A97 and http://www.cedefop.europa.eu/en/events-and-projects/projects/european- qualifications-framework-eqf

[2] Cedefop (2017). Defining, writing and applying learning outcomes - A Europeanhandbook, Luxembourg: Publications Office of the European Union, http://www.cedefop.europa.eu/en/publications-and-resources/publications/4156

[3] The European e-Competence Framework (e-CF) - A common Europeanframework for ICT Professionals in all industry sectors, http://www.ecompetences.eu/