

Overview

This standard identifies the competences you need to assist in the development of machine learning algorithms and their implementation, in accordance with approved procedures.

You will be required to assist in the development of machine learning algorithms, which encompass machine learning workflow; supervised and unsupervised learning; implementation of training and test datasets; fitting a classifier model and interpreting the results.

Machine learning algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to perform the task. Machine learning algorithms are used in a wide variety of applications where it is difficult or infeasible to develop a conventional algorithm to perform the task. This will involve the practical use of software tools for machine learning algorithm development.

Your responsibilities will require you to comply with organisational policy and procedures for assisting in the development of machine learning algorithms, and to report any problems with the development activities or tools used that you cannot personally resolve, or are outside your permitted authority, to the relevant people. You will be expected to work to instructions, alone or in conjunction with others, taking personal responsibility for your own actions.

Your underpinning knowledge will be sufficient to provide a sound basis for your work, and will enable you to adopt an informed approach to applying procedures for the development and testing of machine learning algorithms. You will recognise the importance of good quality data and the role of feature selection in the learning process. You will be able to distinguish between supervised and unsupervised learning and use a classifier algorithm to produce a model for a given dataset. You will be able to judge the performance of a model, and identify the role of training and test datasets in this process.

This role can be increasingly found in any sector or organisation and in particular those associated with implementing automated reasoning systems that can learn to respond based on training datasets provided. This activity is likely to be undertaken by people working as Junior Machine Learning Specialists or Junior Machine Learning Engineers.

Performance criteria

You must be able to:

1. Apply standard machine learning techniques to create and evaluate a classification model for supervised learning for a given scenario
2. Select a classifier algorithm, and use approved software tools to load a dataset and produce a machine learning classifier model
3. Examine datasets in order to identify features relevant to model development through machine learning
4. Assist in producing a classifier model by submitting pre-prepared data to a machine learning classifier algorithm
5. Assist in evaluating the measures of model fit for a classifier model
6. Assist in identifying the features of a classifier model and documenting them in the context of a given dataset
7. Assist in selecting and applying statistical tools to generate descriptive statistics from different datasets for the purpose of feature selection
8. Manipulate datasets with approved software tools to generate descriptive analytics representations

Knowledge and understanding

You need to know and understand:

1. The purpose, key features and applications of machine learning
2. The role that algorithms play in machine learning
3. The importance of good quality data and feature selection in effective machine learning
4. The difference between supervised, unsupervised and reinforcement learning
5. The purpose of training and testing data sets in developing and evaluating a machine learning model
6. The role of feature identification in model development
7. The basic principles of statistical and machine learning methods including how a machine learns through an algorithmic process
8. The main types of algorithm used, including decision trees, nearest neighbour and linear classifier
9. How a decision tree works to develop a classifier model
10. The industry standard packaged solutions for coding machine learning algorithms
11. The characteristics of under-fitting and over-fitting in a classifier model
12. The different approaches to model improvement in a classifier problem
13. The basic measures of performance that can be used in model development, covering classification, prediction and clustering

TECIS805301

Assist in the development and implementation of machine learning algorithms



Developed by ODAG

Version Number 1

Date Approved March 2020

Indicative Review Date March 2023

Validity Current

Status Original

Originating Organisation ODAG Consultants Ltd

Original URN TECIS805301

Relevant Occupations Data Operations; Software Development

Suite IT and Telecoms

Keywords Machine learning, algorithms, artificial intelligence
