

Overview

This standard identifies the competences you need to develop machine learning algorithms, the different approaches to machine learning and their implementation, in accordance with approved procedures.

You will be required to select and apply machine learning algorithms to build models for prediction, classification or clustering. You will develop competence in applying programming to algorithms to undertake the process of training and validation in order to develop machine learning solutions.

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 the development of machine learning algorithms.

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 understand how machine learning works, the tools it deploys, the process of training and validation, and the reliability of the decisions it supports. It will provide an insight into the mechanistic nature of machine learning as one facet of artificial intelligence as well as the role of deep learning. It will also introduce the concepts of error and bias in model development, and their importance in evaluating model performance. You will appreciate the applications and implications of machine learning.

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 Machine Learning Specialists or Machine Learning Engineers.

Performance criteria

You must be able to:

1. Identify and select machine learning algorithms for classification, regression and clustering
2. Design, develop and apply machine learning algorithms (including supervised, unsupervised, or reinforced learning) using approved tools to meet organisational requirements
3. Evaluate machine learning model performance in order to develop models that minimise bias and error
4. Select and apply programming approaches to deliver machine learning solutions, including; to extract data, develop machine learning workflow, for data cleansing and normalisation, data visualisation and for the implementation of machine learning algorithms
5. Build and refine machine learning models, including; training and validation; cross validation; feature transformation and selection; bias, error and precision-recall metrics; over fitting; and performing ensemble learning to improve model performance
6. Develop models for prediction and clustering using a machine learning workbench in line with organisational needs
7. Develop strategies for model improvement as well as improvements to data and re-sampling
8. Create and disseminate reports and presentations to a range of audiences in order to raise awareness of the role, purpose and outcomes of machine learning algorithms

Knowledge and understanding

You need to know and understand:

1. What is meant by artificial intelligence, machine learning and deep learning
2. The components of a machine learning system and how to apply them
3. The characteristics of different machine learning methods and models including; supervised learning; unsupervised learning; reinforcement learning, ensemble learning; predictive modelling; classification models; regression models and clustering models
4. Algorithms, including clustering, decision tree and Bayesian algorithms
5. The classification algorithms for regression, clustering, instance-based and dimensionality reduction and how to apply them
6. Current applications of machine learning in business, engineering, society and science
7. The importance of selecting, cleaning and transforming datasets for machine learning applications
8. Structured and unstructured data sources and the data acquisition process
9. Computational methods for algorithm development
10. Training and testing data sets
11. Model development and how to evaluate machine learning model performance
12. Feature selection and dimensionality reduction
13. Evaluation of Machine Learning outcomes
14. Supervised learning and its use in prediction
15. Unsupervised learning and its use in clustering
16. Reinforcement learning and its use in maximising system performance
17. Common algorithms for supervised, unsupervised and reinforcement learning
18. Relevant programming languages, environments and modern machine learning libraries
19. Neural networks and deep learning
20. The techniques for identifying and reducing bias in datasets and how to apply them
21. How to select and apply machine learning algorithms for classification, regression and clustering

22. Approaches to model improvement in a classifier problem
23. Ethical issues raised by the use of machine learning including algorithmic bias

Develop and implement 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 TECIS805401

Relevant Occupations Data Operations; Software Development

Suite IT and Telecoms

Keywords Machine learning, algorithms, artificial intelligence
