

## **Supplementary material**

### ***Overview of types in AI***

Several types of AI are introduced in this review. Logistic regression (LR) is a popular algorithm used to measure the relationship between the dependent variable and one or more independent variables or predictors by calculating probabilities using a logistic function. LR includes a particular group of models named a generalized linear model. A support vector machine (SVM) is a widely used supervised machine-learning algorithm utilized in classification and regression problems. The SVM algorithm's main objective is to make a perfect decision boundary line that can separate n-dimensional space into the various correct categories. The most appropriate boundary line with fewest errors is called a hyperplane.

### ***Deep learning***

Artificial Neural Networks (ANNs) are one of the main tools used in AI. ANNs are inspired by the neurons of a biological brain and are intended to mimic how humans learn. ANN consists of input, hidden, and output layers. The input layer is the first layer that receives information in numbers, documents, texts, images, and audio files. The middle layer is called the hidden layer, and a single-layer neural network is called a perceptron. However, the hidden layer can be multiple layers and thus gives single or multiple outcomes. A convolutional neural network (CNN) consists of several network layers such as input, convolutional, max pooling, average pooling, and output layers. The total number of layers can be increased or decreased based on how many inputs are used to train the model. The deeper network will perform better with a large dataset. The advantage of using CNN is that it does not need any feature extraction. In the CNN model, the features are automatically extracted hierarchically from the input, and they are further classified using a fully connected layer.

### ***Prediction accuracy of AI model***

The prediction accuracy of an AI model is often evaluated by statistical parameters such as AUC (area under the curve), F-score, and RMSE. AUC (in binary classification only) is used to evaluate how well a binary classification model is able to distinguish between true

positives and false positives. An AUC of 1 indicates a perfect classifier, while an AUC of 0.5 indicates a poor classifier, whose performance is no better than random guessing.

RMSE is the root mean square error. The RMSE will always be larger or equal to the mean average error. The RMSE metric evaluates how well a model can predict a continuous value. RMSE units are the same as those of the parameter of interest, which is useful for understanding whether the size of the error is meaningful or not. The smaller the RMSE, the better the model's performance.

The F-score, also called the F1-score, is a measure of a model's accuracy on a dataset. It is used to evaluate binary classification systems, which classify examples into 'positive' or 'negative'. The F-score is a way of combining the precision and recall of the model, and is defined as the harmonic mean of the model's precision and recall. The F-score is commonly used for evaluating information retrieval systems such as search engines, and also for many kinds of machine learning models. It combines precision and recall relative to a specific positive class. The F-score can be interpreted as a weighted average of the precision and recall, and an F-score reaches its best value at 1 and worst at 0. It is possible to adjust the F-score to give more importance to precision over recall, or vice-versa. Common adjusted F-scores are the F0.5-score and the F2-score, as well as the standard F1-score.

Accuracy is the ratio of correct predictions to the total number of predictions. It is one of the simplest measures of a model. If a model has high accuracy, the model makes correct predictions most of the time. On the other hand, in some cases, many input parameters reduce the prediction accuracy. Therefore, it is important to reduce these inputs and remove the correlations between them. Many techniques have been developed for this purpose, but principal component analysis (PCA) is one of the oldest and most widely used. PCA is a dimensionality reduction method that is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set. Its idea is simple: to reduce the dimensionality of a dataset, while preserving as much 'variability' (i.e. statistical information) as possible, because smaller data sets are easier to explore and visualize, and make analyzing data much easier and faster for machine learning algorithms without extraneous variables to process.