Machine learning methods for optimal compatibility of materials in ecodesign

I. ROJEK† and E. DOSTATNI‡

†Kazimierz Wielki University, Institute of Computer Science, 30 Chodkiewicza St., 85-064 Bydgoszcz, Poland
‡Poznań University of Technology, Faculty of Mechanical Engineering and Management, 5 Maria Skłodowska-Curie Square, 60-965 Poznań, Poland

Abstract. Machine learning (ML) methods facilitate automated data mining. The authors compare the effectiveness of selected ML methods (RBF networks, Kohonen networks, and random forest) as modelling tools supporting the selection of materials in ecodesign. Applied in the design process, ML methods help benefit from the knowledge, experience and creativity of designers stored in historical data in databases. Implemented into a decision support system, the knowledge can be utilized – in the case under analysis – in the process of design of environmentally friendly products. The study was initiated with an analysis of input data for the selection of materials. The input data, specified in cooperation with designers, include both technological and environmental parameters which guarantee the desired compatibility of materials. Next, models were developed using selected ML methods. The models were assessed and implemented into an expert system. The authors show which models best fit their purpose and why. Models supporting the selection of materials, connections and disassembly methods help boost the recycling properties of designed products.

Key words: machine learning methods, classification models, ecodesign, selection of materials, compatibility.

1. Introduction

This innovative study, exploring the selection of materials in ecodesign, is a follow-up on previous research into the application of artificial intelligence (AI) in the selection of materials in product design to provide for their recycling compatibility. The research has been described in [1‒4]. Based on the decision tree induction methods and MLP artificial neural networks, the proposed tools automate the selection of materials in the design process, building upon the designer’s knowledge gained through experience. The expert system applied in previous research featured the following functionalities:

- selection of all additional materials based on the data provided on the main material and the anticipated compatibility,
- selection of connections between the materials.

In this paper, the authors describe the system’s functionality which determines the degree of compatibility between materials. The training examples include data deemed necessary for this task by the designers, i.e., the main material, its technological details, ecological features (cost of recycling), and any added materials. On this basis, the compatibility of materials is assessed. Models based on a neural (RBF or Kohonen) network or a random forest are implemented into the expert system. They support the designer when an additional compatible material, which is to be joined with the main material, needs to be replaced with another (compatible) one. The newly selected added material must be technologically similar to the one used previously in order for the designer to proceed with the design process. Compatible materials need not be separated for recycling. This study is the outcome of continued research into the use of artificial intelligence to support designers in the ecodesign process. The topic is especially relevant in view of the implementation of Industry 4.0. The concept is based on integration of the physical world of manufacturing machines with the virtual reality of the Internet and information technology. Human beings, machines and IT systems have access to tools which enable automated data, information and knowledge exchange [2].

2. Overview of the literature

The concept of sustainable development originated in the final decades of the previous century and is still evolving [5]. However, environmental integrity should not only follow from legal compliance, but should be underpinned by a true belief that protection of the environment is a material aspect of business. Awareness in this area is shaped by the relation of business community with the natural environment, information on and perception of the organisation’s environmental footprint, as well as readiness to take action aimed at mitigating the impact of business operations on the environment [6]. The design process is highly influenced by environmental aspects [7]. Taking them into consideration throughout the product lifecycle minimizes the environmental footprint of products. Ecodesign provides for the manufacturing of products which do not contain hazardous substances or generate hazardous waste. Raw materials used in
manufacturing should be renewable and/or reusable. One of the pillars of ecodesign is the recyclability of products.

Computer systems supporting the environmental assessment of designed products, including recyclability, are coming into common use. Most of them are databases which store data applied in the environmental assessment. However, research is being conducted with the aim to create autonomous systems which suggest the designer the best solutions for optimal environmental parameters of the product. The systems rely on state-of-the-art information technologies implementing artificial intelligence (AI). The challenges of ecodesign fuel the development of increasingly powerful dedicated software tools. Among them are machine learning methodologies which, although not yet widely used in ecodesign, are certainly worthy of attention. Initially developed and explored purely out of scientific curiosity, the methodologies have unexpectedly turned out to be useful in many applications, such as technology, medicine, economics, and even social sciences. They can provide parameter estimates and suggest optimal decisions in the design process.

Recent studies show that businesses have been attempting to incorporate ecodesign into their practice with different levels of success [8, 9], using various ecodesign support solutions, from tools dedicated to ecodesign through ones integrated into other systems, such as Computer Aided Design (CAD) software [10, 11], to tools of general use applied in, e.g., qualitative analyses (such as quality function deployment for environment – QFDE) [12, 13]. A growing number of ecodesign supporting tools use intelligent solutions, such as neural networks or genetic algorithms. One of them is the design for environment (DfE) methodology, which uses the back-propagation neural network (BPNN) model and the technique for order preference by similarity to ideal solution (TOPSIS) method. Based on a BPNN, lifecycle assessment (LCA) models are developed to estimate quantities of hazardous chemical substances and energy consumption for a derivative consumer electronic product throughout the product lifecycle [14]. The solution described in [15] uses the artificial neural network (ANN) for forecasting and performance of product lifecycle assessment (LCA). Any missing data required for the LCA is estimated using the ANN.

Artificial intelligence is also applied in solutions supporting the sorting of retired products in the recycling process. In [16], the authors describe a waste sorting method based on AI combined with an intelligent vision system. The experimental models are pre-trained within VGG-16 (VGG16), AlexNet, support vector machine (SVM), K-nearest neighbour (KNN), and random forest (RF). Another application of AI in waste sorting is the deep learning-based method, implemented by Refined Technologies of Sweden [17]. It recognizes products or product models with a high degree of similarity. The software connects to optical and mechanical systems which sort electronic components based on their material composition. The above-mentioned applications of AI in recycling are first and foremost aimed at streamlining the sorting of waste.

The AI-based methods developed by these authors have been designed so as to support the design of products, boosting their recyclability. Rather than facilitating the sorting of waste, it helps reduce its amount by enabling the design of readily recyclable products made of recoverable and reusable materials. A recyclable product must be made of materials which facilitate disassembly and reuse or re-processing of components at the end-of-life.

Research into the development of legal regulations concerning the consumption of resources in product manufacturing to be implemented in the forthcoming years shows that further restrictions are likely to be imposed on the management of material resources [18].

In view of the above, it seems appropriate to apply tools supporting the environmental analysis of materials as early as at the product design stage.

Many researchers have described the application of ML methodologies across different scientific and practical domains. Numerous research papers discuss Kohonen networks applied for multidimensional data visualization to evaluate classification possibilities of various coal types [19], collision free path planning and control of wheeled mobile robot [20], parametric fault clustering in analog electronic circuits with the use of a self-organizing artificial neural network [21], intrusion detection in software defined networks [22], simulating the milling cutter trajectory [23], or automated monitoring of the surface grinding process [24], with the use of multilayer perceptron (MLP) network, radial basis function network (RBFN), support vector machine (SVM), and the decision tree.

RBF networks are also widely discussed in literature as a tool supporting, among others, rotor fault detection of the converter-fed induction motor [25], local dynamic integration of ensemble in prediction of time series [26], predicting the corrections of the Polish time scale UTC(PL) (Universal Coordinated Time) [27], accurate load forecasting in a power system [28]. Other research papers discuss using random forests to analyse distorted data of an electronic nose for recognising the gasoline bio-based additives [29], or in evaluating the impact of explanatory variables on the accuracy of prediction of daily inflow to the sewage treatment plant [30].

ML methods are widely used in many areas [31–36], ranging from medicine [37] to the estimation of cutting tool wear [38]. However, solutions implementing the RBFN, Kohonen networks or random forest in ecodesign are scarce. Hence the authors’ interest in the application of these ML methodologies.

3. Machine learning methods

The most universal network commonly applied for resolving various problems, including technical ones, has been the MLP [39]. However, RBF networks also have a number of advantages. Firstly, they are able to model any nonlinear function with a single hidden layer, which eliminates the need to decide on the number of layers at the design stage. Moreover, the RBF network typically has one hidden layer with radial neurons, each of which models the Gaussian process-based response surface [40]. Radial networks are composed of neurons whose activa-
Machine learning methods for optimal compatibility of materials in ecodesign


Further research focused on the random forest method, considering its meaningful advantages, such as fast classification, clarity, “mature methodology” and numerous practical applications. Moreover, random forests support the processing of both symbolic and numerical data, which is an important feature. A random forest is a fully functional application of the algorithm developed by Breiman [40, 42]. It is an ensemble of decision trees which predict the value of a dependent variable on the basis of a set of independent variables (predictors). In classification tasks, the result is provided as a value of the qualitative dependent variable. A random forest consists of a number of simple decision trees. Each individual tree in the random forest delivers a class prediction and the class with the most votes becomes the ensemble’s prediction. Ensemble predictions are more accurate than any prediction by an individual, even highly complex decision tree.

4. Creation of models based on the ML methods

The creation of ML method-based models for the selection of materials in eco-design has been organised in the following stages:

- analysis of input data for the selection of materials (based on an analysis of material properties),
- development of the training, testing and validation files, featuring example selections of materials to be used in the creation of the ML methods and assessment of their effectiveness,
- development of models based on neural networks and random forest,
- assessment of the models,
- selection of the most effective material selection models and their implementation in the expert system.

4.1. Data preparation. Recycling-oriented ecodesign relies primarily on the selection of materials and methods of connecting them. The ultimate goal is to design a product made of the largest possible number of standardized and recyclable materials. This has a positive impact on the environment in the last stages of the product’s lifecycle, such as maintenance or withdrawal from use [43]. When selecting product materials, we should also consider their compatibility: materials used in a product should allow for their recycling at the end of the lifecycle without having to be separated [44]. Recycling parameters are shaped primarily by the chemical composition of materials. Matrices of material compatibility have been developed [45, 46]. The matrices list the compatibility of materials regarding, among others, their recyclability. Figure 1 shows a matrix for selected plastics, which compares their recycling compatibility.

For a detailed analysis, selected properties of materials have been added upon consultation with designers. The files have been prepared based on an analysis of properties of materials, such as: name (text value, e.g. PVC), density in grams per cubic centimetre (a real number, e.g. 7.88), tensile strength expressed...
in megapascal (a real number, e.g., 35.5), elongation at yield point (Re) expressed as a percentage value (a real number, e.g., 5.5), processing temperature expressed in degrees centigrade (a real number, e.g., 20.8), the dielectric constant (a real number, e.g., 2.0), Young’s modulus (E) expressed in gigapascal (a real number, e.g., 4.61), water absorbency expressed as a percentage value (a real number, e.g., 22.55), environmental impact (a logical value, e.g., true), the recycling cost expressed in PLN per kilogram (a real number, e.g., 4.25), where a positive value represents a profit from the sale of material, and a negative one – the cost of disposal, and the name of the added material (a text value, e.g. ABS).

The data set includes 980 examples.

4.2. RBF network-based material selection models. The input parameters (10 inputs) for the construction of a neural network include material properties, including eco-friendliness, the added material for the assessment of compatibility with the main material, and one output – the decision class, which in this case is the compatibility of materials (Fig. 2).

Table 1 shows the most important data of the material selection models developed as RBF networks. The models have different numbers of neurons in the hidden layer (from 20 to 60). The activation function in the hidden layer is the Gaussian function, the activation function in the output layer is the Softmax function, and the training algorithm is the RBFT.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>RBF neural networks</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN</td>
<td>20</td>
</tr>
<tr>
<td>Effectiveness [%]</td>
<td>75.06</td>
</tr>
<tr>
<td>TE</td>
<td>0.9637</td>
</tr>
<tr>
<td>TEE</td>
<td>0.8452</td>
</tr>
<tr>
<td>VE</td>
<td>0.9234</td>
</tr>
<tr>
<td>EF</td>
<td>SOS Entropy</td>
</tr>
<tr>
<td>VE</td>
<td>0.8745</td>
</tr>
<tr>
<td>EF</td>
<td>0.8131</td>
</tr>
<tr>
<td>VE</td>
<td>0.6932</td>
</tr>
<tr>
<td>EF</td>
<td>Entropy Entropy</td>
</tr>
<tr>
<td>VE</td>
<td>0.5298</td>
</tr>
<tr>
<td>EF</td>
<td>Entropy Entropy</td>
</tr>
<tr>
<td>VE</td>
<td>0.53333</td>
</tr>
<tr>
<td>EF</td>
<td>Entropy Entropy</td>
</tr>
<tr>
<td>VE</td>
<td>0.3956</td>
</tr>
<tr>
<td>EF</td>
<td>Entropy Entropy</td>
</tr>
<tr>
<td>VE</td>
<td>0.3512</td>
</tr>
</tbody>
</table>


The best RBF network (10–58–1) reached an efficiency of 94.91%. It featured 10 inputs, 58 neurons in the hidden layer, and one output. Measured with cross entropy (CE), the training error was 0.3956, the testing error – 0.2954, and the validation error – 0.3512.

4.3. Kohonen network-based material selection models. Many neural network models were built in the course of the conducted experiments. They all featured the same input layer, whose size resulted from the amount of input data (11 inputs). The neural network models were parameterized by various values: the network topology, number of learning cycles and neighbourhood were changed. The learning process consisted in the assignment of cluster centres to the radial neuron layer. The functioning of a self-organising network during the learning process largely depends on the selection of the measure of distance between the winning neuron and the input vector. The learning coefficient represents the neighbourhood radius, whose value decreases over time.

Table 2 shows the changing parameters of the Kohonen network.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network topology</td>
<td>6×10, 10×20, 15×25</td>
</tr>
<tr>
<td>Neighbourhood</td>
<td>3, 5</td>
</tr>
<tr>
<td>Number of learning cycles</td>
<td>100, 500, 1000</td>
</tr>
</tbody>
</table>

Figure 3 shows a graph representing the learning process of the least effective (3 SOFM 11–60) and the most effective (6 SOFM 11–375) neural network (Fig. 3a and Fig. 3b, respectively), and a graph which enables the visualisation of assignment of cases to trials (Fig. 3c and Fig. 3d). The colour scheme on the right hand side represents the scale of the distance to the
Machine learning methods for optimal compatibility of materials in ecodesign

Bull. Pol. Ac.: Tech. 68(2) 2020

Figure 3. Graphs of Kohonen networks (the best – effectiveness at 99.66%, and the worst – effectiveness at 98.91%). (a) The learning process of the least effective neural network: a) 3 SOFM 11–60, b) 6 SOFM 11–375. A graph which enables the visualisation of assignment of cases to trials for the least effective neural network: c) 3 SOFM 11–60, d) 6 SOFM 11–375)

The best neural network is network no. 6, the least effective – network no. 3.

4.4. Random forest-based material selection models. Random forest-based models were developed based on the same training file as RBF networks. For each random forest-based model, classification parameters were set, which included the cost of incorrect classification and a priori probability. The stopping criterion included the parameter of the minimum number of

Table 3
Models of Kohonen neural networks

<table>
<thead>
<tr>
<th>Net ID</th>
<th>Network name</th>
<th>Error (learning)</th>
<th>Kohonen’s learning algorithm</th>
<th>Neighbourhood</th>
<th>Topology</th>
<th>Effectiveness [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SOFM 11–60</td>
<td>0.942197</td>
<td>100</td>
<td>3</td>
<td>6×10</td>
<td>99.00</td>
</tr>
<tr>
<td>2</td>
<td>SOFM 11–60</td>
<td>0.941915</td>
<td>500</td>
<td>3</td>
<td>6×10</td>
<td>99.00</td>
</tr>
<tr>
<td>3</td>
<td>SOFM 11–60</td>
<td>1.032770</td>
<td>100</td>
<td>5</td>
<td>6×10</td>
<td>98.91</td>
</tr>
<tr>
<td>4</td>
<td>SOFM 11–60</td>
<td>0.994567</td>
<td>1000</td>
<td>5</td>
<td>6×10</td>
<td>98.98</td>
</tr>
<tr>
<td>5</td>
<td>SOFM 11–200</td>
<td>0.553123</td>
<td>1000</td>
<td>3</td>
<td>10×20</td>
<td>99.39</td>
</tr>
<tr>
<td>6</td>
<td>SOFM 11–375</td>
<td>0.342688</td>
<td>500</td>
<td>3</td>
<td>15×25</td>
<td>99.66</td>
</tr>
<tr>
<td>7</td>
<td>SOFM 11–375</td>
<td>0.343706</td>
<td>500</td>
<td>5</td>
<td>15×25</td>
<td>99.63</td>
</tr>
</tbody>
</table>
training examples in the node. The cost of incorrect classification refers to the distribution of examples among classes. Cost minimisation corresponds to minimisation of the percentage of incorrectly classified cases, where a priori probabilities are directly proportional to the class size, and the cost of incorrect classification is equal for each class [40].

Figure 4 shows a summary of the random forest model for the class of compatibility. The model contains 100 decision trees, the maximum size of a tree is 100. The assessed risk reached 0.159226 for the training sample and 0.155844 for the testing sample. The importance of predictors was determined. The most important predictor turned out to be the additional material. Other predictors were lower in rank.

4.5. Example implementation of a neural network into an expert system. An expert system implementing a neural network model supporting the selection of materials and showing their compatibility (Fig. 5) advises designers in the product development process.

<table>
<thead>
<tr>
<th>Predicted class – good</th>
<th>Predicted class – limited</th>
<th>Predicted class – incompatible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed good</td>
<td>220</td>
<td>55</td>
</tr>
<tr>
<td>Observed limited</td>
<td>10</td>
<td>510</td>
</tr>
<tr>
<td>Observed incompatible</td>
<td>25</td>
<td>55</td>
</tr>
</tbody>
</table>

A classification matrix for the values observed and predicted is shown in Table 4.

The random forest proves to be highly efficient in terms of predicting the class with reference to the observed class, with the efficiency at 78.57%, 97.14%, and 54.29% for “good”, “limited”, and “incompatible”, respectively.

On the basis of the input data representing properties of the main and added materials, the system provides information on their compatibility. In the case subject to analysis, the response generated for PVC and ASA is good.

5. Summary

Considering large amounts of numerical input data used in ecodesign, the application of machine learning methods as classification methods seems to be an appropriate choice to support the process. Classification has been conducted using...
RBF neural networks, Kohonen networks, and random forest. All three methods are characterised by excellent classification abilities. It follows from the study that Kohonen networks are the best classifiers.

The selected classification methods have taken ecodesign to the next level. Owing to them, the knowledge which has so far been hidden in human minds or stored in databases can be automatically acquired and used in the design process. The study has proven ML methods to be highly useful and effective in the support of selection of materials in ecodesign.

Being advanced data mining algorithms, ML methods open a wide array of uses of data stored in databases. Among others, they enable automated acquisition of designers’ knowledge, thus becoming a solution which helps utilize what has been learnt through experience. The models developed and implemented into the decision support system aid in the design of new products. The system tells which materials are compatible and which are not, while the designer can take a decision on the materials to be used. What is important, the added material should meet both the environmental criteria and the technological requirements specified in the expert system, so as to maintain technological coherence of the product.

The more compatible the materials in the product are, the less time it takes to separate them. Incompatible materials are those which cannot be recycled or which degrade the secondary raw material – they deteriorate the quality of the recyclate. Therefore, only materials which together form a compatible combination should be used in production. Compatibility matrices enable us to determine which materials can be combined to make recycling most effective. Minimising the diversity of materials helps reduce the risk of their incompatibility.

Acknowledgments. The work presented in the paper has been co-financed under 02/23/DSPB/8716 and the research potential of the Kazimierz Wielki University.

References

206  I. Rojek and E. Dostatni

Bull. Pol. Ac.: Tech. 68(2) 2020


