Machine learning approach to packaging compatibility testing in the new product development process

Norbert Piotrowski[1]

Abstract

The paper compares the effectiveness of selected machine learning methods as modelling tools supporting the selection of a packaging type in new product development process. The main goal of the developed model is to reduce the risk of failure in compatibility tests which are preformed to ensure safety, durability, and efficacy of the finished product for the entire period of its shelf life and consumer use. This kind of testing is mandatory inter alia for all aerosol packaging as any mechanical alterations of the packaging can cause the pressurized product to unseal and stop working properly. Moreover, aerosol products are classified as dangerous goods and any leaking of the product or propellent can be a serious hazard to the storage place, environment, and final consumer. Thus, basic compatibility observations of metal aerosol packaging (i.e. general corrosion, pitting corrosion, coating blistering or detinning) and different compatibility factors (e.g. formula ingredients, water contamination, pH, package material and coatings) were discussed. Artificial intelligence methods applied in the design process can reduce the lengthy testing time as well as developing costs and help benefit from the knowledge and experience of technologists stored in historical data in databases.

Keywords Machine learning · Compatibility testing · New product development · Smart products

Introduction

Every day new products or new formulations of cosmetic, toiletry, household, or technical products are developed and launched into the global markets. Research and development (R&D) departments are constantly outdoing themselves in developing completely new or improved, more sustainable, more cost-effective products and innovations. However, every new product requires a lot of testing before it reaches the store shelves or our homes. These include laboratory and application tests (performed to confirm the various marketing claims appearing on product labels), drop tests (performed to ensure the strength of packaging, to provide information relevant for the packaging's design and to confirm whether it is adequate to withstand the risks of the distribution cycle or not) and transportation tests (performed to simulate possible stresses and strains on the product and pallets during transport from factories to the distribution centers or final

destinations). Most important, however, are stability and compatibility tests, which ensures the shelf life of the product and typically take the longest time in the new product development (NPD) process.

Each finished good that emerges on the market should ultimately be safe to the consumer, durable and, above all, effective throughout the entire period of its shelf life and consumer use. When conducting stability testing for a finished product, the primary goal is to ensure that no detrimental changes in the product's intended physical, chemical, microbiological properties, as well as the product's safety and functional performance, occur during handling, transporting, and storing under conditions that are appropriate for those activities (Williams & Schmitt, 1992). Moreover, the stability of a cosmetic is one of the most important parameters in assessing the quality and safety of a cosmetic, as required by EU law (in accordance with Regulation (EC) No 1223/2009). The primary packaging directly affects the stability of the finished product due to the several reactions that could take place between the package and the product (corrosion, chemical processes, product components migration adsorption into the container) and the packaging with the

✉ Norbert Piotrowski
    norbert.piotrowski@pg.edu.pl

1   Gdańsk University of Technology, Gdańsk, Poland

environment (the ability of a packaging to protect contents from water and/or atmospheric oxygen impacts).

The aim of the compatibility test is to assess chemical migration between the product and its primary packaging. Because there may be physical and chemical interactions between the product, the package, and the external environment, the packaging can have a direct impact on the stability and safety of the product. For this reason, compatibility is one of the mandatory safety tests which confirms that product's performance is not negatively affected during product's specified shelf life. The test should be performed whenever new product is developed, existing products are reformulated in the market, there is a shift to new packaging or modification of the manufacturing process. Compatibility testing is also mandatory for all aerosol packaging. Any mechanical alterations of the packaging can cause the pressurized product to unseal and stop working properly due to can unpressurizing, diminishing product efficacy by metal ion contamination or clogging the valve by loose pieces of coating or laminate film. Moreover, aerosol products are classified as dangerous goods and any leaking of the product and propellent can be a serious hazard to the storage place, environment and final consumer.

As the development cycle of new products is relatively short, real-time compatibly testing is not always practical and feasible. Moreover, due to the great diversity of products and their inherent complexity, regulatory bodies have not established and uniform product shelf-life test criteria. As a result, R&D laboratories must create and implement their own expedited stability and compatibility testing method that is both cost-effective and efficient in addressing the testing requirements. These abbreviated methods usually take several months anyway, which sometimes significantly inhibits NPD projects. Moreover, sometimes testing several packaging or formulations options at once is not possible due to various constraints. Therefore, after negative tests, other solutions and variants are tested which can further extend the duration of tests.

In the paper, basic compatibility observations of metal aerosol packaging (i.e. general corrosion, pitting corrosion, coating blistering or detinning) as well as several compatibility factors (e.g. formula ingredients, water contamination, pH, package material, and coatings) were discussed. The effectiveness of selected machine learning (ML) algorithms as modeling tools supporting the selection of packaging in the product development process were compared. Artificial intelligence (AI) applied in the design process can reduce lengthy testing time, developing costs and help profit from the expertise and experience of technologists stored in historical data. The topic is especially relevant in view of the implementation of Industry 4.0. The idea is based on combining the physical world with the virtual world of the Internet and information technology. The technologies that enable automated data, information, and knowledge exchange are available to people, robots, and information technology systems (Rojek et al., 2019). Smart products development is one of the critical aspects of this new industrial revolution, creating several opportunities for businesses and markets. The disruptive changes that are being addressed by Industry 4.0 will have an impact throughout the whole product lifecycle. This will be made possible by the introduction of advanced computer platforms that are used to create digital tools for product development and prototyping (Nunes et al., 2017).

Related work

A new product development (NPD) can be conceived as a multi-phased process in which the design is detailed gradually. When developing a new product, harmonizing all stages of process is essential since only then can development time be shortened. Several researchers have discussed the activities that take place during the various stages of the NPD process and have concluded that the volume and contents of the activities that take place during each stage are contingent on the quantity and the purpose of the product (Ayağ, 2016; Kušar et al., 2004). Moreover, there is a significant gap between the NPD efforts and knowledge sharing involved in individual production and those involved in mass production (Duhovnik et al., 2001; Gao & Bernard, 2018).

The process of developing a product requires individual adjustments to be made in accordance with the product, the manufacturing process and the designers involved. Within the context of this process, the know-how and approach of the designers as well as their level of expertise are very individual and are frequently incapable of being stated in a manner that is rule-based without significant effort. Despite this, there are a lot of routine tasks that may be identified and that have a significant automation potential. It has been presented that machine learning, and particularly deep learning, has a tremendous capacity to recognize patterns, derive knowledge from complex data sets and support each phase of the complex process of product development (Krahe et al., 2020; Santos et al., 2017). Some examples here may include the use of Naïve Bayes algorithm to map customer requirements to product variants (Wang & Tseng, 2015), Intelligent Support Systems for product design (Figueroa Pérez et al., 2018; Hayes et al., 2011) or AI-based Computer Aided Engineering for automated product design (Krahe et al., 2019). There are some applications of machine learning modelling tools supporting the selection of materials. Rojek and Dostatni (2020) considered both technological and environmental parameters in classification methods to guarantee the desired compatibility of materials used in ecodesign process. Merayoa et al. (2019) examined how the artificial intelligence techniques can assist the manufacturing process by choosing the most

convenient material for the envisaged applications according to their properties and in-service behavior. Silva et al. (2021) used the k-Nearest Neighbor (KNN) algorithm to classify and select biodegradable packaging produced from fish gelatin incorporated with palm oil and clove and oregano essential oils.

Due to the development of many technologies associated with the fourth industrial revolution and in addition to the concepts of Smart Manufacturing, Smart Processes, Smart Factories, the term Smart Product has also become a buzz word in recent years. Moreover, the idea of Smart Product appears in a variety of contextual situations and application fields with diverse meanings (Gutiérrez et al., 2013). Maass and Janzen (2007) outlined three fundamental requirements for Smart Products: adaptability to situational circumstances, adaptation to actors who engage with products or product bundles, and adaptation to underlying business limitations. The authors of the same paper divide these requirements into six characteristics for a fully implemented Smart Product: location-based, personalized, adaptive, proactive, business-aware, and network-capable. Mühlhauser (2008) describes Smart Products as entities (tangible objects, softwares, or services) designed for self-organized integrating into various smart environments throughout their lifecycle, offering improved simplicity and approachability through the use of improved interactions by various means, such as AI and ML. Smart Products may feature various applications form speech recognition systems and production control systems.

Recent studies also introduce the idea of a system for Smart Virtual Product Development (SVPD) that can support the decision-making process of industrial product development cycle at many stages and activities, such as product design, production, and inspection planning. Improvement is accomplished by utilizing knowledge of formal prior decisions events that are recorded, stored, and retrieved as a set of experience (Ahmed et al., 2019). Some researchers have also presented the idea of Lean Product Development but also new criteria for intelligent and Smart Product Development (SPD) by implementing Industry 4.0-related information systems (Rauch et al., 2016). On the basis of the axiomatic design methodology, the authors outline a set of guidelines and principles for designing a lean product development process. These recommendations highlight how cutting-edge technology and tools can be leveraged to create a lean and smart product development process by connecting them to concepts from Industry 4.0.

Most early studies as well as current work focus on the need of defects detection systems for finished goods. As a result of the rapid development of machine vision, image processing and pattern recognition technologies, industrial automation detection has become an inevitable part of many manufacturing processes, as it can significantly improve their precision and efficiency. Wu and Lu (2019) combined in

their study a machine vision and machine learning technologies to examine defects on the surface of the printed packaging box based on support vector machine. In order to rapidly identify printing defects, decrease the cost of human sorting, and increase the production effectiveness, the study has been applied to the printing and packaging carton pipeline sorting manipulator. Park et al. (2022) studied a deep learning-based automatic defect detection system that can train product characteristics and determine defects using open sources. The model was applied to the disposable gas lighter manufacturing process to detect the liquefied gas volume defect of the lighter. Paraskevoudis et al. (2020) developed neural networks-based model for identifying 3D printing defects during the printing (mainly stringing effect) by analyzing video captured from the process. It can help minimize printing costs, as the operator is notified about possible flaws and can stop the process on early stages. Moreover, authors believe that the model can be further developed in order to adjust the printing process. Some recent works have also used machine learning approaches to detect corrosion in metals caused by environmental factors. Atha and Jahanshahi (2018) presented different convolutional neural network-based approaches for corrosion assessment in steel structures. They studied the effect of different model architectures, sliding window sizes and color spaces. A key benefit of convolutional neural networks was its ability to build features without relying on human effort or prior knowledge. The study of Pidaparti (2007) provides an overview of the computational methods developed for the corrosion damage assessment of aerospace structures and artificial neural network-based model for material loss and residual strength predictions.

The existing approaches regarding AI-based product development assists the designers in obtaining a broad picture of the massive volumes of data and, in some cases, reducing NPD time. The real design process and its implementation, however, remain in charge of the technologist and his expertise. Hence, in this study a concept is presented on how past knowledge can be formalized so that it can be transmitted and used even by a very experienced technologist.

Aerosol stability and compatibility testing

Aerosol dispenser is any non-reusable container that contains gas: compressed, liquefied, or dissolved under pressure, with or without a liquid, powder or paste, and equipped with a release device that allows the contents to be expelled as solid or liquid particles suspended in a gas, as a foam, powder, gel, paste, or in a liquid state (Council Directive of the European Union of 20 May 1975). Aerosol products offer a wide range of applications from mass-market products such as personal care, cosmetics, and household

products, to specialized aerosol types designed for medical or industrial purposes. Next to a wide range of potential applications, aerosols combine easiness of usage, resource efficiency and unique performance. Thanks to many benefits that they deliver, the worldwide market has seen a continuous growth in the sales of aerosols finished goods over the past few years (Online document. The FEA (European Aerosol Federation) Statistics Report). Packaging which is an integral part of the aerosol spray consists of a metal (aluminum or steel), plastic or glass can with a permanently attached continuous or metering valve, and actuator (stem-fitting button, applicator or spraycap) designed to dispense products as spray mists or streams, gels, lotions, foam or just gases. According to data from the European Aerosol Federation (FEA), roughly 90% of aerosol cans are made of metal (aluminum 49% of and steel 40%). Plastic or glass containers are still minor, owing to legislative constraints on the allowed filling volumes (Online Document. The FEA (European Aerosol Federation) Industry Standards).

In order to avoid aerosol product failure, the aerosol container and valve system must be compatible with the product to be filled into the aerosol dispenser. An inadvertent discharge of the product or a total breakdown of the aerosol dispenser resulting in an instantaneous rupture of the aerosol dispenser are both possible outcomes of such a failure. Thus, the product should be submitted to proper testing methods to ensure that such failures of the aerosol products do not occur.

Stability and compatibilty testing methods

One of the earliest European regulations related to product safety applicable to aerosol goods is the Directive 75/324/EEC of the European Union of 20 May 1975 on the approximation of the laws of the Member States relating to aerosol dispensers. The document is commonly referred to as the Aerosol Dispensers Directive (ADD) and its main objective is to assure that products covered by the regulation are safe for customers and other users in terms of pressure, flammability, and inhalation hazards (Council Directive of the European Union of 20 May 1975). Additionally, to the restrictions specified in the directives, required recommendations are also included in FEA standards (Online Document. The FEA (European Aerosol Federation) Industry Standards). The complete set of FEA standards, which includes testing methods as well as dimensions and performance criteria, provides technical best practices created by and for the European aerosol industry. Moreover, some international ISO standards have been refined for several FEA standards, and some of them have replaced previous FEA standards. They provide specific information on each part of the packaging, which enables aerosol packaging manufacturers, production line manufacturers, and aerosol

Table 1 Example of stability and compatibility test for aerosols according to FEA 603 standard (Online Document. The FEA (European Aerosol Federation) Industry Standards)

Storage temperature	Position of storage	Minimum number of samples	Minimum storage time
20 °C	Upright	6	1 year
	Inverted	6	
37 °C	Upright	6	4 months
	Inverted	6	
50 °C	Upright	6	2 months
	Inverted	6	

filling supplies to reach unified degree of efficient collaboration.

FEA 603 standard entitled "Guidelines to test long-term preservation and to measure the loss of weight" provides a method to assess the stability and compatibility of the product's contents with the components of the aerosol container and valve, as well as the weight and/or pressure loss that may occur while storage at various temperatures and over a specific period of time. Since the reactions of the substances contained in the aerosol product must not impair its mechanical resistance, even when stored for a long time, the long-term testing approach is recommended for the storage of all aerosol products in metal, glass, and plastic containers. According to the method, the size of the test is always a problem in all circumstances. The greater the size (in terms of number of samples, storage conditions, and storage time), the greater the accuracy of the results. However, in order to fulfill the time and space constraints, compromise is required. The test's size example, as shown in Table 1, is intended to cover most scenarios but should be considered as a minimal size. Moreover, test samples must reflect the containers and valves that will be used in the intended production. If various types of packaging components are to be compared, they must all be subject to the same test, with the same testing conditions, storage temperature, and storage time. All the samples that will be evaluated must be prepared so as to be as close as possible to the industrial conditions. They should be carefully weighed and numbered before being stored and for at least 24 h after the filling. Additionally, the pressure at ambient temperature and other parameters such as spray pattern, discharge rate or organoleptic parameters can be measured and noted.

After the specified storage time, the samples must be taken out of the storage rooms and kept at room temperature. After 24 h, parameters such as weight loss and pressure can be measured. In order to verify spray characteristics, each valve needs to be actuated and after that, samples can be pierced and emptied to recover the product and compare the residue

Fig. 1 Accelerated aging test for shelf-life validation: **a** shelf life and testing time relationship, **b** storage temperature and testing time relationship
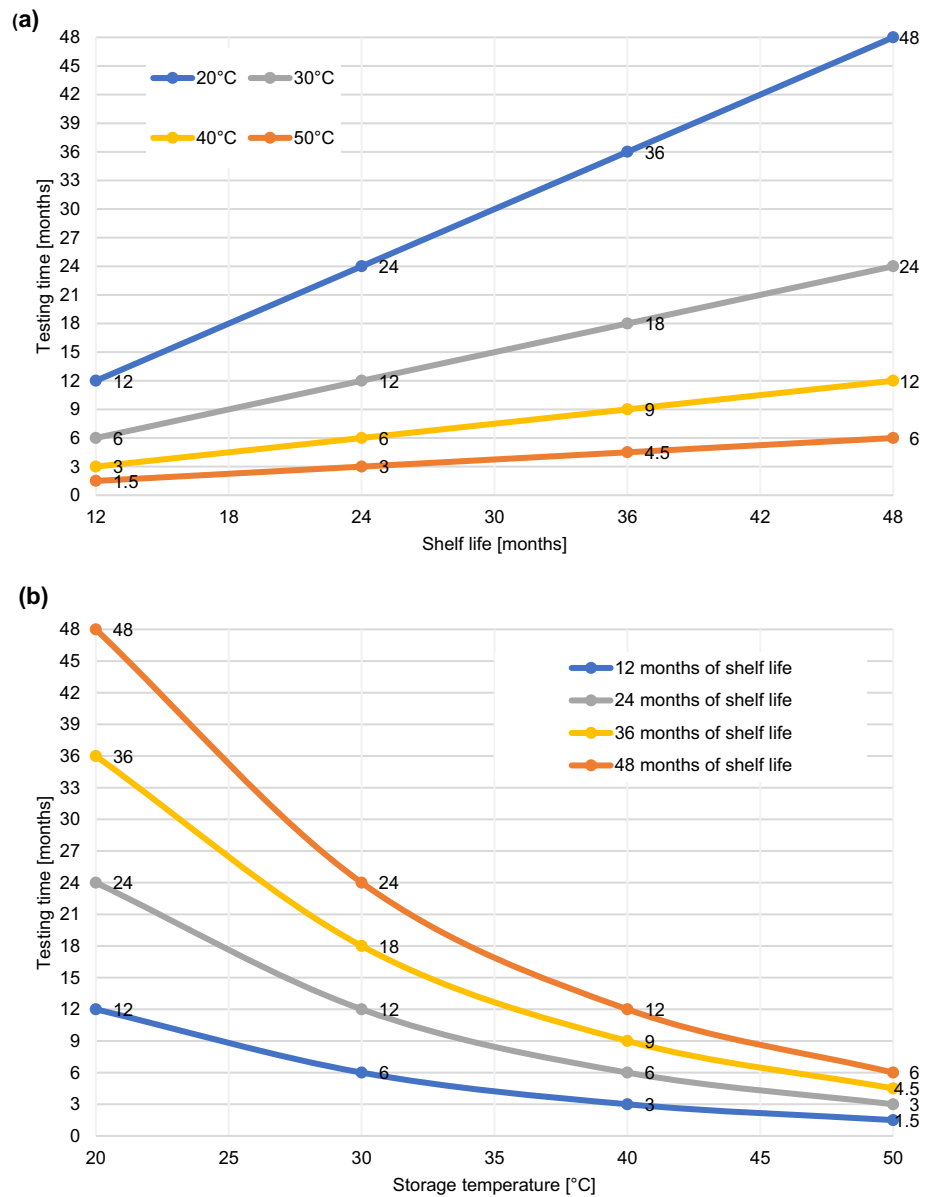
with control samples. Finally, the inner and outer walls of the cans and of the mounting cups in the opened samples can be examined and their compatibility with the formulation can be evaluated.

Despite specific guidelines, the methodology included in the FEA standards does not allow to determine the shelf life expiry date of the product. Test times and samples number are only given there as an example. Therefore, aerosol filling companies have usually their own test methods. Some laboratories, in order to gather information on the corrosion of the aerosol packaging use electrochemical measurements (Root & Maury, 1960; Tait & Maier, 1986). However, most new product development teams still conduct accelerated aging test for shelf-life validation. The test

is based on a van't Hoff thermodynamic temperature coefficient that indicates that the rate of a chemical reaction will double for every 10 °C increase in temperature (Piotrowski, 2022). The time and temperature of sample storage relationship on product shelf life is shown in Fig. 1. It should be noted that for aerosol products the temperature should not exceed 50 °C, and for some formulations even the temperature of 40 °C can degrade some chemical raw materials. Furthermore, any research that involves accelerated aging must also include real-time aging in order to correlate the results.
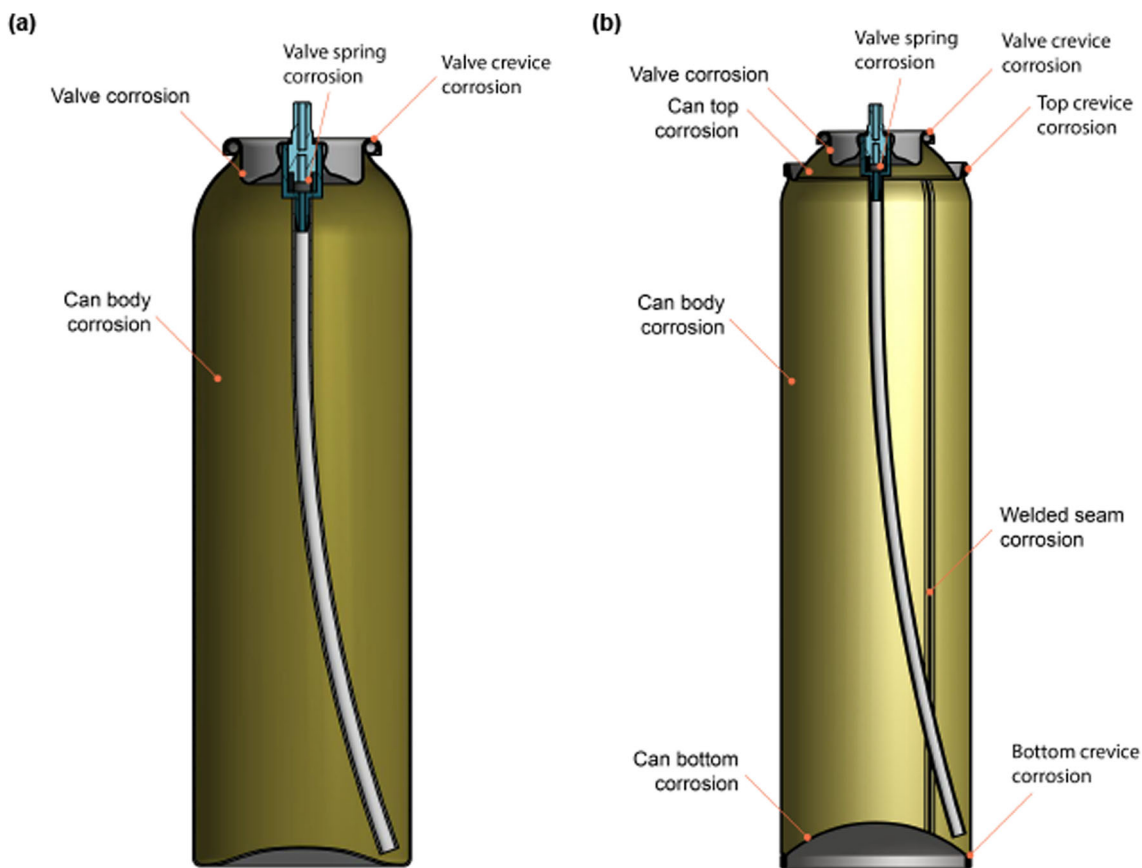
Fig. 2 Aerosol packaging defect's locations: **a** 1-piece aluminum can, **b** 3-pieces tinplate can

Aerosol packaging compatibilty defects

Aerosol containers are manufactured mostly from aluminum (1-piece cans) or tinplated steel (2- or 3-pieces cans). These metals can be either bare (not common with aluminum) or covered with a polymer coating. In addition, various kinds of defects can appear not only on the container, but also on the valve, which can also be made of bare metal or covered with various coatings. Hence, aerosol packaging defects could be a metal corrosion, a polymer corrosion or both. Depending on the packaging used, defects may appear in different areas of the product. The different corrosion locations for both aluminum and tinplated steel containers are shown in Fig. 2.

Moreover, two categories of defects can be observed in aerosol packaging: general and localized. General corrosion (Fig. 3a) occurs over relatively large areas of aerosol packaging and it produces a porous layer of corrosion product on the surface of packaging metals. The porosity and non-uniform thickness cause non-uniform diffusion of materials, which may lead to pitting corrosion under the general corrosion layer. Another example of a general defect is detinning (Fig. 3b) of tinplated steel aerosol containers. In most cases, the process is not considered as troublesome and does not indicate a negative result of compatibility test. However,

sometimes with long shelf life, the removal of a layer of tin can lead to localized corrosion. Pitting corrosion of metals (Fig. 3c) and blistering of internal polymer coatings (Fig. 3d) are the most common forms of localized corrosion, which are typically very small and randomly distributed around the aerosol packaging.

Corrosion is an issue for all types of aerosol containers and packaging materials. However, in order for the process of corrosion to take place, there must be at least two components present: a material with a surface that is prone to corrosion and an environment that contains corrosive chemicals. Any new product development involves the risk of product and packaging incompatibility. Figure 4 contains historical data from 277 compatibility tests. It can be noticed that the risk of product failure is very high as the total number of the tests failed was 41%. However, the compatibility testing can reduce this risk as the defects risk decreases as the testing time increases. Corrosion or other defects in most cases appear very quickly, and the higher risk in first month can be caused, either by the technologist's little experience and poor selection of components, or from the aggressiveness of the product formulation. This risk decreases with time, but still defects can also be observed at the end of testing.

Fig. 3 The most common defect's forms in aerosol containers: **a** general corrosion, **b** detinning, **c** pitting corrosion, **d** polymer coating blistering
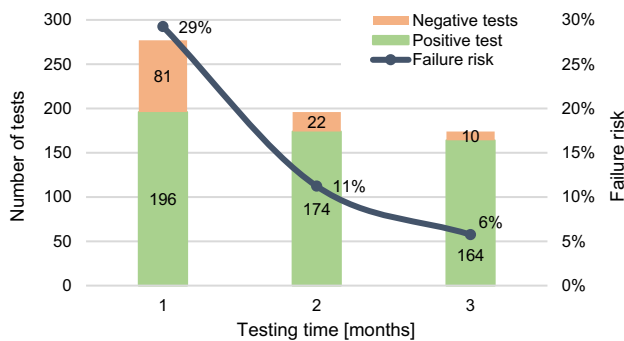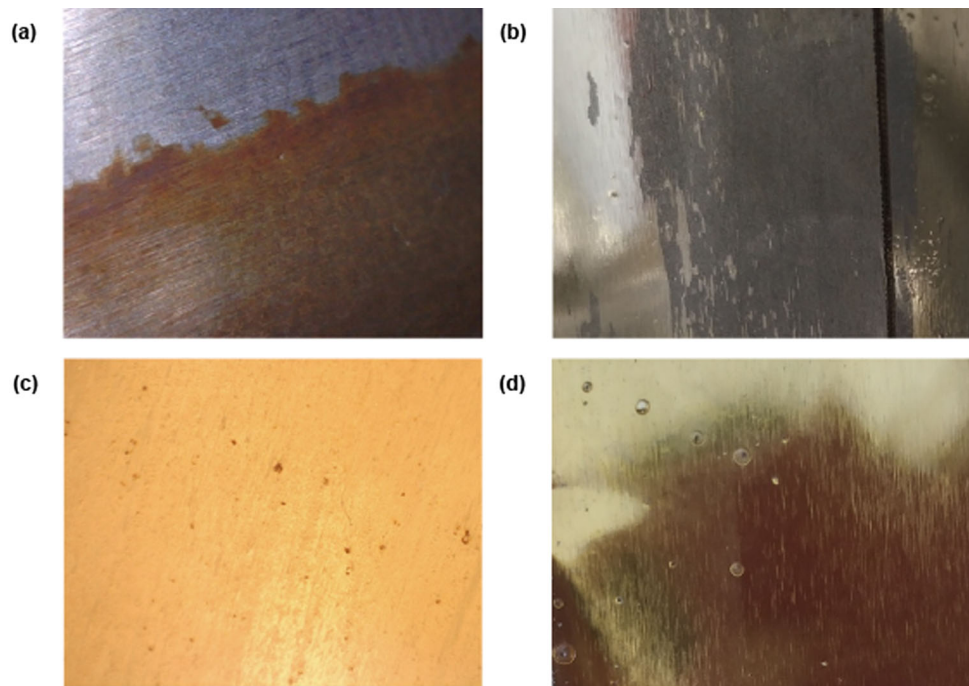
Fig. 4 Accelerated aging test failure risk based on historical data

Data-driven formula and packaging compatibility prediction model

The current state of corrosion science is extensive but not sufficient to predict whether corrosion will occur or not using equations, data tables or chemistry fundamentals. Therefore, compatibility testing is crucial to determine if new and derivative formulations are corrosive to certain types of aerosol cans (1-piece aluminum or 3-pieces tinplated cans), to choose the most corrosion-resistant form of internal can surface treatment (thin tin metal coating or various polymer coatings), and to assess if alternative packaging is appropriate for established formulas. In addition, parallel to factors directly related to packaging, corrosion and other possible defects are influenced by parameters and formula factors such as pH formula and physical form,

content of water, fragrances, corrosion inhibitors, surfactants and acids. In addition, physics-based models can be very effective in predicting some phenomena, yet they are built on strong assumptions that may not hold true under specific circumstances. To tackle these limitations and constraints, data-driven predictive modeling approaches might be employed. They are based on machine learning or statistical methods that can increase their performance with every new dataset. Since machine learning algorithms often require vast amounts of training datasets, they can provide predictions in the absence of any predetermined mechanistic relationships or system behaviors (Piotrowski, 2022).

In this paper, a data-driven approach was employed to predict the compatibility of a new product's formula and its packaging. The model showed in Fig. 5 demonstrates a classification approach which is a subset of supervised machine learning and it draws a conclusion from the input values given for training and generates an output that categorizes a set of data into classes. The input data include information collected from previous experimental compatibility tests as well as details about the formula parameters, packaging factors as well as test results (labels). The dataset consisted of 277 tests and some examples data are presented in Table 2. Additionally, the data was randomly split into two groups: training dataset (80%) and test dataset (20%). Due to the relatively small amount of data and many parameters the study analyzed two approaches, when the model predicts only two classes: compatibility, incompatibility (model 1) or when the model predicts five classes: full compatibility, detinning, general corrosion, pitting, coating blistering (model 2). The

Fig. 5 Supervised classification model for product compatibility prediction
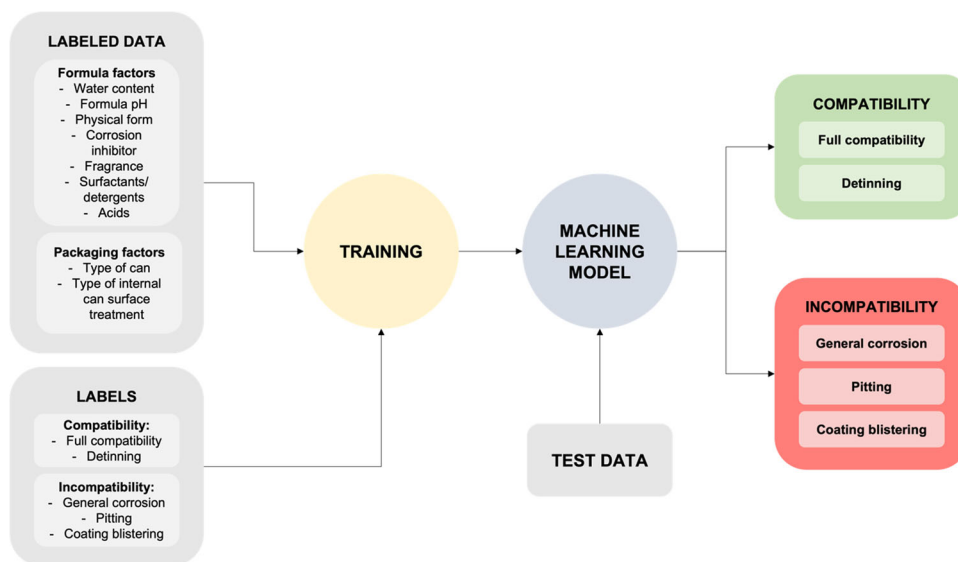
Table 2 Examples of data obtained in experimental tests

Product	Water	pH	Perfumes	Corrosion inhibitor	Surfactants	Acids	Can type	Coating	Label
Air freshener	Yes	7	Yes	No	No	No	Alum	Epoxy gold	Pitting
Hair spray	No	9	Yes	No	No	No	Tinplate	Epoxy gold	Full compatibility
Surface cleaner	No	11	No	Yes	Yes	Yes	Tinplate	2 × epoxy gold	Coating blistering
Deodorant	Yes	6	Yes	No	No	No	Tinplate	None	General corrosion
Biocide	Yes	10	Yes	Yes	Yes	No	Tinplate	Epoxy gold	Full compatibility

more complex is the model, the more it is prone to over-fitting (Ying, 2019). Small datasets need the use of basic classifier models, such as short decision trees, Support Vector Machine (SVM) or k-Nearest Neighbors (kNN). In general, these relatively simple models are less adept at data-driven learning than more complex algorithms e.g., neural networks, hence reducing their susceptibility to overfitting. However, this may be also prevented by minimal tuning, employing cross-validation, regularization, feature selection, and bucketing, all of which attempt to minimize complexity and boost bias.

In this study MATLAB Machine Learning Toolbox and Classifier tool were used to predict the compatibility test results. The built-in classification learner was used to automatically train a selection of different classification models on the experimental data. The choice of classifier type depends mainly on the dataset, but also the trade-off in speed, flexibility, and interpretability must be made. In the first stage, the selection assumed only simple models such as decision trees or discriminant analysis with lower flexibility that provide sufficient accuracy and avoid overfitting. Moreover, the algorithms that were not suitable with both categorical and numeric data (e.g. discriminant analysis or nearest neighbors) were rejected at the very beginning. In addition, automated training allowed to quickly try all selected classifiers (decision trees, support vector machines, logistic regression, naive Bayes, kernel approximation, ensembles and neural networks) and then explore only the promising models interactively. When the training was finished, the corresponding plots and results were explored and compared. Three models with the highest predicted accuracy score namely decision trees, Support Vector Machine (SVM) and Artificial Neural Networks (ANN) were chosen to fitting step and further analysis. For each model, the performance was compared by inspecting the results in the plots and trying to include and exclude different features in the model. To improve the model further, the classifier hyperparameter

Table 3 Overall accuracy of tested ML tools for two predictions models

Model	No of classes	Decision tree	SVM	Neural networks
1	2	86.8%	76.3%	90.2%
2	5	82.6%	71.9%	84.8%

options were changed and then all models were trained with the new options.

The first model that was further analyzed was decision trees, which is one of the most popular approaches for representing classifiers. It classifies instances by arranging them according to their feature values. Each node in a decision tree represents a feature of a classifiable instance, and each branch represents a possible value for that node. Beginning with the root node, instances are classified and ordered depending on their feature values (Kotsiantis, 2007). Decision trees are generally easy to interpret, quick for fitting, prediction and memory efficient, but their predictive accuracy might be sometimes poor. To avoid overfitting a simple decision trees should be grown, and the maximum number of splits must be adjusted to control their depth. The study initially tested all three model options i.e., coarse tree, medium tree and fine trees, but the best accuracy score was obtained with medium trees for finer distinctions between classes, when the maximum number of splits is 20 and when the split criterion options was set by default to Gini's diversity index. Second chosen algorithm was SVM, which was initially developed by Vapnik (Vapnik, 1995) and has since attracted considerable interest in machine learning research. According to several studies, SVMs can be very accurate when it comes to classification performance. However, the performance of the methods is quite sensitive to how the cost and kernel parameters are chosen for certain datasets (Srivastava & Lekha, 2010). As with the first algorithm, each of the nonoptimizable support vector machine options was trained first (i.e., Linear SVM, Quadratic SVM, Cubic SVM, Coarse, Medium or Fine Gaussian SVM). Then the model was improved by feature selection and by changing some advanced options. The highest score of accuracy was achieved with fine gaussian kernel function and when auto scale mode was used in the tested model. The third chosen model, which was analyzed further in the study, was artificial neural networks. ANN is a deep learning method which arose from the concept of the human brain and which consists of large number of units (neurons) joined in a pattern of connections (Zhang et al., 2019). Units in a net are usually segregated in three layers: input layer, hidden layer and output layer. If artificial neural networks are thought of as systems, the structural parameters are the number of hidden layers, the number of neurons in each hidden layer, the parameter in the activated function of each neuron, the weight on the edge, and the parameter bias associated with

each neuron (Wu et al., 2016). The predicted accuracy of neural network models is often high, and they may be employed for multi-class classification, however they are not easy to interpret. The size and number of fully connected layers in the neural network boost the model's flexibility. As the data in the study is less complex and is having fewer dimensions, the MATLAB classifier tool allowed to use one fully connected layer which multiplies the input by a weight matrix and then adds a bias vector. Different neural network model hyperparameter options were tested and the best accuracy was obtained with the first layer size set to 10 units and when default rectified linear unit activation (ReLU) function was applied.

Finally, all three chosen models with optimal options were trained and evaluated ten times and, subsequently, the average accuracy was calculated (Table 3). Neural networks showed the best overall accuracy: 90.2% for model 1 and 84.8% for model 2. Furthermore, the multi-class confusion matrices were shown in Figs. 6 and 7 to further examine the performance of developed models. The number of observation values from the classification model that were correctly and incorrectly categorized is summarized in the confusion matrices. The actual label of classification is represented on the confusion matrix's coordinate axis, while the predicted class is represented by the matrix's horizontal axis. Moreover, to comprehend how the presently chosen classifier succeeded in each category the True Positive Rates (TPR), which is the ration of correctly classified observations to true class and False Negative Rates (FNR), which is the ratio of misclassified observations to true class, were calculated. It can be seen in Fig. 7 that the classifiers did worst in predicting coating blistering, while they did best in detinning.

Conclusion

Time is a crucial parameter in any new product development. Multiple sample prototypes are made during the designing process and it is inefficient and unreasonable to test the stability and compatibility of all samples. However, the research results presented in the article indicate that more than 40% of the compatibility tests performed with a negative result. A negative test requires a change in technology, reformulation or change in packaging and additional testing, which also involves additional time. As any new product development involves the risk of product and packaging incompatibility,

Fig. 6 Confusion matrices summarizing the performance for model 1 of different classifiers: **a** decision tree, **b** SVM, **c** neural networks
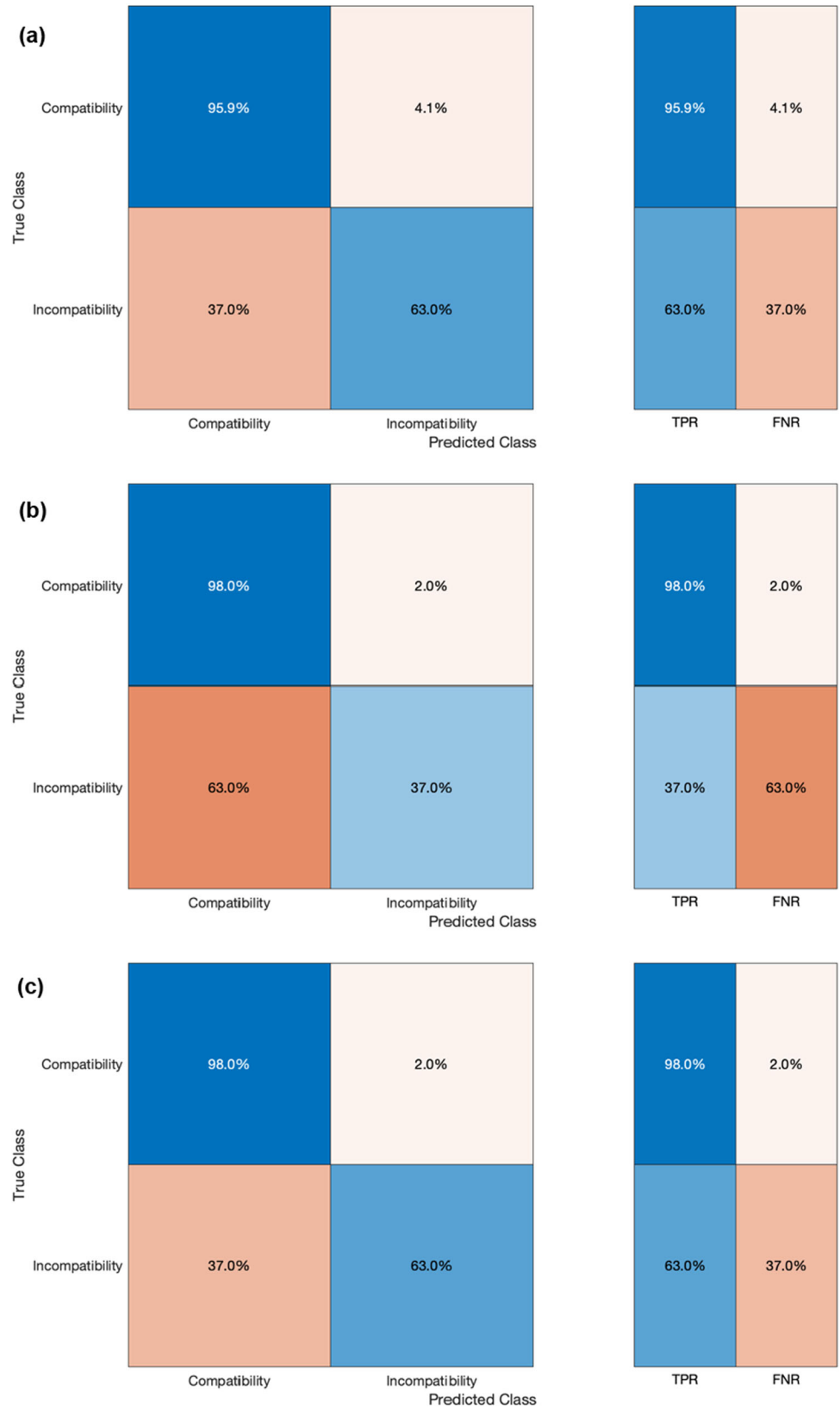
Fig. 7 Confusion matrices summarizing the performance for model 1 of different classifiers: **a** decision tree, **b** SVM, **c** neural networks
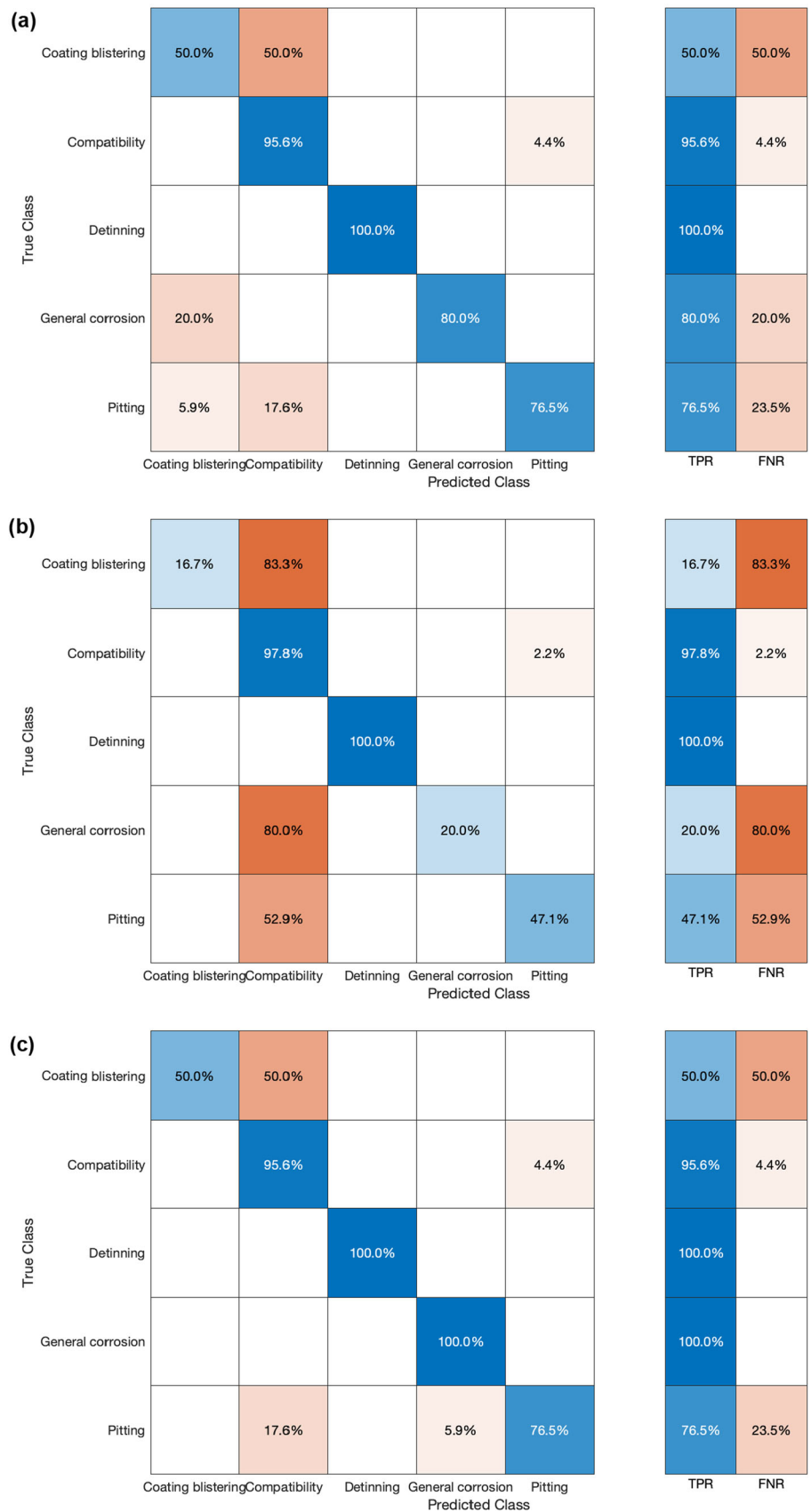
this kind of testing is mandatory to different kind of finished goods such as cosmetics or all aerosol products. In the paper, most common compatibility testing methods for aerosol products were discussed. Accelerated testing at elevated temperatures was presented to show how finished good shelf-time affects the testing time in different temperatures. In addition, basic compatibility observations of metal aerosol packaging (i.e. general corrosion, pitting corrosion, coating blistering or detinning), the defect's locations in the container and valve, as well as several compatibility factors (e.g. formula ingredients, water contamination, pH, package material and coatings) were discussed.

The paper preset a data-driven model to predict the compatibility of formula and packaging. As the current state of corrosion science is extensive, but not sufficient to predict whether corrosion will occur or not on the metal aerosol packaging, classification model which considers various factors can apply here. The input data included information collected from previous compatibility tests of different aerosol products and their formulations. The study analyzed two approaches, when the model predicts two classes (model 1): compatibility, incompatibility or when the model predicts five classes (model 2): full compatibility, detinning, general corrosion, pitting, coating blistering. The effectiveness of three selected machine-learning algorithms were compered for those two models. Neural networks showed the best overall accuracy in the prediction of the compatibility class: accuracy 90.2% for model 1 and 84.8% for model 2. Very similar results were obtained with the decision tree algorithm: 86.8% for model 1 and 82.6% for model 2. The worse result was obtained for SVM algorithm: accuracy 76.3% for model 1 and 71.9% for model 2. All the accuracy results obtained were still higher than the risk of test failure calculated from the experimental data, and for which only 59% pass the tests. It was also noted that the classifiers had the biggest problem in predicting coating blistering, while they did very well in predicting detinning and general compatibility. Moreover, it can be noticed that TPR parameter was much higher with samples describing compatibility compared with the incompatibility, which may be due to the fact that there were more positive tests in the dataset and it is much harder for the algorithm to guess the type of incompatibility than to determine whether the material will be compatible at all.

Although the model does not give full confidence in the success of the test, artificial intelligence applied in the new product development process can reduce lengthy compatibility testing time and developing costs by reducing the risk of failure of the tests carried out. Since for most products the shelf life follows a minimum of 2 years, this means that after 3 months the probability that there will be no need to retest which can double at best scenario the developing time will equals about 90%. Based on a historical data this probability equaled only 59% and was very dependent on the knowledge

of the technologist. This overall can increase development capacity while reducing product developing costs through shortening project lead times and faster sales of finished products. While machine learning algorithms generally need a lot of data for training and testing, much more data is needed to demonstrate statistically and verify that the algorithms are effective. Therefore, future work should address collecting more data from tests, expanding the model with additional factors and considering different locations of defects, with a breakdown of the aerosol container and valve. The additional machine learning algorithms can be applied to not only predict the class of compatibility, but also to estimate the length of a product's shelf life.

Declarations

Conflict of interest The author has no conflict of interest to declare that are relevant to the content of this article.

References

Ahmed, M. B., Sanin, C., & Szczerbicki, E. (2019). Towards experience-based smart product design for Industry 4.0. *Cybernetics and Systems, 50*, 165–175. https://doi.org/10.1080/01969722.2019.1565123

Atha, D. J., & Jahanshahi, M. R. (2018). Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection. *Structural Health Monitoring, 17*(5), 1110–1128. https://doi.org/10.1177/1475921717737051

Ayağ, Z. (2016). An integrated approach to concept evaluation in a new product development. *Journal of Intelligent Manufacturing, 27*, 991–1005. https://doi.org/10.1007/s10845-014-0930-7

Council Directive of the European Union of 20 May 1975 on the approximation of the laws of the Member States relating to aerosol dispensers 1975 Official Journal of the European Communities L147

Duhovnik, J., Starbek, M., Dwivedi, S. N., & Prasad, B. (2001). Development of new products in small companies. *Concurrent Engineering Research and Applications, 9*, 191–210. https://doi.org/10.1177/1063293X0100900302

Figueroa Pérez, J. F., Leyva López, J. C., Santillán Hernández, L. C., & Pérez Contreras, E. O. (2018). An intelligent decision support system for the design of new products. In *Data science and knowledge engineering for sensing decision support* (pp. 995–1002). https://doi.org/10.1142/9789813273238_0126

Gao, J., & Bernard, A. (2018). An overview of knowledge sharing in new product development. *The International Journal of*

Advanced Manufacturing Technology, 94, 1545–1550. https://doi.org/10.1007/s00170-017-0140-5

Gutiérrez, C., Garbajosa, J. Diaz, J., & Yagüe, A. (2013). Providing a consensus definition for the term "Smart Product". In *20th IEEE international conference and workshops on engineering of computer based systems (ECBS)* (pp. 203–211). https://doi.org/10.1109/ECBS.2013.26

Hayes, C. C., Goel, A. K., Tumer, I. Y., Agogino, A. M., & Regli, W. C. (2011). Intelligent support for product design: Looking backward, looking forward. *Journal of Computing and Information Science in Engineering*. https://doi.org/10.1115/1.3593410

Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica, 31*, 249–268.

Krahe, C., Bräunche, A., Jacob, A., Stricker, N., & Lanza, G. (2020). Deep learning for automated product design. *Procedia CIRP, 91*, 3–8. https://doi.org/10.1016/j.procir.2020.01.135

Krahe, C., Iberl, M., Jacob, A., & Lanza, G. (2019). AI-based Computer Aided Engineering for automated product design—A first approach with a Multi-View based classification. *Procedia CIRP, 86*, 104–109. https://doi.org/10.1016/j.procir.2020.01.038

Kušar, J., Duhovnik, J., Grum, J., & Starbek, M. (2004). How to reduce new product development time. *Robotics and Computer-Integrated Manufacturing, 20*, 1–15. https://doi.org/10.1016/S0736-5845(03)00049-8

Maass, W., & Janzen, S. (2007). Dynamic product interfaces: A key element for ambient shopping. In *Bled eConference.*

Merayo, D., Rodríguez-Prieto, A., & Camacho, A. M. (2019). Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0. *Procedia Manufacturing., 41*, 42–49. https://doi.org/10.1016/j.promfg.2019.07.027

Mühlhäuser, M. (2008). Smart products: An introduction. In *Constructing ambient intelligence* (pp. 158–164). Springer, Berlin. https://doi.org/10.1007/978-3-540-85379-4_20

Nunes, M. L., Pereira, A. C., & Alves, A. C. (2017). Smart products development approaches for Industry 4.0. *Procedia Manufacturing, 13*, 1215–1222. https://doi.org/10.1016/j.promfg.2017.09.035

Online document. *The FEA (European Aerosol Federation) Industry Standards.* Retrieved June 14, 2022, from https://www.aerosol.org/publications-news/publications/standards/

Online document. *The FEA (European Aerosol Federation) Statistics Report. 2020 European aerosol production.* Retrieved June 12, 2022, from https://www.aerosol.org/publication/fea-statistics-report-2020/

Paraskevoudis, K., Karayannis, P., & Koumoulos, E. P. (2020). Real-time 3D printing remote defect detection (stringing) with computer vision and artificial intelligence. *Processes, 8*(11), 1464–1478. https://doi.org/10.3390/pr8111464

Park, S.-H., Lee, K.-H., Park, J.-S., & Shin, Y.-S. (2022). Deep learning-based defect detection for sustainable smart manufacturing. *Sustainability, 14*(5), 2697–2711. https://doi.org/10.3390/su14052697

Pidaparti, R. M. (2007). Structural corrosion health assessment using computational intelligence methods. *Structural Health Monitoring, 6*(3), 245–259. https://doi.org/10.1177/1475921707081975

Piotrowski, N. (2022). Tool wear prediction in single-sided lapping process. *Machines, 8*(4), 59–70. https://doi.org/10.3390/machines8040059

Rauch, E., Dallasega, P., & Matt, D. T. (2016). The way from lean product development (LPD) to smart product development (SPD). *Procedia CIRP, 50*, 26–31. https://doi.org/10.1016/j.procir.2016.05.081

Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products

Rojek, I., & Dostatni, E. (2020). Machine learning methods for optimal compatibility of materials in ecodesign. *Bulletin of the Polish Academy of Sciences-Technical Sciences, 68*, 199–206. https://doi.org/10.24425/bpasts.2020.131848

Rojek, I., Dostatni, E., & Hamrol, A. (2019). Automation and digitization of the material selection process for ecodesign. In *Intelligent systems in production engineering and maintenance* (Vol. 835, pp. 523–532). Springer. https://doi.org/10.1007/978-3-319-97490-3_50

Root, M. J., & Maury, M. J. (1960). Corrosion testing of aerosol products. *Anti-Corrosion Methods and Materials, 7*(3), 85–87.

Santos, K., Loures, E., Piechnicki, F., & Canciglieri, O. (2017). Opportunities assessment of product development process in Industry 4.0. *Procedia Manufacturing, 11*, 1358–1365. https://doi.org/10.1016/j.promfg.2017.07.265

Silva, N. D. S. E., Farias, F. D. S., Freitas, M. M. D. S., Hernández, E. J. G. P., Dantas, V. V., Oliveira, M. E. C., Joele, M. R. S. P., & Lourenço, L. D. F. H. (2021). Artificial intelligence application for classification and selection of fish gelatin packaging film produced with incorporation of palm oil and plant essential oils. *Food Packaging and Shelf Life, 27*, 100611. https://doi.org/10.1016/j.fpsl.2020.100611

Srivastava, D., & Lekha, B. (2010). Data classification using support vector machine. *Journal of Theoretical and Applied Information Technology., 12*, 1–7.

Tait, W. S., & Maier, J. A. (1986). A novel corrosion test cell for predicting the long-term corrosion behavior of metal aerosol containers. *Corrosion, 42*(10), 622–628. https://doi.org/10.5006/1.3583033

Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer.

Wang, Y., & Tseng, M. M. (2015). A Naïve Bayes approach to map customer requirements to product variants. *Journal of Intelligent Manufacturing., 26*, 501–509. https://doi.org/10.1007/s10845-013-0806-2

Williams, S. D., & Schmitt, W. H. (1992). *Chemistry and technology of the cosmetics and toiletries industry* (1st ed.). Springer.

Wu, Y., & Lu, Y. (2019). An intelligent machine vision system for detecting surface defects on packing boxes based on support vector machine. *Measurement and Control, 52*(7–8), 1102–1110. https://doi.org/10.1177/0020294019858175

Wu, Z. F., Li, J., Cai, M. Y., Lin Y., & Zhang, W. J. (2016). On membership of black-box or white-box of artificial neural network models. In *IEEE 11th conference on industrial electronics and applications (ICIEA)* (pp. 1400–1404). https://doi.org/10.1109/ICIEA.2016.7603804

Ying, X. (2019). An overview of overfitting and its solutions. *Journal of Physics Conference Series, 1168*, 1–6. https://doi.org/10.1088/1742-6596/1168/2/022022

Zhang, W. J., Yang, G., Lin, Y., Ji, C., & Gupta, M. M. (2019). On definition of deep learning. In *2018 World Automation Congress (WAC)* (pp. 1–5). https://doi.org/10.23919/WAC.2018.8430387