Advanced Engineering Informatics 38 (2018) 41–53

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Semantic weldability prediction with RSW quality dataset and knowledge construction

Kyoung-Yun Kim*, Fahim Ahmed

Department of Industrial and Systems Engineering, Wayne State University, Detroit, MI 48202, USA

ARTICLE INFO

Keywords:
Semantic weldability prediction
Resistance spot welding
Welding quality
Weldability knowledge construction
Decision tree algorithm
Semantic rules

ABSTRACT

This paper presents a semantic Resistance Spot Welding (RSW) weldability prediction framework. The framework constructs a shareable weldability knowledge database based on the regression rules from inconsistent RSW quality datasets. This research aims to effectively predict the weldability of RSW process for existing or new weldment design. A real welding test dataset collected from an automotive OEM is used to extract decision rules using a decision tree algorithm, Classification and Regression Trees (CART). The extracted decision rules are converted systematically into SWRL rules for capturing the semantics and to increase the shareability of the constructed knowledge. The experiments show that the RSW ontology, along with SWRL rules that contains weldability rules constructed from the datasets, successfully predicts the weldability (nugget width) values for RSW cases. The predicted nugget width values are found to be in-close proximity of the actual values. This paper shows that semantic prediction framework construes an intelligent way for constructing accurate and transparent predictive models for RSW weldability verification.

1. Introduction

Resistance Spot Welding (RSW) is considered as the primary sheet metal joining method of the automotive industry. In the RSW process, metal surfaces are joined by the heat obtained from resistance to the electric current. The primary advantage of RSW is its relatively low cost, high speed/high volume operations, and high rate of production. These advantages have helped RSW become the most popular sheet metal welding method [1]. In terms of RSW quality, important welding parameters are weld nugget geometry (e.g., nugget width) and the mechanical behavior of the weld nugget. These parameters provide an idea about weldability performance and the robustness of the process. To obtain an acceptable weld quality, various design and process parameters have been involved.

In the industry, nowadays, system integrators/OEMs working with products having metallic structures rely on the material suppliers and testing service companies to conduct the actual physical testing of materials for new weldment designs. The onus is on the supplier or testing companies to provide the test results at the right time. Otherwise, the selection of new materials, processes, and weldment designs are often hindered. When new (or new combinations of) materials are considered for an assembly, new types of physical tests or numerical simulations are often required. However, the multi-physical nature of RSW and its high temperature material property often obstruct the realistic numerical simulation (e.g., finite element analysis).

Moreover, a large number of parameters involved in RSW are interrelated and the weld quality is operation-dependent [2]. The quality of RSW is inconsistent from weld to weld because of the variability of process conditions, noise, and error [3]. The welding defect can cause certain damage to the welded structures, and in severe cases, may lead to catastrophic accidents. It will also reduce the mechanical properties and shorten the service life of the product [4]. One way to inspect weld quality is destructive testing, such as peel test and shear test; however, significant cost and wastage of time and materials are involved [2]. A number of studies have been conducted to perform nondestructive inspection of the welds. Ye et al. [5] develop a vision inspection system, which recognizes the defects of weld based on neural network. Huang and Kovacevic [6] develop a laser-based vision system for non-destructive weld quality inspection. These vision-based systems have tendency to be successful in a lab environment. However, in industrial environment, there are many difficulties in achieving the same results because of the cost involved, the practical implementation limitation, and the image processing efficiency [4].

Data-driven weldability prediction can pose as an alternative solution for the aforementioned challenges. However, it is underutilized in the industry for multiple reasons. First, legacy data obtained from physical experiments often does not conform to new design

* Corresponding author.
E-mail addresses: kykim@eng.wayne.edu (K.-Y. Kim), fahim.ahmed2@wayne.edu (F. Ahmed).

https://doi.org/10.1016/j.aei.2018.05.006
Received 31 March 2017; Received in revised form 5 April 2018; Accepted 24 May 2018

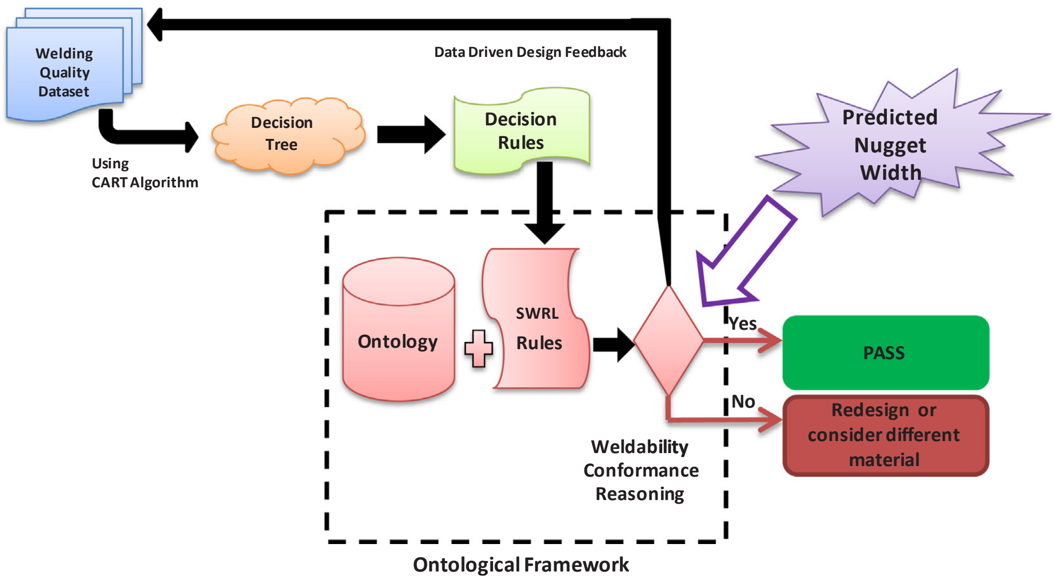

Fig. 1. Semantic RSW weldability prediction framework.

requirements (especially if the material combination is new). Second, often the data is documented in a format (e.g., pdf) that cannot be easily utilized or widely shared. Third, due to the properties of welding materials and the complexity of welding processes, welding test results tend to be inconsistent and form highly-variable datasets [7].

To overcome these issues, in this article, the authors present a semantic RSW weldability prediction framework and construct a shareable weldability knowledge database based on the regression rules from inconsistent RSW quality test datasets. This framework aims to effectively predict the response parameter (i.e., nugget width) of RSW process for existing or new weldment design. This research investigates various input parameters and the correlation of these parameters with the nugget width. Then, this paper employs a decision tree algorithm on the real RSW dataset to plot regression trees (Section 3). From the trees, decision rules are extracted systematically for weld nugget width prediction. These rules are then converted into machine interpretable SWRL (Semantic Web Rule Language) rules with the aid of RSW ontology to realize a streamline of semantic weldability knowledge construction (Section 4).

Fig. 1 illustrates the overall framework for the data-driven RSW weldability prediction and knowledge construction. The extraction of knowledge from the dataset and using the knowledge for semantic rule construction have been shown in Fig. 1. For analyzing the datasets, Classification and Regression Trees (CART) algorithm is employed. CART is relatively simple and is a non-parametric decision tree algorithm (no parameters to be optimized) and it provides an opportunity to realize a systematic rule construction. The constructed decision trees contain decision rules to predict the nugget width. This research assumes that if the dataset collected from the industry is properly analyzed then underlying patterns and influences can be discovered, which is essentially new knowledge about the welding process. To transfer the captured knowledge containing decision rules, the decision rules are mapped systematically to the SWRL rules. To validate the presented framework, multiple experiments are conducted with real RSW quality datasets and the obtained results are discussed in Section 5. Finally, conclusions and future research are discussed in Section 6.

2. Literature review

2.1. Data-driven approach in design and manufacturing

In the design and manufacturing domain, the decision making process is complicated and consists of intricate steps. In an era of information, the data available in each industry is considered as significant, and it is employed to gain competitive advantage in design, manufacturing, analysis, and other decision-making processes. As a result, data science for the industrial sector has emerged. It is the link between data processing technologies (including "big data") and data-driven decision making [8]. Data mining is another important aspect that helps to discover the underlying patterns from large datasets and applies algorithms to extract rules and trends, or to make sense of the dataset. The emergence of data mining has led to the development of many algorithms that extract knowledge (e.g., rules) and features from large datasets [9]. Nowadays, data mining is being applied in various industries, such as semiconductor manufacturing [10], electronic assembly [11], and the health care industry [12]. Data mining techniques are being employed in materials science and engineering as well [13].

2.2. Data mining in manufacturing

Bazan [14] categorizes the learning algorithms, which are of particular interest to industry in four groups: (1) decision tree algorithms [15,16]; (2) decision rule algorithms [17]; (3) inductive logic programming algorithms [18]; and rough set algorithms [19]. The research studies associated with each category are included in the previous statement. According to Chen et al. [21], there are five types of analysis: classification, clustering, association, time series, and outlier. Using the classification analysis, the target class for each test case of the dataset is predicted [22]. There are many methods to classify the data, including decision tree induction, frame-based classification or rule-based classification, expert systems, hierarchical classification, neural networks, Bayesian network, and support vector machines [21].

In the manufacturing sector, classification techniques are employed predominantly to predict quality. Correa et al. [23] conducted high-speed milling tests and collected data to detect quality based on surface roughness. Based on the data, they compared two different machine learning classification methods, Bayesian networks and artificial neural networks. Bayesian networks have been used in various fault diagnosis applications as they combine expert knowledge with probabilistic theory for construction of effective diagnosis methodologies [24]. Kretschmer et al. [25] used data mining techniques to implement a knowledge-based design for assembly in agile manufacturing. In this research, they employed the classification algorithm *k*-nearest neighbor (kNN) with Euclidean distance to identify the objects with similar