European Journal of Operational Research 264 (2018) 294–309

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Computational Intelligence & Information Management

TRANSFORM-ANN for online optimization of complex industrial processes: Casting process as case study

Srinivas Soumitri Miriyala[a], Venkat Subramanian[b], Kishalay Mitra[a,*]

[a] Department of Chemical Engineering, Indian Institute of Technology Hyderabad, Kandi, Sangareddy Telangana 502285, India
[b] Department of Chemical Engineering, University of Washington, 1410 NE Campus Parkway, Seattle, WA 98195, USA

ARTICLE INFO

Article history:
Received 17 July 2016
Accepted 11 May 2017
Available online 17 May 2017

Keywords:
Artificial Intelligence
Multiple objective programming
Neural Networks
Online optimization
Surrogate models

ABSTRACT

Artificial Neural Networks (ANNs) are well known for their credible ability to capture non-linear trends in scientific data. However, the heuristic nature of estimation of parameters associated with ANNs has prevented their evolution into efficient surrogate models. Further, the dearth of optimal training size estimation algorithms for the data greedy ANNs resulted in their overfitting. Therefore, through this work, we aim to contribute a novel ANN building algorithm called TRANSFORM aimed at simultaneous and optimal estimation of ANN architecture, training size and transfer function. TRANSFORM is integrated with three standalone Sobol sampling based training size determination algorithms which incorporate the concepts of hypercube sampling and optimal space filling. TRANSFORM was used to construct ANN surrogates for a highly non-linear industrially validated continuous casting model from steel plant. Multiobjective optimization of casting model to ensure maximum productivity, maximum energy saving and minimum operational cost was performed by ANN assisted Non-dominated Sorting Genetic Algorithms (NSGA-II). The surrogate assisted optimization was found to be 13 times faster than conventional optimization, leading to its online implementation. Simple operator's rules were deciphered from the optimal solutions using Pareto front characterization and *K*-means clustering for optimal functioning of casting plant. Comprehensive studies on (a) computational time comparisons between proposed training size estimation algorithms and (b) predictability comparisons between constructed ANNs and state of art statistical models, Kriging Interpolators adds to the other highlights of this work. TRANSFORM takes physics based model as the only input and provides parsimonious ANNs as outputs, making it generic across all scientific domains.

1. Introduction

Artificial Neural Networks (ANN), one of the efficient data modelling techniques, are finding extensive real world applications including operational research (Sermpinis, Theofilatos, Karathanasopoulos, Georgopoulos, & Dunis, 2013). Well known for their ability to capture nonlinear dynamics of complex data (Barrow & Kourentzes, 2016; Denton & Hung, 1996; Sexton, Dorsey, & Johnson, 1999; Kamini, Vadlamani, Prinzie, & Van denPoel, 2014), ANNs are advantageous over similar class of technologies such as Support Vector Machines (SVMs) and Response Surface Methods (RSMs). SVMs, like ANNs, implement supervised learning techniques to classify the given data. Belonging to the class of machine learning algorithms this method works by creating partition between data points such that the distance from the closest training point to the partition is maximised. Primarily designed for linear classification, this method is also extended for non-linear classification by transforming the working domain to a region of higher dimensions (Sermpinis, Theofilatos, Karathanasopoulos, Georgopoulos, & Dunis, 2017). RSMs on the other hand are statistical models, which try to regress lower order (commonly, second order) polynomials to build data based relationship between explanatory and response variables. Here, the optimizer finds the optimal response of objective function through a sequence of designed experiments. This method has found immense applications in engineering and operational research due to its extreme simplicity and ease in implementation (Shi, Shang, Liu, & Zuo, 2014). However, several disadvantages which creep in due to the heuristic estimation of parameters governing the neural networks create suspicion in their predictability (Wong & Hsu, 2006). This impression is further solidified with the trial and error based determination of training sample size without implementing a formal design of experiments or sample plan. Lack of an intelligent framework for ANN construction has put them in the back step behind robust statistical data modellers such as Kriging Interpolators (Mogilicharla, Mittal, Majumdar, & Mitra, 2015).

* Corresponding author.
E-mail addresses: kishalay@iith.ac.in, kishalay.mitra@gmail.com (K. Mitra).

http://dx.doi.org/10.1016/j.ejor.2017.05.026
0377-2217/

The motivation for the current work is to contribute in operational research a novel ANN building algorithm called TRANSFORM (*TR*ade-off between *A*ccuracy, *N*odes and **S**ample size *FOR M*eta-modelling). This intelligent framework, capable of estimating most of the ANN related parameters, thereby making it parameter free, determines the best configuration and optimal training sample size, simultaneously. While doing this, TRANSFORM ensures a balance between the aspects of over-fitting and prediction accuracy. Further, three novel sample size determination techniques designed using two potential concepts: hypercube (HC) sampling and space filling based single objective optimization (SOOP) formulation, are presented. TRANSFORM is fast enough to be implemented in real time and generic enough to be applied to any physics based model without constraints on dimensionality.

In this work, we aim to explore the scope of TRANSFORM-ANNs (TRANSFORM based ANNs) as surrogate model for online implementation of computationally expensive optimization processes. As an example, we considered a 7-input-2-output industrially validated highly non-linear continuous casting (*concast*) model from steel plants. The main reason behind our decision to use the *concast* model is to demonstrate the potential of TRANSFORM to construct ANNs capable of emulating the complex physics based models. Physics based models are highly robust and accurate owing to rigorous implementation of scientific principles behind the considered phenomena. Often this leads to the comprisal of several Differential Algebraic Equations (DAE), thereby increasing the computational expense in simulating the model (Mogilicharla, Chugh, Majumdar, & Mitra, 2014; Olaf, Barth, Freisleben, & Grauer, 2005; Ruud, Driessen, Hamers, & Hertog, 2005; Uğur, Karasözen, Schäfer, & Yapıcı, 2008). *Concast* is one such model containing a mix of partial different equations, ordinary differential equations and several algebraic empirical correlations, whose details are described in subsequent sections of this paper.

Optimal running of casting plant in steel industries is one of the prime targets of production plant managers. Incorporating the decisions of management, which are mainly driven by the volatile nature of markets, requires the plant wide optimization and control to be implemented in online fashion. In online optimization, the optimizer works in consolidation with a robust controller, thus together forming an effective Advance Process Control (APC) unit. Here, considering the practical changes to be implemented, the optimizer often solves a multiobjective optimization problem (MOOP) in real time. The optimal solutions are then provided to the controller as the set point. However, the working of APC unit in real time requires the inherent model to be computationally efficient. Thus, owing to the complex nature of *concast*, its optimization is always confined to the offline mode.

This enables the implementation of meta-models, also known as surrogate models, which are trained to emulate the physics based model accurately. These surrogates then replace the time consuming physics based model during their optimization to generate simulations fast enough to run it in online mode (Jin, 2011; Tabatabaei, Hakanen, Hartikainen, Miettinen, & Sindhya, 2015).

We thus implemented TRANSFORM in conjunction with the proposed sample size determination techniques to construct parsimonious ANNs capable of predicting *concast* with maximum accuracy. Before moving on to the ANN surrogate assisted online implementation, we optimized the casting model without surrogate using both classical and evolutionary optimizers. Further, a comparative performances between the TRANSFORM based ANN (TRANSFORM-ANN) and Kriging Interpolators is also presented. As it turns out the TRANFORM-ANN outperformed Kriging Interpolators in terms of sample size requirement and statistical accuracy. Significant reduction in computational time due to the implementation of TRANSFORM-ANN assisted optimization of *concast* lead to its online implementation.

Finally, we present a set of operator's rules using the concept of Pareto characterization and *K*-means clustering algorithm. These rules draw the mapping from complex mathematical realization of optimization studies such as Pareto to a set of simple linguistic instructions aiding the ground operators to enable optimal functioning of casting plant.

The rest of the paper is organized as follows – we first present the literature survey of several recent contributions in the field of research which forms the central theme of this paper. This is followed by the continuous casting model description and formulation of its optimization problem. Proposed algorithm TRANSFORM and novel size estimation algorithms are discussed next. The rest of the paper is devoted to surrogate building, comprehensive comparison studies, ANN surrogate assisted optimization, scope for online implementation and discussions on Pareto characterization, all summed up in the results section following which the novelty in current contribution is briefly summarized in conclusions.

2. Literature review

ANNs being potential classifiers can serve as ideal candidates for meta-modelling. However, they suffer with major disadvantages such as those listed below:

- Heuristic based design of architecture.
- No proper guidelines for choosing the transfer function of network.
- No sample plan and measure of optimal sample size.
- ANN model often gets over-fitted.

These drawbacks not only degrade the performance of ANNs but also prevent them from qualifying as potential surrogates for optimization. The objective of this work is to contribute an efficient ANN building algorithm capable of solving all the aforementioned problems simultaneously, within short time frame to ensure its streamlining with online optimization. For this purpose, we present TRANSFORM with two novel sample size estimation techniques. Our proposed methods for architecture estimation and size determination are not implemented yet. However, Dua (2010), reported solving a mixed integer nonlinear programming problem (MINLP) to determine the optimal configuration of ANN. The algorithm being robust, does not address the other issues of ANNs and further, the computational complexity involved in solving MINLP cannot be ignored (Dua, 2010). In another work, configuration of ANN was obtained by solving an optimization problem with the objective of maximizing the prediction accuracy of each of the outputs (Boithias, Mankibi, & Michel, 2012). If the number of outputs are large (say >3), the proposed method might take a huge time to converge (Deb, 2001). Carvalho, Ramos, and Chaves (2011) used the weighted combination of training error, validation error as a single objective function to resolve the problem of architecture design (Carvalho et al., 2011). However, a more generic metric such as the Akaike Information Criteria (AIC) (Akaike, 1971), representing parsimonious nature of the network, might have served as more suitable objective function. Moreover, weighted sum approach has been shown to fail for generating well-distributed Pareto optimal (PO) points while solving MOOPs (Deb, 2001).

The problem of sample size determination (SSD) for black box models in general, is broadly classified into two categories – adaptive and sequential sampling (Eason & Cremaschi, 2014). Sequential sampling is similar to forward marching problem where points are sampled sequentially until the surrogate is trained with desired accuracy. Many researchers contributed in this area using methodologies such as Delaunay triangulations (Davis & Ierapetritou, 2010), Voronoi tessellations (Crombecq, 2011), optimization based approaches and Monte Carlo based random sampling methods (Crombecq, 2011). On the other hand, the adaptive sampling