ATTAINABILITY ESTIMATES IN CHEMICAL REACTOR NETWORKS USING ARTIFICIAL NEURAL NETWORKS

David Ming*
School of Chemical and Metallurgical Engineering, University of the Witwatersrand, Johannesburg, 1 Jan Smuts Avenue, Braamfontein, 2000, Johannesburg, South Africa

Abstract

A method of estimating of the region of attainable states of a reactive system, called the attainable region (AR), using an artificial neural network is proposed. The method is framed as a binary classification problem and trained from a number of reference attainable regions that act as the ground truth in a supervised learning framework. Once trained, the neural network is asked to distinguish between attainable and unattainable states given information about the system such as the feed state. We find that an appropriately trained feed-forward neural network can often estimate regions of attainability with high similarity to reference regions produced by a conventional automated AR construction method.

The method is tested on two systems of varying dimension and complexity, and in both cases the neural network is capable of producing boundary estimations with high confidence. Neural network architecture and training characteristics suitable for attainability estimation are also briefly described.

Keywords

Attainable region; Neural network; Binary classification

* To whom all correspondence should be addressed. Email: david.ming.za@gmail.com

Introduction

The design of a network of chemical reactors that are orchestrated to operate in an optimal manner, for a specified duty, is known as the reactor network synthesis problem. The challenge of reactor network synthesis and, in general, reactor design is that once a design has been formulated, one is often unaware if there are other designs that are superior. Attainable region (AR) theory is seeks to address this problem. The collection of all possible outputs, for all possible reactor configurations, even those that have not been conceptualized, is known as the attainable region.

Whilst the theory of attainable regions is well developed (Feinberg and Hildebrandt, 1997), it is either infeasible or impractical to analytically carry out the computations for many realistic systems. Thus, one must often resort to numerical constructions of candidate regions for practical performance targeting purposes. Over roughly the past two decades, a number of AR construction routines have been formulated that cater toward different construction scenarios. However, the non-linear nature of reaction chemistry, as well as the geometric nature of AR theory, means that it is thus often computationally burdensome to construct an attainable region, which constrains its use in environments where potentially many candidate regions might need to be generated.

The ability of neural networks to generalize and infer trends in data makes them excellent candidates for control and monitoring applications. Machine learning methods have been widely reported in chemical engineering literature over the past three decades, and as a result the scientific literature describing their use is vast. As of yet, no work is known to exist applying neural networks to problems in AR theory.

Attainability estimation using artificial neural networks

We propose a supervised learning problem where an artificial neural network is trained to predict if a test state, **C***i*, is feasible when supplied with feed and kinetic information.

As a demonstration, Figure 1 shows the output of a neural net trained to predict the AR for two-dimensional Van de Vusse kinetics (Van de Vusse, 1964). The shaded region in Figure 1 corresponds to the theoretically correct AR that is supplied at run-time for an arbitrary feed point, which is drawn as a triangle in the figure. The shaded region acts as the ground truth wherefrom an error is computed. Confidence contours outline the certainty held by the neural network in estimating whether states in c_A-c_B space are achievable or not.

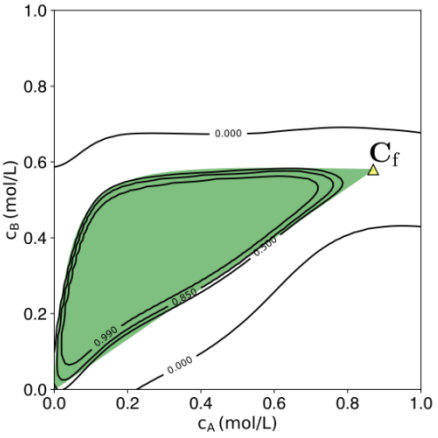

Figure 1: Prediction of 2-D Van de Vusse kinetics

The neural network predicts a large portion of the true region with a greater than 50% confidence. Moreover, whereas the reference region must be generated via a relatively expensive computational procedure (using an automated AR construction algorithm), the neural network prediction is obtained orders of magnitude faster.

The network has never been exposed to the input feed before, yet it is able to output a region of feasibility in agreement with the reference region.

Input representation and network architecture

The core of the method centers on a binary classification problem. By repeatedly querying the trained network to classify unseen test points, a collection of points in state space can be gathered that belong to the AR.

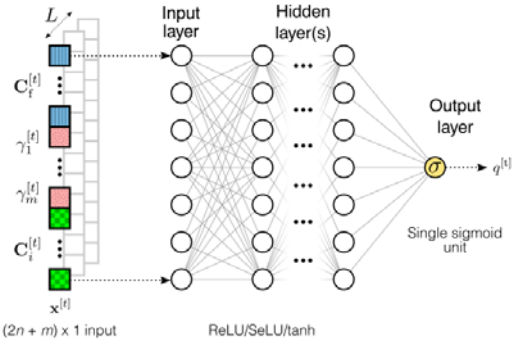

Figure 2: Embedding features

Figure 2 shows the required dimensions of the input and output layers of the neural network. The number of neurons in the input layer must match the number of features in the training problem, which might consist of feed, kinetic and state information. The target layer is a

single neuron that outputs a scalar value between zero and one, mapping the neural network's confidence where values closer to one are considered by the network to be more likely to be attainable.

Training is carried out by feeding L training examples for the input layer and L binary (or boolean) targets for the output layer in conjunction with a suitable training scheme, and appropriate tuning of the algorithm's hyper-parameters.

Since AR boundaries are convex by definition, we can represent a reference region as a convex set of points, which can be converted to a system of inequality constraints that always obeys the above inequality constraints.

Prediction and AR boundary construction

Classification is carried out by constructing a test vector containing feed, kinetic and potentially other significant state information and forward propagating it through a trained network. The output value, q, of the neural network is then a real number, valued between zero and one, representing a confidence estimate of achievability.

Increasing both the quantity of hidden layers and neurons per layer improves classification performance, but the specific strength of their effect will generally be problem dependent. The total number of neurons available in a network is a proxy for the inference capacity of the network, but this capacity could be arranged into different architectures that each distribute capacity differently, resulting in different performance over similar datasets.

To gain a sense of how validation performance is influenced by network capacity for AR boundary constructions, a population of approximately 1000 neural network architectures was created, where each network contained a random permutation of hidden layers and neurons per layer. Each network was then trained over 50 epochs with a subset of 10^4 randomly sampled training examples, recording the final validation accuracy after each run. Preliminary results are shown in Figure 3.

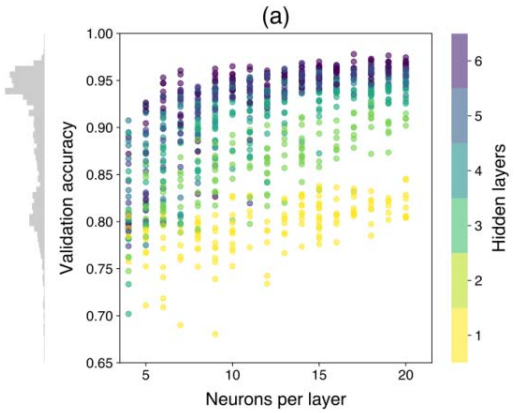

Figure 3: Prediction accuracy and network capacity

We find that validation accuracy improved with an increase in both the quantity of hidden layers and neurons present in each layer. Increasing the number of hidden layers tends to have a stronger influence on validation accuracy than the number of neurons per layer, although even for deep architectures–indicated here by networks with at least four hidden layers–accuracy tends to be limited by approximately seven neurons per layer, which may suggest that overall inference capacity is constrained for these systems.

Conclusion

In this work, we assess the viability of using neural networks as a fast attainability estimator. A number of neural networks were created and trained on three different systems of varying complexity. For the systems currently considered, the neural net appears to learn the connection between changes in feed composition and variations of kinetic parameters. All predicted regions from the neural networks were compared with reference regions that were generated from an automated AR construction method. Given sufficient training data and training time, the neural networks achieved greater than 98% validation accuracy. Overfitting was undetected in the problems considered, and hence the use of regularisation techniques, such as dropout, was optional.

In the examples investigated, the neural networks predicted candidate regions that agreed with the reference regions in terms of shape and size.

Training data requirements and network capacity have also briefly been investigated. Classification efficiency improves with increasing dataset sizes and network capacity, although classification performance tends to increase with an exponential increase in the number of training examples.

References

Feinberg, M., Hildebrandt, D. (1997). Optimal reactor design from a geometric viewpoint- I. Universal properties of the attainable region. *Chemical Engineering Science.*, *52*, 1637–1665.

Van de Vusse, J. (1964). Plug-flow type reactor versus tank reactor. *Chemical Engineering Science.*, *19*, 994–996.