Implementation of Data Mining Techniques and Machine Learning Model in Manufacturing Process

Author: Héctor Vélez Báez
Advisor: Dr. Nelliud Torres Batista
Electrical & Computer Engineering and Computer Science Department

Abstract

Time in manufacturing process is critical. A lot of manufacturing companies rely on their operator's knowledge, experience, or random methods to adjust CNC machines. Sometimes these judgements are ineffective and limited to their own knowledge resulting in an overall decrease in productivity, sales, and profit. Data mining has proven to be an important tool for knowledge acquisition in manufacturing processes. In this paper we will show how data mining techniques and machine learning model could help to effectively adjust CNC machines based on historical data.

Introduction

The use of data mining techniques in manufacturing began in the 1990s and it has gradually progressed by receiving attention from the production community [1]. Data mining is now used in many different areas in manufacturing engineering to extract knowledge for use in predictive maintenance, fault detection, design, production, quality assurance, scheduling, and decision support systems. Data can be analyzed to identify hidden patterns in the parameters that control manufacturing processes or to determine and improve the quality of products. A major advantage of data mining is that the required data for analysis can be collected during the normal operations of the manufacturing process being studied and it is therefore generally not necessary to introduce dedicated processes for data collection. Since the importance of data mining in manufacturing has clearly increased over the last 30 years, it is now appropriate to critically review its history and application [2, 3].

Background

Machine learning is a form of AI that enables a system to learn from data rather than through explicit programming [4]. Machine learning uses a variety of algorithms that iteratively learn from data to improve, describe data, and predict outcomes. As the algorithms ingest training data, it is then possible to produce more precise models based on that data. A machine learning model is the output generated when you train your machine learning algorithm with data. After training, when you provide a model with an input, you will be given an output. For example, a predictive algorithm will create a predictive model. Then, when you provide the predictive model with data, you will receive a prediction based on the data that trained the model.

A neural network consists of three or more layers: an input layer, one or many hidden layers, and an output layer. Data is ingested through the input layer. Then the data is modified in the hidden layer and the output layers based on the weights applied to these nodes. The typical neural network may consist of thousands or even millions of simple processing nodes that are densely interconnected. The term deep learning is used when there are multiple hidden layers within a neural network. Using an iterative approach, a neural network continuously adjusts and makes inferences until a specific stopping point is reached [4].

Problem

In manufacturing, machining is a very complicated process. Adjusting CNC machines for high precision parts could be a very difficult task. These adjustments are made by trial and error or by experiences of the operator and it could take weeks or even months to find out the correct inputs. Companies expends a lot of money in this process and in some cases leads to losses. If this time could be reduced companies could start manufacturing good parts earlier and start generating sales and increasing productivity and profit.

Methodology

A schematic diagram of the proposed procedure is shown in Figure 2. This proposed procedure will be focus on the data mining section but all step for the knowledge discovery process will be applied. The study started with running the CNC machine several times with different input parameters. After that, the parts will be inspected using a Coordinate Measuring Machine (CMM). The results from the CMM will be loaded into a database. Data selection and transformation will be used to get the specific data need it and with the format necessary to input in the neural network algorithm. The neural network algorithm will load the data and will start training. With this training the neural network algorithm will learn about the patterns in the data and then can predict values based on that training set data. Once the training is completed, we will input the target value that we want to reach in the parts and the neural network algorithm will give us the inputs values for the CNC machine. The new values will be introduced in the CNC machine and the process will start again until achieve the nominal dimensions of the part.

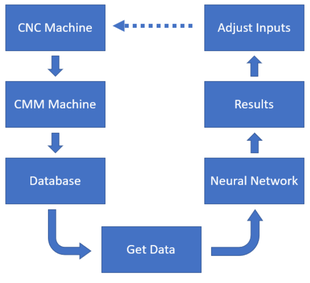

Figure 2
Schematic diagram of the proposed procedure

Results and Discussion

A manufacturing process was chosen from a medical device industry to implement the proposed procedure. Equipment such as CNC machine and Coordinate Measuring Machine are running for a development process of a product. The data collection process is an existing process in the company. They use a program to collect the data from the CMM and then transform that data to load it in their database system.

Neural Network algorithm was created in Python using machine learning libraries such as TensorFlow and Keras. Training set data was obtained from the database and contains 1,000 samples with 10 inputs and 3 outputs. A total of 73 neurons were used for this neural network. Input layer has 10 neurons, 3 hidden layers with 20 neurons each one and output layer with 3 neurons. An example of neural network diagram with 3 hidden layers is shown in Figure 3.

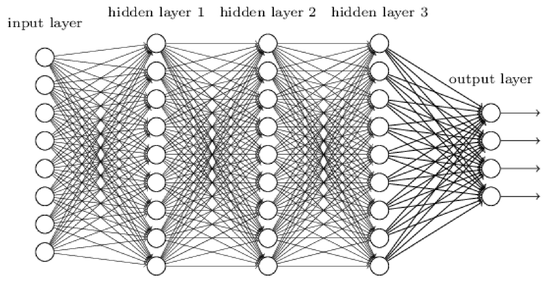

Figure 3
Example of Neural Network Diagram with 3 Hidden Layers

K-fold cross-validation was used to evaluate the neural network model. Cross-validation is a statistical method used to estimate the performance or accuracy of machine learning models. The model evaluation metric used is the Mean Absolute Error (MAE) and Standard Deviation. This metric will predict how off is our model prediction from the real data. Figure 4 shows us that our Mean Absolute value and Standard Deviation is 2.713 and 0.538, respectively.

```
>3.236
>2.428
>3.252
>2.544
>2.040
MAE: 2.713 (0.538)
```

Figure 4
Mean Absolute value and Standard Deviation

Predicted model is trained and cross-validation was completed. Now we can use the neural network algorithm to predict the output values required. Figure 5 and 6 shows the input values selected and the predicted values. Input values represent dimensions of the parts and output values represent input parameters for the CNC machine.

```
[-0.99859353, 2.19284309, -0.42632569, -0.2
1043258, -1.13655612, -0.55671602, -0.63169
045, -0.87625098, -0.99445578, -0.3677487]
```

Figure 5
Input values in Neural Network

```
Predicted: [-147.5597   -84.12978  -92.29834]
```

Figure 6
Predicted values in Neural Network

The predicted values used as input parameters for the CNC machine demonstrated incredibly good results. Adjusting time was reduced to one quarter of the actual time. To calculate the annual saving for a company using this algorithm, we need take in consideration the salary per hours of an operator and how many CNC machines the company has for manufacturing. Assuming that the average salary of an operator is $10/hour, the annual salary based on 2080 hours will be $20,800. If we reduce the time to one quarter, 520 hours, for the same job the company will be paying $5,200. This is a total saving of $15,600 per machine. If we calculate the annual saving for 100 CNC machines, the total saving for the company will be $1,560,000.

Conclusions

A new neural network algorithm was presented for adjusting CNC machine in manufacturing. The algorithm extracts useful knowledge from dataset obtained from the manufacturing process and then identify pattern from this dataset. This will help any operator does not matter the level of experience to adjust CNC machine close to the desire values. The proposed procedure was shown to be capable of giving accurate results and being an asset in the manufacturing process.

Future Work

Preventive Maintenance is very important for manufacturing industry. If we can determine which CNC machine need maintenance before breaking it we can reduce the down time of the machine and increase productivity. Applying Data Mining techniques such as clustering we can determine what machines are in excellent conditions or what machines need maintenance.

Acknowledgements

I would like to thank Dr. Nelliud D. Torres for all the lessons learned about Data Mining and the guidance for complete this paper. As well, I would like to thank Dr. Jeffrey Duffany for his passion about machine learning and the way he teaches the students. One last acknowledgement to my colleague Jaileen Del Valle for her support and always willing to help me in the journey of this master's degree.

References

[1] M. H. Lee, "The knowledge-based factory," Artificial Intelligence in Engineering, vol. 8, no. 2, pp. 109-125, 1993. [Online]. doi: 10.1016/0954-1810(93)90021-7
[2] K. B. Irani, J. Cheng, U. M. Fayyad, and Z. Qian, "Applying machine learning to semiconductor manufacturing," IEEE Expert, vol. 8, no. 1, pp. 41-47, Feb. 1993. [Online]. doi: 10.1109/64.193054
[3] G. Piatetsky-Shapiro, "The data-mining industry coming of age," IEEE Intelligent Systems, vol. 14, no. 6, pp. 32-34, Nov./Dec. 1999. [Online]. doi: 10.1109/5254.809566 \
[4] J. Han, M. Kamber, and J. Pei, J. Data Mining: Concepts and Techniques, 3rd ed. MA, USA: Morgan Kaufmann Publishers, 2012.