Chapter 12

General Model for Metrics Calculation and Behavior Prediction in the Manufacturing Industry:

An Automated Machine Learning Approach

Maria João Lopes
Bosch Termotecnologia SA, Portugal & University of Aveiro, Portugal

Eugénio M. Rocha
University of Aveiro, Portugal

Petia Georgieva Georgieva
https://orcid.org/0000-0002-6424-6590
University of Aveiro, Portugal

Nelson Ferreira
Bosch Termotecnologia SA, Portugal

ABSTRACT

In this chapter, a comprehensive description of a generic framework aimed at solving various predictive data-driven related use cases, occurring in the manufacturing industry, is provided. The framework is rooted in a general mathematical model so called queue directed graph (QDG). With the aid of QDGs and containerized microservices implementations, the typology of the system is analyzed, and real use cases are explained. The goal is for this framework to be able to be used with all use cases which fit in this typology. As an example, a data generation distribution model is proposed, the parameter stability and predictive robustness are studied, and automated machine learning approaches are discussed to predict the throughput time of products in a manufacturing production line just by knowing the processing time in their first stations.

DOI: 10.4018/978-1-7998-6985-6.ch012

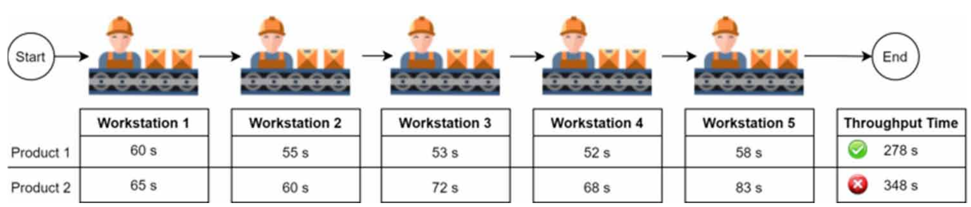
INTRODUCTION

A characteristic aim in the Manufacturing Industry is to optimize the operations in and around production lines (PL). For instance, simulation has been one of the most widely used techniques as shown in Russkikh (2020). Most of these sorts of problems, however, fit into a much broader category and with the right assumptions it is possible to create a model just generic enough to be able to solve a myriad of problems around this abstract entity, being it a PL, a value stream or a material delivery flow.

In order to understand the productive process, it is very common for organizations to resort to KPIs (Key Performance Indicators) and metrics. One of the most well-known is the *throughput time*. The throughput time describes the amount of time that it takes for a product to go through the production line. It can be calculated as the end processing minus the start processing time. Time based metrics are used extensively because they contain within other aspects of the process which would be hidden by other metrics. For example, if the throughput time exceeds some expected value, then it may reveal other underlying factors as the fact that not enough components had being fed to the PL, so the process will be starving at some point; the experience of the workers; a certain product family which is more complex to produce among others.

In Figure 1, a practical example of this calculation is featured. In this image, there is a schematical representation of a PL with a total of 5 workstations. The throughput time is calculated as the sum of all the *processing times* as duration of time the products spent in each workstation. Notice how there might be a trend already at the first 3 stations of the PL where Product 2 takes systematically more time than Product 1. This sort of trend is precisely what a prediction solution would like to catch and to see whether, from a small number of firsts stations of the PL, a significant trend in the throughput time can be inferred. With this information, one could classify the parts prematurely as efficient (such as Product 1) or non-efficient (as Product 2).

Figure 1. Schematic representation of the throughput time calculation in a production line



The focus of this work is to construct a binary classifier where, given a certain effectiveness threshold plus the processing time at the first N stations, i.e., a small amount compared to the total size of the line, the solution is able to predict the outcome of the throughput time class.

In order to have an approach which would fit PLs in a broader sense, and even other types of systems with a similar topology, it is necessary to propose a generic model. The example that was seen above is actually an over-simplification of such manufacturing processes. For instance, it is considered that the throughput time is simply a sum over the different processing times in each workstation. However, there are actually waiting times in between the stations because they are not always available to start processing parts since they may be processing still the previous part. For this reason, the concept of queues will