

Machine Learning Driven Parts Classification

CIMdata Commentary

Key takeaways:

- *Wipro's application of machine learning to parts classification has opened the door for large scale parts classification while reducing subject matter expert effort to reasonable levels.*
- *Wipro's machine learning driven parts classification retains flexibility by leveraging reference standards such as UNSPSC, e-class, and IPC.*
- *Wipro's machine learning driven parts classification can be applied beyond the realm of engineering, for example, into purchasing for supply chain spend analysis and optimization.*
- *Wipro's intelligent parts classification solution can be extended to resolve multi-lingual product development issues such as parts description translation, conversational interfaces (chatbots and voice commands), labels, etc.*

Efficiency of Large-Scale Parts Classification

The growing engineering complexity of products and product systems, coupled with the prevalent global product development and global supply chain strategies, has resulted in undue parts proliferation, which negatively impacts the cost and quality of products. In contrast, effective parts reuse, while delivering on product innovation, is key to competitiveness in the global market place. Though parts reuse may hinder clean-sheet innovation, it helps strike a balance between product innovation and product risk, which is crucial for long-term success.

The effectiveness of parts reuse depends on efficiently capturing and storing all pertinent parts information, coupled with easy and quick search capabilities when and where needed. This, however, has proven to be such a challenge that engineers and product designers often prefer to create new parts instead of searching, thereby wasting time and adding undue cost. Additionally, product performance risk could be elevated due to newly created parts whose behavior is not fully understood. Essentially, the inability to search existing parts quickly leads to increased parts cost, bloated parts inventory, and higher product performance risk, while worsening the search problem further due to the sheer number of parts that need to be searched.

Parts information is needed in several functional organizations within a business. In engineering, the impact is felt in early stages of the product lifecycle where answers to questions about parts availability for reuse, supplier approval status and quality, functional capabilities, cost targets, and compliance are all sought. In sourcing, the questions usually are about alternative suppliers for parts, low cost alternatives, supplier technologies, and supplier prices. In quality and compliance, questions that come up are about consistent poor performance of suppliers, significant changes to manufacturing processes, status of test plans, and exposure to directives such as RoHS and regulations such as REACH, for example.

Lack of availability of consistent parts information across organizational boundaries leads to less than optimal performance at many points along the product lifecycle, and could result in lost productivity, product-launch delays, lost intellectual property, increased warranty costs, slow volume ramp up, missed cost targets, and other negative and costly results.

CIMdata believes that the challenge of reuse is mainly due to inefficient parts libraries, which become worse with every merger or acquisition. CIMdata finds that usually, existing parts are

found in PDM and PLM solutions and are identified with a part number and a small number of part attributes. As a result, designers and engineers need to have prior knowledge of the parts of interest to be able to find them. When parts need to be searched across organizational and geographic boundaries, searches based on a small number of attributes prove inefficient, creating disconnects with far reaching consequences for the business.

To improve the efficiency of classification associated with parts libraries, the following are often mentioned as desirable:¹

- Category management should allow quick changes to the class structure to support an iterative convergence to a robust structure.
- Ability to manage attributes at the category level should be possible so that adding and modifying attributes across multiple categories is possible.
- Tools should be available for automatic cleansing of data at large scales across multiple parts and categories.
- Data quality validation tools should be available for performing initial checks to ensure the whole dataset meets the basic requirements.
- Identification of duplicates and trends in the data should be possible using analytical tools.
- It should be possible to implement workflows so that multiple users can review and approve attributes, items, and categories associated with large volumes of data.

Assuming that such a flexible and robust classification system for parts libraries can be set up, CIMdata is convinced that it would still be a daunting task for a business to undertake parts classification for the first time. The manual classification of parts could stretch over three to four years, involving subject matter experts in the process. The OEMs and suppliers in automotive, aerospace, heavy equipment, and other industries have several million parts in their PDM/PLM solutions, and the task of manually classifying all those parts, while checking for and eliminating duplicates, is challenging. Additionally, the outcome would still be error prone, potentially making the entire effort irrelevant.

CIMdata believes that application of computer based analytics to such large-scale parts classification appears imperative for timely and error-free completion. The algorithms used in machine learning and predictive analytics appear very appropriate for application to parts classification with reference to standards such as UNSPSC, eClass, and IPC.

Wipro's Machine Learning Applied to Large-Scale Parts Classification

Wipro has recently applied the machine learning capabilities available in its Artificial Intelligence platform, HOLMES, to address the challenges of large-scale parts classification. The application involves parts classification based on the reference standard UNSPSC, and can use other standards such as eClass or IPC.

Wipro HOLMES leverages elements of machine learning, natural language processing, genetic and deep learning algorithms, semantic ontologies, pattern recognition, and knowledge modelling technologies provide solutions that are meant to support cognitive enhancement to experience and productivity, accelerate processes through automation, and potentially enable autonomous abilities.

¹ Ten Critical Requirements of a Classification System. Convergence Data. www.convergencedata.com

The following are some of the relevant applications of Wipro HOLMES:

- Digital virtual agents for enhancing user experience with capabilities like speech recognition and natural language understanding.
- Predictive systems for extracting meaning from different forms of data, i.e., using tools and techniques for discovering patterns and predicting future outcomes and trends.
- Cognitive process automation that is defined and executed based on a loose set of instructions or logic. The instructions are largely machine-learned, evolve continuously, and can be user-defined as well.
- Robotic automation powered by a repetitive set of instructions. These instructions are mostly defined by the user and sometimes machine-learned.

In leveraging the machine learning capabilities of Wipro's HOLMES for parts classification, a random sample of the population of all parts is clustered using an unsupervised KMeans based learning algorithm. In each cluster, the subject matter experts manually classify a few parts. This manually classified set is the starting point for the supervised learning model that uses the Multinomial Naive Bayes algorithm for training.

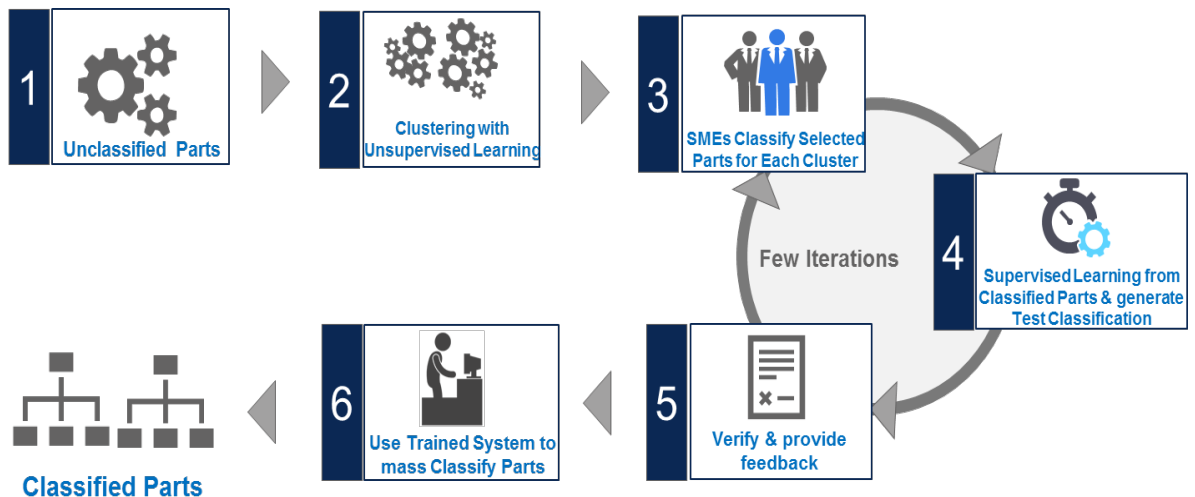


Figure 1—Wipro's Machine Learning Driven Parts Classification

The trained model predicts classifications of the remaining parts. The predictions are validated by the SME, and the errors are identified for correction. The corrected classifications are then used for further fine-tuning of the model. Within a few iterations, Wipro's machine learning based parts classification solution can classify parts to 85% to 90% accuracy. The SMEs then only need to carry out manual classification for 1% to 5% of the parts, which in CIMdata's opinion is a reasonable effort.

Even with multilingual parts descriptions in the mix, Wipro's parts classification solution can classify parts to a high level of accuracy, so long as sufficient training data is provided. When parts attributes are more precise, high accuracy of parts classification can be achieved with a relatively low volume of training data.

In applying its machine learning driven parts classification, Wipro typically has observed a 50% to 60% reduction in time taken to classify the parts, which translates into a proportionate savings in the cost of parts classification.

Additionally, the natural language processing and the recurrent neural network capabilities in Wipro's machine learning offering can help resolve issues related to multiple languages and conversational interfaces, which often create challenges for global product development. Product labelling and packaging art work is another area of application in which the Wipro solution can address translation problems faced by Process/CPG industries.

Summary

Over the years, CIMdata has kept track of several different methods of parts classification that have been proposed. In general, CIMdata has found that parts classification remains a challenging endeavor that progresses slowly while requiring considerable dedicated resources. CIMdata believes that any form of automation that can help the speed and the ease of parts classification will be beneficial to the cost and quality of products.

CIMdata finds that Wipro's machine learning based parts classification solution is a very effective way of resolving the issue of slowness and manpower intensiveness of classifying parts, especially in cases of mergers and acquisitions.

CIMdata has looked at Wipro's machine learning based parts classification solution and the time and effort required of the SMEs for training the model is reasonably small and consequently it is not likely to demand undue time away from their regular occupation for SMEs.

Finally, CIMdata is convinced that Wipro's machine learning driven parts classification solution offers a very effective tool supporting a fast and economic way of classifying parts and expects that Wipro will be helping its customers and prospects very frequently in this effort.

About CIMdata

CIMdata, an independent worldwide firm, provides strategic management consulting to maximize an enterprise's ability to design and deliver innovative products and services through the application of Product Lifecycle Management (PLM). CIMdata provides world-class knowledge, expertise, and best-practice methods on PLM. CIMdata also offers research, subscription services, publications, and education through international conferences. To learn more about CIMdata's services, visit our website at <http://www.CIMdata.com> or contact CIMdata at: 3909 Research Park Drive, Ann Arbor, MI 48108, USA. Tel: +1 734.668.9922. Fax: +1 734.668.1957; or at Oogststraat 20, 6004 CV Weert, The Netherlands. Tel: +31 (0) 495.533.666.