Notes

ML for Manufacturing article

Expert Systems – These were systems in place that required specialized knowledge such having a subject matter expert that required specific training or expertise on topic in which Machine learning was created to assist or work in place of this subject matter expert.

Details from article – Intro summary

Within the semiconductor wafer manufacturing process, tight quality control is of utmost importance. This is due to such factors as the highly competitive nature of the business [2] and the complex nature of the process itself. Therefore, it is imperative that processing prob- lems are recognized and corrected as quickly as possi- ble. To accomplish this, a Parametric Test facility ex- ists where a critical quality check is performed on the wafers. This is accomplished by measuring a number of electrical parameters at existing test sites on the wafer. Each measurement must fall within an acceptable range of values for its associated parameter to pass the crit- ical check. If test results indicate that a parameter is outside of its acceptable range, the wafer may fail the quality check and be scrapped. If this happens an ex- pert must examine all the parametric data associated with that wafer and attempt to determine the reason for failure and where in the manufacturing process the problem may have occurred. The expert's ability to diagnose these failures springs from both a knowledge of semiconductor physics and experience with the wafer manufacturing process. Since this combination of knowledge and experience is valu- able and rare, it was decided that an expert system should be designed and implemented to take the place, or at least reduce the burden, of the human expert in the area of diagnosing wafer failures. As a result, the ADEPT expert system was created through the combined efforts of professors and graduate students in the Computer Science Department at the University of South Florida, in Tampa, Florida and process and de- velopment engineers at Harris Semiconductor, in Pahn Bay, Florida. This expert system was deployed in Au- gust 1989 using a SUN 3/60 workstation.

Automated Learning of Decision Text Categorization" by Chidanand Apte and Fred Damerau

Presents a machine learning-based approach to text categorization. The authors discuss the importance of text categorization in many real-world applications, such as document retrieval and filtering, and argue that automated methods can provide better accuracy and efficiency compared to manual categorization. The article describes a method for automatically learning to classify text documents based on their content. The authors use a training set of pre-labeled documents to create a statistical model that can predict the category of new, unseen documents. They use a technique called "naive Bayes" to calculate the probability of a document belonging to a particular category based on the words it contains. The authors also explore techniques for improving the performance of the classifier, such as feature selection and the use of different statistical models.

The article presents experimental results that demonstrate the effectiveness of their approach on a large dataset of Reuters news articles. The authors compare their method to other existing approaches, such as rule-based and keyword-based methods, and show that their method outperforms these approaches in terms of accuracy.

Overall, the article provides an important contribution to the field of text categorization, demonstrating the potential of machine learning methods for improving the accuracy and efficiency of this important task.

Machine Learning and Data Mining" by Tom M. Mitchell,

Published in the November 1999 issue of Communications of the ACM, discusses the intersection of machine learning and data mining.

The article provides an overview of machine learning, which is the process of training a computer program to learn from data, and data mining, which is the process of discovering patterns and knowledge from large amounts of data. It describes the different types of machine learning algorithms, such as supervised learning, unsupervised learning, and reinforcement learning, and their applications in various fields, including natural language processing, computer vision, and bioinformatics.

The article also discusses the challenges and opportunities in the field of machine learning and data mining, such as the need for large amounts of high-quality data, the importance of feature selection, and the ethical implications of using machine learning algorithms in decision-making.

Overall, the article provides a comprehensive overview of the state of the art in machine learning and data mining in 1999, and remains a useful resource for anyone interested in these fields.

Learning to extract information from text based on user-provided examples.

The article "Learning to extract information from text based on user-provided examples" by Scott B. Huffman describes a system for automatically extracting structured information from text, such as web pages or product catalogs. The system uses machine learning to learn patterns in the text, based on examples provided by the user, and then applies those patterns to new text to extract the desired information.

The article presents the results of an evaluation of the system, which shows that it is effective in extracting structured information from text with high precision and recall. The system is also shown to be robust to changes in the format or wording of the text, which makes it useful in a variety of applications.

The article concludes with a discussion of the potential applications of the system, which include data integration, information extraction, and automated data entry. Overall, the article demonstrates the effectiveness of machine learning in extracting structured information from unstructured text data.