Machine Learning for Process Development for Semiconductor and Nanotechnology Product R&D

M. Mueller,[1] K. Subramanian[1]
[1]Computer Science/OMSCS, Georgia Institute of Technology

Machine Learning algorithms can be utilized to form analytic models for semiconductor or complex nanotechnology manufacturing processes. These serve as a means of mapping the process space for a manufacturing tool (such as metal deposition or oxide etch) within a manufacturing sequence. Methods described may also be extended to a "modules" or sequences of steps, such as deposition-lithography-etch-strip, and even to an entire manufacturing process.

To effectively leverage machine learning for data mining and inference formation, it is important to select variables or "attributes" for input, along with an output or predicted quantity for each set of inputs. This presentation describes how to construct an "X" matrix of inputs, along with how to process a "Y" matrix of outputs, where each row represents a wafer and each column represents an input/output variable.

Predictive modeling is compared to current methods of engineering and experiment execution. Typically, development data belongs to the engineer or group who runs an experiment, and the desired result is a report which provides technical progress or which may guide a business decision. Each experiment is designed around a particular focus, and "old" data lives in engineering notebooks or spreadsheets, while the "wisdom" derived lies in the reports and in professional experience of those involved. In contrast, by creating models based on continual collection of data, it is possible to mine this data with predictive models, such as classifier-based (example: in spec/out of spec) or regression-based machine learning algorithms.

Such models allow a thorough mapping of the process space for addressing challenges and future hardware/process development. Construction and continued extension of a model thus replaces the traditional approach of targeted experiments and design of experiments (DOE). Since the collection and analysis of data is a collaborative effort, this use of machine learning becomes a collaborative tool (such as Google Docs) where multiple team members may contribute.

Machine learning models may be utilized for more applications beyond process development. Such models can be used to measure differences between subsystems, for example, evaluating materials from different vendors or evaluating different power supplies. Machine learning can be leveraged to measure equivalence or non-equivalence between production tools within a high volume factory.