

A novel hybrid under-sampling for semiconductor wafer defect bin classification

You-Jin Park¹, Rong Pan², Douglas C. Montgomery², Kyunghye Joo³

¹ Department of Industrial Engineering and Management, National Taipei University of Technology, Taipei, Taiwan, R.O.C.

² School of Computing and Augmented Intelligence (SCAI), Arizona State University, Tempe, AZ, USA

³ Meta Factory Group, Solution Business Division, Samsung SDS Corporation, Seoul, Republic of Korea

Correspondence

You-Jin Park, National Taipei University of Technology, Taipei, Taiwan, R.O.C.

Email: yjpark@mail.ntut.edu.tw

Funding information

Ministry of Science and Technology (MOST) of Taiwan, R.O.C., Grant/Award Number: MOST 110-2221-E-027-106-MY3

Abstract

Generally, defective chips on semiconductor wafer maps tend to form spatial clusters in unique and distinguishable patterns, and the cluster patterns as well as the defective chips contain considerably important information on specific problems of equipment or process. So, in semiconductor manufacturing, accurate identification and classification of diverse defect patterns is very important in terms of yield enhancement. However, in practice, a serious class imbalance problem in wafer bin data, that is, the number of the defective chips on semiconductor wafer maps is significantly smaller than that of the non-defective chips. Most machine learning approaches designed for classification, however, assume that the number of instances for each observed class is nearly balanced. If the traditional classification algorithms based on this assumption are applied to a class imbalanced dataset, it may lead to incorrect and biased classification results. Thus, in this research, the semiconductor wafer defect bin data combined with wafer warpage information is considered and a novel hybrid under-sampling method to enhance the performances of classification models. From the experimental analysis, it is shown that the proposed method provides better classification performance compared to other data preprocessing methods regardless of classification models.

KEYWORDS

class imbalance, classification, under-sampling, semiconductor wafer bin map, wafer warpage