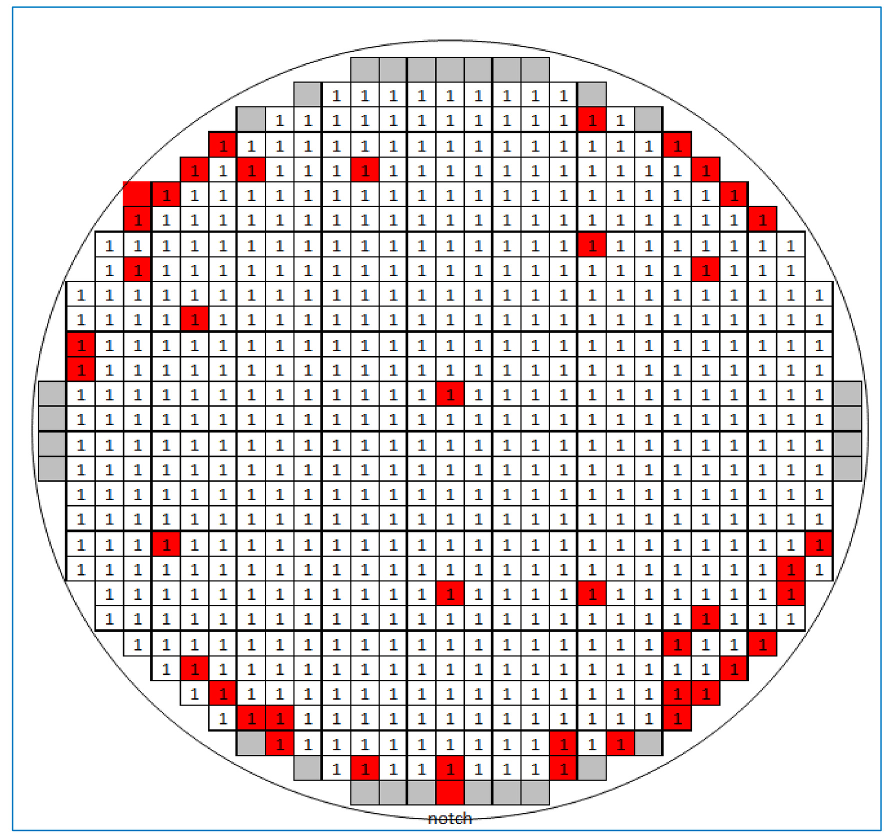
Data analytics and machine learning: root cause searching for a quality issue in semiconductor industry for automotive

Corinne Bergès

NXP Semiconductors

Introduction

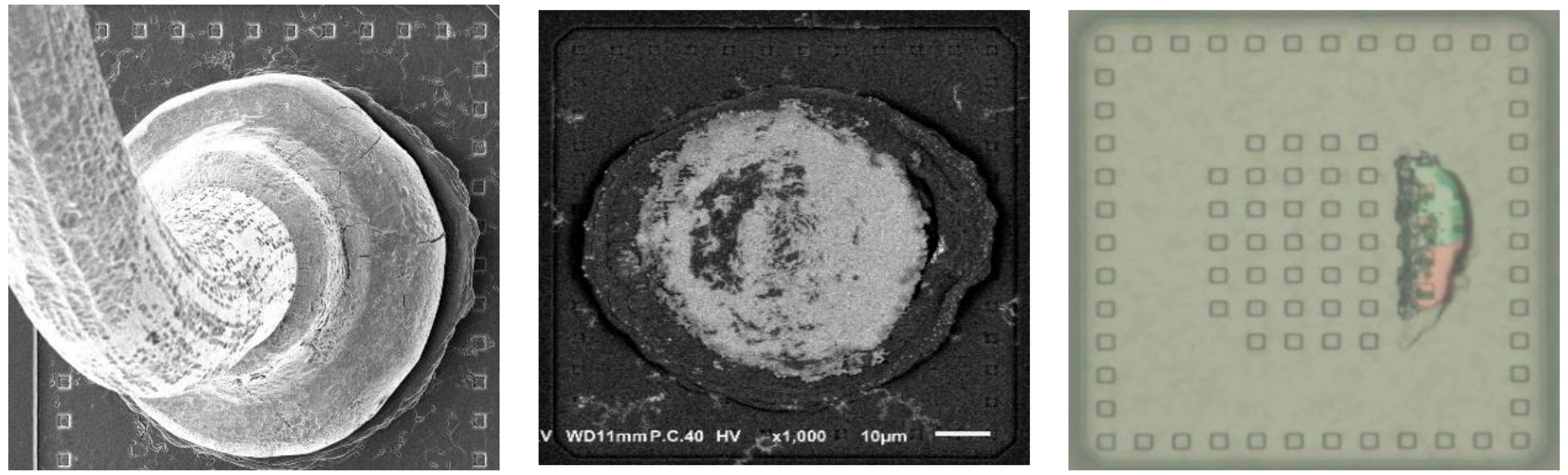
A final customer in automotive industry reported an issue on some parts and the failure analysis highlighted several failing tests in valve management functions. New tests on the products revealed an area in the wafer edge in which the parts were more likely-to-fail. Machine Learning analysis were designed in order to understand the difference between this weak area and the rest of the wafer.



Wafer mapping: the dies at the wafer edge are more likely-to-fail (in red)

Failure Analysis

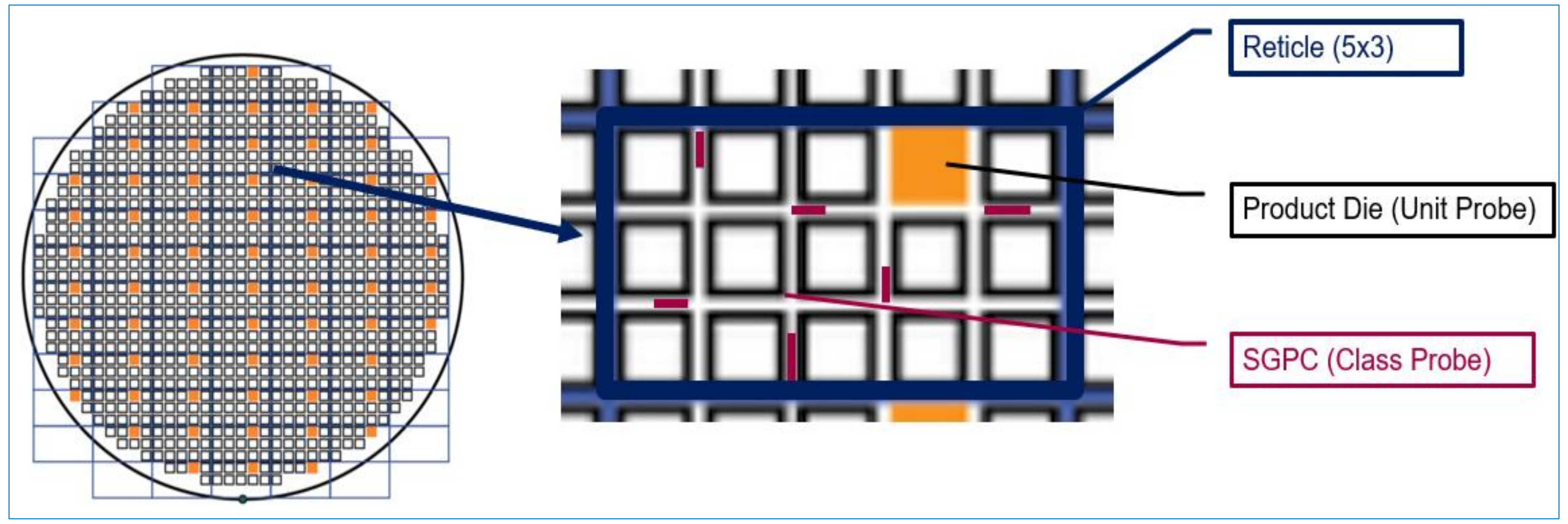
The first step was to perform a failure analysis that revealed pad crack.



Scanning Electron Microscopy pictures of the defect

Which input data for the analysis ?

These analysis used the unit probe and class probe data for 7 wafer lots impacted by the failure (5 wafer lots in which customer returns were reported, and 2 additional wafer lots in which internal tests highlighted defects).
This first data (unit probe) allow to work at die level; the second one (class probe) study the reticles on the wafer. The reticles are specific structures embedded on the wafers between the dies and picture each individual manufacturing step; their test allow to identify the failing step. Class probe tests or unit probe tests are the features for the machine learning analysis implemented.



Reticle vs product die

Data volume

An issue faced was about unit probe data volume: it was not possible to merge all the 7 wafer lots impacted, except sampling it, or focusing the analysis by wafer.
It was not the same case for the reticle data because their number is smaller and the analysis could take all the 7 wafers into account.

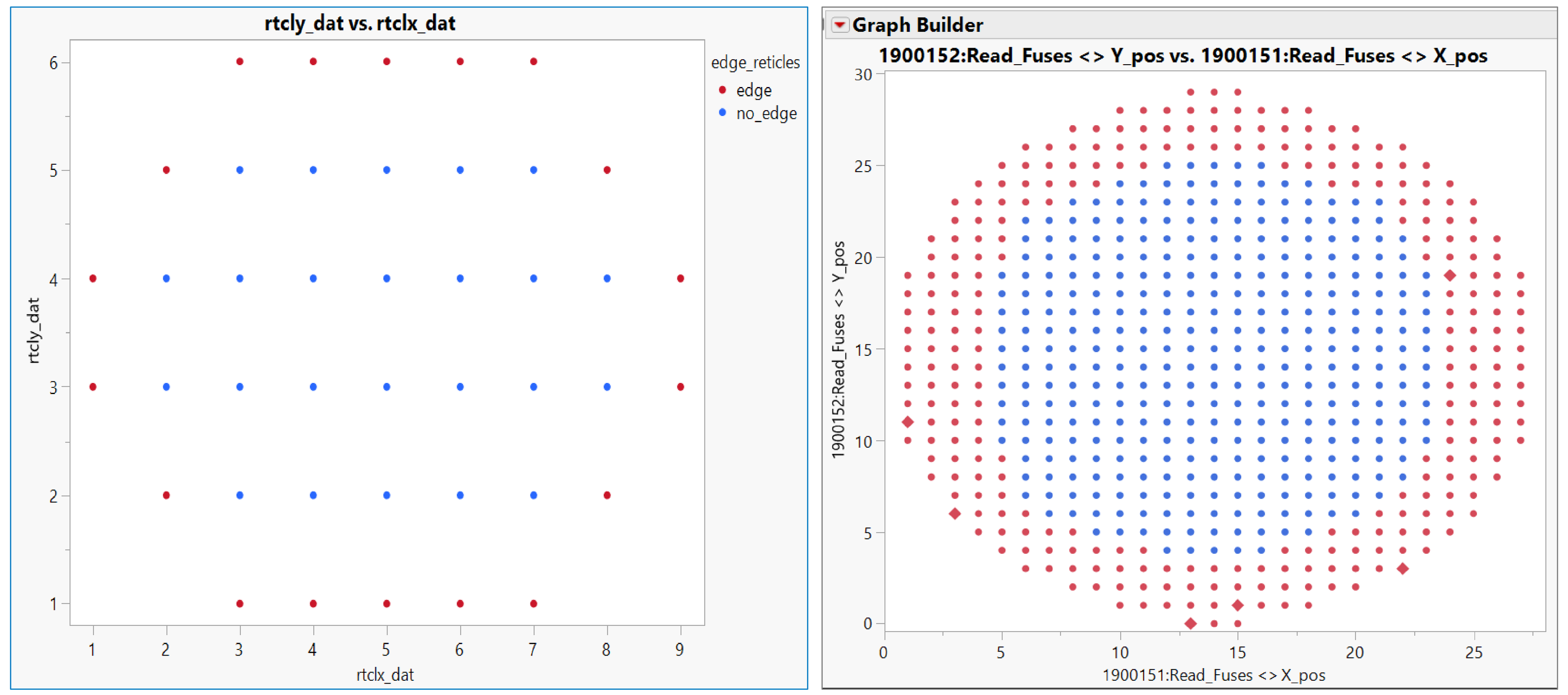
Data analytics and machine learning: root cause searching for a quality issue in semiconductor industry for automotive

Corinne Bergès
NXP Semiconductors

What is modeled ?

There are 2 types of algorithms in machine learning:
-in prediction algorithms or supervised learning, we have elements of different categories or values (training data), and we want to predict categories or values for new elements; we speak about classification when the data to be predicted is a nominal one, or about regression when we model continuous values;
- in clustering or unsupervised learning, we want to discover if there are different categories or if some elements may be distinguished among the majority of them.

In this case, the data that we want to predict is if the part can be detected as in the safe area or in the more likely-to-fail one, based on its class probe or unit probe test values. So, a nominal value is generated:
-0 for the first reticle ring, 1 for the other rings, this data being used for the class probe analysis;
-0 for the 4 first die rings, 1 for the other dies, for the unit probe analysis.



Definition of the more likely-to-fail reticle area (first reticle ring for the class probe analysis, and for first die rings for the unit probe analysis)

Machine learning algorithms available in JMP PRO for classification problems

Some of the algorithms used in these analysis were the following ones:

- Logistic Regression (Fit Model platform)
- Trees (Partition platform)
- Linear classifier: Support Vector Machine (SVM platform)
- Artificial neural network (Neural platform)
- Ensemble methods: Boostrap Forest and Boosted Trees (via Predictive Modeling platform)
- Instance based methods: K-Nearest Neighbords and Naives Bayes (via Predictive Modeling platform, too)

Predictive Modeling platform

Data analytics and machine learning: root cause searching for a quality issue in semiconductor industry for automotive

Corinne Bergès
NXP Semiconductors

Discussion

The best machine learning algorithm doesn't exist …. Some questions to choose it:

−Understand or predict ? (explainability) (ex: regression model vs neural networks)
−Deployability on Big Data frameworks as Hadoop
−Robustness against uncleaned or bad data
−The algorithm has to improve itself when it is trained on more and more data (and no decrease of its performance)
−An algorithm should not need strong expertise to train it and to deploy it
−Algorithm gain vs effort to design it

Conclusion

In this case study, the main result was obtained from the class probe test data: 2 steps were highlighted as the ones generating the difference between the first reticle ring and the other ones: efforts have to be focused on them in order to fix the issue.

Nominal Logistic Fit for edge_reticles

Effect Summary

Source	LogWorth	PValue
168084_param_value_avg	47,695	0,00000
231837_param_value_avg	45,151	0,00000
232296_param_value_avg	10,526	0,00000
299302_param_value_avg	9,719	0,00000
171973_param_value_avg	9,505	0,00000
232282_param_value_avg	8,810	0,00000
299300_param_value_avg	7,723	0,00000
299310_param_value_avg	6,765	0,00000
236258_param_value_avg	5,748	0,00000
231849_param_value_avg	5,511	0,00000
235682_param_value_avg	4,068	0,00009
231844_param_value_avg	2,097	0,00800
168089_param_value_avg	1,667	0,02154

Logistic regression results

References

'CaseStudyJMP_7Jan2021.jmp'

CaseStudyJMP_7Jan2021
- Source
- Fit Model
- Support Vector Machines of edge_reticles
- K Nearest Neighbors of edge_reticles
- Neural of edge_reticles
- Naive Bayes of edge_reticles
- Decision Tree of edge_reticles
- Bootstrap Forest of edge_reticles
- Boosted Tree of edge_reticles

Saved machine learning scripts for this case study