[Prediction of Defect Rate Using Machine Learning in Assembly Process - Yumiko Miyaji]

[Sunsaku Yasuda/ Toshiya Okazaki/ Kurosawa Ryosuke/ Hiroyuki Okawa/ Kinoshita Shinichi]
[Yumiko.Miyaji@sony.com - Shunsaku.Yasuda@sony.com - Toshiya.Okazaki@sony.com- Ryosuke.Kurosawa@sony.com - Hiroyuki.Okawa@sony- Shinichi.Kinoshita@sony.com]
[Sony Semiconductor Manufacturing Corporation]
[Kumamoto TEC 4000-1 Haramizu Kikuyo-machi Kikuchi-Gun, Kumamoto, 869-1102 Japan]
Phone: +[81] -[96 292 6636] Fax: +[81]-[96 292 6818]

I. INTRODUCTION

In semiconductor manufacturing factories, many engineers are actively engaged in productivity improvement, quality maintenance, and yield improvement activities by building statistics and machine learning models using big data accumulated daily. Especially in Wafer process, we have focused on using equipment information to predict product quality. [1]. On the other hand, in assembly processes of image sensors, there are many types of package sizes and structures depending on the application, making it difficult to set up the equipment and causing the defect rate to vary. Furthermore, it is difficult to construct a prediction model because the structure of the data is not standardized, and qualitative and quantitative data are mixed.

Therefore, we report on our efforts to construct the defect rate prediction model using assembly process equipment information and product information for products with different package sizes and structures.

II. DATA-COLLECTION

We selected two assembly processes that have a high impact on product yield. We interviewed engineers in selected processes and chose product information and equipment setting parameters that could affect the occurrence of defects (Fig.1.). Most of the selected equipment setting parameters were structured data already registered in the database, whereas the product information included a lot of unstructured data such as drawings. Therefore, we collected the data manually.

The collected data has four issues. So, we revised them in turn (Fig.2.). First, we unified the character data and the unit of measure that differs from one equipment to another. Next, we selected the optimal statistics that represented the working characteristics of the equipment settings. After visualizing and analyzing the data, we created new explanatory variables from the collected data, then we added them to the data set. If the processing flow differs depending on the materials used, we standardized the corresponding tasks and left the different tasks blank.

III. METHODS

The data set in this study contains quantitative variables, qualitative variables, and blank data. In general, statistical models require a lot of preprocessing for highly accurate prediction, whereas machine learning is difficult to show the basis for decisions, which makes it difficult to improve quality. For these reasons, we used AI predictive analysis tool “Prediction One” (Sony Network Communications Inc., Tokyo, Japan; https://predictionone.sony.biz/). The tool does not require preprocessing of missing value and qualitative variable data, which can reduce prediction accuracy and automatically optimizes the model's layer structure and hyperparameters. It also has the function that ranks the explanatory variables that were considered important in the model construction on a scale called “importance”.

The basic structure of prediction model is an ensemble model of gradient boosting (GBDT) and neural networks (NN) (Fig.3.). Where α is the weight of the model, the predictive model can be expressed by the following equation:

$$Predicted\ value = \alpha \cdot GBDT + (1-\alpha) \cdot NN \qquad (1)$$

We constructed models to predict the defect rate of process Die Bond (DB) and Wire Bond (WB). The DB is a process to bond chips to packages or supports. The WB is the process of connecting electrodes on a chip to electrodes on a package or substrate using metal wires. The objective variable is the defect rate, and the explanatory variables are equipment setting parameters and product information. Table.1 and 2 show details.

IV. RESULTS AND DISCUSSION

Fig.4. shows we predicted of DB defects (coefficient of determination:0.87). The top importance items showed some items related to package size. This is consistent with the engineer's knowledge that package size affects the amount of package warpage and induces defects during chip fixturing. We consider the addition of explanatory variables related to package warpage to further improve accuracy.

In the same way, the predicted results of WB defects (coefficient of determination:0.81). Focusing on the top items of importance, we can see parameters related to the length and number of wires. These parameters were consistent with the engineer's knowledge that the number and arrangement of WB to be joined increased the defect rate.

We also built models for other processes in the same way and attempted to predict yield by combining all models. Fig.5. shows a comparison of the prediction results between proposed method and existing model. The model combined with this method is more accurate than the existing method. (Existing model’s coefficient of determination:0.19/proposed models’ coefficient of determination:0.64).

V. CONCLUSIONS

We appropriately processed the data with the engineer's knowledge when creating the data set, and selected variables when building the model. As a result, we were able to construct a model that allows us to easily examine the results and predict defect rates for products with different package sizes and structures.

In the future, we will use models to determine the optimal package structure and equipment conditions. Furthermore, we will take on the challenge of developing technology and building an environment that will enable us to make consistent yield prediction between wafer process and assembly processes.

REFERENCES

[1] T. Nishimura et al., PM-09, 2020 International Symposium on Semiconductor Manufacturing

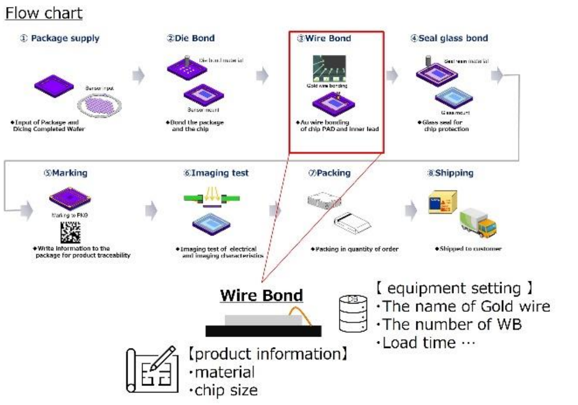

Fig.1. Assembly process flow and example of WB variable selection

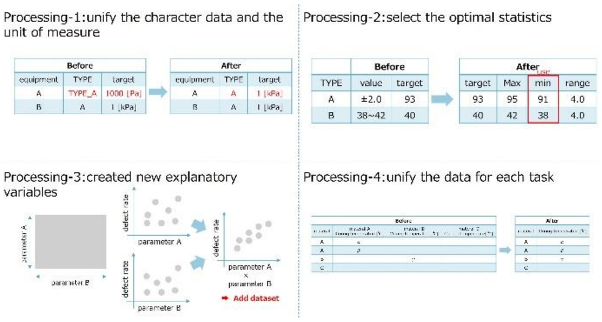

Fig.2. Preprocessing

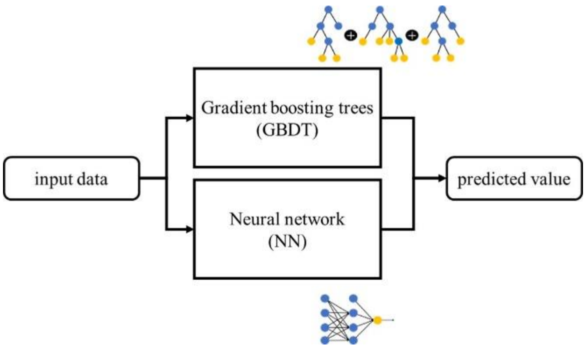

Fig.3. Diagram of gradient boosting trees (GBDT) and neural network (NN) ensemble model

Table.1. Die Bond (DB) process parameter

Model Setting	Selection
Objective variable	defect rate
Explanatory variable	Product/equipment data :43 (Numeric data: 31 / Character data: 12)
Prediction type	Numerical prediction
Number of learning	769
Number of validation	92

Table.2. Wire Bond (WB) process parameter

Model Setting	Selection
Objective variable	defect rate
Explanatory variable	Product/equipment data :43 (Numeric data: 31 / Character data: 12)
Prediction type	Numerical prediction
Number of learning	866
Number of validation	91

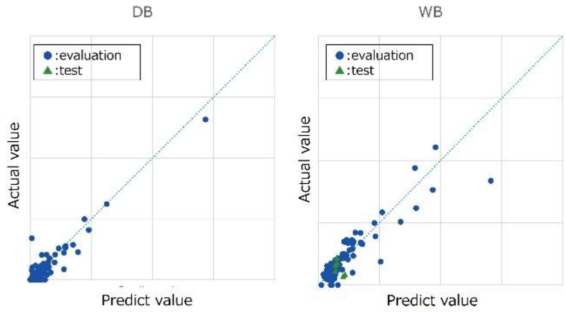

Fig4. Result of Die Bond (DB) and Wire Bond (WB) predictions

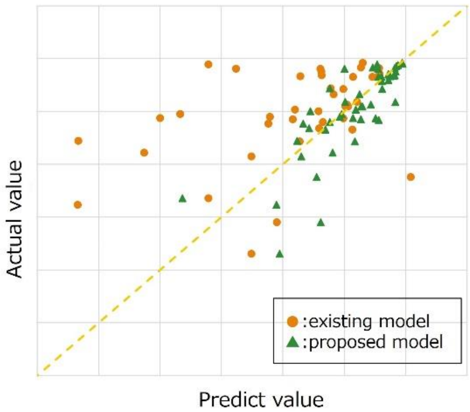

Fig.5. Accuracy comparison of existing and proposed models