

以資料探勘建立工程參數篩選機制於 TFT-LCD 黑色矩陣製程

Data Mining and Feature Selection for TFT-LCD

Black Matrix Process

指導教授：李家岩 教授

專題成員：李季紘

開發工具：R Studio Desktop0.99.489

測試環境：Windows8.1

一、簡介：

For cost reduction and automation, manufacturer uses machine to gradually replace of the manpower. Therefore, if the configuration of machine is not set up well, it will lead to produce a large number of failed products.

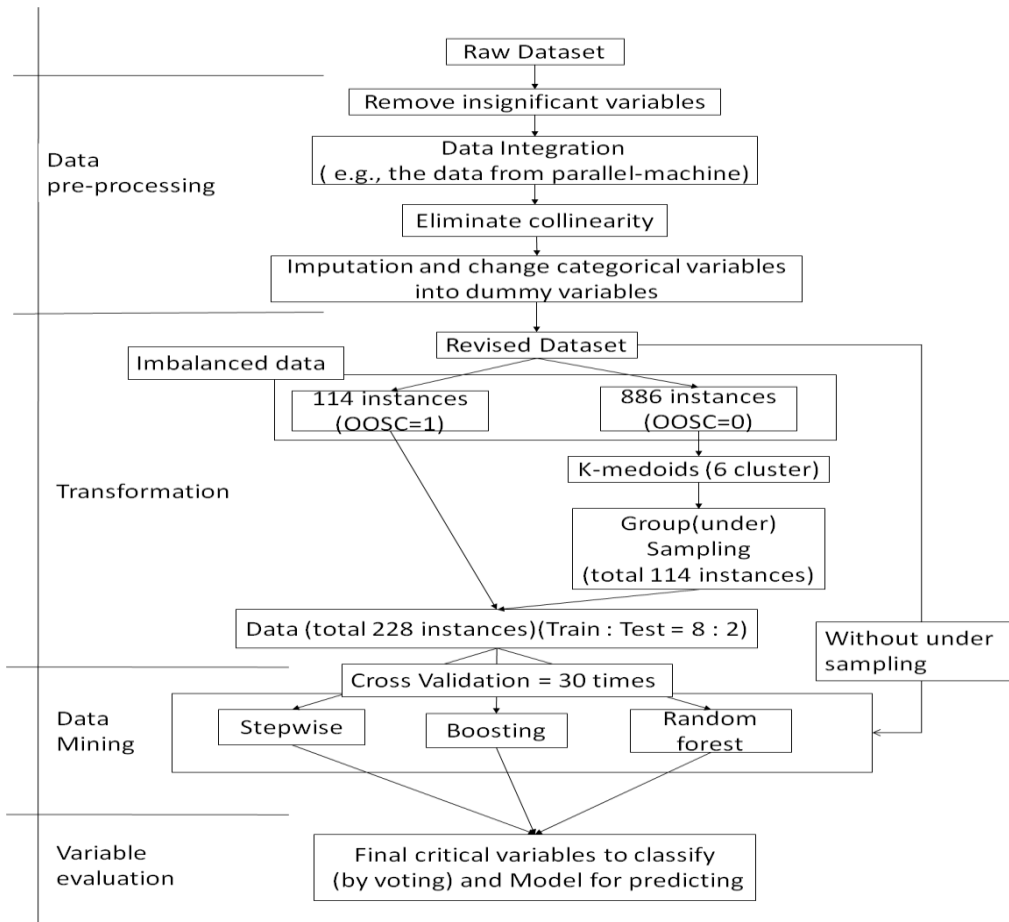
The project goal is to find out what makes the final product out of spec, which may end up unsold. We focus on BM (Black Matrix) process which is one of the CF (Color Filter) process. BM is used for putting antireflective coating on glass substrate.

The dataset we collected is from the black matrix (BM) process and all dataset is transformed via linear or nonlinear simulation due to the confidential issue.

Difficulty and Challenge：

	Encounter problem	Solution
1	Data imbalance	Clustering before under sampling the majority.
2	After handling the imbalanced problem, I come up with the new problem, $p = n$ (The number of sample data is close to the number of variables.)	Boosting Stepwise RandomForest
3	Collinearity exists between the variables.	Variance inflation factor → Continuous Chi-square → Category
4	The definition of the data type should be clarify. (category or continuous)	Sending the email for the detailed definition.

Process flow :



二、測試結果：

變數	類別不平衡			分群抽樣類別平衡			vote
	stepwise	boosting	RF	stepwise	boosting	RF	
X1	0	1	1	4	28	30	64
X2	0	1	1	1	30	27	60
T	0	1	1	0	29	26	57
P1	0	1	0	1	28	26	56
P2	0	1	0	0	28	26	55
L	0	1	1	1	21	30	54
Po	1	1	0	4	29	14	49
R	1	1	0	1	26	16	45
G	0	1	1	0	9	28	39
P3	1	1	0	6	30	0	38
TA	0	1	1	3	6	27	38
PT	0	1	0	2	12	23	38
B	0	1	0	2	30	5	38
C	0	1	0	7	29	0	37
Accuracy	0.85	0.918	0.855	0.862	0.958	0.91	