Pertanika J. Sci. & Technol. 30 (4): 2393 - 2406 (2022)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Robust Linear Discriminant Rule Using Double Trimming Location Estimator with Robust Mahalanobis Squared Distance

Yik Siong Pang*, Nor Aishah Ahad and Sharipah Soaad Syed Yahaya

School of Quantitative Sciences, College of Arts and Sciences, Universiti Utara Malaysia, 06010 UUM, Sintok, Kedah, Malaysia

ABSTRACT

The commonly employed classical linear discriminant rule, based on classical mean and covariance, are highly sensitive to outliers. Therefore, outlier influence on location and scale estimation will affect the accuracy of a discriminant rule and lead to high misclassification rates. The past studies used classical Mahalanobis Squared Distance (MSD) to alleviate the problem. However, the highly sensitive mean and covariance shortcoming can still affect the distance computation, causing masking and swamping effects. In a previous study, researchers proposed a double trimming procedure that adopted MSD-based α-trimmed mean into MSD-based α-trimmed median to construct a robust classifier. However, the proposed procedure has an overlooked flaw because the procedure employed the MSD in the computation. Thus, this study proposed to employ a robust MSD for the distance-based trimmed median procedure. The improvised trimmed median was then used to construct a robust linear discriminant rule and compared with the classical and existing robust rules using a simulation study. The results show that this study's proposed robust linear discriminant rule has better accuracy and consistent performance than the classical linear discriminant rule and two other robust linear discriminant rules.

Keywords: Discriminant analysis, distance-based trimmed median, robust Mahalanobis squared distance

ARTICLE INFO

Article history:
Received: 22 January 2022
Accepted: 20 April 2022
Published: 21 July 2022

DOI: https://doi.org/10.47836/pjst.30.4.05

E-mail addresses:
ys_pang@ahsgs.uum.edu.my (Yik Siong Pang)
aishah@uum.edu.my (Nor Aishah Ahad)
sharipah@uum.edu.my (Sharipah Soaad Syed Yahaya)
* Corresponding author

INTRODUCTION

The classical linear discriminant analysis (CLDA) is a widely employed multivariate classification technique to allocate new observations into one of the dichotomous categories. CLDA is derived by using population mean (μ_1, μ_2) with homoscedasticity assumptions (Σ_1=Σ_2=Σ).

ISSN: 0128-7680
e-ISSN: 2231-8526

Since actual population parameters are usually unattainable, the Classical Linear Discriminant Rule (CLDR) will be constructed based on estimated mean vectors ($\bar{\mathbf{x}}_1$, $\bar{\mathbf{x}}_2$) and scatter matrices ($\mathbf{S}_1$ and $\mathbf{S}_2$). For example, suppose that a set of d-dimensional training data consists of n_1 and n_2 observations corresponding to populations π_1 and π_2, the allocation rule of new unknown observation ($\mathbf{x}_0$) is defined as Equation 1 (Johnson & Wichern, 2013):

If

$$(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^{\mathrm{T}}\mathbf{S}_{pooled}^{-1}\left\{\mathbf{x}_0 - \frac{1}{2}(\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2)\right\} \geq \ln\left(\frac{p_2}{p_1}\right), \tag{1}$$

then $\mathbf{x}_0 \in \pi_1$, otherwise, $x_0 \in \pi_2$.

where the threshold for the decision rule is decided by the prior probability of the original population of an observation, p_1, and p_2.

Notably, the classical mean is extremely sensitive to outliers. Even one outlier distorts the location estimation and influences (co)variance accuracy (Erceg-hurn et al., 2013). Therefore, CLDR's affected mean and covariance will result in a high misclassification rate. Pang et al. (2019) adopted and modified Alloway and Raghavachari's (1990) Mahalanobis Squared Distance (MSD) based α-Trimmed Mean into an MSD-based α-Trimmed Median to construct a robust classifier and obtained practical results compared to the former and classical estimations.

This paper further explores the Distance-based α-Trimmed Median by Pang et al. (2019). The double trimming procedure employed in Pang et al. (2019) used MSD-based trimming as the first layer and Median act as the second trimming has an overlooked flaw, such that the MSD computation itself, which is also based on the outlier sensitive mean vector ($\bar{\mathbf{x}}$) and covariance matrix ($\mathbf{S}$). The notation for MSD for each d-dimensional observation x_i for $i = 1, \dots, n$ is written in Equation 2 (Hadi et al., 2009; Rousseeuw & Van Zomeren, 1990).

$$MSD(x_i) = (x_i - \bar{\mathbf{x}})^{\mathrm{T}}\mathbf{S}^{-1}(x_i - \bar{\mathbf{x}}) \tag{2}$$

Suppose outlier(s) exist in the data. It will result in an inaccurate mean vector and covariance matrix and cause classical MSD to suffer from outliers' secondary effects, *masking,* and *swamping* effects. *Masking* effect or outlier false negative which higher distance measurement outliers masked the supposedly less-far outliers. The *swamping* effect or false positive outlier refers to the shift of the MSD's centroid, causing good observations to appear as outliers. While the high-distance observations were deemed contaminated and trimmed, there are possibilities that outliers may remain in the multivariate data set (Pang et al., 2021). Hence, such a situation may cause any estimation from the trimmed data to be inaccurate. Previous researchers had robustified the distance measurement by using other robust estimators such as M-estimator, Minimum Volume Ellipsoid (MVE), and Minimum Covariance Determinant (MCD) to overcome the masking and swamp problems (Campbell, 1980; Penny & Jolliffe, 2001; Rousseeuw, 1985; Rousseeuw & Van Driessen, 1999; Rousseeuw & Van Zomeren, 1990). Therefore, this study robustified the classical MSD with a robust location estimator, median and robust scale estimator, and Robust Covariance ($\mathbf{S_R}$) to improve the performance of the distance-based procedure in Pang et al. (2019).

The rest of this paper follows: The materials and methods section discusses the formulae, methods, and simulation settings, while the results and discussions section highlights and interprets the simulation results and ends with a conclusion.

MATERIALS AND METHODS

The classical estimators $\bar{\mathbf{x}}$ and $\mathbf{S_{pooled}}$ in CLDR were substituted with robust estimators to construct Robust Linear Discriminant Rules (RLDR). Three robust location estimators were employed in this study, namely (i) α-Trimmed Mean based on classical MSD $(\bar{\mathbf{X}}_{(\mathbf{MSD},\boldsymbol{\alpha})})$, (ii) α-Trimmed Median based on classical MSD $(\widehat{\mathbf{M}}_{(\mathbf{MSD_R},\boldsymbol{\alpha})})$, and (iii) α-Trimmed Median based on robust MSD $(\widehat{\mathbf{M}}_{(\mathbf{MSD},\boldsymbol{\alpha})})$. All three robust location estimators were paired with the same robust scale estimator, Winsorized Robust Covariance ($\mathbf{S_{RW}}$), to reduce unconformity other than the location estimations when comparing the performance of the robust classifiers. The three pairs of robust location and scale estimators were then substituted into Equation 1 to construct RLDR_M, RLDR_T, and RLDR_R. A total of four classifiers (one classical and three robust) were constructed and listed in Table 1. The other formulae and procedures involved in this study are discussed in the subsequent subsections.

Table 1
Classifiers' notation with paired location and scale estimator

	Classifier	Notation	Location Estimator	Scale Estimator
i)	CLDR	CLDR	$\bar{\mathbf{x}}$	$\mathbf{S}$
ii)	$\text{RLDR}_{\text{R}(\alpha)}$	R(α)	$\widehat{\mathbf{M}}_{(\mathbf{MSD_R},\boldsymbol{\alpha})}$	$\mathbf{S_{RW}}$
iii)	$\text{RLDR}_{\text{T}(\alpha)}$	T(α)	$\widehat{\mathbf{M}}_{(\mathbf{MSD},\boldsymbol{\alpha})}$	$\mathbf{S_{RW}}$
iv)	$\text{RLDR}_{\text{M}(\alpha)}$	M(α)	$\bar{\mathbf{X}}_{(\mathbf{MSD},\boldsymbol{\alpha})}$	$\mathbf{S_{RW}}$

Robust Covariance

The Robust Covariance ($\mathbf{S_R}$) is a direct substitution of robust components, Rescaled Median Absolute Deviation (MAD_n) and Spearman Rho (ρ_S), into the components of classical variance-covariance matrix, σ and ρ, as shown in Equations 3 and 4 (Abu-Shawiesh & Abdullah, 2001; Lim et al., 2016; Pang et al. 2019; Yahaya et al., 2016):

$$\mathbf{S_{ij}} = \rho_{ij} \times \sigma_i \times \sigma_j \tag{3}$$

$$\mathbf{S}_{\mathrm{R}_{\mathbf{ij}}} = \rho_{S_{ij}} \times MAD_{n_i} \times MAD_{n_j} \tag{4}$$

Robust Mahalanobis Squared Distance

As mentioned earlier, to alleviate outliers' *masking* and *to swamp* the effect on the distance-based trimming algorithm, Median $(\widehat{\mathbf{M}})$ was chosen as it is a well-known classical location estimator with high outliers tolerance. The classical MSD was robustified as $\mathbf{MSD_R}$ using $(\widehat{\mathbf{M}})$ and $\mathbf{S_R}$ in Equation 5:

$$MSD_R(x_i) = \left(x_i - \widehat{\mathbf{M}}\right)^{\mathrm{T}} \mathbf{S}_{\mathbf{R}}^{-1}\left(x_i - \widehat{\mathbf{M}}\right) \tag{5}$$

Distance-Based Trimming Algorithm

In order to examine the effect of different trimming percentages on robust classifiers' ability to handle outliers, α is set to be 20% and 40%. The selection of these trimming percentages illustrates the difference in using either moderate or high trimming percentages to lower the misclassification rates. The following algorithm was carried out to compute the three robust location and scale estimations:

Step 1: Compute the distance measurement (MSD or $\mathrm{MSD_R}$);
Step 2: Arrange the distance measurement in ascending order;
Step 3: Trim α% of the highest distance for each population (α = 20% or 40%);
Step 4: Compute the mean or median based on the trimmed sample;
Step 5: Winsorize all the trimmed observations (α%) by replacing them with the next α% highest MSD observations in the remaining samples;
Step 6: Compute $\mathbf{S_{RW}}$ based on the winsorized observations.

Simulation Settings

The performance of a classifier depends on classification accuracy, i.e., the lower the misclassification rates, the better the classifier. Therefore, the classical and robust discriminant rules (CLDR, $\mathrm{RLDR_R}$, $\mathrm{RLDR_T}$, and $\mathrm{RLDR_M}$) were examined via simulation study using the Tukey-Huber Contamination Model (THCM), as shown in Equation 6 (Huber, 1964; Tukey, 1962). THCM was employed in various previous robust discriminant studies to generate data for classifier validation (Croux & Dehon, 2001; He & Fung, 2000; Hubert & Van Driessen, 2004; Lim et al., 2016; Pang et al., 2019; Todorov & Pires, 2007; Yahaya et al., 2016).

$$\begin{aligned} &\pi_1: (1-\varepsilon)n_1 N_d(0, I_d) + \varepsilon n_1 N_d(0+\mu, \omega I_d) \\ &\pi_2: (1-\varepsilon)n_2 N_d(1, I_d) + \varepsilon n_2 N_d(1-\mu, \omega I_d) \end{aligned} \tag{6}$$

According to THCM, the underlying distribution for populations π_1 and π_2 are d-dimensional Normal Distributions with locations centered at 0 and 1, respectively, and scaled with d-dimensional Identity Matrices (I_d). The simulation settings used in this study are shown in Table 2.

Table 2
Simulation parameters settings

Contamination Parameters	Controlled Setting
Data Dimensions (d)	2; 6
Contamination Percentage (ε)	0%; 10%; 20%; 40%
Training Sample Size (n_1, n_2)	(20,20); (50,50); (100,100)
Contaminants Location Shift (μ)	0; 3; 5
Contaminants Covariance Amplification (ω)	1; 9; 25; 100

The procedure for the simulation study is as follows:

Step 1: Generate training data according to one of each parameter setting in Table 2;

Step 2: Construct a discriminant rule based on the training data;

Step 3: Generate 2,000 uncontaminated testing samples for each population and record the misclassification rate;

Step 4: Repeat Step 1 to Step 3 for 2,000 iterations;

Step 5: Calculate the mean misclassification rate of the 2,000 iterations.

RESULTS AND DISCUSSION

The average misclassifications of the seven classifiers on the 204 data distributions from 2,000 iterations were recorded and organized according to training samples' size and data dimension in Tables 3, 5, 7, 9, 11, and 13. The performance of the classifiers was compared separately as two categories for different trimming percentages (α), i.e., i) CLDR vs. RLDRs (α=40%) and ii) CLDR vs. RLDRs (α=20%). The lowest misclassification rates were bolded with grey highlights for each category. Additionally, descriptive statistics (Tables 4, 6, 8, 10, 12, and 14) were conducted on the simulation results to illustrate the classifiers' performance better. Similar to the performance comparison, the lowest mean and standard deviation for each category were bolded with grey highlights.

Based on the results in Tables 3, 5, 7, 9, 11, and 13, CLDR undoubtedly performed the best amongst the classifiers when dealing with non-contaminated training data (ε=0, μ=0, ω=1) for all sample size-dimension combinations. Contrarily, when the classifiers deal with contaminated training data, the robust classifiers perform better than the classical classifier as CLDR constantly remains at high misclassification rates. However, throughout the 198 contaminated data distributions, neither robust classifier illustrated an overall dominant performance over one another regardless of trimming percentage.

Firstly, comparing $RLDR_R$ and $RLDR_T$ performance in handling contaminated data (198 distributions), there were 149 and 153 occasions where $RLDR_R$ was better than $RLDR_T$, at a trimming percentage of 40% and 20%, respectively. At the data distributions where $RLDR_R$ outperforms $RLDR_T$, the maximum difference may increase up to 0.2669 (Table 13, n≡100, d=6, ε=40%, μ=5, ω=1, α=40%) and 0.2164 (Table 5, n≡20, d=6, ε=20%, μ=5,

Table 3
Misclassification rate for sample size (20,20) at d=2

ε	μ	ω	**CLDR**	$\mathbf{R}_{(\alpha=0.4)}$	$\mathbf{T}_{(\alpha=0.4)}$	$\mathbf{M}_{(\alpha=0.4)}$	$\mathbf{R}_{(\alpha=0.2)}$	$\mathbf{T}_{(\alpha=0.2)}$	$\mathbf{M}_{(\alpha=0.2)}$
0	0	1	**0.2511**	0.2767	0.2689	0.2641	0.2689	0.2656	0.2614
10	0	9	0.3178	0.2765	0.2747	**0.2704**	0.2671	0.2653	**0.2611**
10	0	25	0.4205	**0.2765**	0.2816	0.2798	0.2666	0.2651	**0.2612**
10	0	100	0.4903	**0.2764**	0.2903	0.2919	0.2665	0.2650	**0.2629**
10	3	1	0.3389	0.2898	0.2936	**0.2880**	0.2723	0.2740	**0.2687**
10	3	9	0.3884	0.2781	0.2798	**0.2756**	0.2677	0.2663	**0.2619**
10	3	25	0.4527	**0.2768**	0.2838	0.2831	0.2669	0.2653	**0.2618**
10	3	100	0.4937	**0.2764**	0.2919	0.2932	0.2666	0.2653	**0.2630**
10	5	1	0.4987	**0.2902**	0.3004	0.2944	0.2698	0.2693	**0.2639**
10	5	9	0.4548	0.2799	0.2827	**0.2783**	0.2680	0.2676	**0.2634**
10	5	25	0.4755	**0.2774**	0.2850	0.2847	0.2669	0.2664	**0.2627**
10	5	100	0.4961	**0.2764**	0.2922	0.2934	0.2667	0.2655	**0.2623**
20	0	9	0.3624	0.2777	0.2761	**0.2735**	0.2628	0.2639	**0.2604**
20	0	25	0.4637	**0.2777**	0.2804	0.2822	0.2615	0.2631	**0.2603**
20	0	100	0.4995	**0.2779**	0.2838	0.2802	**0.2615**	0.2629	0.2638
20	3	1	0.5770	**0.3004**	0.3102	0.3141	**0.2918**	0.3152	0.3454
20	3	9	0.5083	**0.2796**	0.2822	0.2827	0.2627	0.2649	**0.2613**
20	3	25	0.5041	**0.2781**	0.2840	0.2863	0.2617	0.2633	**0.2604**
20	3	100	0.5027	**0.2778**	0.2847	0.2914	**0.2615**	0.2629	0.2638
20	5	1	0.6530	0.2931	0.2955	**0.2929**	**0.2607**	0.3010	0.3871
20	5	9	0.6039	**0.2826**	0.2855	0.2864	**0.2629**	0.2655	0.2630
20	5	25	0.5310	**0.2788**	0.2856	0.2882	0.2617	0.2638	**0.2613**
20	5	100	0.5053	**0.2778**	0.2855	0.2921	**0.2616**	0.2630	0.2640
40	0	9	0.4100	**0.2714**	0.2732	0.2730	**0.2774**	0.2769	0.3465
40	0	25	0.4804	**0.2694**	0.2722	0.2830	0.2775	**0.2774**	0.4523
40	0	100	0.4975	**0.2689**	0.2714	0.3160	0.2781	**0.2777**	0.4971
40	3	1	0.7061	**0.4763**	0.5428	0.6530	**0.5432**	0.5828	0.6756
40	3	9	0.6106	**0.2730**	0.2815	0.3014	**0.2908**	0.2954	0.4685
40	3	25	0.5174	**0.2692**	0.2737	0.2955	**0.2821**	0.2837	0.4872
40	3	100	0.5005	**0.2689**	0.2719	0.3179	**0.2790**	0.2792	0.4995
40	5	1	0.6955	**0.3039**	0.5061	0.6753	**0.4716**	0.5642	0.6941
40	5	9	0.6693	**0.2735**	0.2878	0.3382	**0.3059**	0.3135	0.5694
40	5	25	0.5446	**0.2692**	0.2747	0.3053	**0.2848**	0.2881	0.5099
40	5	100	0.5023	**0.2688**	0.2723	0.3195	**0.2798**	0.2801	0.5011

Table 4
Summary of misclassification rate for sample size (20,20) at d=2

	CLDR	$\mathbf{R}_{(\alpha=0.4)}$	$\mathbf{T}_{(\alpha=0.4)}$	$\mathbf{M}_{(\alpha=0.4)}$	$\mathbf{R}_{(\alpha=0.2)}$	$\mathbf{T}_{(\alpha=0.2)}$	$\mathbf{M}_{(\alpha=0.2)}$
Mean	0.4978	**0.2843**	0.2972	0.3131	**0.2851**	0.2914	0.3514
Std.Dev.	0.0998	**0.0344**	0.0577	0.0891	**0.0572**	0.0719	0.1295
Min	0.2511	0.2688	0.2689	0.2641	0.2607	0.2629	0.2603
Max	0.7061	0.4763	0.5428	0.6753	0.5432	0.5828	0.6941
Range	0.4550	0.2075	0.2739	0.4112	0.2825	0.3199	0.4338

Table 5
Misclassification rate for sample size (20,20) at d=6

ε	μ	ω	CLDR	$R_{(\alpha=0.4)}$	$T_{(\alpha=0.4)}$	$M_{(\alpha=0.4)}$	$R_{(\alpha=0.2)}$	$T_{(\alpha=0.2)}$	$M_{(\alpha=0.2)}$
0	0	1	**0.1409**	0.1948	0.1907	0.1835	0.1748	0.1735	0.1673
10	0	9	0.2108	0.1942	0.1886	**0.1819**	0.1711	0.1711	**0.1645**
10	0	25	0.2543	0.1944	0.1884	**0.1819**	0.1707	0.1712	**0.1645**
10	0	100	0.2725	0.1944	0.1887	**0.1821**	0.1711	0.1714	**0.1648**
10	3	1	0.3915	0.2034	0.2076	**0.1994**	**0.1736**	0.2098	0.2046
10	3	9	0.2679	0.1958	0.1923	**0.1846**	0.1721	0.1715	**0.1651**
10	3	25	0.2655	0.1943	0.1896	**0.1825**	0.1718	0.1715	**0.1651**
10	3	100	0.2733	0.1945	0.1887	**0.1820**	0.1710	0.1713	**0.1648**
10	5	1	0.4998	0.2032	0.2045	**0.1964**	**0.1725**	0.1889	0.1884
10	5	9	0.3253	0.1967	0.1938	**0.1861**	0.1724	0.1718	**0.1656**
10	5	25	0.2783	0.1946	0.1900	**0.1827**	0.1723	0.1714	**0.1650**
10	5	100	0.2742	0.1945	0.1888	**0.1822**	0.1713	0.1713	**0.1649**
20	0	9	0.2514	0.1924	0.1870	**0.1789**	0.1649	0.1678	**0.1613**
20	0	25	0.3613	0.1916	0.1862	**0.1780**	0.1649	0.1678	**0.1613**
20	0	100	0.4694	0.1923	0.1867	**0.1784**	0.1651	0.1676	**0.1610**
20	3	1	0.5365	**0.2139**	0.3438	0.3825	**0.2549**	0.3836	0.4601
20	3	9	0.3933	0.1945	0.1898	**0.1817**	0.1653	0.1666	**0.1606**
20	3	25	0.4204	0.1923	0.1871	**0.1789**	0.1646	0.1676	**0.1611**
20	3	100	0.4780	0.1919	0.1869	**0.1789**	0.1652	0.1677	**0.1610**
20	5	1	0.5668	**0.1994**	0.3348	0.4499	**0.1678**	0.3842	0.5837
20	5	9	0.4956	0.1949	0.1935	**0.1855**	0.1636	0.1652	**0.1596**
20	5	25	0.4625	0.1941	0.1886	**0.1804**	0.1650	0.1670	**0.1607**
20	5	100	0.4846	0.1920	0.1869	**0.1791**	0.1651	0.1676	**0.1610**
40	0	9	0.3240	0.1833	0.1845	**0.1780**	**0.1912**	0.1920	0.2773
40	0	25	0.4563	0.1828	0.1837	**0.1763**	**0.1930**	0.1948	0.4285
40	0	100	0.4991	0.1834	0.1829	**0.1749**	**0.1930**	0.1948	0.4933
40	3	1	0.6433	**0.4665**	0.5349	0.6055	**0.5181**	0.5511	0.6275
40	3	9	0.6382	0.1817	0.1833	**0.1775**	**0.2229**	0.2303	0.4610
40	3	25	0.5355	0.1832	0.1831	**0.1753**	**0.1989**	0.2010	0.4833
40	3	100	0.5035	0.1834	0.1832	**0.1751**	**0.1942**	0.1951	0.4970
40	5	1	0.6137	**0.2644**	0.5186	0.6406	**0.4709**	0.5303	0.6589
40	5	9	0.7232	0.1809	0.1838	**0.1800**	**0.2623**	0.2805	0.5876
40	5	25	0.5805	0.1829	0.1828	**0.1745**	**0.2069**	0.2119	0.5185
40	5	100	0.5070	0.1832	0.1832	**0.1748**	0.1959	**0.1955**	0.4992

Table 6
Summary on misclassification rate for sample size (20,20) at d=6

	CLDR	$R_{(\alpha=0.4)}$	$T_{(\alpha=0.4)}$	$M_{(\alpha=0.4)}$	$R_{(\alpha=0.2)}$	$T_{(\alpha=0.2)}$	$M_{(\alpha=0.2)}$
Mean	0.4235	**0.2023**	0.2173	0.2209	**0.2005**	0.2166	0.2961
Std.Dev.	0.1413	**0.0481**	0.0853	0.1152	**0.0774**	0.0965	0.1751
Min	0.1409	0.1809	0.1828	0.1745	0.1636	0.1652	0.1596
Max	0.7232	0.4665	0.5349	0.6406	0.5181	0.5511	0.6589
Range	0.5823	0.2856	0.3521	0.4661	0.3545	0.3859	0.4993

Table 7
Misclassification rate for sample size (50,50) at d=2

ε	μ	ω	CLDR	$R_{(\alpha=0.4)}$	$T_{(\alpha=0.4)}$	$M_{(\alpha=0.4)}$	$R_{(\alpha=0.2)}$	$T_{(\alpha=0.2)}$	$M_{(\alpha=0.2)}$
0	0	1	**0.2442**	0.2550	0.2515	0.2496	0.2516	0.2499	0.2483
10	0	9	0.2759	0.2548	0.2558	**0.2547**	0.2507	0.2503	**0.2488**
10	0	25	0.3863	**0.2548**	0.2641	0.2664	0.2506	0.2509	**0.2500**
10	0	100	0.4842	**0.2549**	0.2765	0.2833	**0.2505**	0.2511	0.2506
10	3	1	0.2960	0.2628	0.2658	**0.2623**	0.2531	0.2545	**0.2519**
10	3	9	0.3610	**0.2556**	0.2585	0.2576	0.2509	0.2507	**0.2492**
10	3	25	0.4441	**0.2550**	0.2662	0.2696	0.2506	0.2509	**0.2502**
10	3	100	0.4916	**0.2549**	0.2774	0.2848	**0.2505**	0.2513	0.2508
10	5	1	0.4986	**0.2633**	0.2717	0.2677	0.2524	0.2523	**0.2502**
10	5	9	0.4732	**0.2568**	0.2615	0.2610	0.2511	0.2512	**0.2494**
10	5	25	0.4870	**0.2552**	0.2672	0.2716	0.2506	0.2510	**0.2502**
10	5	100	0.4963	**0.2549**	0.2779	0.2856	**0.2505**	0.2513	0.2510
20	0	9	0.3055	**0.2552**	0.2561	0.2559	0.2491	0.2503	**0.2483**
20	0	25	0.4277	**0.2551**	0.2623	0.2651	**0.2487**	0.2500	0.2479
20	0	100	0.4911	**0.2551**	0.2671	0.2731	**0.2486**	0.2499	0.2487
20	3	1	0.6202	**0.2699**	0.2770	0.2769	**0.2634**	0.2814	0.2967
20	3	9	0.5334	**0.2566**	0.2588	0.2591	0.2493	0.2508	**0.2485**
20	3	25	0.5062	**0.2554**	0.2629	0.2662	0.2487	0.2502	**0.2479**
20	3	100	0.4993	**0.2552**	0.2675	0.2740	**0.2486**	0.2500	0.2487
20	5	1	0.6911	0.2669	0.2693	**0.2656**	**0.2484**	0.2705	0.3282
20	5	9	0.6795	**0.2578**	0.2590	0.2608	0.2492	0.2508	**0.2487**
20	5	25	0.5590	**0.2559**	0.2629	0.2662	0.2487	0.2505	**0.2482**
20	5	100	0.5048	**0.2554**	0.2680	0.2746	0.2487	0.2500	**0.2486**
40	0	9	0.3491	0.2528	0.2536	**0.2523**	0.2548	**0.2545**	0.2897
40	0	25	0.4571	**0.2521**	0.2534	0.2527	0.2551	**0.2550**	0.4075
40	0	100	0.4965	**0.2520**	0.2533	0.2640	**0.2549**	0.2551	0.4925
40	3	1	0.7328	**0.4645**	0.5662	0.6920	**0.5686**	0.6159	0.7046
40	3	9	0.6767	**0.2537**	0.2550	0.2580	**0.2603**	0.2638	0.4598
40	3	25	0.5499	**0.2522**	0.2537	0.2546	**0.2568**	0.2579	0.4854
40	3	100	0.5035	**0.2520**	0.2534	0.2658	**0.2552**	0.2557	0.4999
40	5	1	0.7252	**0.2759**	0.5070	0.7049	**0.4599**	0.5954	0.7175
40	5	9	0.7172	**0.2537**	0.2576	0.2822	**0.2688**	0.2773	0.6317
40	5	25	0.5992	**0.2522**	0.2536	0.2582	**0.2581**	0.2599	0.5379
40	5	100	0.5080	**0.2520**	0.2534	0.2677	**0.2556**	0.2563	0.5042

Table 8
Summary on misclassification rate for sample size (50,50) at d=2

	CLDR	$R_{(\alpha=0.4)}$	$T_{(\alpha=0.4)}$	$M_{(\alpha=0.4)}$	$R_{(\alpha=0.2)}$	$T_{(\alpha=0.2)}$	$M_{(\alpha=0.2)}$
Mean	0.5021	**0.2626**	0.2784	0.2913	**0.2680**	0.2755	0.3409
Std.Dev.	0.1266	**0.0355**	0.0654	0.1022	**0.0631**	0.0829	0.1427
Min	0.2442	0.2520	0.2515	0.2496	0.2484	0.2499	0.2479
Max	0.7328	0.4645	0.5662	0.7049	0.5686	0.6159	0.7175
Range	0.4886	0.2125	0.3147	0.4553	0.3202	0.3660	0.4696

Table 9
Misclassification rate for sample size (50,50) at d=6

ε	μ	ω	CLDR	$R_{(\alpha=0.4)}$	$T_{(\alpha=0.4)}$	$M_{(\alpha=0.4)}$	$R_{(\alpha=0.2)}$	$T_{(\alpha=0.2)}$	$M_{(\alpha=0.2)}$
0	0	1	**0.1214**	0.1449	0.1420	0.1385	0.1364	0.1359	0.1329
10	0	9	0.1812	0.1448	0.1423	**0.1385**	0.1355	0.1343	**0.1311**
10	0	25	0.2696	0.1450	0.1430	**0.1393**	0.1354	0.1342	**0.1309**
10	0	100	0.4413	0.1448	0.1434	**0.1397**	0.1355	0.1336	**0.1305**
10	3	1	0.3286	0.1522	0.1523	**0.1475**	**0.1353**	0.1524	0.1487
10	3	9	0.2757	0.1460	0.1450	**0.1408**	0.1355	0.1340	**0.1308**
10	3	25	0.3288	0.1449	0.1435	**0.1398**	0.1356	0.1341	**0.1308**
10	3	100	0.4572	0.1449	0.1433	**0.1397**	0.1354	0.1337	**0.1305**
10	5	1	0.5004	0.1524	0.1499	**0.1451**	**0.1348**	0.1419	0.1402
10	5	9	0.3809	0.1470	0.1463	**0.1421**	0.1354	0.1340	**0.1308**
10	5	25	0.3812	0.1451	0.1440	**0.1401**	0.1355	0.1339	**0.1307**
10	5	100	0.4675	0.1449	0.1433	**0.1398**	0.1354	0.1337	**0.1305**
20	0	9	0.1980	0.1442	0.1421	**0.1388**	0.1316	0.1335	**0.1304**
20	0	25	0.3534	0.1440	0.1454	**0.1426**	0.1316	0.1333	**0.1302**
20	0	100	0.4871	**0.1442**	0.1505	0.1484	0.1316	0.1329	**0.1298**
20	3	1	0.5611	**0.1596**	0.2614	0.2981	**0.2079**	0.3228	0.4302
20	3	9	0.4948	0.1455	0.1437	**0.1398**	0.1312	0.1326	**0.1296**
20	3	25	0.4977	0.1444	0.1449	**0.1419**	0.1315	0.1334	**0.1301**
20	3	100	0.5036	**0.1440**	0.1499	0.1479	0.1317	0.1330	**0.1299**
20	5	1	0.6101	**0.1486**	0.2497	0.3910	**0.1325**	0.3268	0.6372
20	5	9	0.6776	0.1464	0.1444	**0.1403**	0.1308	0.1319	**0.1290**
20	5	25	0.5911	0.1447	0.1444	**0.1411**	0.1314	0.1333	**0.1300**
20	5	100	0.5146	**0.1443**	0.1494	0.1472	0.1317	0.1330	**0.1298**
40	0	9	0.2487	0.1391	0.1400	**0.1353**	0.1476	**0.1465**	0.2029
40	0	25	0.4247	0.1390	0.1397	**0.1351**	0.1483	**0.1475**	0.3767
40	0	100	0.4949	0.1391	0.1398	**0.1352**	0.1487	**0.1479**	0.4909
40	3	1	0.7165	**0.4905**	0.5410	0.6529	**0.5432**	0.5722	0.6814
40	3	9	0.7623	0.1383	0.1382	**0.1334**	**0.1733**	0.1798	0.5195
40	3	25	0.5995	0.1387	0.1395	**0.1345**	**0.1540**	0.1546	0.4967
40	3	100	0.5101	0.1389	0.1398	**0.1351**	0.1495	**0.1488**	0.5015
40	5	1	0.6793	**0.2704**	0.5172	0.6895	**0.4747**	0.5486	0.7132
40	5	9	0.8173	0.1376	0.1377	**0.1329**	**0.2155**	0.2345	0.7144
40	5	25	0.6701	0.1385	0.1388	**0.1337**	**0.1599**	0.1625	0.5800
40	5	100	0.5195	0.1388	0.1398	**0.1350**	0.1503	**0.1501**	0.5089

Table 10
Summary on misclassification rate for sample size (50,50) at d=6

	CLDR	$R_{(\alpha=0.4)}$	$T_{(\alpha=0.4)}$	$M_{(\alpha=0.4)}$	$R_{(\alpha=0.2)}$	$T_{(\alpha=0.2)}$	$M_{(\alpha=0.2)}$
Mean	0.4725	**0.1581**	0.1728	0.1830	**0.1651**	0.1787	0.2830
Std.Dev.	0.1664	**0.0618**	0.0930	0.1318	**0.0885**	0.1062	0.2104
Min	0.1214	0.1376	0.1377	0.1329	0.1308	0.1319	0.1290
Max	0.8173	0.4905	0.5410	0.6895	0.5432	0.5722	0.7144
Range	0.6959	0.3529	0.4033	0.5566	0.4124	0.4403	0.5854

Table 11
Misclassification rate for sample size (100,100) at d=2

ε	μ	ω	CLDR	$R_{(\alpha=0.4)}$	$T_{(\alpha=0.4)}$	$M_{(\alpha=0.4)}$	$R_{(\alpha=0.2)}$	$T_{(\alpha=0.2)}$	$M_{(\alpha=0.2)}$
0	0	1	**0.2420**	0.2482	0.2462	0.2452	0.2461	0.2452	0.2442
10	0	9	0.2587	**0.2482**	0.2492	0.2488	0.2456	0.2454	**0.2446**
10	0	25	0.3447	**0.2481**	0.2563	0.2588	**0.2455**	0.2460	0.2459
10	0	100	0.4800	**0.2481**	0.2689	0.2773	**0.2455**	0.2469	0.2475
10	3	1	0.2741	0.2531	0.2548	**0.2522**	0.2471	0.2478	**0.2464**
10	3	9	0.3270	**0.2484**	0.2511	0.2506	0.2457	0.2458	**0.2450**
10	3	25	0.4234	**0.2482**	0.2569	0.2607	**0.2456**	0.2460	0.2460
10	3	100	0.4929	**0.2481**	0.2696	0.2790	**0.2455**	0.2469	0.2474
10	5	1	0.5010	**0.2533**	0.2594	0.2563	0.2467	0.2467	**0.2454**
10	5	9	0.4804	**0.2490**	0.2523	0.2522	0.2458	0.2460	**0.2451**
10	5	25	0.4917	**0.2483**	0.2567	0.2609	**0.2456**	0.2460	0.2459
10	5	100	0.5012	**0.2482**	0.2698	0.2797	**0.2455**	0.2469	0.2474
20	0	9	0.2745	**0.2479**	0.2490	0.2493	0.2446	0.2453	**0.2442**
20	0	25	0.3929	**0.2479**	0.2549	0.2580	0.2443	0.2452	**0.2439**
20	0	100	0.4896	**0.2480**	0.2617	0.2668	0.2443	0.2451	**0.2439**
20	3	1	0.6542	**0.2573**	0.2628	0.2615	**0.2528**	0.2650	0.2732
20	3	9	0.5678	**0.2488**	0.2498	0.2500	0.2446	0.2454	**0.2442**
20	3	25	0.5237	**0.2481**	0.2537	0.2575	0.2443	0.2453	**0.2440**
20	3	100	0.5042	**0.2481**	0.2617	0.2678	0.2443	0.2452	**0.2439**
20	5	1	0.7124	0.2554	0.2577	**0.2548**	**0.2442**	0.2584	0.2931
20	5	9	0.7158	0.2499	**0.2493**	0.2497	0.2446	0.2454	**0.2442**
20	5	25	0.6061	**0.2483**	0.2521	0.2557	0.2443	0.2454	**0.2439**
20	5	100	0.5144	**0.2481**	0.2613	0.2678	0.2443	0.2453	**0.2440**
40	0	9	0.3063	0.2468	0.2470	**0.2461**	**0.2476**	**0.2476**	0.2645
40	0	25	0.4346	**0.2463**	0.2471	0.2464	**0.2476**	0.2478	0.3655
40	0	100	0.4940	**0.2462**	0.2469	0.2479	**0.2474**	0.2476	0.4788
40	3	1	0.7442	**0.4503**	0.5969	0.7147	**0.5946**	0.6475	0.7236
40	3	9	0.7162	**0.2471**	0.2472	**0.2471**	**0.2503**	0.2523	0.4543
40	3	25	0.5867	**0.2463**	0.2472	**0.2463**	**0.2483**	0.2490	0.4791
40	3	100	0.5076	**0.2462**	0.2470	0.2480	**0.2475**	0.2479	0.4949
40	5	1	0.7389	**0.2605**	0.5148	0.7231	**0.4514**	0.6276	0.7312
40	5	9	0.7372	**0.2470**	0.2481	0.2561	**0.2554**	0.2611	0.6729
40	5	25	0.6453	**0.2464**	0.2472	0.2468	**0.2491**	0.2500	0.5635
40	5	100	0.5159	**0.2462**	0.2469	0.2485	**0.2478**	0.2481	0.5065

Table 12
Summary on misclassification rate for sample size (100,100) at d=2

	CLDR	$R_{(\alpha=0.4)}$	$T_{(\alpha=0.4)}$	$M_{(\alpha=0.4)}$	$R_{(\alpha=0.2)}$	$T_{(\alpha=0.2)}$	$M_{(\alpha=0.2)}$
Mean	0.5059	**0.2549**	0.2718	0.2833	**0.2628**	0.2710	0.3367
Std.Dev.	0.1456	**0.0342**	0.0720	0.1093	**0.0674**	0.0918	0.1501
Min	0.2420	0.2462	0.2462	0.2452	0.2442	0.2451	0.2439
Max	0.7442	0.4503	0.5969	0.7231	0.5946	0.6475	0.7312
Range	0.5022	0.2041	0.3507	0.4779	0.3504	0.4024	0.4873

Table 13
Misclassification rate for sample size (100,100) at d=6

ε	μ	ω	**CLDR**	$\mathbf{R}_{(\alpha=0.4)}$	$\mathbf{T}_{(\alpha=0.4)}$	$\mathbf{M}_{(\alpha=0.4)}$	$\mathbf{R}_{(\alpha=0.2)}$	$\mathbf{T}_{(\alpha=0.2)}$	$\mathbf{M}_{(\alpha=0.2)}$
0	0	1	**0.1157**	0.1279	0.1261	0.1243	0.1235	0.1228	0.1213
10	0	9	0.1505	0.1274	0.1270	**0.1254**	0.1222	0.1220	**0.1204**
10	0	25	0.2252	**0.1275**	0.1294	0.1282	0.1223	0.1221	**0.1207**
10	0	100	0.4310	**0.1275**	0.1382	0.1384	0.1223	0.1228	**0.1217**
10	3	1	0.2740	0.1315	0.1316	**0.1287**	**0.1224**	0.1313	0.1288
10	3	9	0.2414	0.1279	0.1282	**0.1263**	0.1223	0.1223	**0.1206**
10	3	25	0.3142	**0.1274**	0.1295	0.1281	0.1223	0.1223	**0.1208**
10	3	100	0.4562	**0.1274**	0.1380	0.1383	0.1223	0.1229	**0.1218**
10	5	1	0.4991	0.1316	0.1301	**0.1274**	**0.1222**	0.1256	0.1244
10	5	9	0.4000	0.1285	0.1290	**0.1267**	0.1223	0.1221	**0.1204**
10	5	25	0.4072	**0.1275**	0.1294	0.1278	0.1223	0.1223	**0.1208**
10	5	100	0.4736	**0.1274**	0.1377	0.1380	0.1224	0.1229	**0.1217**
20	0	9	0.1587	0.1273	0.1263	**0.1247**	0.1213	0.1225	**0.1207**
20	0	25	0.2921	**0.1273**	0.1298	0.1292	0.1213	0.1223	**0.1204**
20	0	100	0.4684	**0.1272**	0.1357	0.1355	0.1212	0.1219	**0.1202**
20	3	1	0.5866	**0.1367**	0.2081	0.2341	**0.1783**	0.2737	0.3957
20	3	9	0.5381	0.1280	0.1266	**0.1248**	0.1211	0.1221	**0.1202**
20	3	25	0.5044	0.1274	0.1282	**0.1271**	0.1213	0.1223	**0.1204**
20	3	100	0.4960	**0.1274**	0.1351	0.1348	0.1212	0.1220	**0.1202**
20	5	1	0.6526	**0.1298**	0.1968	0.3281	**0.1216**	0.2800	0.6842
20	5	9	0.7669	0.1288	0.1268	**0.1249**	0.1208	0.1217	**0.1198**
20	5	25	0.6490	0.1276	0.1267	**0.1253**	0.1212	0.1224	**0.1203**
20	5	100	0.5147	**0.1274**	0.1343	0.1339	0.1213	0.1220	**0.1202**
40	0	9	0.1893	0.1250	0.1255	**0.1232**	0.1285	**0.1281**	0.1585
40	0	25	0.3682	0.1250	0.1255	**0.1232**	0.1290	**0.1287**	0.3053
40	0	100	0.4853	0.1250	0.1253	**0.1229**	0.1291	**0.1288**	0.4745
40	3	1	0.7677	**0.5069**	0.5553	0.6932	**0.5673**	0.5954	0.7254
40	3	9	0.8194	0.1245	0.1238	**0.1216**	**0.1445**	0.1484	0.5534
40	3	25	0.6495	0.1249	0.1249	**0.1225**	**0.1323**	0.1326	0.5001
40	3	100	0.5128	0.1250	0.1253	**0.1230**	0.1296	**0.1295**	0.4960
40	5	1	0.7300	**0.2669**	0.5338	0.7277	**0.4822**	0.5756	0.7511
40	5	9	0.8526	0.1240	0.1235	**0.1213**	**0.1781**	0.1905	0.7862
40	5	25	0.7379	0.1248	0.1243	**0.1219**	**0.1356**	0.1373	0.6380
40	5	100	0.5306	0.1250	0.1253	**0.1229**	**0.1300**	**0.1300**	0.5119

Table 14
Summary on misclassification rate for sample size (100,100) at d=6

	CLDR	$\mathbf{R}_{(\alpha=0.4)}$	$\mathbf{T}_{(\alpha=0.4)}$	$\mathbf{M}_{(\alpha=0.4)}$	$\mathbf{R}_{(\alpha=0.2)}$	$\mathbf{T}_{(\alpha=0.2)}$	$\mathbf{M}_{(\alpha=0.2)}$
Mean	0.4782	**0.1427**	0.1577	0.1707	**0.1513**	0.1635	0.2802
Std.Dev.	0.1985	**0.0677**	0.0984	0.1403	**0.0949**	0.1118	0.2293
Min	0.1157	0.1240	0.1235	0.1213	0.1208	0.1217	0.1198
Max	0.8526	0.5069	0.5553	0.7277	0.5673	0.5954	0.7862
Range	0.7369	0.3829	0.4318	0.6064	0.4465	0.4737	0.6664

ω=1, α=20%). On the contrary, in the 49 and 45 instances where $RLDR_T$ outperforms $RLDR_R$, only a small difference margin with maximum of 0.0060 (Table 5, n≡20, d=6, ε=10%, μ=0, ω=25, α=40%) and 0.0019 (Table 9, n≡50, d=6, ε=10%, μ=0, ω=100, α=20%) are observed. It implied that employing robust MSD yields better results than the classical MSD during the distance-based trimming procedure.

Moving on to the comparison between $RLDR_R$ and $RLDR_M$, the performance of both robust classifiers in terms of occasions are comparable in both α=40% and α=20%. A total of 88 and 99 per 198 occasions were found where $RLDR_M$ outperforms $RLDR_R$ for trimming percentages of 40% and 20%, respectively. When 40% trimming was applied, the maximum difference in misclassification rates when $RLDR_M$ achieved lower misclassification was 0.0139 (Table 5, n≡20, d=6, ε=20%, μ=0, ω=100), while $RLDR_R$ outperforms with a maximum difference of 0.4626 (Table 11, n≡100, d=2, ε=40%, μ=5, ω=1). On the other hand, in 20% trimming category, the maximum differences in misclassification rates are 0.0073 ($RLDR_M < RLDR_R$; Table 5, n≡20, d=6, ε=10%, μ=5, ω=25) and 0.6081 ($RLDR_M > RLDR_R$; Table 13, n≡100, d=6, ε=40%, μ=5, ω=9). The comparison between these two robust classifiers shows that despite the high number of occasions where $RLDR_M$ is better, the misclassification rates yielded are not substantially lower than $RLDR_R$, whereas $RLDR_R$ yielded noticeably better performance.

In addition to evaluating the classifiers' performance, the descriptive statistics in Tables 4, 6, 8, 10, 12, and 14 can help the classifiers' assessment. Even though CLDR was the best classifier in non-contaminated data distribution, poor overall descriptive statistics are expected. CLDR had the highest mean of misclassification rates. Notably, regardless of trimming percentages, $RLDR_R$ achieved the lowest mean, standard deviation, and range compared to the other robust classifiers. The comparisons of descriptive statistics between the three classifiers in both trimming categories showed that $RLDR_R$ yielded averagely lower and relatively consistent misclassifications.

In summary, applying Robust MSD in distance-based trimming procedure to the method in Pang et al. (2019) showed improvement with a much lower misclassification rate and stable performance presented in $RLDR_R$.

CONCLUSION

The established THCM considered contamination from both location and covariance aspects. Based on the simulation results in this study, it is noticeable that the increment of location and/or covariance contaminations led to an increment in misclassification rates in the classical rule. Despite CLDR showing optimal results when dealing with non-contaminated data, the CLDR should be applied with caution as non-contaminated data are hardly attainable in real-life applications. On the other hand, the simulation study has shown promising outcomes from $RLDR_R$ as it outperforms CLDR, $RLDR_{T,\ \text{and}}$ $RLDR_M$ when dealing with outliers in 2- and 6-dimensional contaminated distributions where n=20, 50, and 100. One should note that the performance of a classifier should not be determined by the number of occasions it bests the others but should also consider the overall performance. Based on the performance of $RLDR_R$ in the simulation study, researchers could consider employing $RLDR_R$ if data contamination is suspected.

The main goal of this study is to determine whether employing robust MSD in distance-based trimmed median $(\widehat{\mathbf{M}}_{(\mathbf{d},\boldsymbol{\alpha})})$ can yield a lower misclassification rate in constructing a robust classifier. It is shown that the robust classifier employed $\widehat{\mathbf{M}}_{(\mathbf{MSD}_\mathbf{R},\boldsymbol{\alpha})}$ noticeably outperformed $\widehat{\mathbf{M}}_{(\mathbf{MSD},\boldsymbol{\alpha})}$ with a lower and consistent misclassification rate. However, the study was limited to comparing classical MSD (with $\bar{\mathbf{x}}$ and $\mathbf{S}$) and robust MSD (with $\widehat{\mathbf{M}}$ and $\mathbf{S}_\mathbf{R}$) in $\widehat{\mathbf{M}}_{(\mathbf{d},\boldsymbol{\alpha})}$. Therefore, future studies can explore ways to robustify MSD in the distance-based trimmed median.

ACKNOWLEDGEMENT

This research was supported by the Ministry of Higher Education (MoHE) of Malaysia through Fundamental Research Grant Scheme (FRGS/1/2019/STG06/UUM/02/8). The authors also want to express gratitude toward Universiti Utara Malaysia for the support and facilities provided to conduct this study.

REFERENCES

Abu-Shawiesh, M. O., & Abdullah, M. B. (2001). A new robust bivariate control chart for location. *Communications in Statistics-Simulation and Computation*, *30*(3), 513-529. https://doi.org/10.1081/SAC-100105076

Alloway, J. J., & Raghavachari, M. (1990). Multivariate control charts based on trimmed means. In *ASQC Quality Congress Transactions* (Vol. 44, pp. 449-453). American Society for Quality Control.

Campbell, N. A. (1980). Robust procedures in multivariate analysis I: Robust covariance estimation. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *29*(3), 231-237. https://doi.org/10.2307/2346896

Croux, C., & Dehon, C. (2001). Robust linear discriminant analysis using S-estimators. *Canadian Journal of Statistics*, *29*(3), 473-493. https://doi.org/10.2307/3316042

Erceg-Hurn, D. M., Wilcox, R. R., & Keselman, H. J. (2013). Robust statistical estimation. In T. D. Little (Ed), *The Oxford Handbook of Quantitative Methods: Foundations* (Vol. 1, pp. 388-406). Oxford University Press.

Hadi, A. S., Imon, A. H. M., & Werner, M. (2009). Detection of outliers. *Wiley Interdisciplinary Reviews: Computational Statistics*, *1*(1), 57-70. https://doi.org/10.1002/wics.6

He, X., & Fung, W. K. (2000). High breakdown estimation for multiple populations with applications to discriminant analysis. *Journal of Multivariate Analysis*, *72*(2), 151-162. https://doi.org/10.1006/jmva.1999.1857

Huber, P. J. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics, 35*(1), 73-101.

Hubert, M., & Van Driessen, K. (2004). Fast and robust discriminant analysis. *Computational Statistics & Data Analysis*, *45*(2), 301-320. https://doi.org/10.1016/S0167-9473(02)00299-2

Johnson, R. A., & Wichern, D. W. (2013). *Applied Multivariate Statistical Analysis* (6th Ed.). Pearson.

Lim, Y. F., Yahaya, S. S. S., & Ali, H. (2016). Winsorization on linear discriminant analysis. In *AIP Conference Proceedings* (Vol. 1782, No. 1, p. 050010). AIP Publishing LLC. https://doi.org/10.1063/1.4966100

Pang, Y. S., Ahad, N. A., & Yahaya, S. S. S. (2021). Robust multivariate location estimation in the existence of casewise and cellwise outliers. *Mathematics and Statistics*, *9*(5), 653-663. https://doi.org/10.13189/ms.2021.090505

Pang, Y. S., Ahad, N. A., Yahaya, S. S. S., & Lim, Y. F. (2019). Robust linear discriminant rule using novel distance-based trimming procedure. *Journal of Advanced Research in Dynamical and Control Systems, 11*(05-Special Issue), 969-978.

Penny, K. I., & Jolliffe, I. T. (2001). A comparison of multivariate outlier detection methods for clinical laboratory safety data. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *50*(3), 295-307. https://doi.org/10.1111/1467-9884.00279

Rousseeuw, P. J. (1985). Multivariate estimation with high breakdown point. *Mathematical Statistics and Applications*, *8*, 283-297.

Rousseeuw, P. J., & Van Driessen, K. (1999). A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, *41*(3), 212-223. https://doi.org/10.1080/00401706.1999.10485670

Rousseeuw, P. J., & Van Zomeren, B. C. (1990). Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association*, *85*(411), 633-639. https://doi.org/10.1080/01621459.1990.10474920

Todorov, V., & Pires, A. M. (2007). Comparative performance of several robust linear discriminant analysis methods. *REVSTAT Statistical Journal*, *5*, 63-83. https://doi.org/10.2307/3316042

Tukey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics, 33*(1), 1-67.

Yahaya, S. S. S., Lim, Y. F., Ali, H., & Omar, Z. (2016). Robust linear discriminant analysis with automatic trimmed mean. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *8*(10), 1-3.