CHAPTER 3

METHODOLOGY

3.1 Experimental Design

Profiling on lard was conducted in two conditions: first, profiling of the samples obtained from different regions and body parts, and second, profiling of the lard upon heating-process and comparing them among the selected fats. First, the sample collection was carried out from three regions, each with three different body parts of pig fats. A total of 30 samples of each body part, making a total of 90 different samples per region, were collected, thus resulting in a total of 270 samples from three regions. The second design is profiling lard and four types of selected fats (three different animals & one plant) after the heating-process. Heating-process protocols were conducted at three elevated temperatures; at each temperature, four different time durations of the heating-process were used and thus resulting in a total of 60 treatments.

Lard samples collected from three regions were measured by using FTIR. After heat-processing, samples were measured using four analytical tools; FTIR, ^{1}H & ^{13}C-NMR, GC-FID, and LC-MS/MS. The resulting analytical data were then subjected to PCA, PCR, PLS, and SVM for in-depth investigation and model development for lard profiling after heating-process to classify and predict lard profiles at different conditions by focusing on multivariate calibration and validation analysis in multivariate classification and multivariate regression for pattern recognition and prediction.

Some equations were presented to describe the procedure of chemometrics because the multivariate has high data throughput. In addition,

the data could be simplified using theorems to explain the chemometrics method. Figure 3.1 illustrates the framework of the methodologies in the thesis.

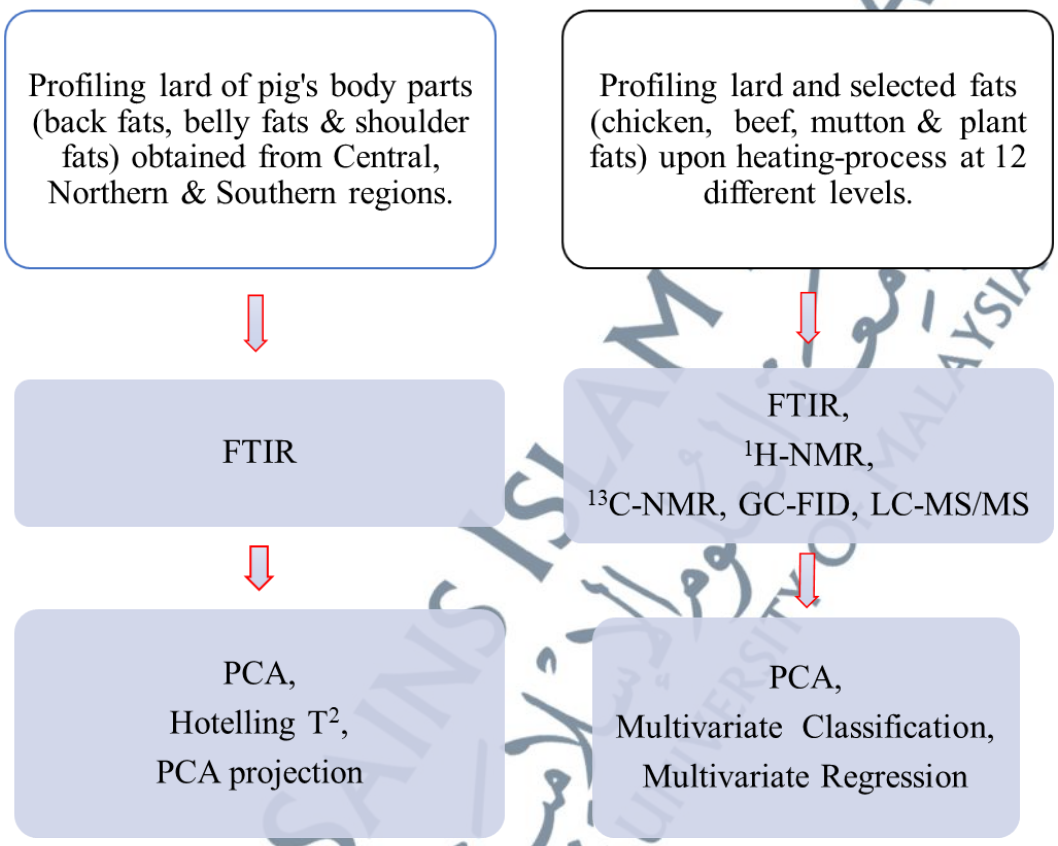

Figure 3.1: Framework of Lard Profiling.

3.2 Profiling of Lard from Collected Pig Samples at Northern, Central, and Southern Malaysia

3.2.1 Chemicals and Materials

3.2.1.1 Chemicals

Table 3.1 shows the list of chemicals used for profiling lard samples collected from northern, central, and southern regions of Peninsular Malaysia. Distilled water used in the experiment was produced in the International Fatwa and *Halal* Centre (iFFAH) laboratory, USIM.

Table 3.1: Chemicals.

	Chemicals	**(Manufacturer)**
1.	Chloroform ($CHCl_3$) GC grade, 99.9 %	(Sigma Aldrich)
2.	Methanol (MeOH) GC grade, 99.9 %	(Sigma Aldrich)
3.	Sodium chloride (NaCl) GC grade, 99.9 %	(Sigma Aldrich)
4.	Anhydrous sodium sulphate ($NaSO_4$) GC grade, 99.9 %	(Sigma Aldrich)

3.2.1.2 Lard Samples

About 100 g of fats were purchased from farms and wet markets in Penang, Perak, Selangor, Negeri Sembilan, Melaka, and Johor to represent different regions of Peninsular Malaysia as follows; northern region (Penang and Perak.), southern region (Melaka and Johor) and central region (Selangor and Negeri Sembilan). From each location, the samples selected consisted of belly fats (BL), back fats (BK), and shoulder fats (SF) to represent fats of different parts of the pig (Franco et al., 2014). The numbers of total lard samples are presented in Table 3.2.

Table 3.2: Total Pigs Fat of the Body Parts and Regions.

	Regions	States	BL	BK	SF	Total
A	Northern	Penang	10	10	10	30
		Perak	20	20	20	60
B	Southern	Melaka	18	18	18	54
		Johor	12	12	12	36
C	Central	Selangor	10	10	10	30
		Negeri Sembilan	20	20	20	60
				Total		270

The crude fats were neatly wrapped in a polystyrene container containing ice to maintain freshness along transportation. During the transportation, the ice cubes were replaced with new ice every 3 hrs. Then all the collected fats were stored in the laboratory freezer at -4 °C before use.

3.2.2 Instruments

Fourier Transform Infrared Spectrometer (FTIR), Model 640-IR, Varian Inc. was equipped with ATR crystal and deuterated triglycine sulphate (DTGS) detector. The IR wavenumber ranges from 650 cm^{-1} to 4000 cm^{-1}.

3.2.3 Samples Preparation

3.2.3.1 Preparation of Lard Samples

About 100 g of fats were cleaned from blood, mucus, pieces of meat and other tissues. Then the fats were finely chopped and dried before being weighed at 10 g ± 0.05 in a 100 mL beaker. The remaining samples were kept subsequent for future use.

3.2.3.2 Extraction of Lard Samples

Fats were extracted based on gravimetric methods using non-polar solvents (Folch et al., 1957). A total of 10 g of fat tissues were homogenized with 100 mL of chloroform and 50 mL of methanol. After dispersion, the whole mixture was agitated for 15-20 min in a shaker at room temperature (27 °C). Then, the homogenate was filtrated using a funnel with folded filter paper (Whatman™ Grade 1) to recover the liquid phase.

The liquid phases that contained $CHCl_3$ / MeOH were washed with two volumes of distilled water containing 0.9 % (w/v) of NaCl (0.9 g in 100 mL distilled water) and left standing for 30 minutes until the mixture was fully separated into two layers where the lower layer of $CHCl_3$ phase containing lipids was that collected in a flask.

An anhydrous $NaSO_4$ was added to the flask to absorb any moisture using a small spatula. Then, the $CHCl_3$ portion was filtrated again and carefully dried under vacuum in a rotary evaporator (Rotavapor® R-300) at 40 ºC for 20 minutes. Following this protocol, about 3-4g of final fats could be extracted successfully. They were kept in the freezer before FTIR analysis.

3.2.3.3 Samples Preparation for FTIR Analysis

The extracted fats were removed from the freezer and thawed for 10 minutes in a water bath (Memmert GmbH™) at 60 °C before being placed in a centrifuge tube and cooled to room temperature (25 ° C) for FTIR analysis.

The obtained fats were placed carefully on a multibounce plate at a controlled ambient temperature (25 °C) using a Pasteur pipette to ensure direct contact with the attenuated total reflectance (ATR) crystal. FTIR

spectra were recorded upon 16 scans at a resolution of 8 cm^{-1} from 4000–650 cm^{-1}. Upon each analysis, the ATR plate was carefully cleaned with MeOH twice and dried with soft tissue before filling in with the subsequence sample. A new reference air background scan was performed before sample analysis.

All FTIR spectra were also recorded as transmittance values at each data point. Samples heated at different temperatures and heating times, samples were measured using FTIR at room temperature. As part of data quality assurance, the sample cell was washed and dried every time after use to avoid contamination from previous samples (Mueller et al., 2013).

3.3 Profiling on Pork (Lard), Chicken, Beef, Mutton, and Plant Fats after Heating-Process

3.3.1 Chemicals and Materials

3.3.1.1 Chemicals

Table 3.3 shows the list of chemicals used for profiling pork (lard), chicken, beef, mutton, and plant fats subjected to the heating-process.

Table 3.3: Chemicals.

	Chemicals	**(Manufacturer)**
1.	Chloroform ($CHCl_3$) GC grade, 99.9 %	(Sigma Aldrich)
2.	Methanol (MeOH) GC grade, 99.9 %	(Sigma Aldrich)
3.	Isopropyl Alcohol (IPA) LC grade, 99.9 %	(Sigma Aldrich)
4.	Deuterated Chloroform (*d*-$CHCl_3$) 99.9 % of	(Sigma Aldrich)
5.	Deuterated Tetramethylsilane (*d*-TMS) 99.9 %	(Sigma Aldrich)
6.	Acetone GC grade, 99.9 %	(Sigma Aldrich)
7.	n-Hexane GC grade, 99.9 %	(Sigma Aldrich)
8.	Toluene GC grade, 99.8 %	(Sigma Aldrich)
9.	Potassium bicarbonate GC grade, 99.8 %	(Sigma Aldrich)
10.	Methanolic Hydrogen Chloride 5 % solution GC grade, 99.8 %	(Sigma Aldrich)
11.	Sodium Chloride Sulphate Anhydrous GC grade, 99.8 %	(Sigma Aldrich)
12.	Certified Reference Standards (CRM) of FAMEs Mix GC grade, 99.9 %	(Sigma Aldrich)
13.	Acetonitrile (ACN) LC/MS-grade 99.9 %	(Thermo Scientific)
14.	Formic Acid LC/MS-grade 99.8 %	(Thermo Scientific)
15.	Ammonium Formate LC/MS- grade 99.8 %	(Fisher Chemical)

3.3.1.2 FAMEs Standard

The Cerified Reference Material (CRM) of FAMEs Mix were purchased from Supelco®, Sigma–Aldrich, USA for GC-FID external standard that contained 37 FAMEs; Butyric (C4:0), Caproic (C6:0), Caprylic (C8:0), Capric (C10:0), Undecanoic (C11:0), Lauric (C12:0), Tridecanoic (C13:0), Myristic (C14:0), Myristoleic (C14:1), Pentadecanoic (C15:0, *cis*-10-Pentadecenoic (C15:1), Palmitic (C16:0), Palmitoleic (C16:1),

Heptadecanoic (C17:0), *cis*-10-Heptadecenoic (C17:1), Stearic (C18:0), Elaidic (C18:1n9t), Oleic (C18:1n9c), Linolelaidic (C18:2n6t), Linoleic (C18:2n6c), Arachidic (C20:0), Gamma Linolenic (C18:3n6), *cis*-11-Eicosenoic (C20:1), Linolenic (C18:3n3), Heneicosanoic (C21:0) *cis*-11,14 Eicosadienoic (C20:2), Behenic (C22:0), *cis*-8,11,14-44-Eicosatrienoic (C20:3n6), Erucic (C22:1n9), *cis*-11,14,17-Eicosatrienoic (C20:3n3), Arachidonic (C20:4n6), Tricosanoic (C23:0), *cis*-13,16-Docosadienoic (C22:2), Lignoceric (C24:0), *cis*-5,8,11,14,17-Eicosapentaenoic (C20:5n3), Nervonic (C24:1), and *cis*-4,7,10,13,16,19-Docosahexaenoic (C22:6n3).

3.3.2 Samples of Lard and Selected Fats

A part from pigs fats, chicken fats, beef, and mutton fats were selected for comparison. Chicken and beef fats were selected due to their close similarities with lard based on TAGs of lard (Codex Alimentarius, 2003). The fats of pigs, chickens, cows, and mutton (represents goat/sheep) which was obtained from the local and wet market. All the animal fats were taken from the belly part. Plant fats are palm oil-based shortening purchased from the Lam Soon brand. Shortening was used as a control fat to differentiate between plant and animal fats.

3.3.3 Apparatus and Instruments

1. Hot plate

   The heating-process of the collected samples was operated on a digital hotplate (DAIHAN MaXtir™ 500) with adjustable and controllable temperatures with the probe's maximum value at 500

°C ± 0.3 °C. The temperatures were measured by a probe that dipped into fats.

2. FTIR

FTIR spectrometer (Varian 3100 Excalibur™ series 300) was equipped with deuterated triglycine sulphate (DTGS) detector and single-bounce ATR crystal attachment. The IR wavenumber range from 400 cm^{-1} to 4000 cm^{-1}.

3. NMR (^{1}H & ^{13}C)

A Varian (VNMRS-500 MHz 5 mm AutoX DB (Dual Broadband) probe 1H) operating with VNMRJ 2.3 software under LINUX Red Hat 5 was performed by Automatic tuning for all nuclei by ProTune accessory.

4. GC-FID

An Agilent Technologies 7890N Network GC (Agilent Technologies USA) with 7683B Autosampler and FID detector was analysed and identified FAs by Agilent Chemstation B.03.01 software.

5. LC-MS/MS

An AB Sciex 3200 (AB Sciex, Darmstadt, Germany), triple quadrupole and hybrid triple quadrupole linear ion trap (QTRAP®) with Multiple Reaction Monitoring (MRM) detection mode

controlled was used to analyse and identified lipids by Analyst® software version 1.5.2.

3.3.4 Sample Preparation of Animals Fats

The crude animal fats were cleaned from blood, mucus, pieces of meat, and other tissues using a clean cutter. The fats were finely chopped, minced, dried, and strained. Then, the dried samples were weighted at 100g ± 0.05 and labelled. Minced fats were put in the quartz beaker 500 mL. Then the weighted fats were put on the hotplate to undergo the heating-process protocols.

3.3.4.1 Heating-Process Protocols

The temperature and duration time was determined according to the frying and baking process commonly used. The typical temperatures of heating protocols on foods were reported to be operated between 180 °C - 220 °C. (Aniołowska & Kita, 2016; Coelho et al., 2015; Wang et al., 2016). In this study, the selection temperature was conducted at 120 °C, 180 °C, and 240 °C. Time of heating was applied from 30, 60, 120, and 180 min (Krokida et al., 2001) until the fats changed darker than their original colour.

It was observed that the rude fats are very resistant to temperatures below 90 °C. Then, the selection of the lowest temperatures (i.e., 120 °C) was made because it is higher than the boiling point of water to evaporate and followed by the melting point for fat in general. Therefore, the heating-process protocols included the sum of heat (temperature sums with duration hours). Upon temperatures sum with the duration of heating, volumes of 10

mL of heated fat were collected and cooled at room temperature. Then heated fats were bottled, labelled and stored in the freezer below -4 °C for further analysis. For the convenience of the labelling in this study, the duration times were changed from minutes (min) to hour scales (0.5, 1, 2 & 3 hrs) as per labelling (Table 3.4).

Table 3.4: Heating-Process Protocols.

Temperature. T (°C)	Heating times, hours (hrs)	Sum-heat (°C x hrs)	Beef fats	Chicken fats	Lard	Mutton fats	Plant fats
120	**0.5**	**60**	B-01	C-01	L-01	M-01	V-01
120	**1**	**120**	B-02	C-02	L-02	M-02	V-02
120	**2**	**240**	B-03	C-03	L-03	M-03	V-03
120	**3**	**360**	B-04	C-04	L-04	M-04	V-04
180	**0.5**	**90**	B-05	C-05	L-05	M-05	V-05
180	**1**	**180**	B-06	C-06	L-06	M-06	V-06
180	**2**	**360**	B-07	C-07	L-07	M-07	V-07
180	**3**	**540**	B-08	C-08	L-08	M-08	V-08
240	**0.5**	**120**	B-09	C-09	L-09	M-09	V-09
240	**1**	**240**	B-10	C-10	L-10	M-10	V-10
240	**2**	**480**	B-11	C-11	L-11	M-11	V-11
240	**3**	**720**	B-12	C-12	L-12	M-12	V-12

3.3.4.2 Extraction of Fats after Heating-Process

The heated fats samples from different types were extracted for further analysis, as mentioned in Section 3.2.3.2.

3.3.4.3 Sample Preparation of FTIR Analysis

The extracted fats were removed from the freezer and thawed for 15 min in a water bath (Memmert GmbH™) at 60 °C, left to stand at room temperature, and then transferred into a centrifuge tube for FTIR for analysis.

3.3.4.4 Sample Preparation of ^{1}H-NMR Analysis

The thawed fat samples were weighed at 0.5 g ± 0.01 using a microbalance (Mettler Toledo, USA) and 1 mL of deuterated chloroform (*d*-$CHCl_3$) containing 0.01 mg of *d*-TMS as an internal standard in a test tube. Then, 600 µL of the solution was pipetted using a micropipette (Eppendoft Research®) and transferred in 5 mm NMR tubes (NORELL®). Then the mixtures were agitated vigorously for 30 sec with a vortex (Lolli et al., 2018) prior to ^{1}H-NMR analysis.

3.3.4.5 Sample Preparation of ^{13}C-NMR Analysis

The thawed fat samples were weighed at 0.2 g ± 0.01 using a microbalance (Mettler Toledo, USA) mixed well with 1 mL of *d*-$CHCl_3$ containing 0.01 mg of *d*-TMS as an internal standard in a test tube. Then 600 µL of the well-mixed solution were pipetted using a micropipette (Eppendoft Research®) and transferred in 5 mm NMR tubes (NORELL®). Then the mixtures were agitated vigorously for 30 sec with a vortex before ^{13}C-NMR analysis.

3.3.4.6 Sample Preparation of GC-FID Analysis (FAs Methylation)

The extracted lard was used for FAs analysis. FAMEs were prepared according to the IUPAC (1982) method with slight modifications by replacing sulfuric acid with methanolic hydrogen chloride 5% as recommended (Christie, 1993). Methanolic hydrogen chloride is most widely used as an acid catalyst because it is a relatively mild reagent and gives higher yields.

The thawed fat samples were weighed at 0.5g ± 0.01 using a microbalance (Mettler Toledo, USA) and dissolved in toluene (1 mL) in a test tube with Teflon cap. 2 mL of chloride acids in methanol (5 %) were added before the mixtures were refluxed at 80 °C in a water bath (Memmert GmbH™) for 2 hrs in a stoppered tube and sealed with paraffin film. The mixtures were then cooled at room temperature. Finally, 5 mL of deionised water containing sodium chloride (5 %) was added to stop the reaction.

The required esters were extracted with hexane (2 x 5 mL). The extractants have water-containing acids removed using a Pasteur pipette to separate the layers. The hexane at the upper layer was washed with 4 mL of distilled water containing potassium bicarbonate (2 %) to neutralise any excessive acids. This step was crucial because the acidic solution could damage the GC column. Next, the methyl esters were dried using a small amount of sodium sulphate anhydrous. The solutions were filtered using a syringe filter (Whatman™ 50 mm / 0.2 µm), and FAMEs were extracted with 1.5 mL of hexane. Before GC analysis, the hexane portion was transferred to 2 mL vials and kept in the freezer at -4 °C.

3.3.4.7 Sample Preparation of LC-MS/MS Analysis

Samples of fats that had thawed were pipetted accurately at 20 µL using a micropipette (Eppendoft Research®) into a 20 mL capped test tube at then 10 mL of isopropyl alcohol (IPA) was added (McIntyre, 1999). Then, the mixtures were agitated vigorously on a vortex mixer for 30 sec. Then 1 mL of the mixtures was transferred into 2 mL vials and kept in the freezer at -4 °C before LC-MS/MS analysis.

3.3.5 Fats Analysis

3.3.5.1 FTIR Analysis

The obtained fats from section 3.3.4.2 were placed carefully on a multibounce plate at a controlled ambient temperature (25 °C) using a Pasteur pipette in direct contacted with attenuated total reflectance (ATR) crystal. FTIR spectra were recorded from 16 scans at a resolution of 8 cm^{-1} from 4000–650 cm^{-1}. The sample cell was washed and dried upon analysis prior to the subsequent analysis to avoid cross-contamination.

3.3.5.2 NMR (^{1}H & ^{13}C) Analysis

^{1}H-NMR and ^{13}C-NMR spectra were collected at optimum parameters such as pulse angle, pulse delay, and relaxation time (T1). ^{1}H-NMR was utilised to investigate any significant changes in the overall chemical shifts, and then FAs correlations were determined by ^{13}C-NMR. Single-dimensional, the ^{1}H NMR spectra were measured at 300K. A pulse sequence was applied to suppress the residual water signal. Then the 128 transients were collected in 32 K data points with a relaxation delay of 2 sec. A spectral width of 9615.4 Hertz (Hz) and an acquisition time per scan of 1.70 sec were used. Before Fourier pre-processing, an exponential line broadening function of 0.30 Hz was applied to the free induction decay. The chemical shifts for all samples were referenced to tetramethylsilane (*d*-TMS) at 0.00 ppm.

3.3.5.3 GC-FID Analysis

GC-FID operating conditions were as follows: the column was an HP Agilent 88 capillary column (100 m× 0.25 mm i.d., 0.25-µm film thickness;

Agilent, USA); ultra-high purity nitrogen was used as the carrier gas at a flow rate of 1 mL/min, and separation was conducted in constant flow mode; a split liner with glass wool was installed in the injector; the injector temperature was set at 250 °C for split injection at a split ratio of 100:1; the FID temperature was set at 280 °C; ultra-high purity hydrogen and scientific-grade air were the fuel gases for the FID with the flow rate of 40 mL/min and 450 mL/min, respectively; the initial oven temperature was set at 140 °C and held for 5 min, then ramped up at 4 °C/min to 240 °C; this temperature was maintained for an additional 15 min.

3.3.5.4 LC-MS/MS Analysis

In the AB Sciex, 3200 systems (AB Sciex, Darmstadt, Germany), a Phenomenex Synergi Fusion 100 mm x 2.1 mm x 3 um (USA) column was used for the analysis. Gradient chromatography was performed with 60 % water, 40 % (acetonitrile) ACN 10 mM ammonium formate, 0.1 % with formic acid (A) and 80 % isopropanol, 20 % ACN with 10 mM ammonium formate and 0.1 % formic acid (B) at a flow rate of 0.25 mL/min. The gradient used was 0–0.5 min, 5 % B; 1.5 min, 70 % B; 2.5 min, 90 % B; 3.0–3.5 min, 5% B; 3.6 min, stop. The injection volume was 10 L. Isocratic elution was applied at 6% B with 0.1% acetic acid. The column was equilibrated for 1 min before each run.

Positive ion electrospray tandem mass spectra were recorded by Analyst® software version 1.5.2. The ion spray voltage was set to 5000 V, and the probe temperature was set at 250 °C. Nitrogen was used as the collision gas. The nebulizer (GS1), curtain, and turbo gas (GS2) were set to

be 40, 10, and 20 psi, respectively. The selected run time for total detection was fixed to around 30 minutes.

3.4 Data Analysis

The analysis's FTIR and ^{1}H-NMR data files were required to convert into the x-y format. ^{1}H-NMR data were transferred from *.fid* format to excel in Microsoft Excel 2010®. The FTIR data were transferred from *.sp* files into ASCII files before analysis. Similarly, GC-FID data originally by *.d* files were transformed into ASCII files. These processed and normalised data were subjected to multivariate analysis using The Unscrambler® X version 10.3 (CAMO software, USA).

3.4.1 Data Pre-processing

All data sets (X) were applied pre-processing or transformation based on the type of spectra. Different pre-processing were chosen according to the instrument type and some previous works. The data were employed for FTIR and ^{1}H-NMR because both are continuous (wavelet) types. The orthogonal signal correction (OSC) is a pre-processing method that removes variations from the X-data unrelated to some responses. The assumption is that a multivariate regression model trained on modified data will be more sparse, robust, and interpretable. OSC was initially designed for use with data from near-infrared reflectance spectroscopy. The implementation in The Unscrambler® X is based on the work of Fearn (2000).

3.4.1.1 Data of FTIR

Pre-processing assessments were emphasized on infrared (IR) data output because they are more complex spectra than the other instrument's data output. The obtained results *bsp*. format files were transferred into ASCII files using The Unscrambler® X. The assessments were conducted in more detail for FTIR spectra based on the most commonly used ones, especially in oil and fats research. The pre-processings of the datasets were selected, such as normalisation, multiplicative scatter correction (MSC), second derivation Savitzky-Golay (2^{nd} DSG), baseline correction, and correlation optimized warping (COW).

3.4.1.2 Data of ^{1}H-NMR.

The spectral region of 0.5–9 ppm containing the metabolite peaks information was binned or bucketed into segments of equal width of 0.04 ppm (255 bin variables) using MestReNova software (version 11.0). Then, the binned data were organized into a matrix of 255 variables and subjected to sum normalisation, cube root pre-processing, and auto-scaling. Data pre-processing and normalisation are two strategies to make features more comparable for multivariate analysis using The Unscrambler® X.

3.4.1.3 Data of ^{13}C-NMR

The spectral region of 0.5–220 ppm containing the metabolite peaks information was binned or bucketed into segments of equal width of 0.02 ppm using MestReNova software (version 11.0). The binned data were organized into a matrix of 550 variables and subjected to sum normalisation.

Whittaker Smoother corrections were processed for denoises. For further correction, the full ^{13}C-NMR spectra were exported and converted to ASCII files and then Excel files before being transferred for multivariate analysis using The Unscrambler® X.

3.4.1.4 Data of GC-FID

The identification of FAs was compared to 37 component FAMEs standard mixtures. The external standard was run in the GC prior to methylated fat samples. FAs ratios of the internal standard were extrapolated and calibrated by Agilent Chemstation b.04.03 version software. The data was transformed into ASCII files, then into Microsoft Excel 2013 before being transferred for multivariate analysis using The Unscrambler® X.

3.4.1.5 Data of LC-MS/MS

The LipidView™ Software automatically produced lipid identifications that accompanied peak intensities and areas. The peaks could be reliably identified using library searching of full-scan MS/MS spectra that are acquired automatically when a peak exceeds the specified threshold at coefficient variation values less than 20%. The results were exported into Microsoft Excel 2013 before being transferred for multivariate analysis using The Unscrambler® X.

3.4.2 Chemometrics Techniques

PCA was used as a preliminary observation of the sample's distribution. PCA and Hotelling T^2 were conducted to test homogeneity on FTIR lard spectra from the various

regions before PCA projection. Then, the profiling lard and selected fats after the heating-process were investigated for similarities or differences in each class of fats. For the spectral data such as FTIR and NMR, multivariate classification and multivariate regression were used. This study also applied the support vector machines (SVMDA). SVMDA is implemented as classify function, using the default setting the type option to linear is *Nu* and *C* in the Unscrambler® X. The function of SVMDA is almost similar to the ANN. The linear discriminant algorithms were exploited by SVM as *C* values: 1 and 3, *Nu* (υ) value: 0.5 as the smallest values.

Multivariate regression aims to predict whether the model can differentiate lard from selected fats. This approach determines the relationship with the variables studied. The OSC-PLSR, PCR, and PLSR algorithms compared their performances by observing the values of the *Y*-binary matrix. *Y*-binary matrix non-*halal* represented by lard (pig fats), denominated as 1, and represented by the other edible fats, denominated as 0. The distribution of the scores plot displaying factors component (*X*, *Y*) and PCR denotes by PC is to summarize more variation in the data.

FTIR and ^{1}H-NMR dataset in profiling the heating-proses samples, two data samples, namely the training set and test set, were divided according to the Kennard-Stone (K-S) selection. After the division, the training set or calibration set group went through the pre-processing, and then PCA was applied to multivariate classification and multivariate regression. The process is called chemometrics modelling. Then, the test set was validated through projected chemometric modelling respectively.

3.4.2.1 Kennard-Stone (K-S) Algorithm Selection

The K-S algorithm selected the training set: test set at 70:30 ratios. However, in this study, the 2:1 ratio was used. The test set was used to validate the prediction ability of the suggested models (Li et al., 2016). The K–S algorithm chooses a subgroup of samples from X that provide a demonstrative and identical coverage over the data matrix. The method initiates by discovery the two samples that are the farthermost in the ordered distance of PCA scores plot of the raw data before the pre-processing.

Overall, suppose data set from analytical platform FTIR and ^{1}H-NMR have a training data set consisting of pairs (x_i, y_i) where, $i = 1, \dots N$, $x_i = (x_{i1} \dots \dots x_{ip})$ contains measurements of p variables on subject i, i.e. wavenumber and chemical shift. Then the y_i is symbolised as the value of each variable, i.e., classes of collected lard samples and types of heated fats. Likewise, suppose a test (x_i^*, y_i^*), where $i* = 1, \dots, N^*$. $x_i^* = (x_{i1}^* \dots \dots x_{ip}^*)$. The training and test sets assumed that both samples are from the same universe as Equations 3.1 and 3.2. The training and test sets partition are according to the K-S selection. The chemometrics model, namely PCA projection, multivariate classification, and multivariate regression, are function ƒ that maps x vectors into one-dimensional scores,

$$\hat{y}_i = f(x_i) \tag{3.1}$$

$$\hat{y}_i^* = f(x_i^*) \tag{3.2}$$

where $\hat{y}_i$and $\hat{y}_i^*$ are the scores vector for the training set and test set at i^{th}, column respectively. If larger values of ƒ are associated, then the better classification of lard.

Scores plots of each model are fitted using the training data, and next the models are evaluated using the test data. These strategies were used on FTIR and ^{1}H-NMR data after the heating-process. First, multivariate classification is used to group the edible fats accordingly. Second, multivariate regression is used to discriminate lard and other fats. Then, OSC-PLSR was employed to find the important variables or chemical constituents contributing to the lard identification.

Identification of the specific FAs subjected to the heating-prosess were employed by GC-FID, ^{13}C-NMR, and LC-MS/MS datasets, then only PCA was applied onto these data because samples were limited, and then one left out of the cross-validation was performed. In addition, heatmap and *S*-plot or *S*-graph were also performed to simplify the visualisation of the *X*-correlation loading plots.

3.4.2.2 Flowchart of Chemometrics Evaluation

FTIR generated the profiling of lard from pigs from different body parts and regions to create a database (Figure 3.2). Data were divided into training and test sets before pre-pre-processing and PCA analysis. First, the Hotelling T^2 was utilised as boundary measures of FTIR spectra. Then the test set was resampled onto the PCA.

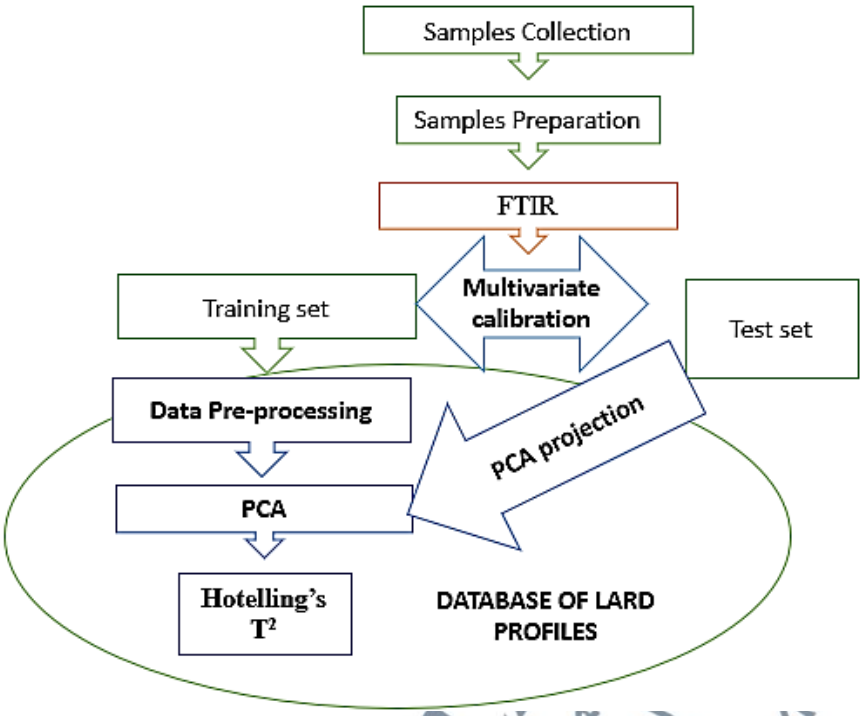

Figure 3.2: Chemometrics Techniques for Lard Profiles.

Multivariate classification and multivariate regression evaluated FTIR and ^{1}H-NMR data of samples subjected to the heating-process (Figure 3.3). Both instruments produce continuous data points suitable for conducting multivariate calibration by dividing them into the training and test sets for chemometrics model development. Since the ^{13}C-NMR, GC-FID, and LC-MS/MS data were limited, and only PCA techniques were applied.

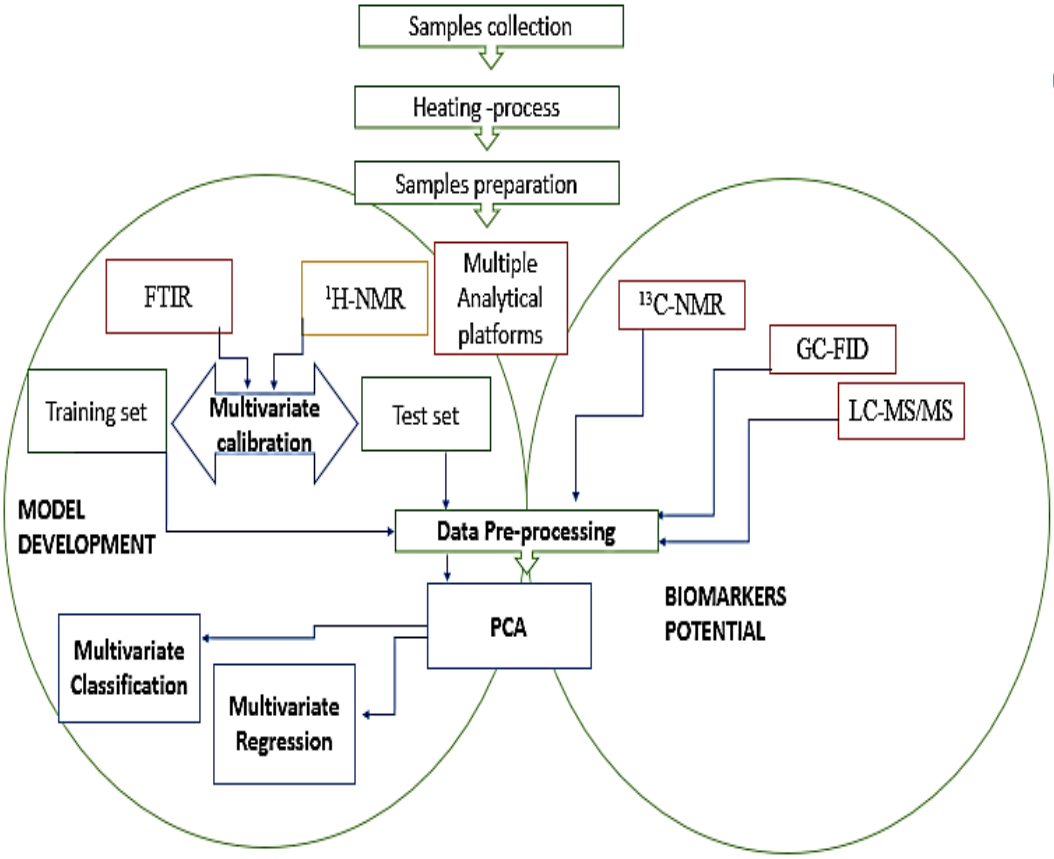

Figure 3.3: Chemometrics Techniques for the Heated Lard Profiles.