IoT based Milk Quality Analysis System for Fat and Protein Detection

By

Ch. Muhammad Shaheer Yasir (286021) (BEE 11B)

Nihal Ahmad (303249) (BEE 11B)

Muhammad Hamza Khalid (298694) (BEE 11B)

A project report submitted in partial
fulfilment of the requirement for the
degree of

Bachelor of Electrical Engineering

School of Electrical Engineering and Computer Science

(SEECS)National University of Sciences and Technology

(NUST)

Islamabad,

Pakistan

(2023)

Approval Page

It is certified that the contents and form of project report entitled **"IoT-based Remote Monitoring and Control System for Smart Dairy & Poultry Farm Housing"** submitted by **Ch. Muhammad Shaheer Yasir, Nihal Ahmad** and **Muhammad Hamza Khalid** has been found satisfactory for the requirement of the degree.

Advisors:

Dr. Mansoor Ali Khan _____

Engr. Nasir Mahmood _____

Acknowledgements

We thank Dr Mansoor Ali Khan Assistant Professor SEECS NUST and Engr. Nasir Mahmood Assistant professor SEECS NUST for their kind support from the idea to design and development for their precious time spent in the completion of the project. We also thank Dr Imran Tarique Assistant professor ASAB NUST for giving us access to Industrial Biotechnology Lab (IBT) and use spectrophotometer for our project.

List of Figures

List of Tables

Chapter 1

Introduction

Most of Pakistan's economy today is dependent on the agriculture and dairy industry. The biggest problem that consumers face regarding the dairy industry is the adulteration of pure milk with water and detergents. We propose a solution to measure the quality of milk through attenuation of light waves as they pass through a milk sample. The degree of attenuation is dependent on the amount of fat and protein present in the sample. The higher the fat and protein content, the more light is reflected, and we get more attenuation.

We have divided our project in three parts:

1. In the first part, we identified the top seven adulterants that can be used in our country. Then we identified the top six techniques which can be used to measure the quality of milk. We cross analysed the adulterants with these techniques to determine which techniques are more effective for certain adulterants. In the end, we made a comparison analysis which listed the techniques against the adulterants along with their figure of merits.
2. In the second part, we chose a specific technique called spectrophotometric techniques to measure milk quality. We chose different samples of milk having different protein and fat content and analysed them on Spectrophotometer. In the end, we build a PLS Regression model which can predict the fat and protein content in milk.
3. In the third part, we build a small device based on raspberry Pi to predict the fat content. This device consists of an LDR which measures the absorption of light passing through milk samples. We chose different milk samples available in the market having different fat content. We collected the LDR readings for each sample and trained a model on this data. This model is also able to predict fat content in a milk sample effectively.

Properties of Milk

Before moving to literature review and further work, the reader must know some of the properties in milk. Milk is a white liquid abundant in nutrients produced by mammals. In general cow milk contain the following constituents:

1. **Water Content:** The major constituent of milk is water, which constitutes 87% of its weight.
2. **Protein Content:** There are various types of protein present in milk such as whey, Casein and others. Casein is the dominant protein in milk.
3. **Carbohydrates:** Milk contains different types of carbohydrates among which Lactose is the major. It usually constitutes around 4-5% of its weight.
4. **Fat:** The fat content in milk is around 3-4%.
5. **Others:** Milk is a good source of vitamins and minerals, including calcium, vitamin D, vitamin B12, and others.

These properties are of cow milk. A detailed analysis of these properties in milk from different breeds in Pakistan is discussed in [1].

Chapter 2

Literature Review

The constituents in milk can be determined by three basic techniques/methods:

1. Chemical methods
2. Optical Methods
3. Ultrasonic Methods

1. Chemical Methods:

Chemical methods are commonly used to measure milk quality, allowing producers to assess factors such as fat content, protein content, and somatic cell counts. One of the most important measurements in milk quality is milk fat content. Fat content affects the flavour and texture of dairy products, and accurate measurement is crucial for ensuring that products meet the required standards. The most widely used methods for determining milk fat content are the Gerber method and Rose-Gottleib Method [2].

Protein content is another important factor in milk quality. Protein content affects the nutritional value of milk and the properties of dairy products such as cheese and yogurt. There are various methods for measuring protein content, including the Kjeldahl method and the Bradford assay [3]. The Bradford assay involves adding a dye to the milk sample, which binds to the protein. The amount of dye bound to the protein is then measured spectrophotometrically, allowing the protein content to be determined.

Somatic cell count is another important measurement in milk quality. Somatic cells are cells that are naturally present in milk, including white blood cells and epithelial cells from the udder. High somatic cell counts can indicate the presence of mastitis, a common udder infection in dairy cows. Mastitis can reduce milk yield, affect the quality of milk, and increase the risk of antibiotic residues in milk. One common method for measuring somatic cell counts is the California Mastitis Test (CMT). This method involves adding a CMT reagent to a milk sample and observing the formation of a gel-like substance. The degree of gel formation is proportional to the somatic cell count, with higher counts resulting in thicker gels. Other methods for measuring somatic cell counts include flow cytometry and electronic cell counting.

The predominant carbohydrate found in the milk of many species is lactose, which is considered a readily available energy source for new-borns. Determining the percentage of lactose in milk can be done using methods such as the HPLC method or the DNS method. High Performance Liquid Chromatography (HPLC) is usually applied as a reference method for lactose analysis. A method for measuring lactose using HPLC-ELSD was developed by Schuster-Wolff-Buhring, Michel, and Hinrichs in 2011. The method was able to detect lactose at a minimum level of 3.8 mg L1 and quantify it at a minimum level of 17.3 mg L1 [4].

Other methods like ELISA, atomic absorption spectroscopy and ICP Analysis are also used to measure milk quality [5]. Enzyme-linked Immunosorbent Assay (ELISA) is a test kit available commercially for plasmin detection in milk [6].

Optical Methods:

Spectroscopy is a useful tool to investigate the interaction between light and milk components. Through this method, various phenomena related to light-milk interactions can be analysed [7]. It has been shown that the properties of light like absorption, scattering, and refraction are dependent on its wavelength. Therefore, analysing different portions of the light spectrum can lead to different responses. Thus, spectroscopy is divided into different type based on different wavelength range.

- **UV Spectroscopy:**

The UV range spans from 100 to 400 nm in wavelength. UV light was initially used to pre-homogenize raw milk [8]. In this process, fat globules were dispersed using a combination of chemicals and a spectrophotometric method was used to determine fat and protein content.

In 1994, Kuaye [9] introduced a method that shifted the absorption-centred wavelength of tyrosine to the wavelength ranging from 248 – 256 nm. They changed the optical properties of this protein using a strong basic solution.

In 1999, Lüthi-Peng and Puhan [10] created a straightforward and fast method to measure protein and casein levels in cow milk using spectrophotometry. Their technique utilized the fourth derivative at 283.5 and 294.5 nm wavelengths.

In 2005, Forcato et al. [11] introduced another approach to determine milk fat content using spectrophotometry within the range of 208-215 nm. Spectral analysis in UV does not provide sufficient information on the primary components of raw milk, leading to a limited number of available UV-based methods. Although UV techniques are a dependable option, they are time-consuming and require sample preparation.

- **Visible Spectroscopy:**

The visible spectrum ranges from 500 nm to 700 nm. Xin et al. [12] presented a quick and uncomplicated technique in 2006 to detect the fat and protein content in milk using a red light at 632.8 nm. While this method may serve as a substitute for laboratory analysis, it may be challenging to implement it online.

Muniz et al. in 2009, emphasized the advantages of Visible range (VIS) in fat and protein detection due to its low cost. They also suggested using PLS regression or PCR Analysis to enhance the accuracy of results obtained through these statistical techniques.

Bogomolov et al. [13] introduced a method in 2012 that utilized the visible range to determine the total fat and protein content in milk. The absorbance spectra were collected within the 400-1000 nm range. However, milk samples used had a narrow range, with

protein varying from 2.60% to 3.20% and fat varying from 3.02% to 3.98%. As a result, the practical application of this technique is not considered reliable.

In 2014, Kucheryavskiy et al. [14] proposed a distinct method in which they used image processing to analyse the light transmitted through milk samples and establish a correlation between fat and protein content. A digital camera and RGB LEDs were used to capture images in a dark environment. This approach offers a practical and cost-effective solution for the development of online systems. According to the authors, the results obtained from this method closely resemble those from spectroscopic analysis.

In 2019, Gowri et al. [15] determined the fat content in milk using an optical fibre probe. They conducted experiments to demonstrate that the changes in refractive index (RI) of milk diluted with water can be detected using this method. Although this technique has potential for conveniently determining the fat content, it is invasive and requires cleaning of the probe for 3 minutes after each determination.

- **IR Spectroscopy:**

In 2001, Šašić and Ozaki [16] conducted a study in which they analysed raw milk samples in the 700-1100 nm range, also known as the short-wave near-infrared (SWNIR) range. The potential of this wavelength range to perform quantitative and qualitative analysis of milk is shown in this paper.

Woo et al. [17] created the MilkSpec-1 unit in 2002, utilizing transmittance spectroscopy techniques to measure the fat, lactose, and protein content of milk non-destructively. The NIR spectra obtained from the system were subjected to pre-treatments such as multiplicative scatter correction and other derivatives to reduce the effects of scattering from fat globules and casein micelles. PLS calibrations were performed, and the results showed that the MSC spectrum was the most effective in predicting fat with 11 factors and protein content 13 factors, respectively, while lactose was predicted with 14 factors with no pre-treatment.

Etzion et al. in 2004 [18] measured the protein content in raw milk by analysing absorbance in the wavelength ranges of 5882-6666 and 9090-9433 nm. Although the prediction error was less than 0.5%, the calibration samples' protein content range was from 2.47 to 3.95%. Hence, these calibration models are incapable of providing accurate estimations in real environment where readings from milk samples can fall outside this range.

Kawamura et al. in 2007 [19] developed an online monitoring system that uses a wavelength range of 600-1050 nm to estimate the fat, protein and lactose content in milk. The system yielded high accuracy with coefficient of determination and standard error of prediction (SEP) for fat, protein, and lactose of r2 = 0.95, SEP = 0.42%; r2 = 0.91, SEP = 0.09%; and r2 = 0.94, SEP = 0.05%, respectively.

In 2008, Kawasaki et al. [20] introduced a milking robot that utilizes NIR spectra and PLS analysis to estimate milk quality in real-time. However, calibration models need to be updated regularly to maintain the accuracy of this robot, which makes it impractical in a real environment.

A study conducted by Aernouts et al. [21] in 2011 aimed to determine the most suitable technique for estimating fat, protein, lactose, and urea content in raw milk and used reflectance and transmittance in the VIS/NIR wavelength range. The authors concluded that fat content wasn't predicted accurately in this wavelength range as enough information is not provided. It may be due to the reason that the film used is very thin that is 1mm of size. Therefore, this method cannot be used for online analysis.

Muñoz-Ossa et al. [22] proposed a technique to determine fat content in milk using a tapered optical fibre sensor. The tapered section of the fibre was submerged in milk samples with varying concentrations of fat content. The fat globules in the taper region caused attenuation of the power signal. Although this method is invasive, it has the potential to improve linearity by using a larger taper, but then it becomes less sensitive.

Melfsen et al. in 2012 [23] created and tested a milking process system that can estimate the fat, protein, lactose, urea, and somatic cell count concentration in raw milk. The milking process system utilizes NIR spectroscopy to analyse the wavelength between 851 to 1649 nm. However, the authors claim that the laboratory analysis standards will not be met by the systems based on NIR due to low accuracy.

Feng et al. in 2013 [24] used non-dispersive short-wave near-infrared spectrometry to determine fat, protein, and lactose in non-homogenized raw milk samples within the wavelength range from 600 to 1100 nm, using PLS regression models. While this approach is suitable for estimating the principal components of milk as an off-line system, it is too time-consuming to heat the samples.

Aernouts et al. (2015) [25] designed an optical system measured total transmittance and reflectance of a milk sample within a 600-um path length cuvette. It utilized a supercontinuum laser with a monochromator. The wavelength used was 500 – 2250 nm. Although the correlation coefficients were very low, still the analysis of the backscattered optical power provides valuable information for testing different sensor arrangements and understanding the interaction of the light source with the turbid media, such as raw milk.

Zhu et al. [26] proposed a method to measure fat content in raw milk using a W-type optical fibre sensor system. The system was tested at different milk temperatures, and 40 °C was found to be the optimal temperature for the test. The light source used by the system was centred at the wavelength of 1060nm. It was then divided and travelled into a W-type fibre. Although this approach can provide a rapid estimate of fat content, it is important to note that significant variations in fat concentration need to be evaluated to develop a practical system.

Niero et al. [27] used PLS analysis to investigate the potential of the MIR to measure the protein content of cow milk. They acquired spectra from 110 milk samples using the MilkoScan FT6000 (Foss Electric A/S, Hillerød, Denmark), and calculated the absorbance. The best prediction model obtained with the system had a coefficient of determination R^2 equal to 0.88 for protein content. The method proposed can be used for practical purposes, but it is only limited to offline analysis.

In 2016, Ragni et al. [28] introduced an inexpensive and straightforward NIR-based system light detection using a photodiode. The system includes an ANN that reports predictions with an R2 value of 0.992. This method can be used for both on-line and off-line systems; still, the system's accuracy may be affected by temperature fluctuations.

Li et al. [29] introduced a method to estimate the fat and protein content in raw milk by applying spectral denoising, derivative denoising, genetic algorithm, and PLS regression. A calibration set was constructed from selected samples and validated their method, which showed an R2 of 0.8010 for protein and 0.9100 for fat content. However, samples had a very narrow range of fat and protein content, and the method used the MID IR range from 2564–25,000 nm. In this range the presence of water and other components of milk affects the spectrum.

Regular calibration of IR systems is essential. For NIR-based systems, calibration should be conducted using representative samples prior to any measurement, and a separate validation set should also be used. This calibration typically involves the use of multiple linear regressions (MLR), principal component regression (PCR), or partial least squares (PLS) methods.

Ultrasonic Methods:

Ultrasonic methods have also been explored as a means to measure milk quality. One example is the use of ultrasound to estimate milk fat content. The principle behind this technique is that ultrasonic waves can travel through the milk and are reflected by the interface between the fat and water. By measuring the time, it takes for the waves to travel through the milk and be reflected back, the thickness of the fat layer can be estimated, and from there, the fat content can be calculated. Other parameters, such as protein content and somatic cell count, can also be estimated using ultrasound, although these methods are less established than fat measurement. One of the advantages of ultrasonic methods is that they are non-destructive and can be used on live animals, making it a promising option for on-farm applications.

The acoustic properties of milk have been exploited to use the ultrasonic milk analyser for monitoring fat globules, SNF, and protein in raw milk [30]. In addition, the analyser has also been employed for the detection of adulteration in fluid food items or for monitoring of microbial growth through the use of a constant phase element sensor [31, 32]. However, the major drawbacks of these analysers are their high cost, prolonged measurement time, and the need for skilled personnel to operate them.

In their study, Das et al. [33] introduced a method to measure milk adulteration using Constant Phase Element (CPE) sensor. They measured changes in the pH level of the milk sample to detect various impurities such as tap-water, liquid-whey, and urea content. Different changes in pH level were analysed to identify the type of impurity added.

Chapter 3

Problem Statement

In Pakistan, milk is one the most consumed products with over 90% of people relying on it. However, the milk consumed is of low-quality with mixed adulterants. According to the WHO, Pakistan's annual milk production is 140 million litres, while the amount sold is 500 million litres. Drinking contaminated milk can lead to a number of health complications including indigestion, vomiting, diabetes and kidney issues. Natural pure milk has a foamy look. The foamy look is reduced by adding water to milk. To restore this natural look, vendors will often add detergent to milk. Adding a touch of urea to milk will give it a natural white look which helps to sell the lie that it is pure milk.

Our objective is to decrease milk adulteration by offering a cost-effective and user-friendly solution that can be used anywhere. Our smart LDR device produces instant reliable results that consumers can check at home, farms, dairies, or any location. Our primary aim is to provide a cost-effective quality monitoring system for milk products that allows any consumer, including farmers and laymen, to check milk quality on their premises without having to visit specialized labs.

Market Research:

While there are devices like that to measure milk quality, however none of them are present in our local markets. Also, they are primarily marketed towards the industries and labs rather than a common milk consumer. Also, they are very expensive and are out of reach of a common man. One of such devices is LactoScan Milk Analyzer which is imported in Pakistan from Bulgaria. This device in very commonly used in industries as well as sophisticated laboratories. Government also set up a mobile laboratory in Lahore two years ago and they are also using the same device.

Figure 1: LactoScan Milk Analyzer

However, this device is very expensive and the cheapest model of it cost nearly 700$. It also costs 115$ for the shipping costs. So, total costs for this are around 230K PKR. It is a very hefty amount, and a common man can't even think to buy such a device.

However, there are cheaper alternatives available as well. One of these alternatives available in the market are color testing strips which cost around 1000 PKR for measuring quality of 3 milk samples.

Figure 2: Milk Quality Testing Strips

However, these testing strips are not much accurate, and have very limited characteristics. Also, these testing strips also have to be used in a laboratory with someone who is specialized in such knowledge. Thus, it cannot serve as a portable and easy-to-access solution for milk adulteration in out country.

Customer Survey:

To gauge the customer interest in the product we are developing, we created a google form and circulated it on different social media platforms. We have received 45 responses out of which majority are from the metropolitan cities of Lahore, Karachi and Islamabad.

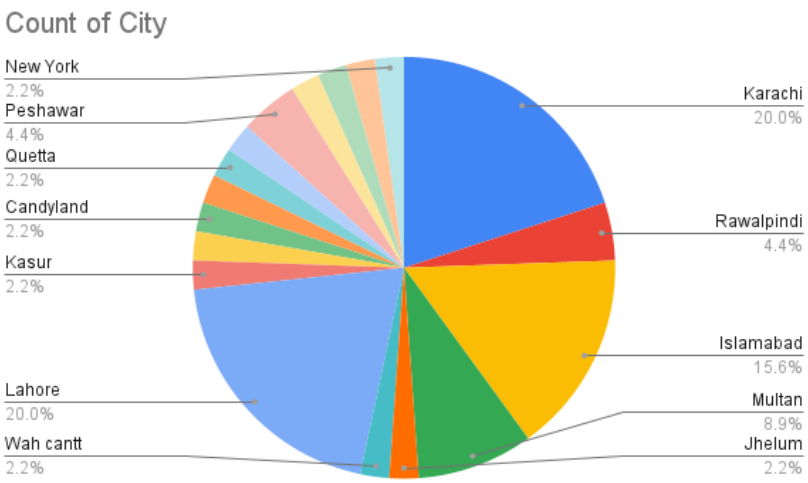

Figure 3: Responses Division by City

According to our survey, 43% of people buy milk from their local milk shops and 32% buy Olpers milk commercially available.

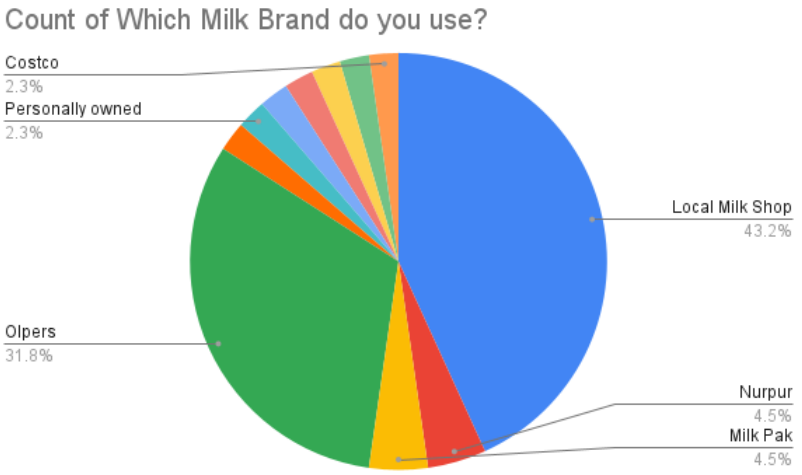

Figure 4: Division of responses by the type of milk they are consuming

When we asked them to rate the quality of milk, they are consuming out of 10, majority of people rated them between 6.25 to 9.08. However, they were still not satisfied and wanted to know if any adulterants are being added in their dairy products.

Figure 5: Quality of Milk Consumed by Users

Then, we asked whether they are willing to buy such a device which can measure their quality instantaneously. The response was very positive and 60% of people were willing to buy such a device. 25% chose that they will think about buying it if needed. Also, 90% of the respondents were willing to buy it if the device costs below 10,000 PKR.

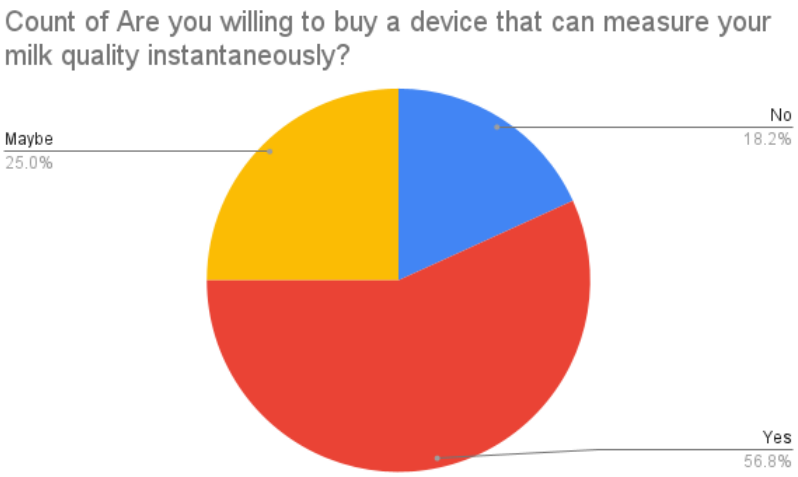

Figure 6: Willingness of customers to buy a milk quality analyzer

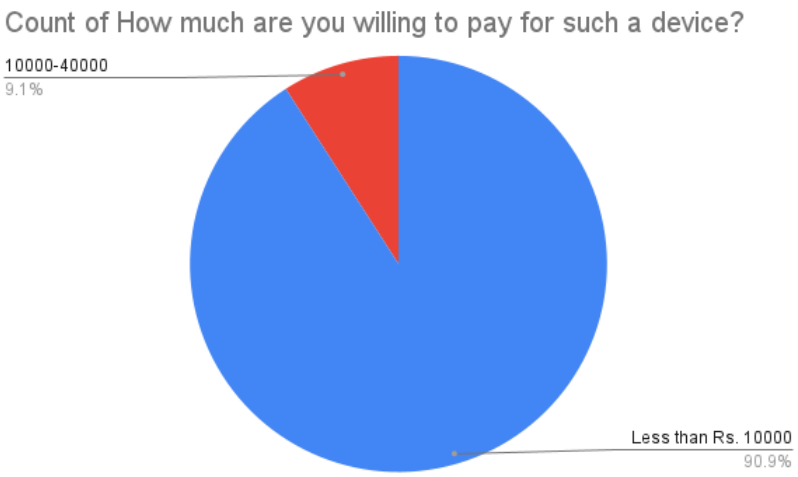

Figure 7: Cost of milk Analyzer desirable by the survey

This survey proves that there is a huge gap in our market which needs to be filled by the academic research. Such a device will not only help the consumers take care of their hygiene, but it will also expose the fraudulent who are playing with the lives of people by adding different dangerous adulterants in Milk.

Chapter 4

Methodology and Simulation

As it has already been told in the introduction chapter, that we have divided our project into three sections. Now we will discuss the methodology for each section one by one.

Comparison Analysis:

In order to start working on a given principle, first all the principles should be studied and analysed against the adulterants we are defining. Then, the most suitable method can be chosen which have the best figure of merits. For this purpose, we had selected six basic principles to measure milk quality. To define milk quality, we also identified seven major parameters of milk which can be used to measure milk quality.

The six major adulterants or parameters for milk quality are as follows:

1. **Water Level:** Water level in milk refers to the amount of water present in a sample of milk. This can be an important factor in determining the quality and composition of milk, as well as its suitability for certain uses such as cheese or yogurt making. The water content of milk can be affected by a variety of factors including breed of cow, stage of lactation, and environmental factors such as temperature and humidity. Inaccurate measurement or control of water levels in milk can lead to reduced product quality, decreased yield, and potential health risks for consumers.
2. **Fat Content:** The fat content in milk affects the sensory properties of the milk, such as flavor, texture, and color. It is also a key factor in determining the value of milk in the dairy industry, as milk with higher fat content is often more desirable for the production of dairy products such as butter, cheese, and cream.
3. **Protein Content:** Protein content is one of the important parameters used to assess the quality of milk. Milk with high protein content is generally considered to be of better quality compared to milk with low protein content. To ensure the quality of milk, regulatory authorities in many countries have established minimum standards for protein content in milk. For example, in the United States, the minimum required protein content for fluid milk is 3.2% for whole milk, 2.9% for reduced-fat milk, and 2.5% for low-fat milk.
4. **Detection of Baking Soda:** Adding baking soda to milk can cause a reaction that results in the milk becoming slightly alkaline. It can result in an unpleasant taste and texture of milk. Also, consuming too much alkaline milk can disrupt the natural pH balance of the body, potentially causing digestive issues.
5. **Detection of Salts:** Adding salts to milk is a common practice in the dairy industry for various reasons. One of the main reasons is to control the growth of bacteria and other microorganisms, as salt can act as a preservative. However, adding excessive amounts of salt can negatively affect the taste and quality of the milk. Therefore, it is important to carefully monitor and control the amount of salt added to milk.

6. **Detection of Urea:** Adding urea to milk is a prohibited practice as it is considered hazardous to human health. Urea is a nitrogen-containing compound that is commonly used as a fertilizer in agriculture. However, when added to milk, it can result in the formation of biuret, a compound that is toxic to humans and can lead to kidney damage. In addition, the addition of urea to milk can be used as an adulterant to increase the apparent protein content of milk. This fraudulent practice can lead to economic losses for consumers and also pose health risks.
7. **Detection of Detergents:** Adding detergents to milk is not recommended as it can have harmful effects on human health. Detergents are designed to remove dirt and grease from surfaces and can be toxic if ingested. Ingestion of detergents can cause irritation of the digestive system, nausea, vomiting, and in severe cases, damage to the liver, kidneys, and other organs. Additionally, detergents can affect the flavor and quality of milk, making it unsuitable for consumption.

Now, we will define six methods that are defined to measure milk quality in research and industry. We will use each method to determine each parameter to milk and then decide whether this method is suitable for that parameter. If that method is suitable, we will measure its figure of merit as well.

1. LDR Method:

The LDR (light-dependent resistor) method has been proposed as a low-cost and rapid method to measure the quality of milk. The LDR is a type of resistor that changes its resistance based on the amount of light it receives. In this method, a light source is passed through a sample of milk, and the intensity of the transmitted light is measured by an LDR. The amount of light that passes through the sample is dependent on the quality of the milk, such as its fat and protein content. By measuring the intensity of the transmitted light, the LDR method can provide an estimate of the quality of the milk.

We have used LDR method to measure different milk parameters. The LDR readings for different samples of milk is shown below:

Sample	LDR Reading	Fat(in 100 ml)	Protein (in 100 ml)
Without Anything	21000-25000	-	-
With Test Tube	66000-70000	-	-
With Water	25000- 29000	-	-
Milk Pak	900000-1000000	3.5g	2.7g
Dayfresh Milk	1200000-1400000	3.5g	3.3g
Dairy Omung	750000-800000	1.7g	2.7g
Milk Pak With Baking Soda Added	900000-1010000	3.5g	2.7g
Milk Pak With Detergent Added	900000-1000000	3.5g	2.7g
Milk Pak With Salts Added	900000-1000000	3.5g	2.7g
Milk Pak With Water Added	900000-950000	3.5g	2.7g

Figure 8: Use of LDR Method to measure milk quality parameters

As it can be seen in Figure 8, LDR can successfully distinguish between milks of different fat and protein content. It has given different readings for Milk Pak, Dayfresh Milk and Dairy Omung which has different fat and protein content. However, when we added different adulterants to Milk Pak, it failed to detect them and thus, gives the same reading as in the last 4 columns.

For fat content, there is an increase of 150000 in LDR readings with an addition of 1.7g of fats. In the case of protein, an addition of 0.6 g proteins leads to almost 200000 increase in LDR Reading.

Ultrasonic Method:

We have already discussed ultrasonic method to measure milk quality in Chapter 2: Literature Review. For the purpose of analysing these milk quality parameters, we have made use of LactoScan Operations Manual. LactoScan is a device commonly used in Pakistan and is imported from Bulgaria. Figure 9 shows the accuracy and range of milk parameters Lactoscan can measure.

Accuracy:

Fat .....±0.06%
SNF .....±0.15%
Density .....±0.3 kg/m3
Proteins .....±0.15%
Lactose .....±0.20%
Water content .....±3.0%
Temperature of milk ....±1oC
Freezing point.....±0.005oC
Salts .....±0.05%
PH.....±0.05
Conductivity.....±0.05
Total solids .....±0.17%

Measuring Ranges:

Fat .....from 0.01% to 25% (45%*)
SNF.....from 3% to 40%
Density …….....from 1000 to 1150 kg/m3
Proteins .....from 2% to 7%
Lactose .....from 0.01 % to 20 %
Water content .....from 0 % to 70 %
Temperature of milk …from 1oC to 40oC(if measurement is 30 sec, then t° is from 15 to 40°)
Freezing point.....from –0,400 to –0,700oC
Salts .....from 0,4 to 4%
PH.....from 0 to 14
Conductivity ……from 3to 14[mS/cm]
Total Solids.....from 0 to 50 %

Figure 9: Accuracy and Range of Milk Quality Parameters measured by Lactoscan

As we can see, Lactoscan can measure all the seven parameters we have discussed above. Baking Soda can be detected by the change in pH levels of milk. Similarly, detergents can be detected through the change of conductivity of milk.

Spectroscopic Method:

We have already discussed different UV, VIS and IR spectroscopy in the literature review section of this report. As the second part of this project is concerned with the detection of fat and protein content with spectrophotometer, we can say that we can measure fat and protein content through these methods. However, as we will see accuracy in protein content measurement is somewhat compromised.

We will discuss the other parameters which can be determined through these methods. [34] studied the detection of several adulterants with NIR spectrometers whose wavelength lies from 700 nm to 1100 nm. Figure 10 shows the best wavelength ranges to detect these adulterants in milk.

TABLE 2. RANGES OF WAVELENGTHS FOR THE BEST CALIBRATION MODEL OF EACH COMPONENT

Components	Wavelengths range, nm
Milk	926.634 – 939.409
Urea	996.599 – 1021.85
NaOH	945.788 – 977.599
Oil	926.634 – 961.708
Shampoo	933.024 – 945.788

Figure 10: Best wavelength ranges to measure several adulterants in milk

The NIR calibration and validation statistics for different adulterants are shown in Figure 11.

TABLE 3. NIR CALIBRATION AND VALIDATION STATISTICS OF COMPONENTS IN MILK SAMPLES

Components	Elements	Calibration			Validation		
		R*	SEC	Bias	R	SEP	Bias
Milk	58	0.89	4.33	-0.00	0.89	4.32	-0.34
Urea	58	0.98	0.76	-0.00	0.98	0.78	0.29
NaOH	58	0.95	0.69	-0.00	0.86	0.88	0.06
Oil	58	0.89	1.99	-0.00	0.74	2.53	-0.55
Shampoo	58	0.69	4.24	-0.00	0.58	3.83	-0.33

*: multiple correlation coefficient; SEC: standard error of calibration; SEP: standard error of prediction.

Figure 11: NIR Calibration and Validation Statistics of Components in milk samples

Figure 12 shows the regression model for urea detection.

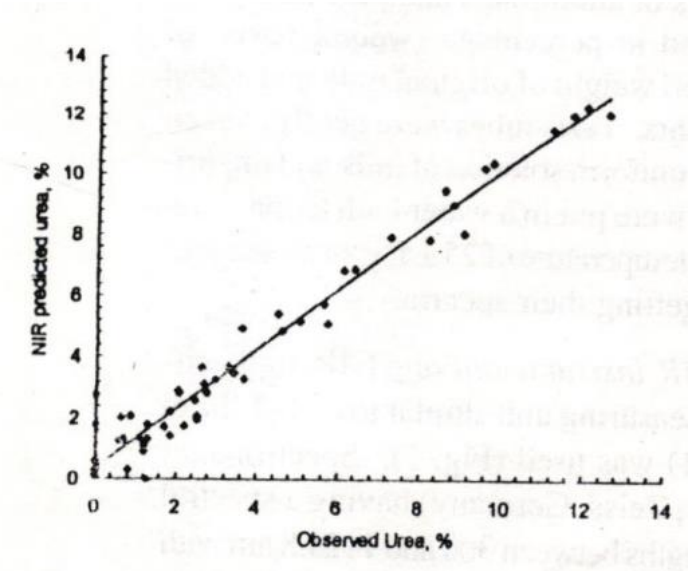

Figure 12: Regression Model for Urea Detection in milk

As you have seen several important adulterants can be measured with spectroscopic analysis. However, these methods cannot measure water content in milk. Similarly, we haven't any work where these methods were employed to detect baking soda in milk.

Density Method:

Density method refers to the measurement of density to predict milk quality parameters. We have measured density of the milk through Archimedes principle. The principle of Archimedes states that when an object is submerged in a fluid, it experiences an upward force equivalent to the weight of the fluid that it displaces. The force exerted by the milk on the metal ball equal equals the upthrust. This force is measured by the load sensor and HX711 amplifier.

The experimental apparatus setup by us consists of a load sensor, HX711 amplifier, a metal spherical ball and Arduino Uno.

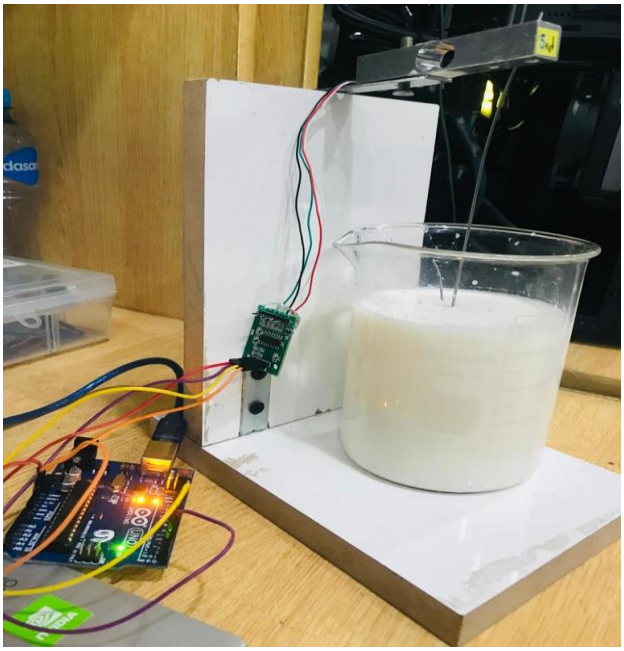

Figure 13: Experimental Apparatus set up to measure density of milk

Measurement of density for different milk samples is as follows:

Sample	Density Reading	Fat(in 100 ml)	Protein (in 100 ml
With Water	1	-	-
Milk Pak	1.034	3.5g	2.7g
Dayfresh Milk	1.027	3.5g	3.3g
Dairy Omung	1.035	1.7g	2.7g
Milk Pak With Baking Soda Added	1.013	3.5g	2.7g
Milk Pak With Detergent Added	0.98	3.5g	2.7g
Milk Pak With Salts Added	0.99	3.5g	2.7g
Milk Pak With Water Added	1.01	3.5g	2.7g

Figure 14: Use of density method to measure milk parameters

As it can be seen, water content, baking soda, salts and detergents can be measured through density method. It can also be used to measure fat content however it isn't much reliable. Protein content and Urea cannot be detected through this method.

pH Method:

pH method refers to the measurement of pH to predict different milk quality parameters. pH can be measured through pH sensor available commercially in the market. pH method cannot be used to measure fat, protein and water content in milk.

It can also not be reliably used to measure Urea contamination. Figure 15 shows that despite adding Urea, the milk pH stays in the pure milk region. However, Sodium Bicarbonate, or baking soda increases pH well above 7.1. So it can be detected using pH.

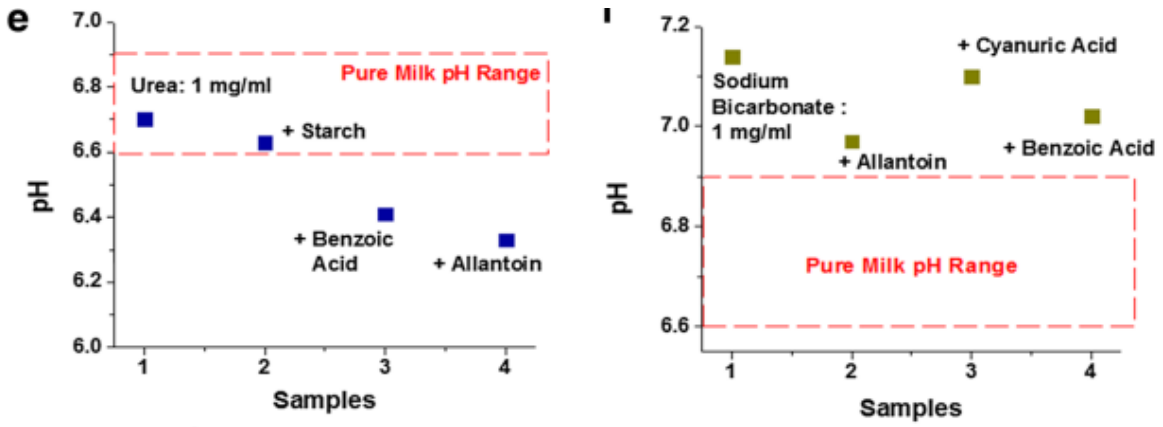

Figure 15: Urea and Baking Soda Detection with pH measurement

Salts and detergents both reduce the pH significantly. In figure 16, The reference value of pH in fresh milk is 160. However, both of the contaminants significantly reduce it to 120-130.

Sr. no	pH	Air quality	Conductivity	Urea
1.	114	150	193	0%
2.	120	152	200	0%

Table 3.Fresh Milk +Salt

Sr. no	pH	Air quality	Conductivity	Urea
1.	128	255	333	0%
2.	138	253	326	0%

Table 2.Fresh Milk +Detergent

Figure 16: Salts and Detergent detection with pH measurement

Chemical Methods:

There is not a single chemical method to measure all the milk parameters. However, every milk parameter has a reference chemical method. As chemical methods are used as reference, it is not possible to give figure of merits regarding their measurements. Although most of the chemical methods requires observation of changes in color, it cannot be quantified. Now, we will discuss the reference method for each milk parameter briefly.

- **Gerber Method:**

This is the chemical method commonly used in laboratories for fat detection. When sulfuric acid is added to the sample, the protein breaks down and the fat is released. Amyl Alcohol is then added, which creates a fat layer. The height of the layer indicates the fat percentage.

- **Kjedahl Method:**

This method is most commonly used to determine protein contents in milk. The modern Kjeldahl method [35] involves the catalytic mineralization of organic material in a boiling mixture of sulfuric acid and sulfate salt at high temperatures above 400 °C. This process converts organically bonded nitrogen into ammonium sulfate. Then titration is used to determine protein content.

Sample quantity (ml)	Nitrogen %	Protein %
5	0.533	3.399
5	0.534	3.409
5	0.537	3.427
5	0.541	3.449
5	0.533	3.398
5	0.536	3.419
5	0.533	3.401
5	0.538	3.431
5	0.535	3.410
5	0.537	3.423
Average ± SD%	**0.536 ± 0.003**	**3.417 ± 0.016**
RSD% *	**0.490**	**0.479**
Protein Labeled Value: 3.25 g/100 ml		
Protein Factor: 6.38		
* RSD% = (Standard Deviation * 100) / Average		

Figure 17: Measurement of Protein Content in Milk using Kjedahl Method

- **Freezing Point Osmometry:**

The determination of the osmolality of aqueous samples using a freezing point osmometer is a well-established, routine laboratory method [36]. The osmolality value of milk is used in milk processing to control the water content, based on the German Food Control Regulations for Milk.

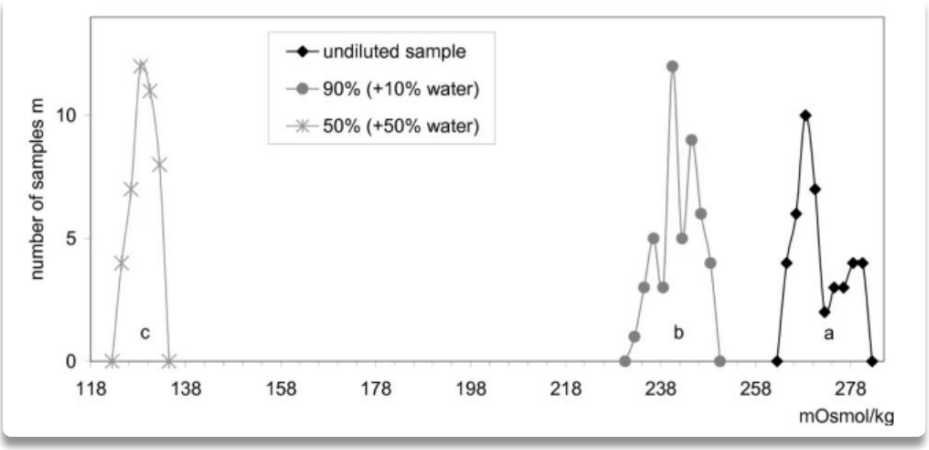

Figure 18: Determination of water content in milk using freezing point osmometry

- **Rosalic Acid Test:**

In this test, an equivalent amount of alcohol is added to the milk sample followed by a few drops of rosalic acid. If the color of the sample changes to pink red, it indicates that the milk is contaminated with sodium bicarbonate.

- **Urea Detection in Milk:**

An equivalent volume of paradimethyl amino benzaldehyde is added to the milk sample. If the solution turns yellow, then the given sample of milk is adulterated with urea.

- **Detergent Detection in Milk:**

Add a few drops of bromocresol purple solution to the milk sample. After mixing it well, if a violet colour appeared in the sample, then detergent is added to the milk sample. A faint violet colour will appear in case of unadulterated milk.

- **Detection of Salts in Milk:**

Add 5 ml of silver nitrate chemical and a few drops of potassium dichromate in a test tube. Then add 1 ml of milk in the above solution. If the colour of sample become yellow, it will indicate the presence of salts in milk. A chocolate or reddish brown will appear in case of unadulterated milk.

This was a comprehensive analysis of major milk parameters and the major techniques used to measure them. Now, we will move to the second part of our project.

Detection of Fat and Protein content with Spectrophotometer:

After careful analysis of all the methods, we have chosen two specific milk parameters (Fat and protein content) and spectroscopy analysis to measure them. As discussed earlier, optical spectroscopy involved the use of light absorption through milk to measure these parameters. Very light absorption readings are obtained at different wavelengths through the same milk sample. We have used the spectrum of 200 – 1100 nm. This spectrum includes UV, VIS and NIR Spectra. The reason we haven't included MIR and beyond is because light at such wavelengths is very difficult to generate. As indicated in different research papers, the results in MIR and beyond spectra is not much encouraging and included a lot of noise.

In order to build a model which can predict fat and protein content successfully in milk, the first step was to build data for such a model. For that purpose, we have prepared separated milk samples of fat and protein content.

For fat content, we have prepared total 36 milk samples which covered the fat range of 0.9 – 3.5 g. We used the milk available commercially and used their fat readings to label the output. The milk samples used for fat prediction are summarised in the following table:

Number of Milk Samples	Name of the Milk Brand Used	Fat Content (Per 100 ml)
12	Olpers Low Fat Milk	0.9 grams
12	Dairy Omung	1.7 grams

12	Milk Pak Full Cream Milk	3.5 grams

Table 1: Milk Samples used for Fat prediction

We took four samples every day from each category to ensure repeatability. We stored the milk samples in refrigerator at our home. Their temperature became quite normalized by the time we reached university and took the readings.

For protein content, we have prepared total 36 milk samples which covered the protein range of 2.2 – 3.2 g. We used the milk available commercially and used their protein readings to label the output. The milk samples used for protein prediction are summarised in the following table:

Number of Milk Samples	Name of the Milk Brand Used	Protein Content (Per 100 ml)
12	Asli Full Cream Milk	2.2 grams
12	Milk Pak Full Cream Milk	2.7 grams
12	Day Fresh Milk	3.2 grams

Table 2: Milk samples used for Protein prediction

Similar to fat content, we have taken four readings every day from each category to ensure repeatability in our results.

The reading was then taken on Spectrophotometer available in Industrial Biotechnology Lab, ASAB NUST. The warm-up time of spectrophotometer was 15 mins, and a blank reading was required before taking other readings. The blank reading required two cuvettes of distilled water in the spectrometer. After the blank, one of the cuvettes has to be taken out and filled with our milk sample. This was due to the reason that the spectrophotometer was measuring the light absorption in milk sample by making the light absorption in distilled water as a reference point. After every reading, the cuvette must be rinsed with distilled water before taking the next reading.

The spectrometer gave a graph of absorption readings against the wavelengths (from 200 to 1100 nm). These graphs are stored in the format of sca files. Later, a python script was written to remove the metadata from these files and convert them to csv format.

```
f = open("C:/Users/SHAHEER/Desktop/Shaheer-28/Readings/fat0.9protein3.4-4.sca", "r")
newfile = open("C:/Users/SHAHEER/Desktop/Shaheer-28/Readings/CSV/fat0.9protein3.4-4.csv", "w")
content = f.read().splitlines()
content = content[6:1807]

for a in content:
    newfile.write(f"{a}\n")
newfile.close()
f.close()
```

Figure 19: Python Code for data cleaning

The code is very simple. It opens two files (one is the sca file where the data is stored and the other is csv file where csv data has to be stored). Then in line 3 and 4, we read the data from the file and removed the meta data. Finally, we used a for loop to write the new data in a new csv file.

After the csv files have been created, we needed to import the data into MATLAB for further processing. However, the content in csv file was a four-dimensional matrix which couldn't be exported. Thus, we created a python script to create two vectors: absorption (it contained the absorption readings) and wavelengths (it contained the wavelength values). These vectors were then stored in csv format which can easily be imported into MATLAB.

```
import os
import csv
import numpy as np
import pandas as pd
```

Figure 20 (a)

First, we imported the required libraries.

```
files = os.listdir()
files = files[1:]
```

Figure 20 (b)

Then we imported all the files present in the directory using the os module. We have removed the first file which was the python file itself and all other files were csv files.

```
X = list(range(len(files)))

for index, file in enumerate(files):
    f = open(file, 'r')
    filename = f.read().splitlines()

    for idx, i in enumerate(filename):
        filename[idx] = i.split(' ')

    X[index] = filename
```

Figure 20 (c)

Then we created a X matrix which stored our absorption and wavelength readings. Then, we used a for loop to iterate over every file and store its content in X. Remember every file is a milk sample reading who took from our spectrometer.

```
20
21   wavelengths = list(range(1801))
22
23   for idx, i in enumerate(X[0]):
24       wavelengths[idx] = i[0]
```

Figure 20 (d)

Then we separated the wavelengths from the X vector using a for loop.

```
25
26   absorption = [[i for i in range(1801)] for j in range(36)]
27
28   for idx, i in enumerate(X):
29       for index, j in enumerate(i):
30           absorption[idx][index] = X[idx][index][1]
```

Figure 20 (e)

Similarly, we separated the absorption readings from the X vector using the for loop.

```
31
32   wavelengthData = np.asarray(wavelengths)
33   absorptionData = np.asarray(absorption)
34   dataframe = pd.DataFrame(absorptionData)
35   dataframe.to_csv(r"C:/Users/SHAHEER/Desktop/MatlabData/absorptionData.csv")
36
37   with open("C:/Users/SHAHEER/Desktop/MatlabData/wavelengthData.csv", "w") as f:
38       writer = csv.writer(f)
39       writer.writerow(wavelengths)
```

Figure 20 (f)

Finally, we stored these vectors as numpy arrays and then exported them in csv format. Note that we have to pandas dataframe to export three-dimensional data in absorption matrix reliably.

Then, this data was imported into MATLAB for further processing.

```
1   clear all
2   close all
3   wavelengths = csvread("wavelengthData.csv");
4   absorptionFat = csvread("absorptionData.csv");
5   absorptionProtein = csvread("proteinData.csv");
```

Figure 21 (a)

A y vector was also created for labelling the output using a for loop.

```
6
7   outputFat = zeros(0);
8   outputProtein = zeros(0);
9   tempFat = [0.9, 1.7, 3.5];
10  tempProtein = [2.2, 2.7, 3.2];
11  for i=1:3
12      for j = 1:12
13          outputFat(end+1) = tempFat(i);
14          outputProtein(end+1) = tempProtein(i);
15      end
16  end
17
```

Figure 21 (b)

The raw absorption data taken from the spectrophotometer wasn't suitable for data analysis. A graph of Milk Pak (3.5 g fat, 2.7 g protein) as taken from spectrophotometer is shown as follows:

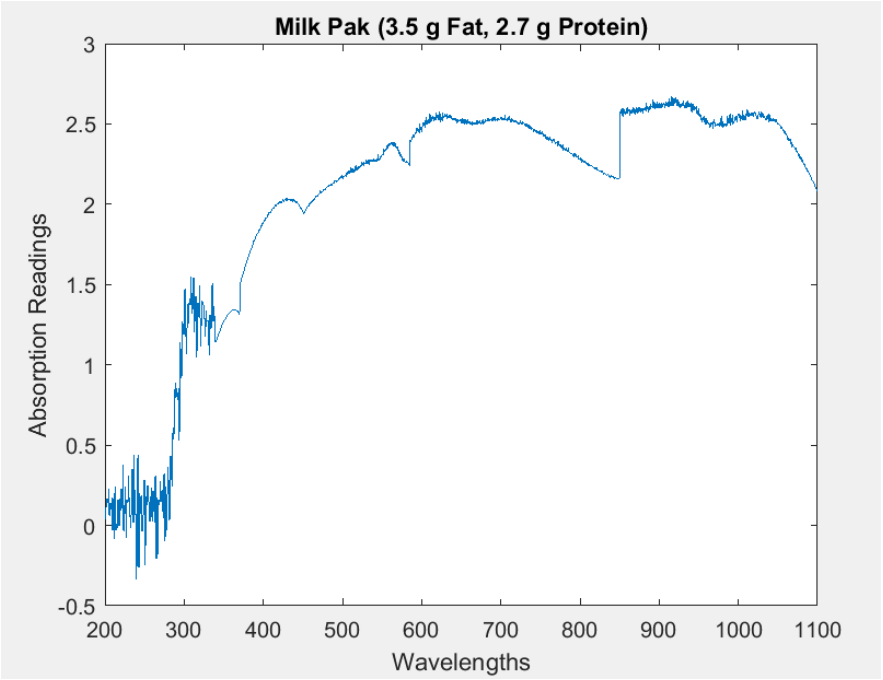

Figure 22: Raw Absorption Data (Milk Pak) From Spectrometer

As it can be seen, there is too much noise in the data which will deter the performance of our model. Thus, data was then smoothed using the MATLAB built-in function smoothdata().

```
21  for i = 1:length(absorptionFat(:, 1))
22      smoothedFat(i, :) = smoothdata(absorptionFat(i,:), 2);
23      smoothedProtein(i, :) = smoothdata(absorptionProtein(i,:), 2);
24  end
25
```

Figure 21 (c)

A graph of smoothed data of Milk Pak (3.5 g Fat, 2.7 g Protein) is shown below:

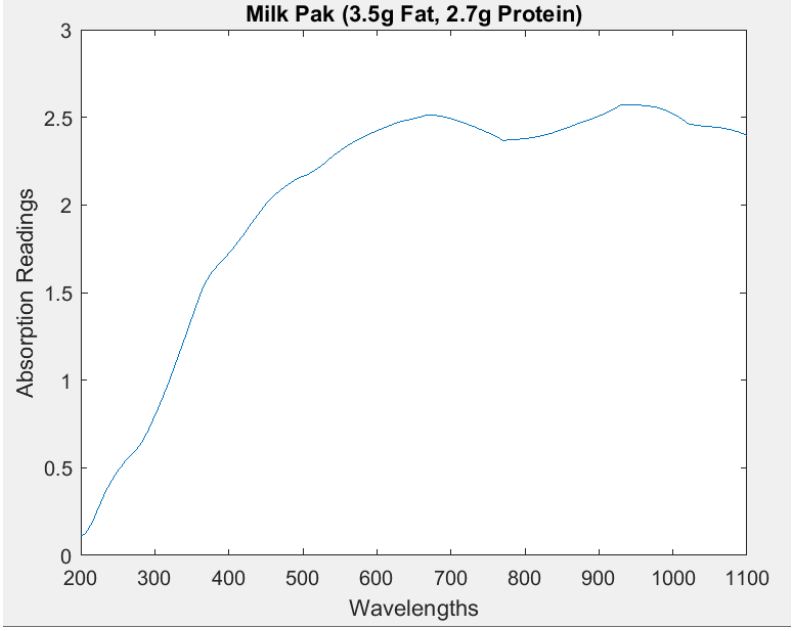

Figure 23: Graph after smoothing the data

After smoothing the data, it was converted to train and test data using the Kennard Stone Algorithm. The algorithm divides the data into train and test data by 80:20 ratio and in a very efficient way to help the model train better.

```
26    [ trainXData, testXData ] = kennardstonealgorithm(smoothedFat, 24);
27
28    sampleFat = smoothedFat(trainXData,:);
29    trainOutputFat = outputFat(trainXData);
30    testFat = smoothedFat(testXData, :);
31    testoutputFat = outputFat(testXData);
32
33    sampleProtein = smoothedProtein(trainXData,:);
34    trainOutputProtein = outputProtein(trainXData);
35    testProtein = smoothedProtein(testXData, :);
36    testoutputProtein = outputProtein(testXData);
```

Figure 21 (d)

Now that, our data is cleaned, we can further analyse it to identify patterns and train a model to predict fat and protein content. We have used PLS Toolbox build by Eigenvector Research for our data processing. The interface of PLS Toolbox is shown as follows:

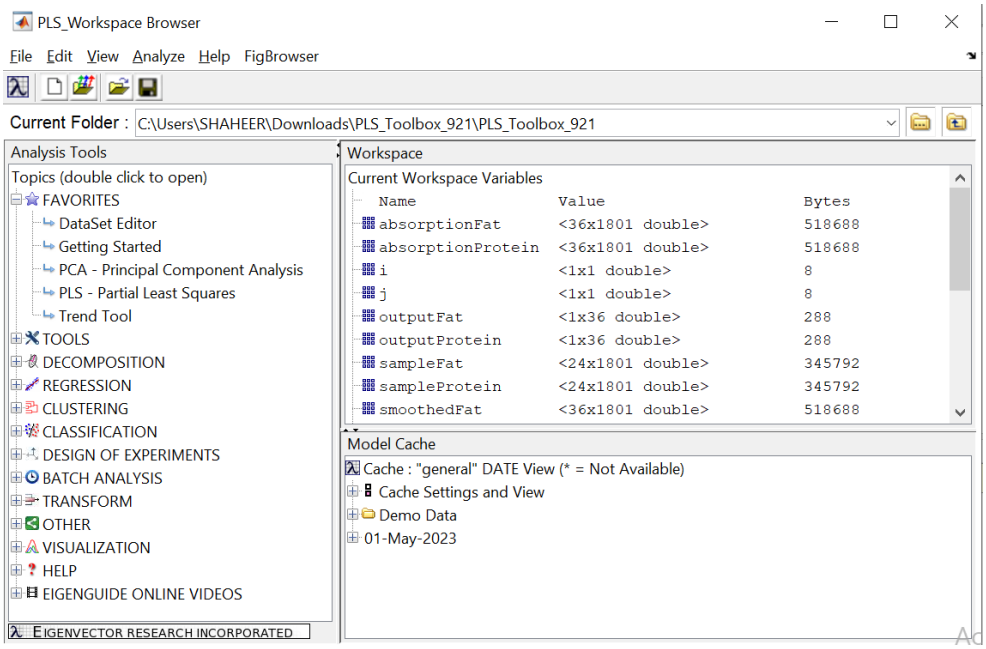

Figure 24: PLS Toolbox workspace interface

We will discuss further findings in the next chapter of Results.

LDR Method:

In this method, we have chosen an HW-483 LED module that emits red light at a wavelength of 650nm. The light falls on a light dependent resistor connected to a capacitor.

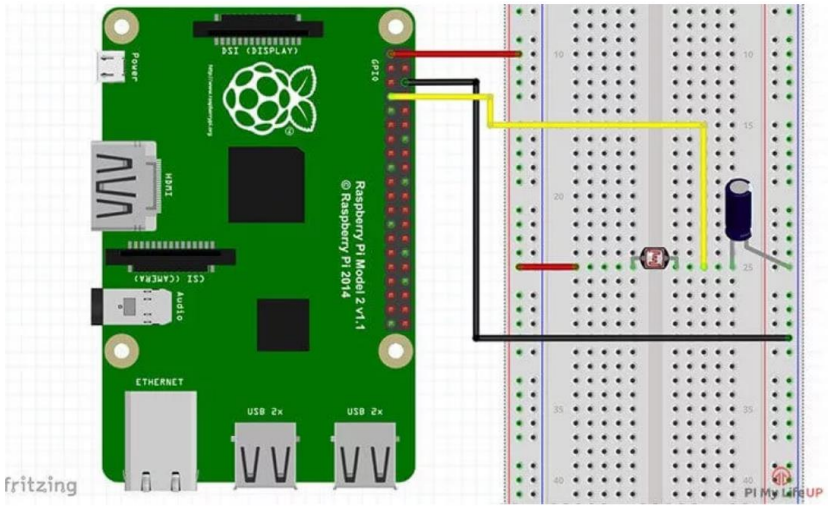

Figure 25: Circuit Diagram

The resistance of the LDR decreases with an increase in light intensity. By measuring the time it takes to charge the capacitor, we can measure the LDR resistance and thus measure the amount of light passing through our milk sample. The longer it takes to charge the capacitor, the higher the resistance of the LDR. The picture below shows the sealed box we use to test our samples.

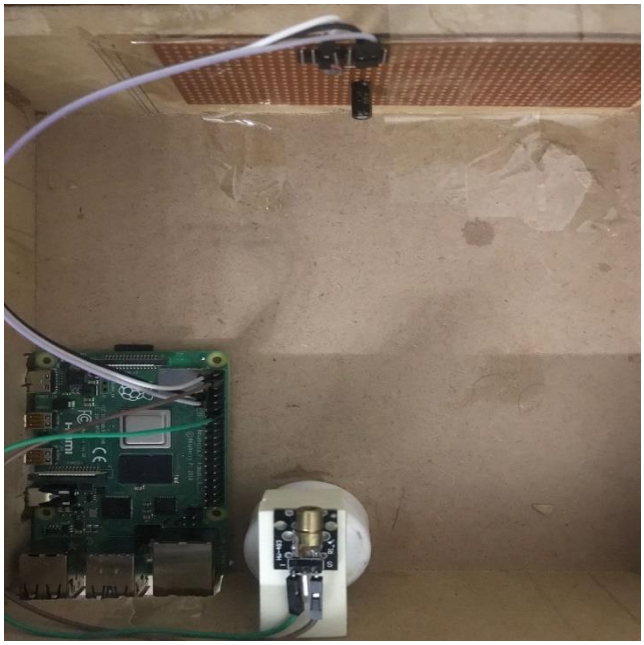

Figure 26: The LDR setup

We have chosen different categories of milk samples of different fat content ranging from 0.9 – 3.5 g. The milk samples belonged to the commercially available in the market and we used their fat readings to label our output. A summary of the milk samples is shown in Table 1. We had total nine milk samples, and we took 30 LDR readings from each milk sample. These readings were stored in csv format and then exported to the MATLAB. The MATLAB code to process this data is shown below:

```
clear all
close all

ldrReadings = (csvread("LDR.csv"));

output = zeros (0);
temp = [3.5, 1.7, 0.9];

for i=1:3
    for j = 1:90
        output(end+1) = temp(i);
    end
end
output = output';

[ trainXData, testXData ] = kennardstonealgorithm(ldrReadings, 180);

trainSample = ldrReadings(trainXData);
trainOutput = output(trainXData);
testSample = ldrReadings(testXData);
testOutput = output(testXData);
```

Figure 27: MATLAB Code to process LDR Readings

We have used the same method to process LDR readings as was discussed previously. Although here, the difference is that the ldrReadings is a one-dimensional vector instead of two-dimensional absorption readings in our previous part. A scatter plot of our LDR readings is shown below:

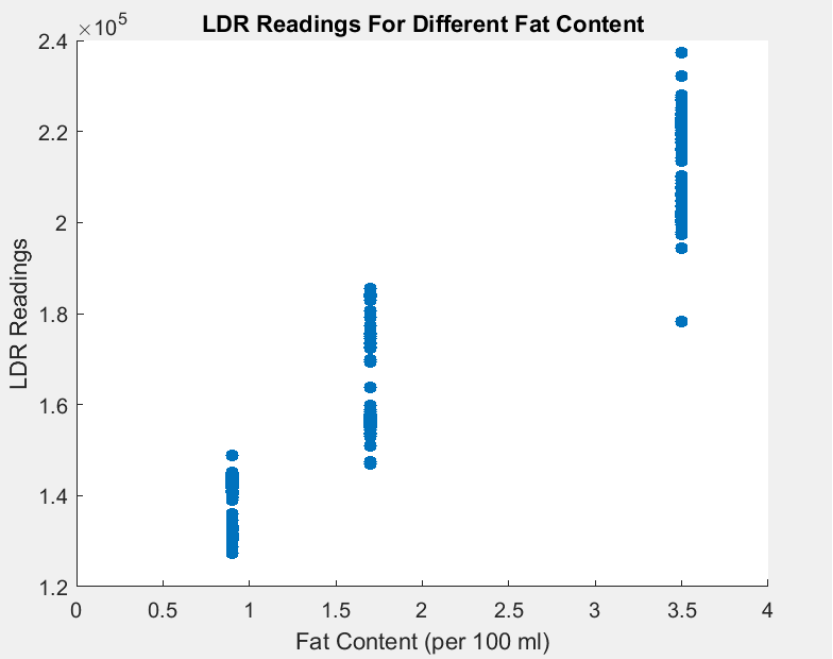

Figure 28: A scatter plot of LDR Readings for different fat content

The PLS Regression was performed in PLS Toolbox to build a model which can predict fat content in a milk sample. Further results will be shown in the Results sections.

Chapter 5

Results and Discussion

In this chapter, we will discuss the results of each part of our project.

Comparison Analysis:

As we have already discussed seven major milk quality parameters and six major techniques to measure them in detail in the previous chapter. We will summarize the whole discussion in the form of a table with their figure of merits. This table will serve to guide the new students who will start research in this domain in the future.

Five Methods and Five Important Characteristics of Milk to Review	LDR Method	Ultrasonic Method	Spectroscopic Method	Density Method	pH Method	Chemical Methods
Water Level	No	Yes (±3% from 0 to 70%)	No	Yes (1.034 changed to 1.01)	No	Yes (Freezing Point Osmometry)
Fats	Yes An addition of 1.7g fats leads to almost 150000 increase in LDR reading	Yes (±0.06% from 0.01 % to 25%)	Yes (0.376 g/100 ml)	Yes, But Not Much Accurate (1.034 changed to 1.027)	No	Yes (Gerber Method)
Proteins	Yes An addition of 0.6g proteins leads to almost 200000 increase in LDR reading	Yes (±0.15% from 2% to 7%)	Yes (0.655 g/100 ml)	No	No	Yes (Kjeldahl Method)
Detection of Baking Soda	No	Yes (±0.05 from 0 to 14)	No	Yes (1.034 changed to 1.013)	Yes (pH increases to 7.1, Pure milk has pH ranging from 6.6 to 6.9)	Yes (Rosalic Acid Test)
Detection of Salts	No	Yes (±0.05% from 0.4 to 4%)	Yes (R=0.86 and SEP=0.88)	Yes (1.034 changed to 0.99)	Yes (Voltage drop from 160mV to 120mV)	Yes (Silver Nitrate Reagant)
Detection of Urea	No	Yes	Yes (R=0.98 and SEP=0.78)	No	No	Yes (Paradimetyl amino)

Detection of Detergents	No	Yes (±0.05 from 3 to 14mS/cm)	Yes, But Not Much Accurate (R=0.58 and SEP=3.83)	Yes (1.034 changed to 0.98)	Yes (pH falls to 5.9)	Yes (Bromocresl Purple Sol)

Table 3: A summary of Comparative Analysis

Spectroscopic Method:

A PCA Score Analysis was carried out on smoothed data of absorption readings of fat content. Four PCs were required to accomplish the analysis. The number of PCs were automatically selected by the PLS Toolbox software. First two PCs were able capture 99.59% of data and remaining was of PC3 and PC4. A PCA Scores graphs is shown as follows:

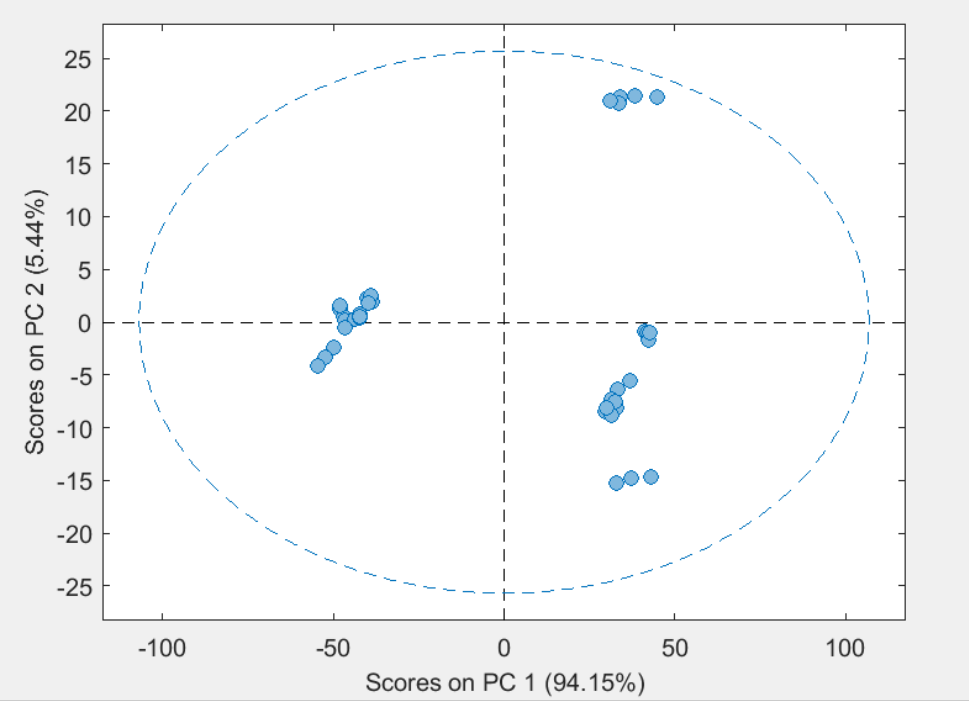

Figure 29: PCA Scores Graph of Smoothed Spectroscopic Data of Different Fat Content

As it can be seen in the figure that the PCA analysis can successfully segment the three categories of fat content.

Then, in the next step, we applied PLS regression with 8 LVs on trained fat data. Very good results were obtained with R^2 = 0.982, RMSEC = 0.147, RMSECV = 0.359. The PLS Regression model is shown as follows:

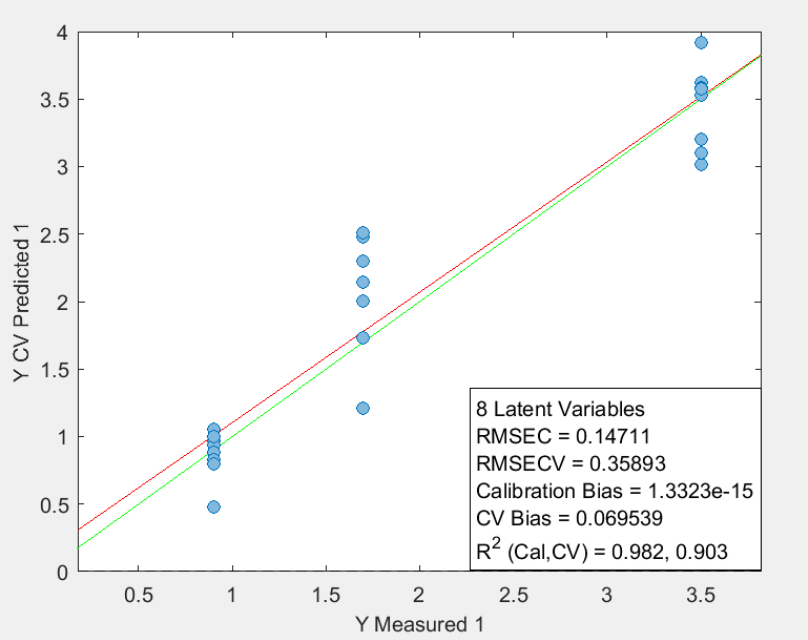

Figure 30: PLS Regression Model for Absorption readings for different fat content

After the model was trained, model was applied to the test data. The figure of merits of this PLS model is shown below:

RMSEC	0.147
RMSECV	0.359
RMSEP	0.136
R^2 Cal	0.982
R^2 CV	0.903
R^2 Pred	0.983

Table 4: Figure of Merits of PLS Regression Model for Fat Content Prediction.

Limit of Detection (LOD) of this model is found to be 0.436g/100ml.

Next, we done the same with samples of different protein content. First, a PCA score analysis was done on smoothed data of different protein content. Three PCs were required to accomplish the analysis. The first two PCs captured about 99.96% of data and the remaining by PC3. A PCA Scores graph is shown as follows:

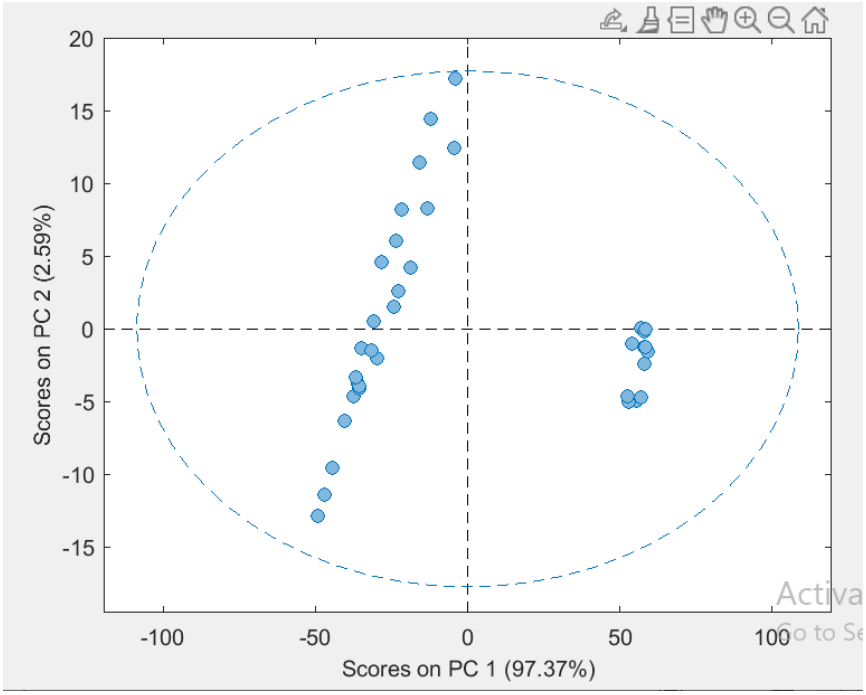

Figure 31: PCA Score Analysis on smoothed data of different protein content

As it can be seen from the graph, that PCA Score analysis only segmented the data into two categories. The second category is quite distinct, but it failed to separate the first category.

Next, we applied the PLS regression model on train protein data. The software automatically chose 5 LVs and trained the model. The PLS regression graph for protein content prediction is shown below:

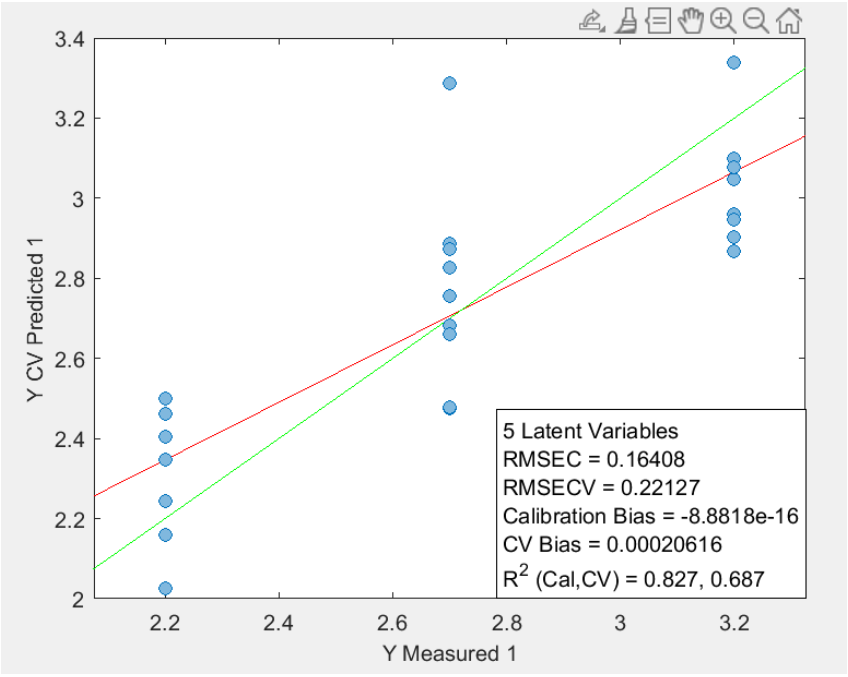

Figure 32: PLS Regression Model for Protein Content Prediction

As it can be seen, the results are not as good as fat content prediction model. But still, they are satisfactory with R^2 = 0.807, RMSEC = 0.164, and RMSECV = 0.221.

Then we applied our model on test protein data. The results weren't satisfactory and although the error is not much high, still the model failed to find relatability among the data given. The results of PLS regression model for protein detection is as follows:

RMSEC	0.164
RMSECV	0.221
RMSEP	0.390
R^2 Cal	0.827
R^2 CV	0.687
R^2 Pred	0.399

Table 5: Results of PLS Regression Model for Protein Prediction

So, our model was able to predict the fat content with good results, but it failed to predict accurate results for protein. There can be a lot of reasons. The spectroscopic data we got for different protein content have a lot of variability. This is due to the reason that the protein micelles are very small in size as compared to fat globules. Further research is required to predict protein content accurately.

Although spectroscopic method is established as a technique to measure protein content, but not all spectrometers can achieve it. For example, [37] discussed in their research paper than SCiO spectrometer was able to predict protein content effectively but NeoSpectra Spectrometer couldn't predict it and had an R^2 close to 0. The spectrometer we used was primarily UV-VIS Spectrometer thus other spectrometers especially made for NIR (and beyond) range should be explored.

LDR Method:

As we have already prepared our data of LDR readings in the methodology chapter, now we will apply PLS regression on it to predict the fat content. A single LV was selected by the software automatically to train the PLS regression model. The regression model for fat prediction is shown as follows:

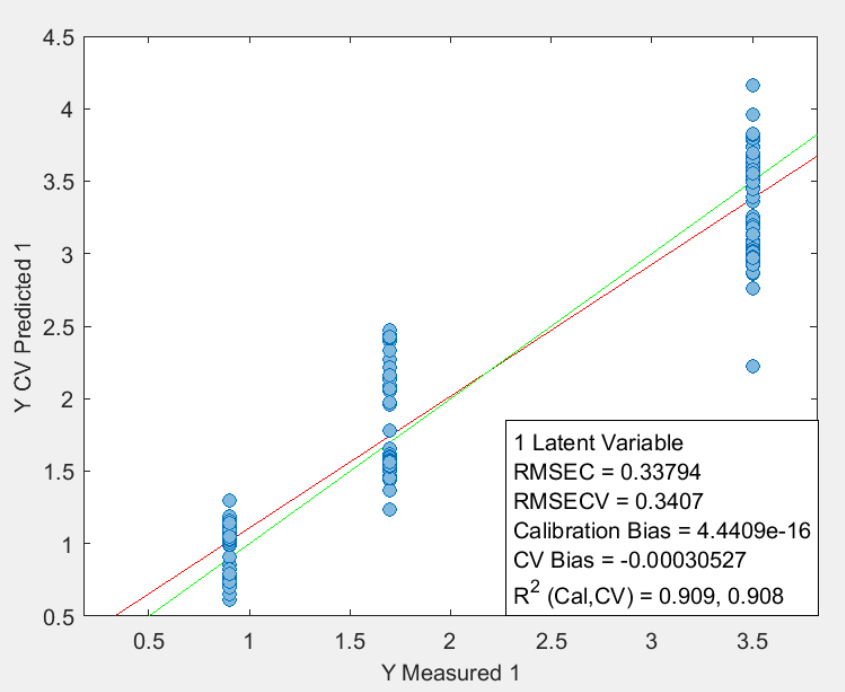

Figure 33: PLS Regression for Fat Prediction with LDR

As it can be seen, good results are shown by PLS Regression model during the training. Although the model is not as good as model trained from spectroscopic method in the previous part. But still this model is satisfactory and can be used for fat prediction. In the next part, we applied this model to the test data. The results shown by PLS regression model are as follows:

RMSEC	0.338
RMSECV	0.341
RMSEP	0.244
R^2 Cal	0.909
R^2 CV	0.908
R^2 Pred	0.932

Table 6: Results of PLS Regression Model for fat prediction with LDR

Our model can predict 75% of the data accurately. Overall, the model is satisfactory and can be improved through increasing the data samples.

Chapter 6

Failures, Future Recommendations, and Conclusion

We have carried out this project in almost 6 months. We have experimented with a lot of things and faced many challenges along the way. Now, we will discuss some of our failures. We have tried to build an ultrasonic based milk quality analyser. We used ultrasonic receiver and transmitter. The circuit used for this purpose is as follows:

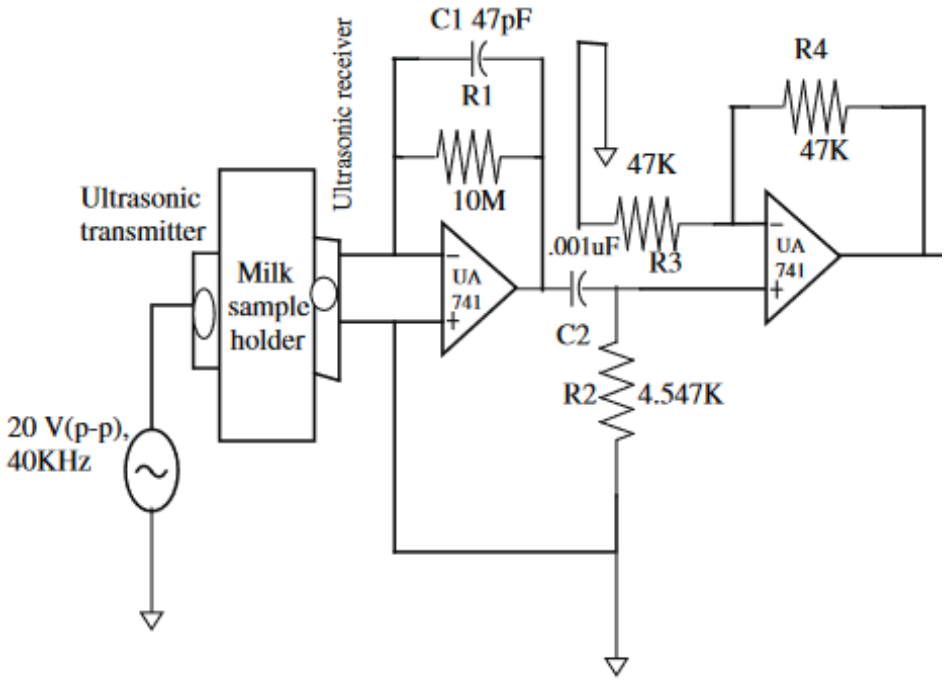

Figure 34: Ultrasonic Milk Analyser Circuit

We implemented this circuit, but the results were not satisfactory. The biggest problem we faced with this circuit was that the ultrasonic receiver and transmitter required proper isolation which was quite expensive. So, we had to leave this implementation.

We also tried to build our own spectrometer, however the cost of building it was around 60-70k PKR. We used cheap alternatives on Daraz to build it in 4K PKR. But the resolution of our device was very poor, and the device wasn't suitable for analysis.

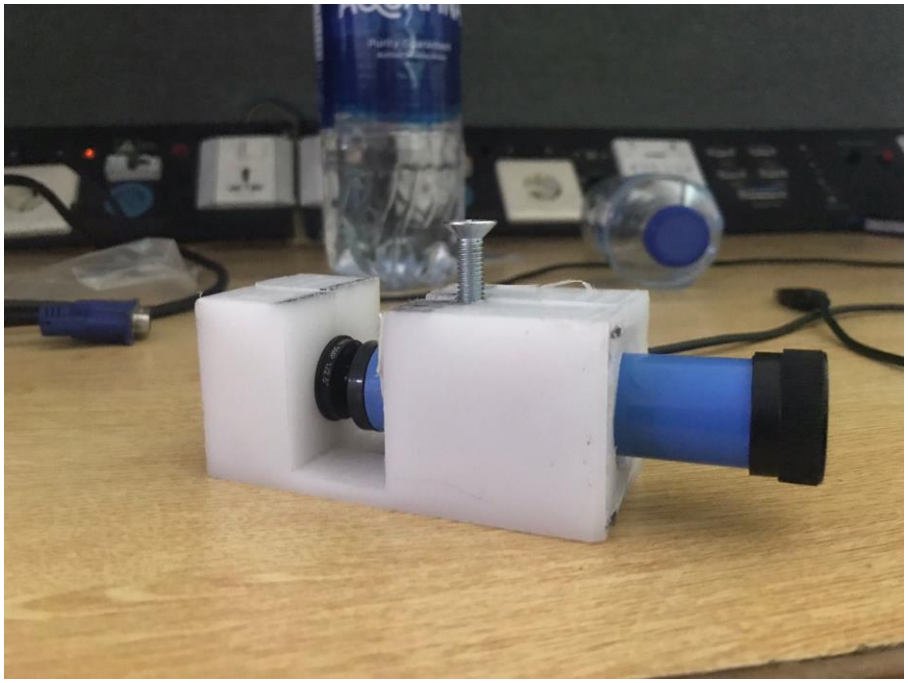

Figure 35: A small portable spectrometer

With these failures, we have spent quite a lot of time exploring different methods to measure milk quality.

Future Recommendations:

As we have already told in the results section, that our spectrometer wasn't able to measure protein content effectively. Other spectrometers can be used to collect similar data and see if the models built on their data can predict protein content. Other spectrums (like MIR, and beyond) can also be explored and see if they respond to changes in protein effectively. Other milk constituents like carbohydrates (lactose), calcium and other vitamins should also be quantified by similar process. We have already mentioned the major milk adulterants used in our country and their respective research papers are also mentioned. Spectroscopy can be used to detect detergents, salts, shampoo and oil in milk. Other methods can also be explored like ultrasonic, chromatography and dielectric techniques.

The milk samples we have used were commercially available milk samples like Milk Pak, Olpers etcs., and the fat and protein readings on their labels were used as output labels. However, milk from local vendors can also be used for analysis. For this purpose, we need to have a reference method to measure fat and protein prior to measuring them on spectroscope or LDR. Reference methods are usually chemical methods like Gerber method for Fat and Kjedahl method for protein and they must be performed in a laboratory under the supervision of a person with specialized knowledge.

In summary, there is a huge gap of research in this field in our country which needs to be filled. For this purpose, funding should be made from concerned stakeholders so that the researchers can deep dive into these techniques and find the best method to measure milk constituents efficiently.

Conclusion:

As milk is one of the highly consumed products in our country, it is usually adulterated which leads to impure milk and comprised health of consumers. Our project aims to solve this problem by building a Milk Quality Analyzer which can predict the fat and protein content instantaneously. We divided our project into three parts.

In the first part, we made a comparative analysis of major milk quality parameters with major techniques to measure milk quality. We have identified seven milk quality parameters and six major techniques to measure them. We have designed a table which provided information about which techniques are suitable for which milk quality parameters along with their figure of merits.

In the second part, we collected milk samples of different fat and protein content and analysed them on a spectrophotometer. After obtaining their spectra, we cleaned them and smoothed the noise in MATLAB. Then we applied PCA Score Analysis to identify the patterns among the spectra. A PLS Regression model was built from Spectra and then predicted the fat and protein content. The model worked really well for fat content prediction, but it couldn't produce satisfactory results for protein prediction.

In the final part, we made a portable device which used an LDR to measure fat content. We used similar milk samples and took LDR readings. Those readings were used to build a linear regression model which predicted the fat content with good accuracy. This device can be used for fat prediction but still it needs a lot of improvement.

List of References

[1] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7492176/

[2] https://www.sciencedirect.com/science/article/abs/pii/S0143816605000655?via%3Dihub

[3] https://www.sciencedirect.com/science/article/abs/pii/S0924203111000907?via%3Dihub

[4] https://hal.science/hal-00868673/file/art_3A10.1051_2Fdst_2F2010034.pdf

[5] http://www.eemj.icpm.tuiasi.ro/pdfs/vol7/no6/33_Semaghiul%20Birghila.pdf

[6] https://www.cambridge.org/core/journals/journal-of-dairy-research/article/abs/differential-titration-of-plasmin-and-plasminogen-in-milk-using-sandwich-elisa-with-monoclonal-antibodies/12620838A8A5C9DD3C4E0F5994C2908B

[7] https://link.springer.com/chapter/10.1007/978-3-319-45776-5_6

[8] https://www.journalofdairyscience.org/article/S0022-0302(70)86197-5/pdf

[9] https://www.sciencedirect.com/science/article/abs/pii/0308814694901600

[10] https://www.sciencedirect.com/science/article/abs/pii/S000326709800823X?via%3Dihub

[11] https://www.journalofdairyscience.org/article/S0022-0302(05)72709-0/fulltext

[12] https://www.sciencedirect.com/science/article/abs/pii/S0143816605000655?via%3Dihub

[13] https://www.sciencedirect.com/science/article/abs/pii/S0308814612002658?via%3Dihub

[14] https://www.sciencedirect.com/science/article/abs/pii/S0039914013010485?via%3Dihub

[15] https://www.sciencedirect.com/science/article/abs/pii/S1068520018305078?via%3Dihub

[16] https://pubs.acs.org/doi/10.1021/ac000469c

[17] https://journals.sagepub.com/doi/abs/10.1366/0003702021955150?journalCode=aspc

[18] https://linkinghub.elsevier.com/retrieve/pii/S0022030204734050

[19] https://eprints.lib.hokudai.ac.jp/dspace/bitstream/2115/20500/1/NIRforMilkQualityMonitoring.pdf

[20] https://www.sciencedirect.com/science/article/abs/pii/S0168169908000100?via%3Dihub

[21] https://www.journalofdairyscience.org/article/S0022-0302(11)00558-3/fulltext

[22] https://www.academia.edu/download/75273797/Medicin_Del_Porcentaje_De_Grasa_En_Leche20211127-1995-6d2z6y.pdf

[23] https://www.journalofdairyscience.org/article/S0022-0302(12)00650-9/fulltext

[24] https://link.springer.com/content/pdf/10.1007/s40242-013-2191-y.pdf

[25] https://www.journalofdairyscience.org/article/S0022-0302(15)00513-5/fulltext

[26] https://journals.sagepub.com/doi/abs/10.1177/0142331215593994?journalCode=tima

[27] https://www.journalofdairyscience.org/article/S0022-0302(16)00040-0/fulltext

[28] https://www.sciencedirect.com/science/article/abs/pii/S0260877415300182?via%3Dihub

[29] https://gvpress.com/journals/IJMUE/vol11_no10/12.pdf

[30] https://www.researchgate.net/profile/Abba-Priev/publication/5323767_Acoustic_radiation_forces_in_monitoring_of_milk_composition/links/55d9d84108aec156b9ae1c46/Acoustic-radiation-forces-in-monitoring-of-milk-composition.pdf

[31] https://ieeexplore.ieee.org/abstract/document/508696

[32] Biswas, Karabi, Siddhartha Sen, and Pranab Kumar Dutta. "A constant phase element sensor for monitoring microbial growth." *Sensors and Actuators B: Chemical* 119.1 (2006): 186-191.

[33] Das, Siuli, et al. "A low cost instrumentation system to analyze different types of milk adulteration." *ISA transactions* 56 (2015): 268-275.

[34] https://www.researchgate.net/publication/285568587_Detection_of_adulterants_in_milk_using_near_infrared_spectroscopy

[35] https://www.velp.com/public/file/nprotein-determination-in-milk-kjeldahl-method-206359.pdf

[36] https://bmcchem.biomedcentral.com/articles/10.1186/1752-153X-2-6