Artificial Intelligence Methods for Raman Spectroscopy

Project submitted to the

SRM University – AP, Andhra Pradesh

For the partial fulfilment of the requirements to award the degree of

Bachelor of Technology in

Computer Science and Engineering

Submitted by

AP20110010023 – Raghavendra Narahari

AP20110010038 – Yuva Kishore kalyanam

AP20110010016 – Rohan Gokul

Under the Guidance of

Dr Raja Pandiyan J P

Certificate

Date: 13-Dec-22

This is to certify that the work present in this Project entitled "**Artificial Intelligence Methods for Raman Spectroscopy**" has been carried out by **Raghavendra Narahari, Yuva Kishore, Rohan Gokul** under my/our supervision. The work is genuine, original, and suitable for submission to the SRM University – AP for the award of **Bachelor of Technology in School of Engineering and Sciences**.

Supervisor

(Signature)

Dr. Raja Pandiyan J P

Assistant Professor,

Department of Chemistry,

SRM University – AP, Andhra Pradesh

Acknowledgements

We would like to extend our sincere gratitude and appreciation to everyone who gave us the chance to finish this report. We notably thank Prof. Dr. Rajapandiyan J P, our supervisor, whose contribution of energising recommendations helped us to plan our effort, particularly in preparing this report. On the other side, we'd also want to express our sincere gratitude to the SRM AP Library Department, which gave us permission to utilise all required materials, provided us with the direction we needed to finish the project, and offered us access to prior research papers and relevant books.

Table of Contents

Certificate……3

Acknowledgment……..5

Table of contents……7

Abstract……..9

List of figures……11

List of tables……13

Introduction……15

Methodology……22

Concluding remarks……33

Future works……35

References ……37

Abstract

Raman spectroscopy named after Indian physicist C.V.Raman is a spectroscopic technique typically used to determine vibrational modes of molecules, although rotational and other low-frequency modes of systems may also be observed. Here we have to see the compound via an app that will process by giving the coordinates which is provided by the Raman Machine .what if we make it easy?

So we solved this issue by using coding and some functions by simply inserting the coordinates into the code as a file and getting the graph as output and we check the graph is the given known sample by putting the peak values of the sample into the code and if we get the sample correctly then we came to know that the given sample is correct.

By this we can make the work easier for researchers who are working on this area by simply inserting the coordinates.

LIST OF FIGURES:

Figure (1) ……A modern Raman Microscope System

Figure (2) ……Methanol Raman Spectroscopy

Figure (3) ……Ethanol Raman Spectroscopy

Figure (4) ……Rhodamine 6G

Figure (5) ……Code and Graph of Rhodamine 6G

Figure (6) ……Code and output of Rhodamine 6G

Figure (7) ……Graph of Rhodamine 6G

Figure (8) ……Crystal Violet

Figure (9) ……Code and Graph of Crystal Violet

Figure (10) ……Code and Output of Crystal Violet

Figure (11) ……Graph of Crystal violet.

List of Tables:

1……Methanol peaks table

2……Ethanol peaks table

3……Rhodamine6G peaks table

4……..Crystal Violet peaks table.

INTRODUCTION

Our lives are being shaped in numerous ways by machine learning. Machine learning offers an unmatched possibility in the analytical sciences to extract data from difficult or large datasets in chromatography, mass spectrometry, NMR, and spectroscopy, among other analytical techniques. This is particularly true for Raman and surface-enhanced Raman scattering (SERS) methods, which acquire the vibrational spectra of intricate chemical mixtures as sizable datasets for chemical system research or imaging. Raman spectroscopy with machine learning algorithms can detect low-concentration materials with a very high degree of precision.

Raman scattering, also known as inelastic photon scattering, is the basis of Raman spectroscopy. Although X-rays can also be employed, monochromatic light is often produced by lasers in the visible, near infrared, or near ultraviolet spectrum. The energy of the laser photons is pushed up or down as a result of the laser light's interactions with phonons, molecular vibrations, or other excitations in the system. The energy shift reveals details about the system's vibrational modes. Usually, infrared spectroscopy provides comparable but additional data.

Raman Spectroscopy is a non-destructive chemical analysis technique which provides detailed information about chemical structure, phase and polymorphy, crystallinity and molecular interactions. It is based upon the interaction of light with the chemical bonds within a material.

Raman is a light scattering technique, whereby a molecule scatters incident light from a high intensity laser light source. Most of the scattered light is at the same wavelength as the laser source and does not provide useful information – this is called Rayleigh Scatter. However a small amount of light (typically 0.0000001%) is scattered at different wavelengths, which depend on the chemical structure of the analyte – this is called Raman Scatter.

A Raman spectrum features a number of peaks, showing the intensity and wavelength position of the Raman scattered light. Each peak corresponds to a specific molecular bond vibration, including individual bonds such as C-C, C=C, N-O, C-H etc., and groups of bonds such as benzene ring breathing mode, polymer chain vibrations, lattice modes, etc.

Here, we done this project by inserting the x and y coordinate of a compound into our code and checking the given output is matching with the compound or not by checking the bond peaks of that compound.

Information provided by Raman spectroscopy:

Raman spectroscopy probes the chemical structure of a material and provides information about:

->Chemical structure and identity
->Phase and polymorphism
->Intrinsic stress/strain
->Contamination and impurity

Typically a Raman spectrum is a distinct chemical fingerprint for a particular molecule or material, and can be used to very quickly identify the material, or distinguish it from others. Raman spectral libraries are often used for identification of a material based on its Raman spectrum – libraries containing thousands of spectra are rapidly searched to find a match with the spectrum of the analyte.

Raman is used for microscopic analysis:

Raman spectroscopy can be used for microscopic analysis, with a spatial resolution in the order of 0.5-1 µm. Such analysis is possible using a Raman microscope.

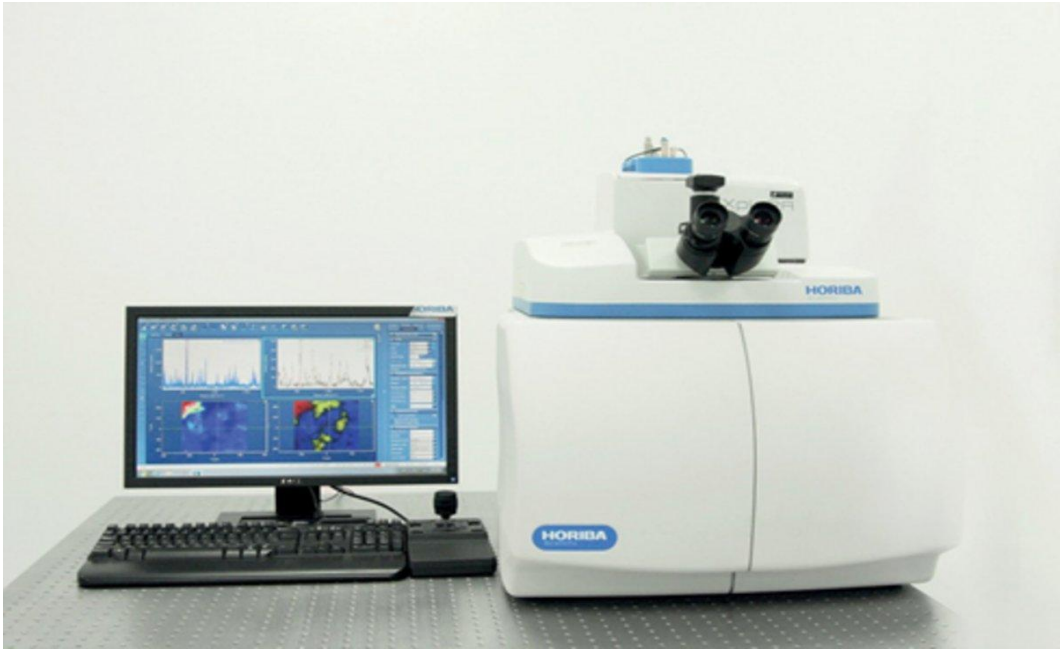


Figure 1: A modern Raman microscope system

A Raman microscope couples a Raman spectrometer to a standard optical microscope, allowing high magnification visualization of a sample and Raman analysis with a microscopic laser spot. Raman micro-analysis is easy: simply place the sample under the microscope, focus, and make a measurement.

Methanol Raman spectroscopy

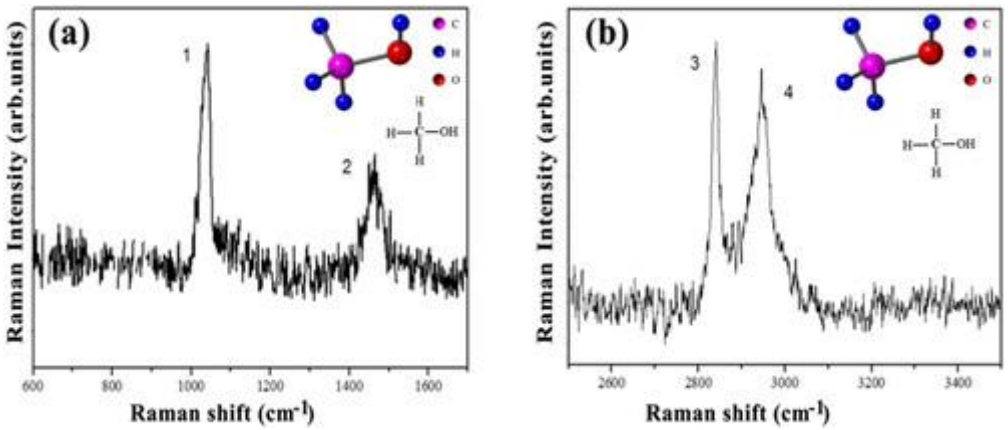


Figure 2(a) 2(b)

Figure 2(a) and 2(b) show the Raman spectra of methanol, which has four distinct characteristic peaks, and the molecular formula of methanol is CH3OH. The four main characteristic peaks correspond to methanol molecular vibrations. The wave number positions of the four peaks of

methanol are shown in Table I. Peak 1: 1035.9 cm−1 is caused by the stretching vibration of C–O outside the methanol plane, and it is generally considered that the C–O stretching vibration wave number of the monohydric alcohol is from 1000 cm−1 to 1075 cm−1; peak 2: 1461.9 cm−1 is the CH3 antisymmetric deformation vibration because the CH3 deformation vibration frequency is about 1450 cm−1;10 and the characteristic peaks of 3 (2840.2 cm−1) and 4 (2949.7 cm−1) are generated by CH3 symmetrical and asymmetrical stretching vibrations, respectively; peaks 3 and 4 are very intense peaks. In this way, the characteristic vibration peak is a response to the structural characteristics of methanol (CH3OH), that is, the fingerprint vibration peak identifying methanol.

TABLE FOR METHANOL:

	Raman shift	Methanol vibration
Numbers	(cm^{-1})	modes
1	1035.9	C–O stretching vibration
		modes ν^{CO}
2	1461.9	CH_3 anti-symmetric deformation
		vibration modes ν^{CH_3}
3	2840.2	CH_3 symmetric stretching vibration
		modes ν^{CH}
4	2949.7	CH_3 asymmetric stretching vibration
		modes ν^{CH}

Table(1)

Ethanol Raman spectroscopy

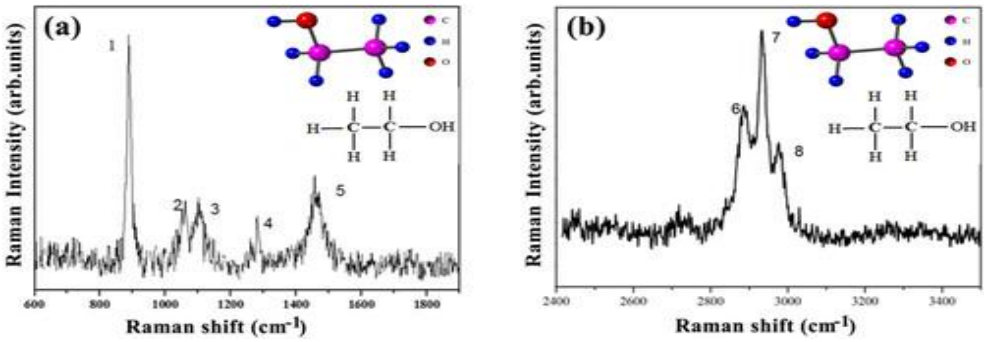


Figure 3(a) 3(b)

The Raman spectrum of ethanol is shown in Figure3(a) and 3(b), from which it can be seen that the ethanol Raman spectrum has eight typical characteristic peaks. The wave number positions of the eight peaks of ethanol are shown in Table-2. The molecular characteristic of ethanol is CH_3CH_2OH, and the eight characteristic peaks are structural responses of ethanol molecules. Peak 1: 888.8 cm^{-1} is the CCO skeleton symmetric stretching vibration, which is obviously unique to ethanol; peak 2: 1054.7 cm^{-1} is CO scaling; peak 3: 1104 cm^{-1} is the CCO skeleton stretching vibration; peak 4: 1287.3 cm^{-1} is ethanol molecule CH_2 deformation; peak 5: 1462.7 cm^{-1} is CH_3 antisymmetric deformation; peak 6: 2887.8 cm^{-1} is superposition of CH_3 and CH_2 symmetric stretching; peak 7: 2934.0 cm^{-1} is scaling for asymmetric CH_2; peak 8: 2975.0 cm^{-1} is scaling for asymmetric CH_3. In summary, these characteristic vibration peaks are the response of the alcohol CH_3CH_2OH molecular structure characteristics.

TABLE FOR ETHANOL:

	Raman shift	Ethanol vibration
Numbers	(cm^{-1})	modes
1	888.8	CCO symmetric stretching vibration
		modes ν^{CC}
2	1054.7	C–O scaling
		modes ν^{CO}
3	1104	CCO skeleton stretching
		vibration modes ν^{CH_3}
4	1287.3	CH_2 deformation vibration
		modes ν^{CH_2}
5	1462.7	CH_3 anti-symmetric deformation
		vibration ν^{CH_3}
6	2887.8	Superposition of CH_3 and
		CH_2 symmetric stretching ν^{CH}
7	2934.0	CH_2 asymmetric stretching
		vibration modes ν^{CH_2}
8	2975.0	CH_3 asymmetric stretching
		vibration modes ν^{CH_3}

Table(2)

METHODOLOGY:

Rhodamine 6G:

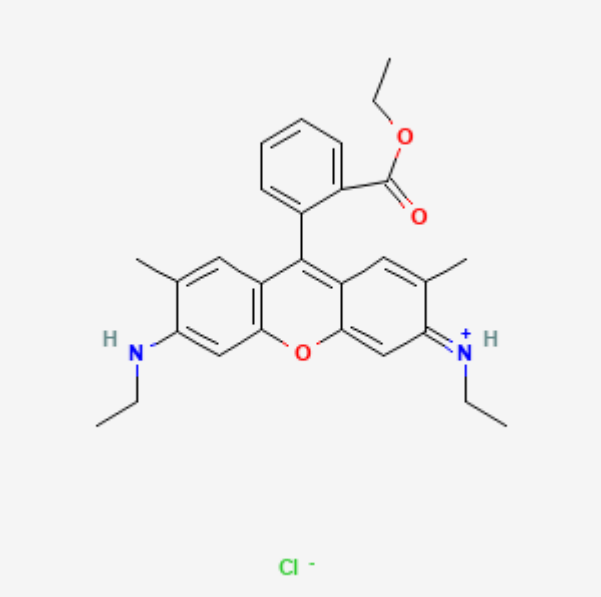
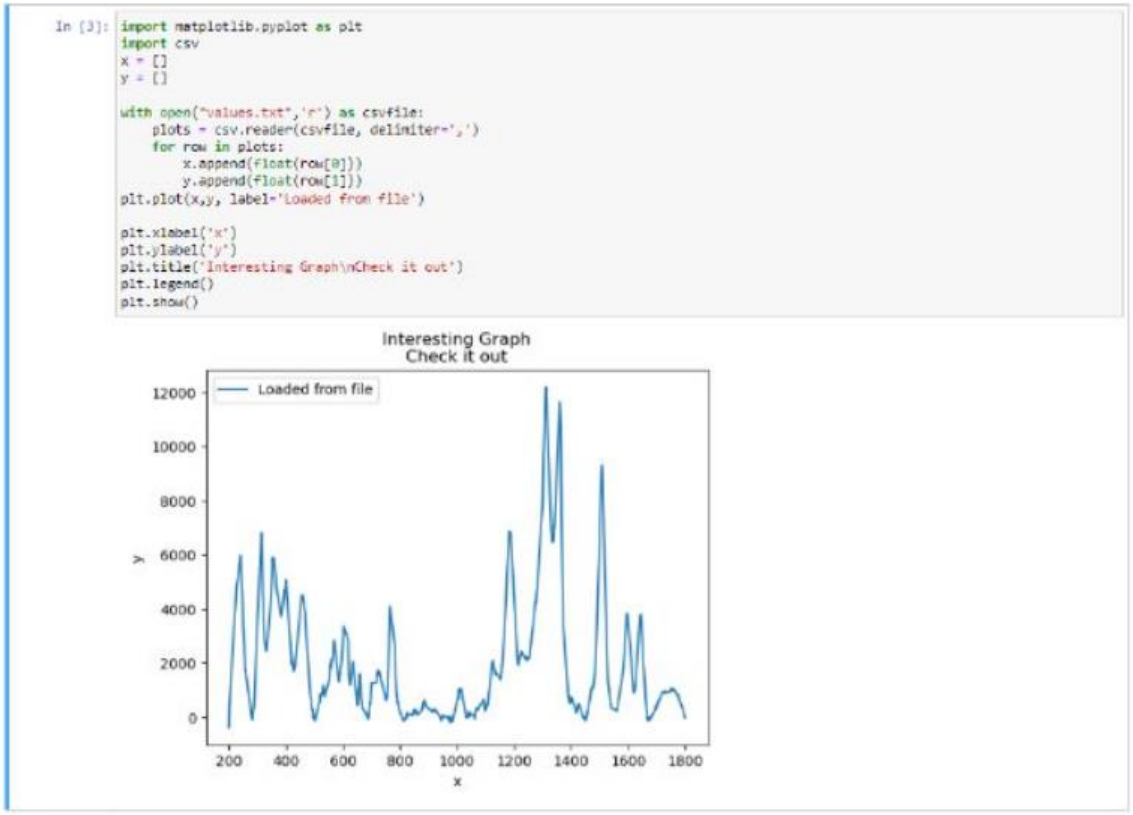


figure (4)

Rhodamine 6G has been used extensively as a model dye to probe the nature of the SERRS effect. It is an extremely strong fluorophore when excited by visible radiation. Hence normal Raman is not observed except with near-infrared excitation. However, when SERRS is used the dye adsorbs strongly to the roughened metal surface and consequently this strong fluorescence is quenched and an extremely strong, enhanced Raman signal is observed. Figure 3 illustrates the resonance Raman and SERRS spectra collected from Rhodamine 6G. Attomolar levels (10^{-18} M) of detection have been reported for this system, which is approaching single molecule detection. The fluorescence-quenching properties of surface enhancement coupled with the additional sensitivity obtained from SERRS have been exploited by several researchers. Rhodamine 6G adsorbs very effectively on the roughened silver surface. However, the detection of single adsorbates of dopamine or phthalazine on colloidal clusters, with a limit of detection at picogram levels, illustrates that ultrasensitivity of this technique for other adsorbates is possible.

We use to find this sample in Raman Spectrometer—

First we take a sample of this Rhodamine 6G and put this sample under the Raman Machine and get the readings/coordinates from the the machine and by getting those coordinates we simply put those in our code and we will get the graph……..



Figure(5)

X axis – Raman shift (cm-1) and y axis Raman intensity

Here we are going to put those coordinates in a file as ‘values’ and place that file in the code and by using some functions we use to get the graph as shown in figure(5).

And we have a task to check that the given graph representing the given known sample Rhodamine 6G or not by simply putting the peak values of Rhodamine 6G…as follows…….

Peak Position	Tentative Peak Assignment
608 cm-1	C-C-C ring in plane bending
769 cm-1	C-H out of plane bending
1127 cm-1	C-H in plane bending
1183 cm-1	C-H in plane bending
1312 cm-1	Aromatic C-C stretching
1360 cm-1	Aromatic C-C stretching
1508 cm-1	Aromatic C-C stretching
1599 cm-1	Aromatic C-C stretching
1643 cm-1	Aromatic C-C stretching

Table(3)

And the code and output is……

```
In [8]: lst = [0,0,0,0,0,0,0,0,0]
        for i in x:
            if(i in range(600,615)):
                lst[0] = 1
            elif(i in range(760,775)):
                lst[0] = 1
            elif(i in range(1120,1135)):
                lst[0] = 1
            elif(i in range(1175,1190)):
                lst[0] = 1
            elif(i in range(1305,1320)):
                lst[0] = 1
            elif(i in range(1355,1365)):
                lst[0] = 1
            elif(i in range(1500,1515)):
                lst[0] = 1
            elif(i in range(1595,1603)):
                lst[0] = 1
            elif(i in range(1640,1650)):
                lst[0] = 1
        ans = 1
        for i in range(0,len(lst)):
            if(lst[i] == 1):
                ans = 0
                break
        if(ans):
            print("rhodamine - 6g")

rhodamine - 6g
```

Figure(6)

Here we came to know that the given sample is Rhodamine 6G..

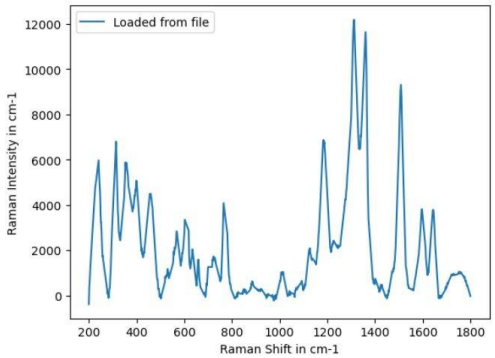
CODE:

```
import matplotlib.pyplot as plt
import csv
x = []
y = []

with open("values.txt",'r') as csvfile:
    plots = csv.reader(csvfile, delimiter=',')
    for row in plots:
        x.append(float(row[0]))
        y.append(float(row[1]))
plt.plot(x,y, label='Loaded from file')

plt.xlabel('x')
plt.ylabel('y')
plt.title('Interesting Graph\nCheck it out')
plt.legend()
plt.show()
```

Output:



Figure(7)

X axis – Raman shift (cm-1) and y axis Raman intensity

And then

CODE:

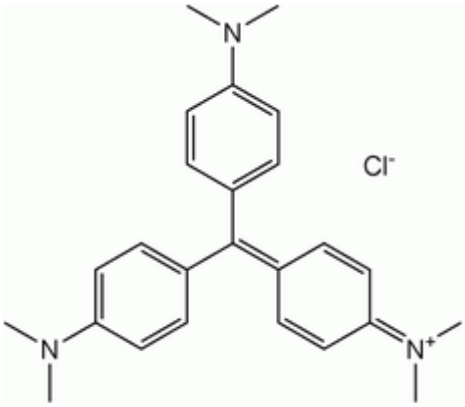
```
lst = [0,0,0,0,0,0,0,0,0]
for i in x:
    if(i in range(600,615)):
        lst[0] = 1
    elif(i in range(760,775)):
        lst[0] = 1
    elif(i in range(1120,1135)):
        lst[0] = 1
    elif(i in range(1175,1190)):
        lst[0] = 1
    elif(i in range(1305,1320)):
        lst[0] = 1
    elif(i in range(1355,1365)):
        lst[0] = 1
    elif(i in range(1500,1515)):
        lst[0] = 1
    elif(i in range(1595,1603)):
        lst[0] = 1
    elif(i in range(1640,1650)):
        lst[0] = 1
ans = 1
for i in range(0,len(lst)):
    if(lst[i] == 1):
        ans = 0
```

```
        break
if(ans):
    print("rhodamine - 6g")
```

OUTPUT:

Rhodamine 6G.

CRYSTAL VIOLET:



Figure(8)

a monochloride salt of the crystal violet cation that is an organic chloride. Although it was effective against some Gram-positive bacteria, particularly Staphylococcus species, and some pathogenic fungi, including Candida species, it was no longer used in creams for the topical treatment of bacterial and fungal diseases after reports that it was carcinogenic to animals. Along with being employed as a histology stain, it has also been used to dye wood, silk, and paper.

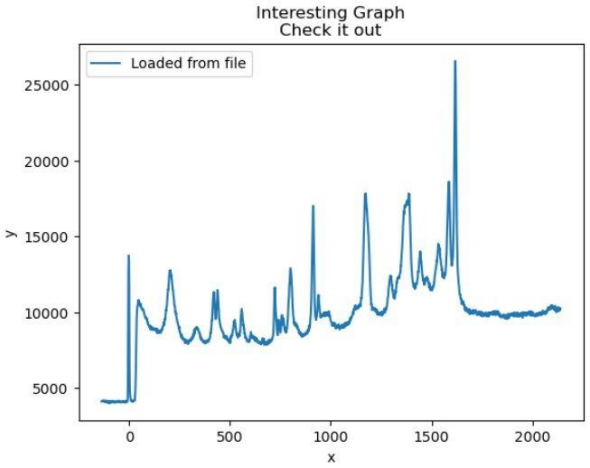
Same as in the above experiment we just have to put a sample of the crystal violet in the Raman Machine and take the reading/coordinates and put them in our code and get the graph…

In [1]:

```
import matplotlib.pyplot as plt
import csv
x = []
y = []

with open("CV_values.txt",'r') as csvfile:
    plots = csv.reader(csvfile, delimiter=',')
    for row in plots:
        x.append(float(row[0]))
        y.append(float(row[1]))
plt.plot(x,y, label='Loaded from file')

plt.xlabel('x')
plt.ylabel('y')
plt.title('Interesting Graph\nCheck it out')
plt.legend()
plt.show()
```



Figure(9)

After that we check the graph that it is crystal violet or not, By checking the peaks of the compound is present in the graph or not…

In [3]:

```
lst = [0,0,0,0,0,0,0,0,0,0]
for i in x:
    if(i in range(415,425)):
        lst[0] = 1
    elif(i in range(435,445)):
        lst[0] = 1
    elif(i in range(795,805)):
        lst[0] = 1
    elif(i in range(910,920)):
        lst[0] = 1
    elif(i in range(1210,1225)):
        lst[0] = 1
    elif(i in range(1170,1177)):
        lst[0] = 1
    elif(i in range(1360,1375)):
        lst[0] = 1
    elif(i in range(1610,1620)):
        lst[0] = 1

ans = 1
for i in range(0,len(lst)):
    if(lst[i] == 1):
        ans = 0
        break
if(ans):
    print("Crystal violet")
```

```
Crystal violet
```

Figure(10)

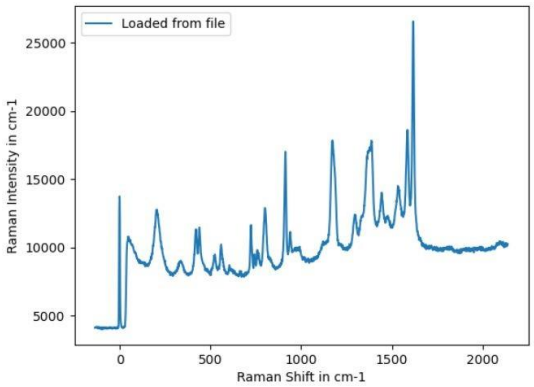
Here we came to know that the given sample is Crystal Violet.

CODE:

```
import matplotlib.pyplot as plt
import csv
x = []
y = []

with open("CV_values.txt",'r') as csvfile:
    plots = csv.reader(csvfile, delimiter=',')
    for row in plots:
        x.append(float(row[0]))
        y.append(float(row[1]))
plt.plot(x,y, label='Loaded from file')
plt.xlabel('x')
plt.ylabel('y')
plt.title('Interesting Graph\nCheck it out')
plt.legend()
plt.show()
```

Output:



Figure(11)

And then we adding the peak values to check the sample…

Peak Position(cm-1)	Tentative Peak Assignment
419	C–N bending vibration
440	Out-of-plane deformation vibrations of the phenyl
800	C–H bending vibrations
914	radical-ring skeletal vibration
1217	
1173	C–H bending vibrations
1367	C-C center strectching
1617	C-C strectching

Table(4)

```
lst = [0,0,0,0,0,0,0,0,0,0]
for i in x:
    if(i in range(600,615)):
        lst[0] = 1
    elif(i in range(760,775)):
        lst[0] = 1
    elif(i in range(1120,1135)):
        lst[0] = 1
    elif(i in range(1175,1190)):
        lst[0] = 1
    elif(i in range(1305,1320)):
        lst[0] = 1
    elif(i in range(1355,1365)):
        lst[0] = 1
```

```
    elif(i in range(1500,1515)):
        lst[0] = 1
    elif(i in range(1595,1603)):
        lst[0] = 1
    elif(i in range(1640,1650)):
        lst[0] = 1
ans = 1
for i in range(0,len(lst)):
    if(lst[i] == 1):
        ans = 0
        break
if(ans):
    print("Crystal violet")
```

Output:

Crystal violet

CONCLUSION

Here we have done a work that to describe the known compound using coding by representing graph and proved. But if we do the same work for unknown compound we can really do some miracles in medical sciences and many other fields. But I'm glad to do this project that I have learnt a lot about Raman sir's work for scattering light on photons and I have gain some experience regarding how to do a project….

FUTURE WORK:

As I mentioned in the earlier conclusion section that if we do the same project on identifying the unknown compounds we can really do some miracles in science , nowadays we are suffered a big problem named corona we know that scientists have done the research on this problem using Raman spectroscopy by testing the blood serums and saliva…we have to develop this technology and in future it will definitely help to solve so many problems.

References:

Diagnosing COVID-19 in human serum using Raman spectroscopy, National library of medicine.

CHEBI:41688 - crystal violet ,Chebi website.

Handbook of Surfaces and interfaces of materials,2001, science direct.

Rhodamine 6G,scott prahl,omlc.org.

Rhodamine 6G,pubchem,natinal library of medicine.

Raman study of mixed solutions of methanol and ethanol, aip.scitation.org.

Introductory Raman Spectroscopy by John R. Ferrero, Kazuo Nakamoto;1st chapter basic theory.

Advantages and limitations of Raman spectroscopy for molecular diagnostics: an update , taylor & Francis online.

Reference database of Raman spectra of biological molecules,wiley analytical science.

Infrared and Raman Spectroscopy by Peter J .Larpin 1-3.

Raman Spectroscopy , D.A.Long, The Characterization of chemical purity organic compounds.

Introduction to Raman Scattering, Springer link.

Modern Raman Spectroscopy: A Practical Approach , Ewen Smith ,Geoffrey Dent.

Detecting RNA virus in saliva samples using Raman spectroscopy, The Hindu.

IIsc researchers working on covid 19 detection using raman spectroscopy artificial intelligence.

Hyperglycemia and cancer in human lung carcinoma by means of Raman spectroscopy and imaging,Nature.com.