“

**SYNOPSIS**

**“BREAST CANCER DETECTION USING**

**NEURAL NETWORKS”**

**Submitted to the**

**Savitribai Phule Pune University**

**In the partial fulfillment for the award of the degree of**

**Bachelor of Engineering**

**In**

**Information Technology**

**By**

**ANKIT KUMAR SINGH**

**KARAN NAINWAL**

**PANKAJ MEHRA**

**RAKESH KUMAR**

**Under the guidance of**

**Prof. Ashwini Sapkal**



ABSTRACT

Breast Cancer is one of the main reasons of the death among women due to diseases. It accounts for the second-most deaths in woman every year. Mammography is the most popular detection method and there are several CAD (Computer Aided Design) techniques for the detection using mammogram images.

Based on World Health Organization (WHO) statistics, cancer was responsible for 13% of all deaths in the world by 2004, an estimated 12 million people dying from cancer in 2030.

Mammography offers high-quality images at a low radiation dose, and is currently the only widely accepted imaging method used for routine breast cancer screening.

One of the difficulties with mammography is that mammograms generally have low contrast. This makes it difficult for radiologists to interpret the results. So a need of CAD was experienced to provide a second opinion to the radiologist using Machine Learning techniques.

The new system uses three algorithms for the detection of cancer using mammogram images. We have used Back propagations, SVM (Support Vector Machine), and the advanced algorithm SVM with MGSA (Binary Gravitational Search Algorithm).

The system uses the above mentioned three algorithms and compares the performance and efficiency of each algorithm on a particular dataset of mammogram images and detects whether the cancer is present on a given sample.

Firstly the BP and SVM are applied and then the SVM is improved using BGSA.

The database used is the Wisconsin Breast Cancer Database and the languages used are Python and PyQt with different machine learning libraries. The system aims to provide a detection of the mammogram images and provide a comparison between different algorithms.

In a country like India where the medical facilities are slow in their delivery there is a shortage of work force and medical experts so there is a urgent need to use the Information Technology for fast delivery off medical services.

**INTRODUCTION**

According to the World Health Organization, breast cancer is the most common cancer among women worldwide, claiming the lives of hundreds of thousands of women each year and affecting countries at all levels of modernization.

There are several imaging techniques for examination of the breast, including magnetic resonance imaging, ultrasound imaging, and X-ray imaging.

Mammography is a specific type of imaging that uses a low-dose X-ray system to examine the breast, and is currently the most effective method for detection of breast cancer before it becomes clinically palpable.

Mammography offers high-quality images at a low radiation dose, and is currently the only widely accepted imaging method used for routine breast cancer screening.

* 1. **Need**

One of the difficulties with mammography is that mammograms generally have low contrast. This makes it difficult for radiologists to interpret the results.

Double reading of mammograms has been advocated to reduce the proportion of missed cancers. The basic idea of double reading is to have two radiologists read the same mammograms.

There are following benefits and needs of the CAD:

* **Better Result -** CAD provides a second opinion to the radiologist and it eliminates the need of second opinion form other radiologist. It also improves the detection capability of a system as whole
* **Cost Efficiency -** The cost of a CAD system is comparatively low when compared to the cost of radiologist as it requires special technical skills.

Also, these systems have low maintenance cost .

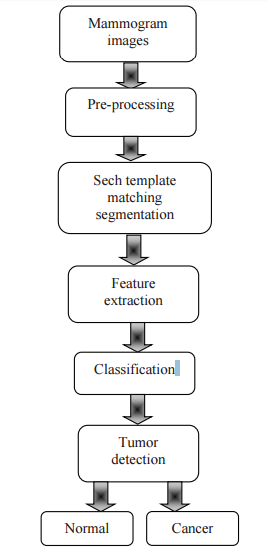
* **Faster Result –** Using the several machine learning algorithms and advanced hardware, faster detection and diagnosis of the mammogram images is possible.
* **Sustainable -** These systems are quite sustainable as they are very much adaptive to new changes in the algorithm and new techniques. So these systems can be modified very quickly.
  1. **The Proposed System – Introduction**

The system uses three algorithms to detect the present of cancer in a given sample of mammogram image :

* Back Propagation
* Support Vector Machine(SVM)
* SVM with BGSA (Binary Gravitational Search Algorithm)

The basic system has the following steps involved only the algorithm used changes in each case :

* Image Preprocessing
* Segmentation
* Feature Extraction
* Classification
* Tumor Detection



*Fig 1:Working of a Detection system*

We have worked on the following nine features of a mammogram image:

- Clump Thickness

- Marginal Adhesion

- Uniformity of Cell Size

- Uniformity of Cell Shape

- Normal Nucleoli

- Single Epithelial Cell Size

- Bland Chromatin

- Bare Nuclei

- Mitoses

We have used the Wisconsin Breast Cancer Database provided by the UCI Repository.

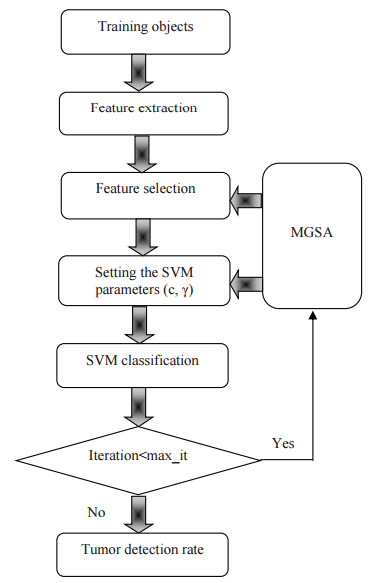
* 1. **Improving the System**

We apply the Back propagation and SVM to the database and compare these algorithms on the basis of efficiency and accuracy. Then we found the need to improve the SVM as the existing system with SVM has low efficiency so SVM with BGSA is introduced in the system.

If all of these extracted features (such as mean, variance, entropy, Correlation and etc.) from images are used, the processing time is increased and the training of the SVM would be complicated.

The aim is to optimize the evaluation function that is the precision of the SVM classifier. After training, the selected features are optimized and tuned SVM are used for detecting target objects.

BGSA improves two parameters of the SVM. These parameters are the kernel Parameter, gamma (ϒ, (for the radial basis functions (RBF) kernel and penalty parameter (C).



*Fig 2: Improving SVM with MGSA and BGSA*

All the three algorithms are applied and the results of all are compared and the detection is done and the result is displayed. Also the result of each feature is also displayed and the final result is also displayed.

**PROBLEM STATEMENT**

**3.1 PROBLEM STATEMENT**

Develop an Artificial Intelligent system to detect cancer in mammogram images to provide support to the manual diagnosis system and also provide comparison between different algorithms on the basis of accuracy and efficiency.

**3.2 PROJECT OVERVIEW**

The development of the proposed system would bring into consideration the following user scenario wherein:-

1. The user (radiologist) will login into the system.
2. The radiologist will provide the preprocessed data of the digital mammograms.
3. The input can be in unformatted form or in form of feature matrix.
4. The user will then command the system to perform detection.
5. The system will then display the result. The result will be in two parts :
6. Yes or No : Cancer Present
7. Analysis of every feature
8. Now the radiologist can use the above result as a second opinion.

The system also provides research opportunities to the developer in the following ways:

* Stores the result of every detection and stores the data of each algorithm.
* This data can be further used to compare the different algorithm on regular basis.
* On the basis of comparison the best method can be selected and that method is selected as primary methods
* Remarks from the radiologist are also collected so as to improve the system.

The system can also be useful to the patient in following ways:

* The data of all the patients is stored in the database.
* Data is saved in both the data form also in digital form.
* The data can also be used to track the health of a patient
* Patient can directly collect their reports from the online system.

**CONCLUSION**

The proposed system can be implemented using Back propagation and SVM. The system now can be implemented after literature survey and requirement analysis.

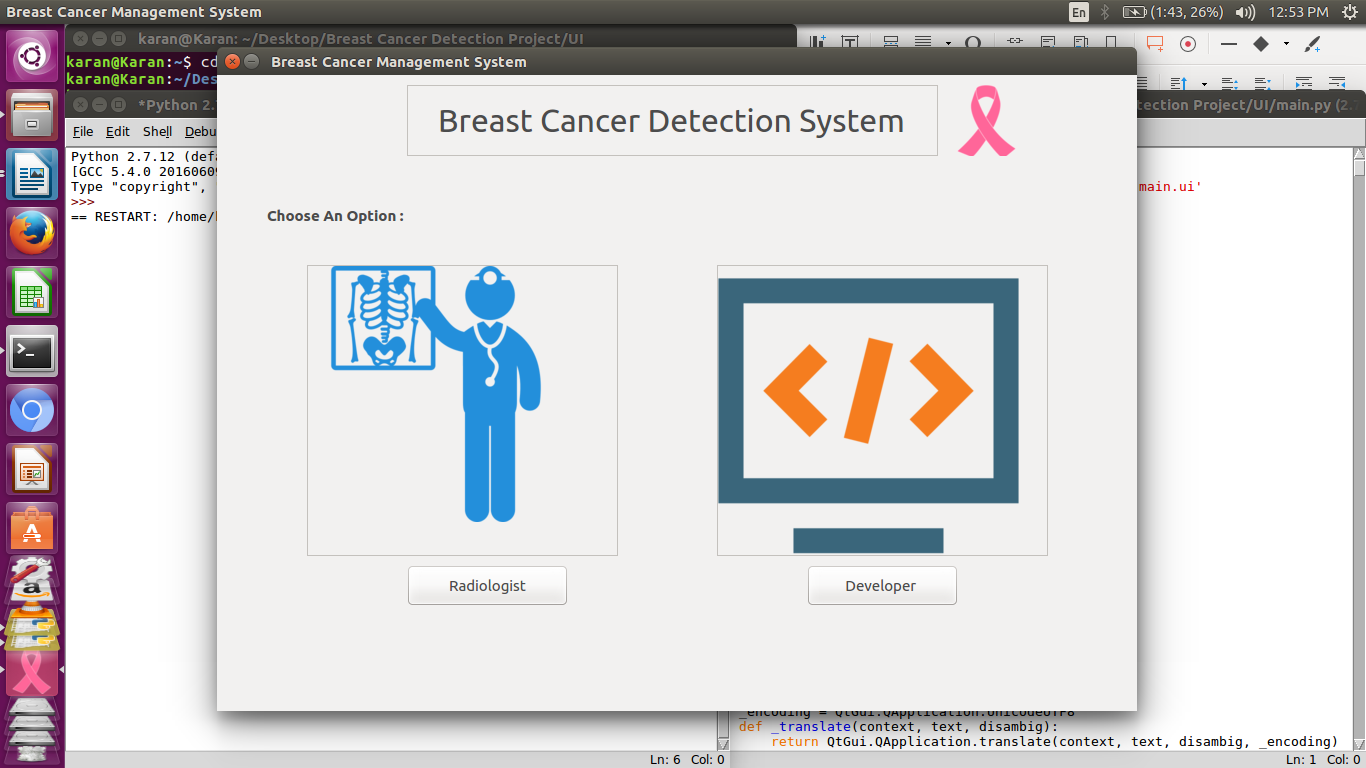
Now, on the basis of required information we can start the design and coding phase of the system using different platforms and languages.

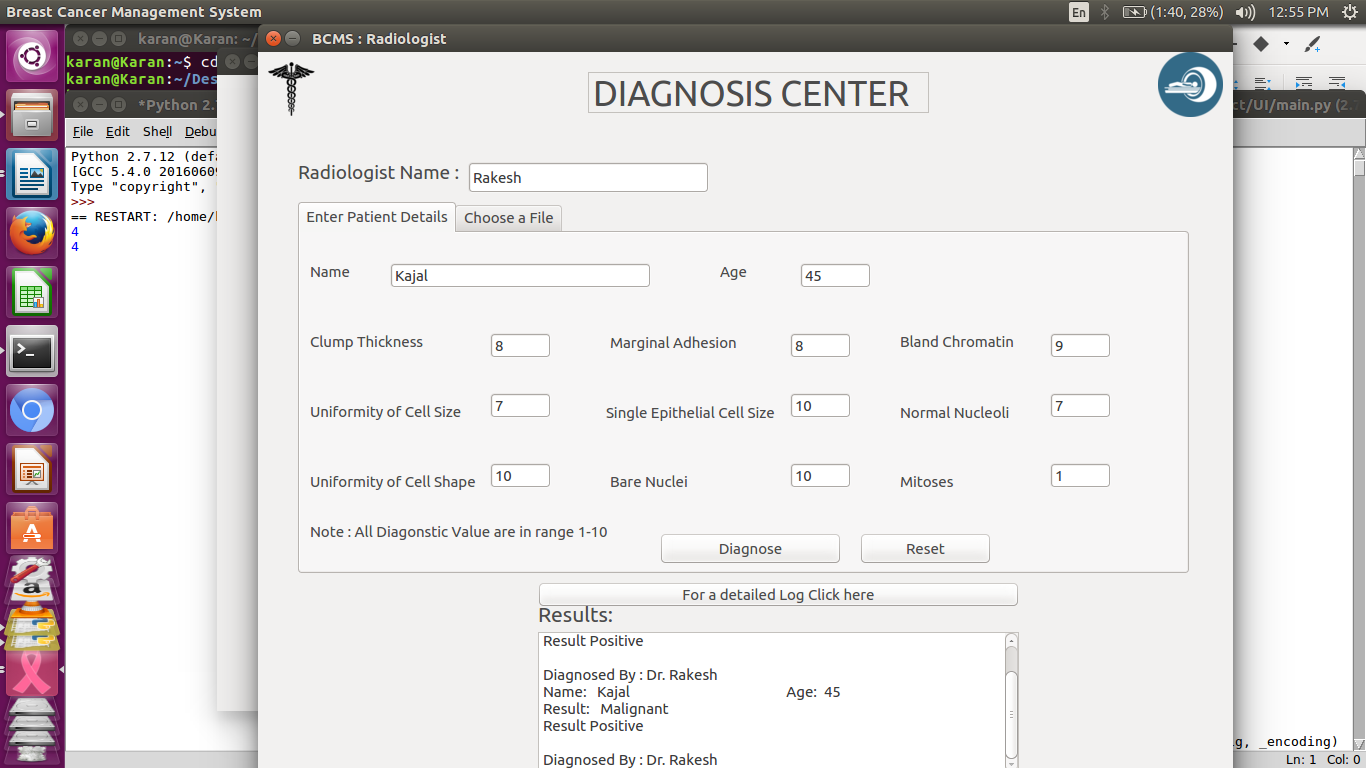
This system will be very useful for the detection of Breast Cancer , it will provide a quick second opinion to the radiologist ,It will also prove to be cost efficient and will be able to manage the reports and data of the patient and will also be helpful in tracking the health of different patients.

The system can be utilized by different diagnosis laboratory and can provide a second opinion .It will make the process efficient and faster in a country like India where the medical facilities are very slow in their deliveries.

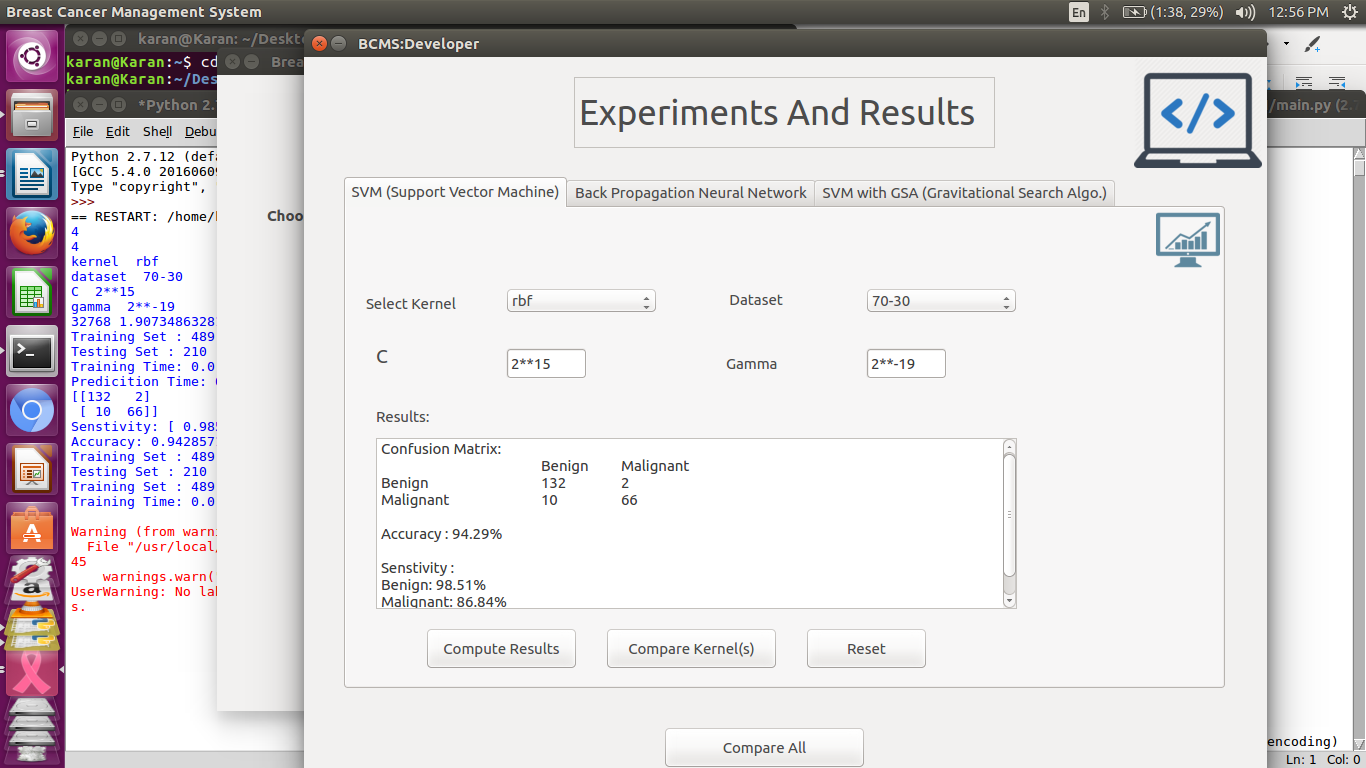
This system also provides research opportunities to the developers also as it stores the data and the result and provides a comparison between different methods of detection.

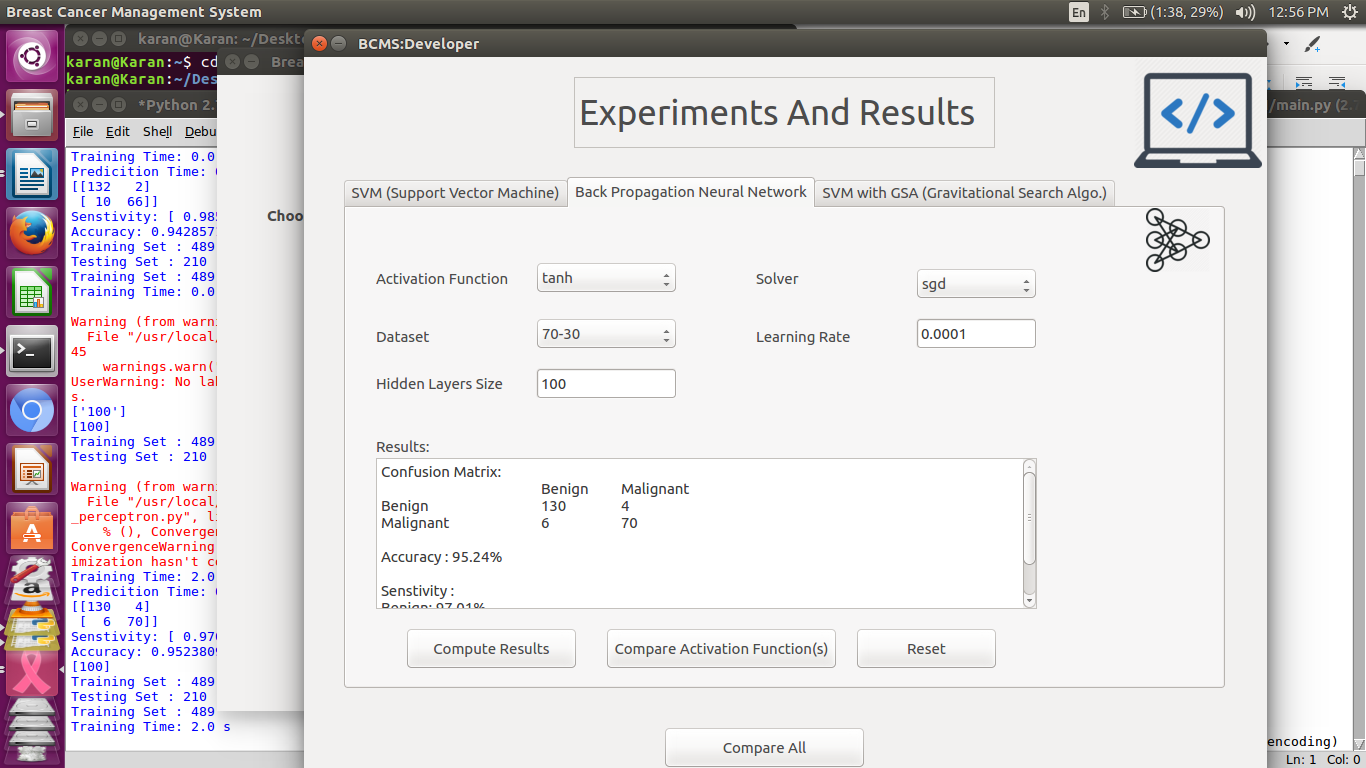
GRAPHS AND RESULTS

 *Home Screen*

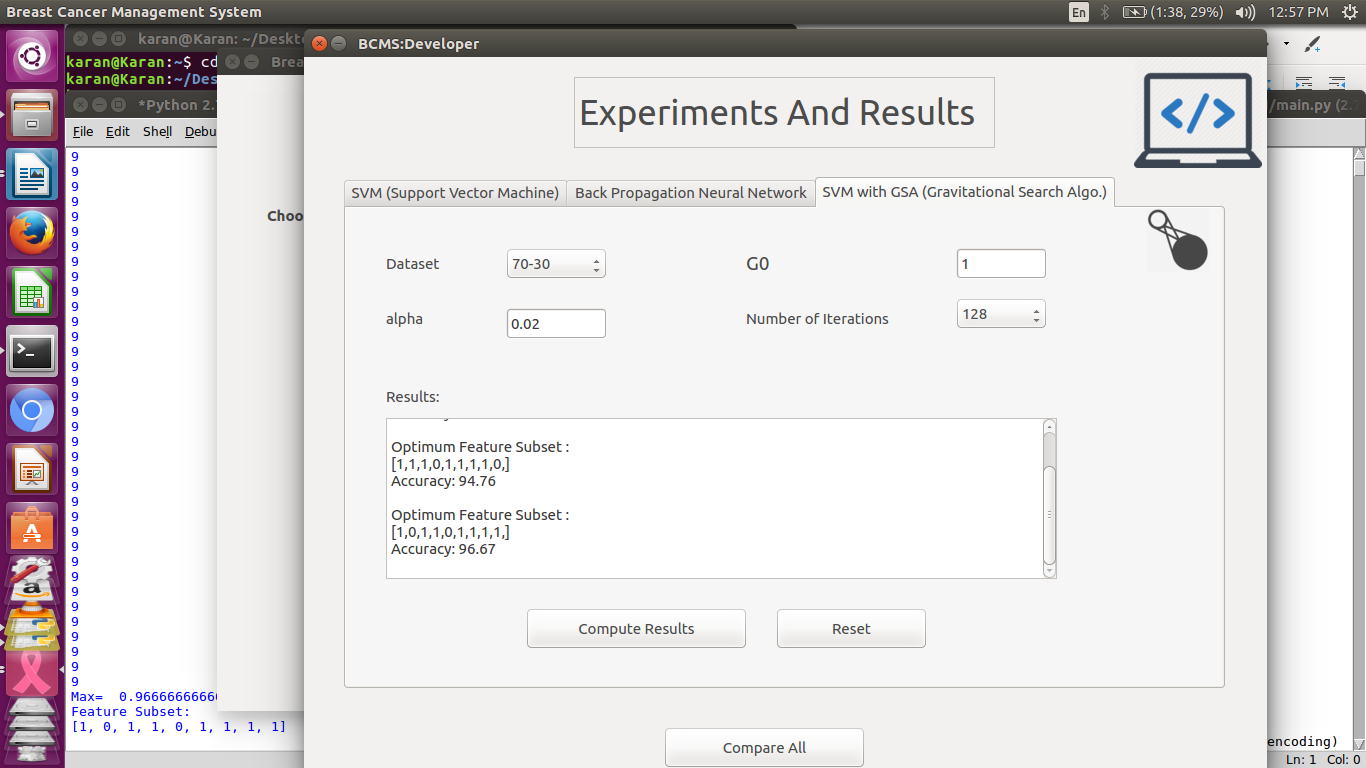
**

*Diagnosis Center*

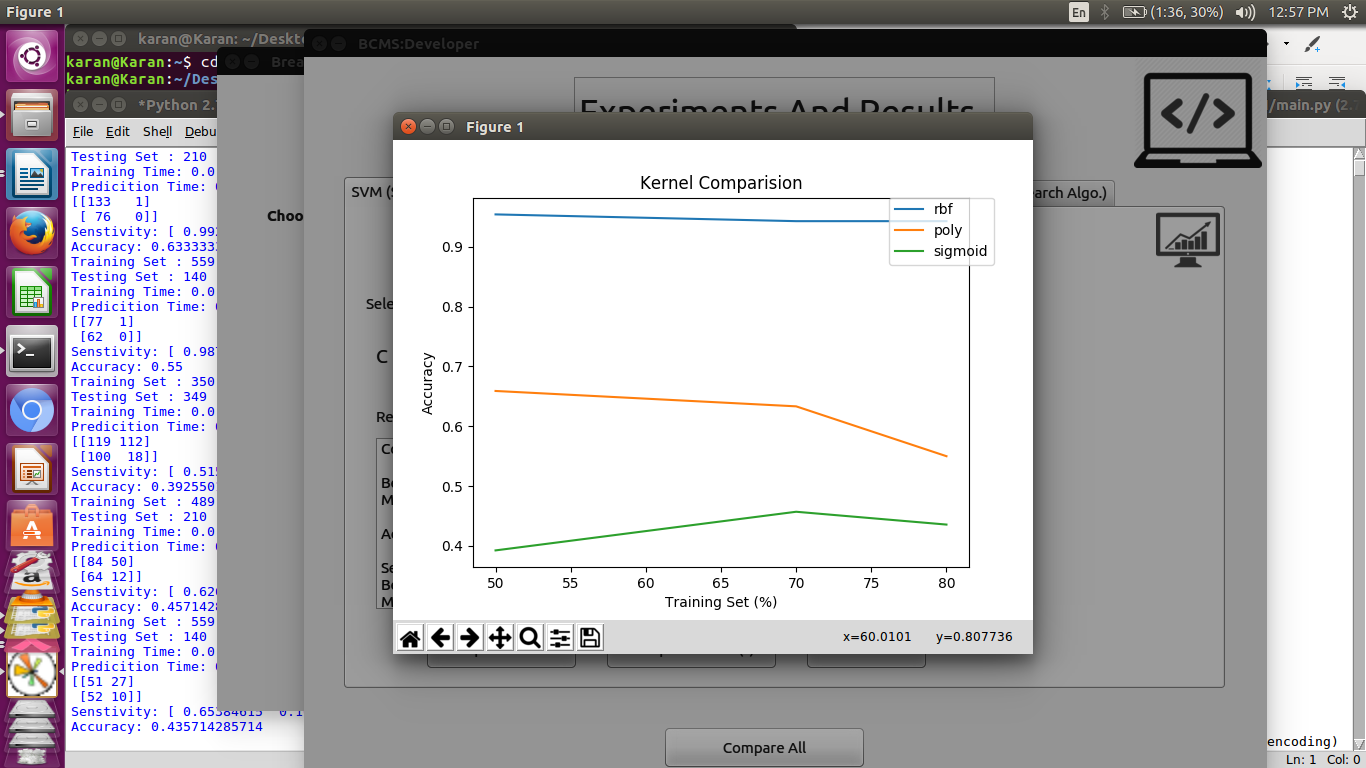
 *SVM Results*

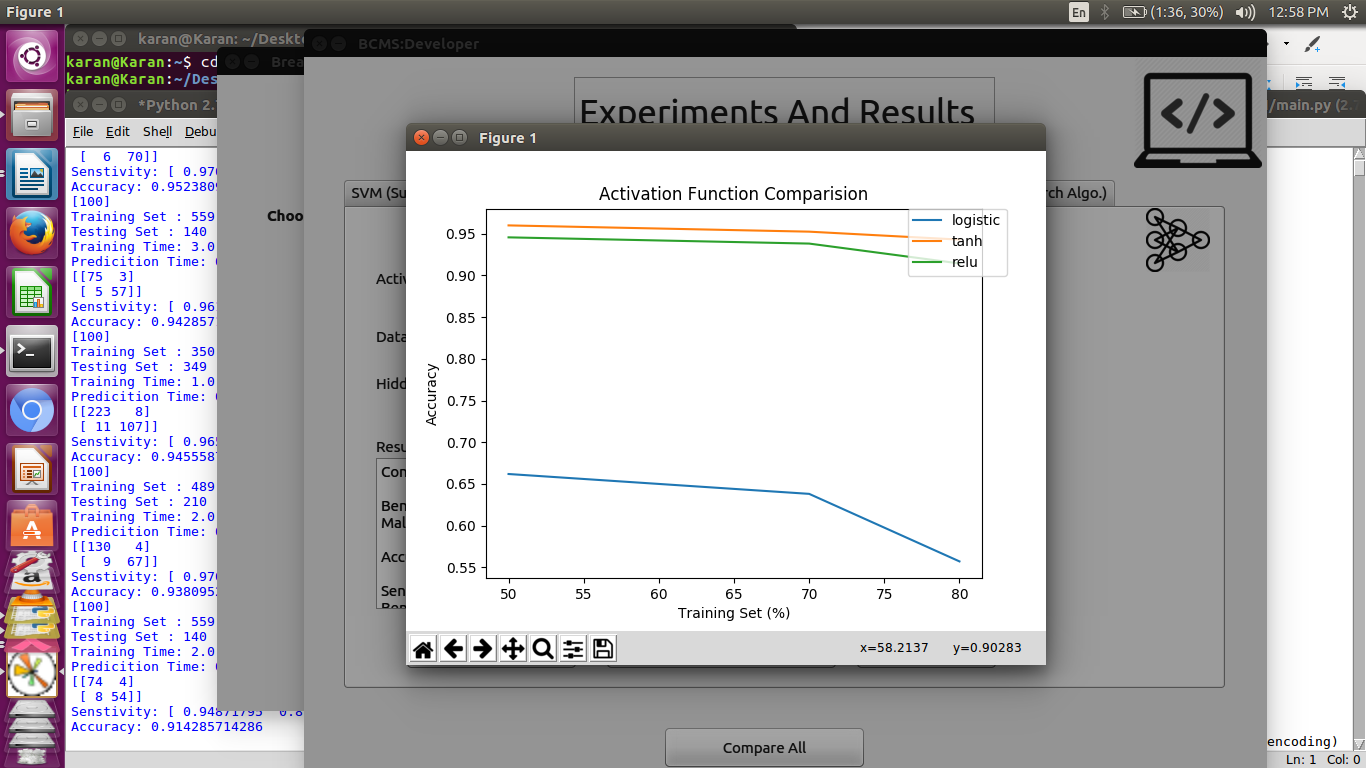


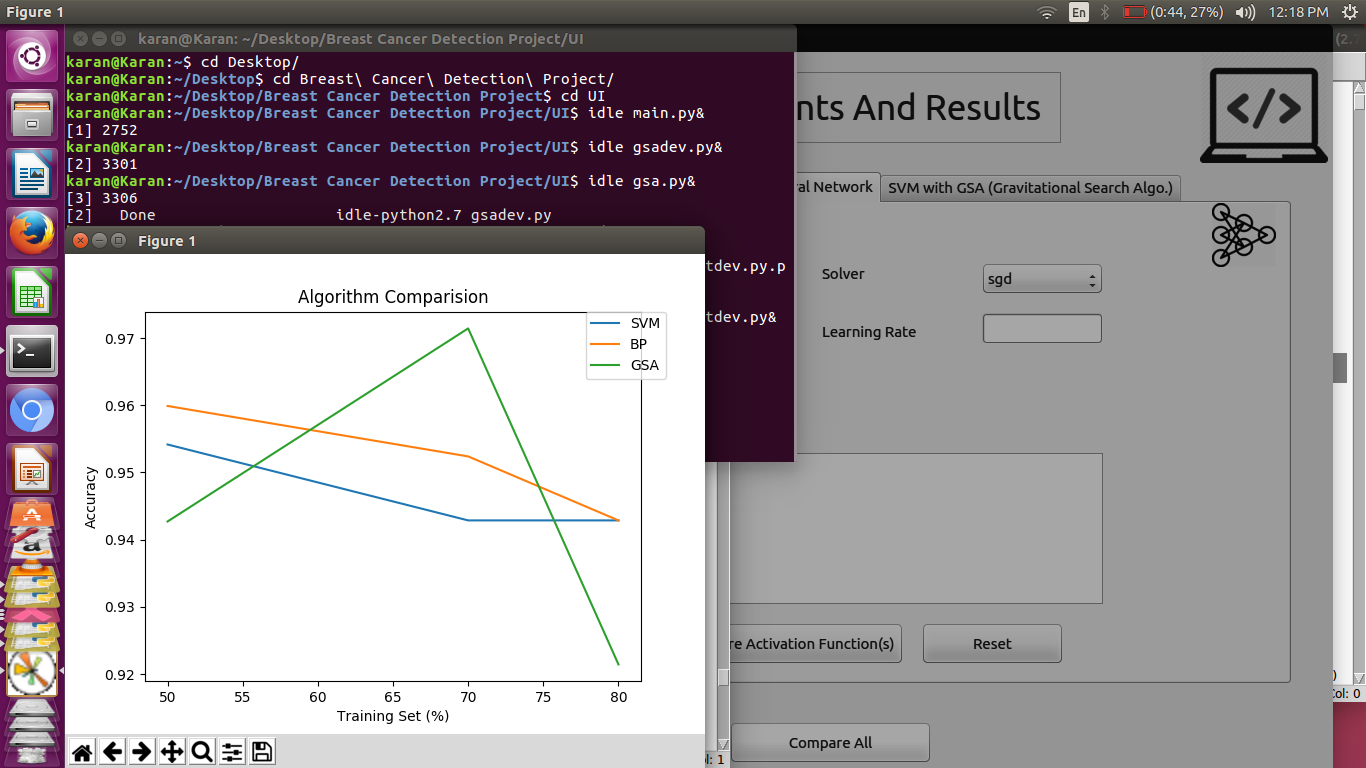
*Backprop Results*



*GSA Results*

 *SVM Kernel Comparision*

 *BP Activation Functions Comparision*

 *Comparision of Algorithms*