

# Recent Impact of Allergy In Bangladesh

Md, Raihan

171-15-9134  
Computer Science &  
Engineering  
Daffodil International  
University  
Gazipur, Bangladesh  
[raihan15-9134@diu.edu.bd](mailto:raihan15-9134@diu.edu.bd)

Prity Ali

172-15-9630  
Computer Science &  
Engineering  
Daffodil International  
University  
Bogra, Bangladesh  
[prity15-9630@diu.edu.bd](mailto:prity15-9630@diu.edu.bd)

Sha. Istiaque Ahammed Torun

171-15-8671  
Computer Science &  
Engineering  
Daffodil International  
University  
Sherpur, Bangladesh

## Abstract

Bangladesh is a developing country with 160 million population. About 25% to 30% of population is suffering from different types of allergic diseases. Few studies were carried out on asthma, allergic rhinitis and atopic conditions in Bangladesh. There are many indoor, outdoor, occupational and food allergens that trigger allergy. More over many pollutants, allergens are still unidentified in the poor and developing countries due to fund constrains for research activities. In addition to multiple known and unknown allergens, over population, un-hygienic living, poverty, lack of education and awareness, negligence to take treatment leads to increasing incidence of these diseases. Diagnostic facilities are yet depending on history and clinical examination for majority of our patients. Skin prick tests and IgE estimation are not available outside capital Dhaka as well as these are expensive too. Majority of our population can't

afford them. Asthma and allergies are affecting the quality of life that has impact on national economy and development a lot. Treatment modalities and drugs available with the costs in our country are discussed. More evidence based studies and dissemination of ARIA, WHO and other guidelines to health professionals will be needed to improve the situation. While most allergies can't be cured, treatments can help relieve your allergy symptoms.

In recent days we saw the rising rate of allergic patients. Mostly cold allergy, sinus infection and food allergy. Our research tried to find out where people are affected mostly. And what type of allergy are common now a days. And what is the relation between the road construction and allergy.

**Keywords—Asthma, Allergic Rhinitis, Atopic Dermatitis**

## I. INTRODUCTION

Allergic diseases are prehistoric and very common. None is immune to this condition. Every person has an experience of one form of allergies in their lifetime. Incidence and prevalence of allergic lesions are increasing day by day. About more than 150 million populations are affected by asthma disorder globally<sup>1</sup>. This is causing a tremendous strain on health resources in many part of the world. It is also a main cause of hospital admission for chronic diseases in children as well as in adults. There is also a huge financial involvement to treat the allergic diseases. Epidemiology and Statistics - Allergy, Asthma, Food Allergy, Occupational Allergy: Bangladesh is a developing country with about 140 million population and per capita income of 470 USD. per annum. In our country, about 45 – 51% people live in poverty, of these 20 – 25% are in extreme poverty . Population density in our country is 928 per square kilometer. Hospital bed and population ratio is 1: 3500, and doctor patient ratio is 1: 3866. Total health expenditure is only 6.6% of our total budget . First National Asthma Prevalence Study (FNAP) in Bangladesh in 1999, showed that about 7 million people (5.2% of the population) are suffering from current asthma in at least three episodes of asthma attack in one year; more than 90% of whom do not take modern treatment. . Half of these patients are innocent children in 7.4% of the total paediatric population (1-15 years of age group). It is estimated that over 20 - 50% of the world's population suffer from IgE mediated allergic disease such as asthma, rhinoconjunctivitis, eczema & anaphylaxis. Epidemiological, immunological and clinical studies showed that there is strong correlation between allergic rhinitis and asthma. Epidemiologically, up to forty per cent of allergic rhinitis patients also have asthma, and up to eighty per cent of cases with asthma are related to nasal condition.

Our goal was to find out where the allergy has more impact. Who are affected mostly. What is their living place. Which type of allergy mostly affecting.

For this we interviewed random people with respect to their age, gender, living place etc. Our sample data was collected by using google form and we collected data by sending the form on social media and email.

We collected 150 samples. And almost all of the sample owners age were between 20-30. And we got some interesting result.

Then using this data we made a machine learning model, where giving some inputs we can assume if a person is affected or not.

**Keywords: Machine learning, dataset, allergy, population, health, IgE**

## II. TOOLS USED

**1.Google form:** Google form is a popular medium for collecting data from remote user via internet.

**2.Microsoft Excel:** Microsoft Excel is the industry leading spreadsheet software program, a powerful data visualization and analysis tool.

**3.Python:** Python is an interpreted high-level general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant indentation. For this we used some libraries of python they are

1. Numpy: For working with array.
2. Pandas: For data analysis and manipulation.
3. Matplotlib: For plotting graphs, barchart, histogram etc.
4. Scikit learn: For building machine learning model.

**4.Jupyter notebook:** This software is used for data mining purpose.

## III. DATA VISUALIZATION

**Dataset:** In our dataset we collected 150 samples. We collected these sample mainly from facebook and email.

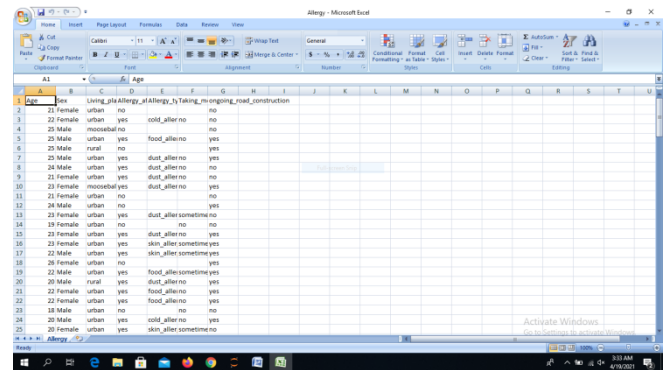


Fig 1: Samples Dataset

In our data set we used six columns. They are sex, age, living place, allergy affected, allergy type and on going road construction.

**Age:** Age columns represents the age of the sample.

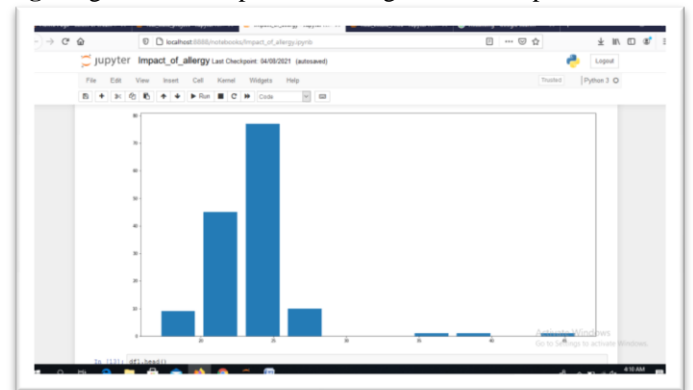


Fig 2: Age histogram

From this histogram we can see most of our sample has age of 25 years. Our main target was to see the impact of allergy in young people.

**Gender:** Here is our gender piechart.

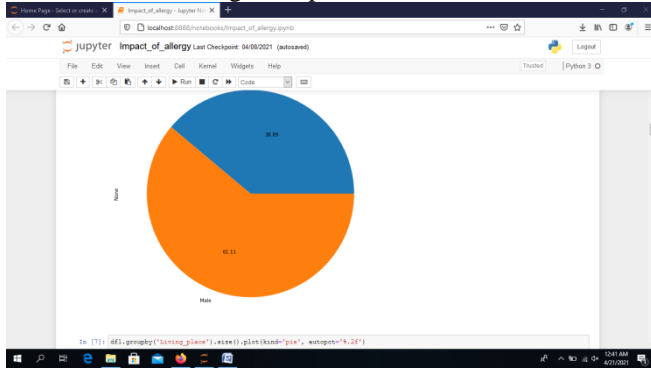


Fig 3: Gender piechart

In this piechart we found that 61.11% are male & 38.89% are female. And female portion is blue colored and male portion is orange colored.

**Living Place:** Here we can see the piechart of living place feature.

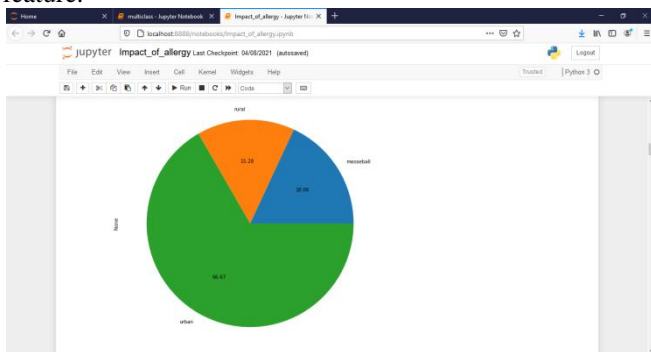


Fig 4: Sample's living place piechart.

Here 66.67% data lives in urban area, 15.28% data lives in rural area and rest of sample lives in moose ball area.

**Allergy Affected:** Here our samples 72.92% are affected in allergy and 27.03% are not affected in allergy.

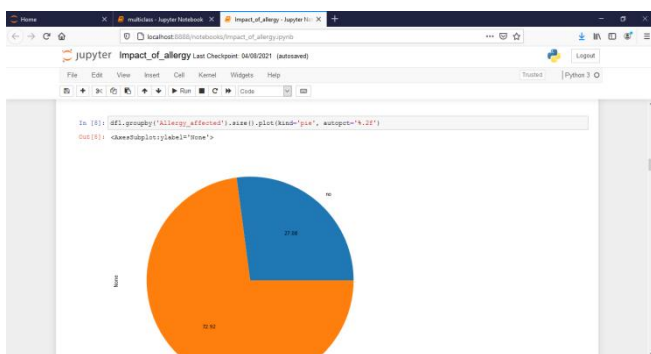


Fig 5: Allergy affection

**Allergy type:** There are 6 types of allergy. In this pie chart we got all type of allergies are affected. Here, 37% has dust allergy(colored orange), 14.15% has cold allergy(colored red), 23.58% has skin allergy(colored brown), 4.72% has sinus infection, 15.09% has food allergy

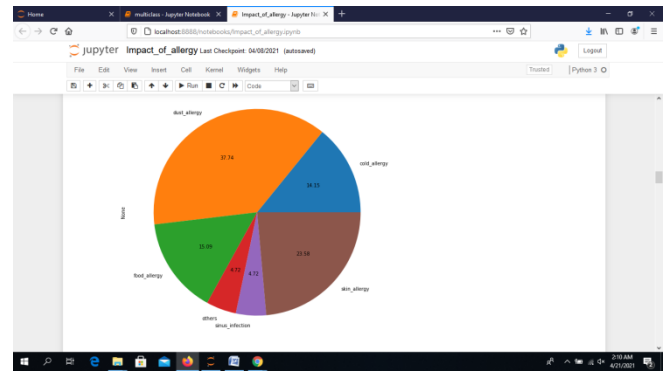


Fig 6: Types of allergy

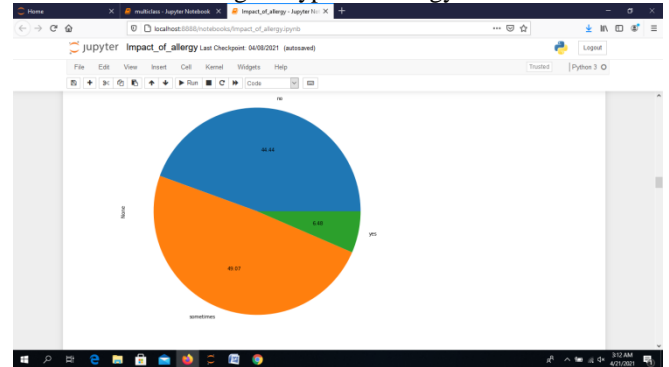


Fig7:Medicine taking

Here we can analyze how much percentage of our sample taking medicine and how much are not.

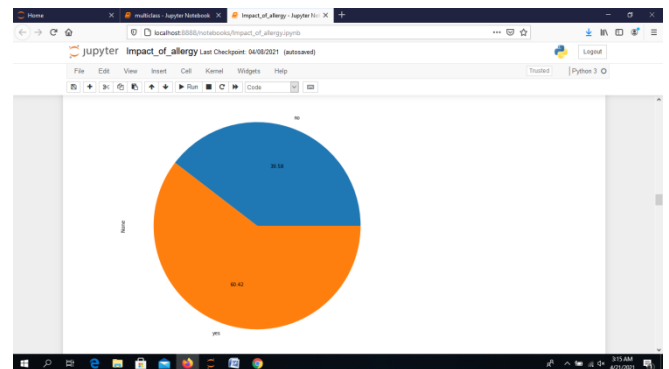


Fig8:Road construction

In this figure we can find out that how much of our samples living in that place where road is constructing. We needed this feature because we know that cold and dust allergy has a relation with dusts and constructing roads produces more dust than any other things.

## IV. WORKING PROCEDURE

To visualize our data we mostly used pie chart. Because pie chart is the most easy diagram to show the condition of the data to general public.

**Data Cleaning:** In this part we find out how much null values our data contains and where null values are most in our data set. Then we removed all null values. In this project if we drop null values our model wouldn't be affected so much. That's why we dropped all null values.

```
[3]: df1.shape
```

```
Out[3]: (144, 7)
```

```
[4]: df1.isnull().sum()
```

```
Out[4]: Age                0
Sex                  0
Living_place         0
Allergy_affected     0
Allergy_type        38
Taking_medicine      36
ongoing_road_construction 0
dtype: int64
```

```
[5]: df2=df1.dropna()
```

```
[6]: df2.isnull().sum()
```

```
Out[6]: Age                0
Sex                  0
Living_place         0
Allergy_affected     0
Allergy_type         0
Taking_medicine      0
ongoing_road_construction 0
dtype: int64
```

```
[7]: df2.shape
```

```
Out[7]: (106, 7)
```

Fig 9,10: Dropping null values

All of our data was in string format. But as we know to build a ml model we need data to be in integer or in float format. There is many method to convert strings into integer format such as creating dummy variables, label encoding , one hot encoding etc.

We used label encoding for this particular model build. Scikit learn has a method called label encoder we used this method for label encoding.

```
from sklearn.preprocessing import LabelEncoder
le_sex=LabelEncoder()
le_living_place=LabelEncoder()
le_affected=LabelEncoder()
le_type=LabelEncoder()
le_med=LabelEncoder()
le_road=LabelEncoder()

df2['sex']=le_sex.fit_transform(df2['Sex'])
df2['place']=le_living_place.fit_transform(df2['Living_place'])
df2['affected']=le_affected.fit_transform(df2['Allergy_affected'])
df2['type']=le_type.fit_transform(df2['Allergy_type'])
df2['med']=le_med.fit_transform(df2['Taking_medicine'])
df2['road_construction']=le_road.fit_transform(df2['ongoing_road_construction'])
```

Input

	Age	Sex	Living_place	Allergy_affected	Allergy_type	Taking_medicine	ongoing_road_construction	sex	place	affected	type	med	road_construction
1	22	Female	urban	yes	cold_allergy	no	no	0	2	1	0	0	0
3	25	Male	urban	yes	food_allergy	no	no	1	2	1	2	0	1
5	25	Male	urban	yes	dust_allergy	no	yes	1	2	1	1	0	1
6	24	Male	urban	yes	dust_allergy	no	no	1	2	1	1	0	0
7	21	Female	urban	yes	dust_allergy	no	no	0	2	1	1	0	0

Output

Fig 11,12: Label encoding inputs and out puts

After label encoding we dropped unnecessary features mainly that columns which contains string type datas. After doing all of these our data frame looks like this.

```
Age sex place affected type med road_construction
1 22 0 2 1 0 0 0
3 25 1 2 1 2 0 1
5 25 1 2 1 1 0 1
6 24 1 2 1 1 0 0
7 21 0 2 1 1 0 0
```

Fig 13: New data frame

Then again we deleted some columns which has no benefit for our model.

After doing that we called train test split method from scikit learn. To divide our data into train samples where ml algorithm will train our sample and then would test our samples . It would give us better accuracy.

```
In [32]: df4=df3.drop(['type','med','road_construction'],axis='columns')
df4
```

```
Out[32]: Age sex place affected
1 22 0 2 1
3 25 1 2 1
5 25 1 2 1
6 24 1 2 1
```

```
X=df4.drop(['affected'],axis='columns')
X
```

```
]: y=df4.affected
y
```

```
]: 1 1
3 1
5 1
6 1
7 1
..
```

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.2)
```

Fig13,14,15,16:Train test split, Deleting columns

For this project we used three algorithms for finding out which model can give us the best result. Three algorithms are Support vector machine, logistic regression and decision tree.

**Key words:** Feature, Scikit learn, Label encoding, Null values , Data cleaning

## V. RESULT ANALYSIS AND CONCLUSION

We got higher accuracy in all algorithms for this data set. The accuracy was almost 1 in all algorithms. That means 100% accurate is our work. And also we got these result without tweaking any kernels and parameters of these algorithms.

```
In [60]: from sklearn.svm import SVC
         model2=SVC()

In [61]: model2.fit(X_train,y_train)
Out[61]: SVC()

In [41]: model2.score(X_test,y_test)
Out[41]: 1.0

In [42]: from sklearn.ensemble import RandomForestClassifier
         model3=RandomForestClassifier()

In [43]: model3.fit(X_train,y_train)
Out[43]: RandomForestClassifier()

In [44]: model3.score(X_test,y_test)
Out[44]: 1.0

In [45]: import pickle
         saved_model = pickle.dumps(model3)
         model3_from_pickle = pickle.loads(saved_model)
```

Fig 17: Results.

In our country we couldn't find many data for allergy affection. But day by day it is becoming one of the major

problems in our day to day life. Our research will help people to understand why allergy is rising high so quickly. And It will help doctors find out the allergy cluster in our country.

## References

1. Allergic rhinitis, asthma and atopic diseases: Bangladesh perspective.

(M. Alamgir Chowdhury, ARM Lutful Kabir, Md. Monjurul Alam )

2. Food allergy in gastroenterologic diseases

(Mansueto P, Montalto G, Esposito-Pellitteri)

3. Using ML approaches for prediction of the types in allergy in Turkey.( Sevinç İlhan Omurca, Ekin Ekinci, Bengisu Çakmak, Selin Gizem Özkan)

4.ALLERGY ATTACK AGAINST AUTOMATIC SIGNATURE GENERATION.(Simon P. Choung, Aloysious K Mok)