Prediction of Asthma

**Algorithms Used:**

* K-Nearest Neighbors
* GaussianNB (Gaussian Naive Bayes is like Naive Bayes but for numbers, not categories)

**Colab Notebook:** <https://colab.research.google.com/drive/1o7JIbhnWYEMa69nntI3whKHbOrcYq33U?usp=sharing>

**Source:**

<https://www.kaggle.com/code/naniruddhan/asthmaprediction-gnb-balanced-smote/notebook>

**Dataset:**

<https://www.kaggle.com/code/naniruddhan/asthmaprediction-gnb-balanced-smote/input>

**Rows, columns:** (2392, 27)

**About**

Asthma is a chronic respiratory condition that affects millions worldwide. It is characterized by inflammation and narrowing of the airways, leading to symptoms such as wheezing, shortness of breath, chest tightness, and coughing. These symptoms can vary in intensity and are often triggered by environmental or lifestyle factors. Asthma is a complex condition influenced by genetic, environmental, and physiological factors, making accurate diagnosis and prediction essential for effective management and treatment.

**Objective of the Project**

The goal of this project is to leverage machine learning techniques to predict asthma presence based on various demographic, environmental, and physiological factors. Early and accurate prediction can help identify at-risk individuals and enable timely intervention, ultimately improving patient outcomes and quality of life.

**Dataset Overview**

The dataset used for asthma prediction comprises 2,392 samples, with features covering a wide range of factors:

* **Demographics**:
* **Age**: Age of the individual.
* **Gender**: Male or Female.
* **Ethnicity**: Categorical representation of the individual's ethnicity.
* **Education Level**: Represents socioeconomic factors linked to asthma prevalence.
* **Lifestyle and Environmental Factors**:
* **BMI**: Body Mass Index, an indicator of overall health.
* **Smoking**: History or current smoking habits.
* **Physical Activity**: Level of regular exercise.
* **Diet Quality**: Nutritional habits influencing overall health.
* **Sleep Quality**: Correlation between poor sleep and asthma symptoms.
* **Pollution Exposure**: Contact with environmental pollutants.
* **Pollen and Dust Exposure**: Common asthma triggers.
* **Allergy and Medical History**:
* **Pet Allergy**: Allergy to pet dander, a potential asthma trigger.
* **Family History of Asthma**: Genetic predisposition to asthma.
* **History of Allergies**: General allergic conditions that may overlap with asthma.
* **Eczema**: Skin condition often associated with asthma in atopic individuals.
* **Hay Fever**: Common allergic reaction linked to asthma risk.
* **Gastroesophageal Reflux**: GERD can exacerbate asthma symptoms.
* **Lung Function Parameters**:
* **Lung Function FEV1**: Forced expiratory volume in the first second.
* **Lung Function FVC**: Forced vital capacity, measures overall lung health.
* **Symptom Profiles**:
* **Wheezing**: A common symptom indicating airway obstruction.
* **Shortness of Breath**: Difficulty in breathing during normal or physical activity.
* **Chest Tightness**: Pressure or discomfort in the chest.
* **Coughing**: Persistent cough, particularly at night or early morning.
* **Nighttime Symptoms**: Asthma symptoms worsening at night.
* **Exercise-Induced Symptoms**: Asthma triggered by physical exertion.
* **Target Variable**:
* **Diagnosis**: Binary classification indicating the presence or absence of asthma.

**Algorithms Model and Accuracy**

|  |  |
| --- | --- |
| **Models** | **Scores** |
| K-Nearest Neighbors | 0.8298 |
| GaussianNB | 0.8308 |

Best: **GaussianNB**