**Project Name:** Cervical Cancer Predection

**Github Link:** https://github.com/projectsforstudents2022/Cervical\_Cancer\_Prediction.git

**Why was this project created?**

One of the most fatal diseases to affect women worldwide is cervical cancer. It is brought on by a protracted infection of the skin and mucous membrane cells in the vaginal region. The concerning element of this malignancy is that when it first develops, there are no signs. After breast cancer and lung cancer, it is the third most common type of cancer in women. Only when it is discovered and treated at an early stage can it be cured.

**What problem is it solving?**

The reduction of cervical cancer incidence and mortality is the goal of cervical cancer screening for prediction. An effective prediction model is possible to reduce fatalities by detecting early, preinvasive lesions during the pre-clinical detectable phase and preventing the development of invasive cancer.

**Entire explanation of project**

* **PROPOSED APPROACH**

The information is gathered and added to our database. Data collection is the process of collecting information from various sources and analyzing it. 36 columns make up the information organization, four of which serve as the targets for the four tests performed to determine whether the patient has cancer. The data will be cleaned up by removing any columns from the dataset that are undesirable or are null or not applicable. Data mapping is used to divide the acquired dataset into two parts: training data, which make up 80% of the dataset, and testing data, which make up 20% of the dataset.

The data has been divided into training and testing sets in order to assign data points to the former and latter in the modelling dataset. As a result, a model is developed using a training set before being used on a test set. Our programme will eventually be able to forecast the early identification of cervical cancer. utilising machine learning techniques with the provided dataset

Algorithm for creating next word prediction model :

**Step 1:** Import Libraries & Load Dataset

**Step 2:** Data Preprocessing

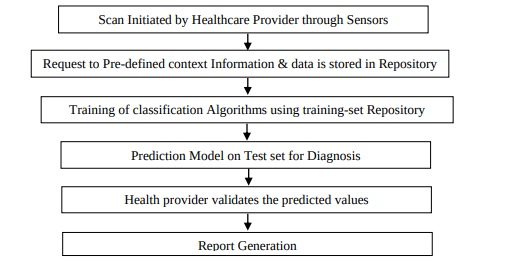
**Step 3:** Label Encoding

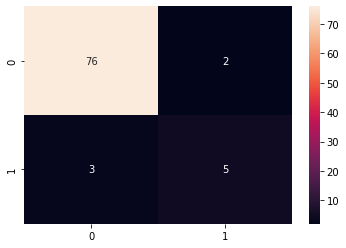
**Step 4:** Build XGBoost Classifier

**Step 5:** Train Model

**Step 6:** Testing & Visualization

* **DATA FLOW DIAGRAM**



* **RESULT**
* **CONCLUSION**

Today, cervical cancer is a widespread disease, and screening frequently includes drawn-out clinical tests. The diagnosis procedure can be sped up with the help of ML. This project might be used as a template for creating a future healthcare system for people with cervical cancer. Outcomes for algorithms based on accuracy show that XGBoost Classifier 94.3%