**GOOGLE AI FOR GOOD**

**MICROSOFT: AI FOR GOOD**

**Links for Datasets**

[Open Data for Society | Microsoft AI](https://www.microsoft.com/en-us/ai/data-for-society?activetab=pivot1:primaryr7)

[Microsoft Data for Society Catalog](https://microsoft.github.io/data-for-society/)

[Azure Open Datasets](https://azure.microsoft.com/en-us/services/open-datasets/#overview)

<https://idd.insaan.iiit.ac.in/dataset/details/>

CARBON FOOTPRINT RELATED

[2022 sustainability consumer research: Sustainability and profitability | IBM](https://www.ibm.com/thought-leadership/institute-business-value/report/2022-sustainability-consumer-research)

[Let’s build a better world with a smaller footprint - AI for Good](https://aiforgood.itu.int/lets-build-a-better-world-with-a-smaller-footprint/)

DISEASE (MALARIA & DENGUE)

[The top 10 causes of death](https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death#:~:text=The%20world's%20biggest%20killer%20is,8.9%20million%20deaths%20in%202019.)

[Crime in India | Kaggle](https://www.kaggle.com/datasets/rajanand/crime-in-india)

[India Water Quality Data | Kaggle](https://www.kaggle.com/datasets/venkatramakrishnan/india-water-quality-data)

[Indian water quality data | Kaggle](https://www.kaggle.com/datasets/anbarivan/indian-water-quality-data)

2 Interface - regular and BMC

Regular -   
Check water quality

#### PCOS (POLYCYSTIC OVARIAN SYNDROME) DETECTION USING DEEP LEARNING

Predicting Polycystic ovarian syndrome (PCOS) is a combination of symptoms caused by high levels of androgens in women. PCOS is caused by a mix of hereditary and environmental factors, which are common disorders such as atherosclerosis, hirsutism, acne, and hyperandrogenism, as well as recurrent infertility. According to recent studies, almost 18 percent of Indian women are affected by this disease.

Doctors manually inspected ultrasound scans to establish which ovary was affected, but they couldn't tell if the cyst was benign, PCOS-related, or cancerous.

DCNN-based algorithms are proposed in this study, and code for PCOS categorization is created using Python programming, and they are filled with blood or fluid using ultrasound images.

The work employs DCNN-based image classification to classify PCOS in the dataset.

That is, the study is based on a dataset of PCOS-related illnesses that have been trained.

Finally, using performance settings, the test dataset is used to execute feature extraction and measure accuracy. PCOS (Polycystic Ovary Syndrome) is a hormonal disorder that affects many women in their reproductive years and has been associated with infertility, diabetes, and cardiovascular disease.

To diagnose the condition, the bulk of imaging parameters are used.

Ultrasound imaging has become an important tool for diagnosing PCOS.

The normal appearance of the image becomes increasingly difficult because of overlapping follicles, inherent noise of the equipment, and a lack of operator comprehension because it is primarily an experience-based operation, making the diagnosis procedure time demanding. As a result of the aforementioned situations, cyst detection accuracy is affected. Early and accurate diagnosis of anomalies in the female reproductive system is critical prior to the treatment process to avoid infertility.

This paper covers several methodologies proposed so far in terms of reducing speckle noise, extracting regions of interest using segmentation, and picture classification in order to achieve maximum accuracy in cyst diagnosis in a short period of time.

#### Real-time Traffic Management in Emergency using Artificial Intelligence

Heart Disease/Attack Prediction Project

Wrist Band to monitor Heart rate, Blood Pressure, Sugar levels (if possible), and electrocardiographic values.

Application aimed toward Old People, Medically ill people.

Updation of data from time to time if they think that they are under the threat of a heart attack.

Data predicting

App giving specified tips and suggestions.

SOS.

Text to voice, Languages,

Heart failure signs and symptoms may include:

* Shortness of breath with activity or when lying down
* Fatigue and weakness
* Swelling in the legs, ankles, and feet
* Rapid or irregular heartbeat
* Reduced ability to exercise
* Persistent cough or wheezing with white or pink blood-tinged mucus
* Swelling of the belly area (abdomen)
* Very rapid weight gain from fluid buildup
* Nausea and lack of appetite
* Difficulty concentrating or decreased alertness
* Chest pain if heart failure is caused by a heart attack

Heart disease is the leading cause of death worldwide -- accounting for one-third of deaths in 2019 -- and the death toll continues to rise, a new paper says. China had the highest number of heart disease deaths last year, followed by India, Russia, the United States, and Indonesia.

High blood pressure, high blood cholesterol, and smoking are key risk factors for heart disease. Several other medical conditions and lifestyle choices can also put people at a higher risk for heart disease, including Diabetes. Overweight and obesity.

Heart Disease in the United States

About 697,000 people in the United States died from heart disease in 2020—that's 1 in every 5 deaths. Heart disease cost the United States about $229 billion each year from 2017 to 2018.15-Jul-2022

Includes Diseases: Coronary artery disease; Stroke

Includes causes of death: Coronary artery disease

The age-adjusted death rate attributable to cardiovascular disease (CVD), based on 2018 data, is 217.1 per 100,000. On average, someone dies of CVD every 36 seconds in the US. There are 2,380 deaths from CVD each day, based on 2018 data.

Number 1 – Heart Disease. Heart disease is a term that includes many specific heart conditions. According to the CDC, coronary artery disease (CAD), which can lead to heart attacks, is the most common heart disease in the United States.

[**Heart Disease Dataset | Kaggle**](https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset)

[**Heart Attack Analysis & Prediction Dataset | Kaggle**](https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset)

* Age: Age of the patient
* Sex: Sex of the patient
* exang: exercise-induced angina (1 = yes; 0 = no)
* ca: number of major vessels (0-3)
* cp: Chest Pain type chest pain type
  + Value 1: typical angina
  + Value 2: atypical angina
  + Value 3: non-anginal pain
  + Value 4: asymptomatic
* trtbps : resting blood pressure (in mm Hg)
* chol : cholesterol in mg/dl fetched via BMI sensor
* FBS : (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)
* rest\_ecg : resting electrocardiographic results
  + Value 0: Normal
  + Value 1: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV)
  + Value 2: showing probable or definite left ventricular hypertrophy by Estes' criteria
* thalach: maximum heart rate achieved
* target: 0= less chance of heart attack 1= more chance of heart attack

[**Heart Disease Prediction Project**](https://nevonprojects.com/heart-disease-prediction-project/)

**Natural Calamities Prediction Application**

Flood Prediction

Earthquakes -

[LANL Earthquake Prediction | Kaggle](https://www.kaggle.com/competitions/LANL-Earthquake-Prediction)

[LANL Earthquake EDA and Prediction | Kaggle](https://www.kaggle.com/code/gpreda/lanl-earthquake-eda-and-prediction) ,

Volcanic Eruptions - [INGV - Volcanic Eruption Prediction | Kaggle](https://www.kaggle.com/competitions/predict-volcanic-eruptions-ingv-oe) ,

Hurricanes

, tsunamis, floods,

Wildfires - [AI Geospatial Wildfire Risk Prediction | Kaggle](https://www.kaggle.com/code/thjaquenoud/ai-geospatial-wildfire-risk-prediction) ,

Droughts - [Predict Droughts using Weather & Soil Data | Kaggle](https://www.kaggle.com/datasets/cdminix/us-drought-meteorological-data) ,

Landslides - [Landslides After Rainfall, 2007-2016 | Kaggle](https://www.kaggle.com/datasets/nasa/landslide-events)

Disaster - <https://www.kaggle.com/code/aswin1871/disaster-prediction/data>

[Cyclone,Wildfire,Flood and Earthquake](https://www.kaggle.com/code/aswin1871/disaster-prediction/data)

Weather Prediction - [WEATHER PREDICTION | Kaggle](https://www.kaggle.com/datasets/ananthr1/weather-prediction)

Post Disaster - [https://www.kaggle.com/code/thomaskonstantin/disaster-tweets-analysis-and-prediction](https://www.kaggle.com/code/thomaskonstantin/disaster-tweets-analysis-and-prediction/notebook)

[Acea Smart Water Analytics | Kaggle](https://www.kaggle.com/competitions/acea-water-prediction)

[Air pollution impact | Kaggle](https://www.kaggle.com/datasets/cscomali/air-pollution-impact)

1. pH value: PH is an important parameter in evaluating the acid–base balance of water. It is also the indicator of acidic or alkaline condition of water status. WHO has recommended maximum permissible limit of pH from 6.5 to 8.5. The current investigation ranges were 6.52–6.83 which are in the range of WHO standards.
2. Hardness: Hardness is mainly caused by calcium and magnesium salts. These salts are dissolved from geologic deposits through which water travels. The length of time water is in contact with hardness producing material helps determine how much hardness there is in raw water. Hardness was originally defined as the capacity of water to precipitate soap caused by Calcium and Magnesium.
3. Solids (Total dissolved solids - TDS): Water has the ability to dissolve a wide range of inorganic and some organic minerals or salts such as potassium, calcium, sodium, bicarbonates, chlorides, magnesium, sulfates etc. These minerals produced un-wanted taste and diluted color in appearance of water. This is the important parameter for the use of water. The water with high TDS value indicates that water is highly mineralized. Desirable limit for TDS is 500 mg/l and maximum limit is 1000 mg/l which prescribed for drinking purpose.
4. Chloramines: Chlorine and chloramine are the major disinfectants used in public water systems. Chloramines are most commonly formed when ammonia is added to chlorine to treat drinking water. Chlorine levels up to 4 milligrams per liter (mg/L or 4 parts per million (ppm)) are considered safe in drinking water.
5. Sulfate: Sulfates are naturally occurring substances that are found in minerals, soil, and rocks. They are present in ambient air, groundwater, plants, and food. The principal commercial use of sulfate is in the chemical industry. Sulfate concentration in seawater is about 2,700 milligrams per liter (mg/L). It ranges from 3 to 30 mg/L in most freshwater supplies, although much higher concentrations (1000 mg/L) are found in some geographic locations.
6. Conductivity: Pure water is not a good conductor of electric current rather’s a good insulator. Increase in ions concentration enhances the electrical conductivity of water. Generally, the amount of dissolved solids in water determines the electrical conductivity. Electrical conductivity (EC) actually measures the ionic process of a solution that enables it to transmit current. According to WHO standards, EC value should not exceeded 400 μS/cm.
7. Organic\_carbon: Total Organic Carbon (TOC) in source waters comes from decaying natural organic matter (NOM) as well as synthetic sources. TOC is a measure of the total amount of carbon in organic compounds in pure water. According to US EPA < 2 mg/L as TOC in treated / drinking water, and < 4 mg/Lit in source water which is use for treatment.
8. Trihalomethanes: THMs are chemicals which may be found in water treated with chlorine. The concentration of THMs in drinking water varies according to the level of organic material in the water, the amount of chlorine required to treat the water, and the temperature of the water that is being treated. THM levels up to 80 ppm is considered safe in drinking water.
9. Turbidity: The turbidity of water depends on the quantity of solid matter present in the suspended state. It is a measure of light emitting properties of water and the test is used to indicate the quality of waste discharge with respect to colloidal matter. The mean turbidity value obtained for Wondo Genet Campus (0.98 NTU) is lower than the WHO recommended value of 5.00 NTU.
10. Potability: Indicates if water is safe for human consumption where 1 means Potable and 0 means Not potable.

Acc to mann whitney u

The sample distributions are not equal (reject H0)

the final model will include all the variables.

there are no mutually correlated signs, outliers can be removed, and class imbalance can also be eliminated.