

Health Risk Checker

Problem Area

- I am interested in creating a tool that helps people understand the health risks of diabetes and heart disease.
- Many people don't know they might be at risk of serious health issues until it's too late – Like me when i was diagnosed with diabetes, its level was 22. This project can give them a quick and easy way to check their risk and take action sooner.

The User

- As we all know, Canada has the shortage of family doctors. People who don't have access to regular medical checkups or want to know if they should see a doctor.
- They will get an idea of their health risks by simply entering basic health information like age, glucose levels, and blood pressure. If they are at risk, they will get simple advice to help them take care of their health.

The Big Idea

- Machine learning can look for patterns in health data and predict if someone is at risk of diabetes or heart disease. For example, based on data like high glucose or cholesterol levels, it can show if the user might need medical attention.
- I can find some projects which are only for either diabetes or Heart. I want to build the same tool to predict both. Later, if time allows, it can include Natural Language Processing (NLP) to give personalized advice in simple language.

The Impact

- This tool will help people take care of their health early, which will help them to avoid the major health risks in the future.
- It can also help doctors by reducing their stress by helping people decide when they really need a checkup.

How big is the problem?

- According to data from the Canadian Chronic Disease Surveillance System (CCDSS) for 2017–2018, approximately 1 in 12 Canadian adults aged 20 and over (or 2.6 million people) live with diagnosed heart disease. Alarming, every hour, about 14 Canadian adults in this age group with diagnosed heart disease lose their lives. [link](#)
- Between 2007 and 2019, data for adults aged 20–79 revealed that 1.9% had undiagnosed diabetes (2.3% males, 1.5% females), and 22.5% of individuals who met the criteria for diabetes were unaware of their condition. [link](#)
- These statistics highlight the urgent need for tools that help individuals monitor their health and catch potential issues early. This is why I want to work on the Health Checker project, to create a system that empowers people to better understand and manage their health before it's too late.

The Data

I need to decide on the dataset from the below options

- 1. Diabetes Dataset - [Almost Confirmed]**
 - a. **Source:** Kaggle
 - b. **Link:** <https://www.kaggle.com/datasets/mathchi/diabetes-data-set>
- 2. Heart Disease Dataset [Almost Confirmed]**
 - a. **Source:** Kaggle
 - b. **Link:** <https://www.kaggle.com/datasets/alexteboul/heart-disease-health-indicators-dataset>
- 3. Heart Disease Dataset [Optional]**
 - a. **Source:** UCI
 - b. **Link:** <https://archive.ics.uci.edu/dataset/45/heart+disease>
- 4. Heart Disease Dataset [Backup]**
 - a. **Source:** OpenML
 - b. **Link:** <https://www.openml.org/search?type=data&status=active&id=45950>
- 5. Diabetes Dataset [Backup]**
 - a. **Source:** OpenML
 - b. **Link:** <https://www.openml.org/search?type=data&status=active&id=46254>

6. Diabetes Dataset [Optional]

- a. **Source:** UCI
- b. **Link:** <https://archive.ics.uci.edu/dataset/891/cdc+diabetes+health+indicators>

The Alternative

- I had considered focusing on raising awareness about traffic rules as another area of interest. However, for this capstone project, the scope of such a topic is quite broad, and finding a relevant and comprehensive dataset to support the work has proven to be a significant challenge.