Diabetes Determinator

Project Proposal

Mihail Kenarov

Table of Contents

[1. Introduction 3](#_Toc170600953)

[What is the goal? 3](#_Toc170600954)

[Why should such a model exist? 3](#_Toc170600955)

[Who would benefit from such a system? 3](#_Toc170600956)

[When will the project be realized? 3](#_Toc170600957)

[How will the project be done? 4](#_Toc170600958)

[2. Domain Understanding 4](#_Toc170600959)

[2.1 What is diabetes? 4](#_Toc170600960)

[2.2 Spread of diabetes. 4](#_Toc170600961)

[2.3 History of diabetes and findings. 5](#_Toc170600962)

[2.4 Treatment for diabetes. 6](#_Toc170600963)

[2.4.1 Treatments for type 1 Diabetes 6](#_Toc170600964)

[2.4.2 Treatments for type 2 Diabetes 8](#_Toc170600965)

[2.5 Causes, risks and symptoms 9](#_Toc170600966)

[2.5.1 Risks and causes 10](#_Toc170600967)

[2.5.2 Symptoms 10](#_Toc170600968)

[2.6 Diagnosis 11](#_Toc170600969)

[2.7 The healthcare system. 12](#_Toc170600970)

[2.7.1 What is the healthcare system? 13](#_Toc170600971)

[2.8 World Health Organization 13](#_Toc170600972)

[2.8.1 What is the role of WHO 13](#_Toc170600973)

[2.8.2 Ethical guidelines for AI in healthcare 14](#_Toc170600974)

[2.8.3 Investments and fundings 17](#_Toc170600975)

[2.9 Pros and cons 17](#_Toc170600976)

[2.9.1 Pros 17](#_Toc170600977)

[2.9.2 Cons 18](#_Toc170600978)

[2.8.3 Conclusion 18](#_Toc170600979)

[3. Data sourcing 19](#_Toc170600980)

[3.1 Objective 19](#_Toc170600981)

[3.2 Data requirements 19](#_Toc170600982)

[3.3 Data sources 20](#_Toc170600983)

[3.4 Data Legality and Ethics 20](#_Toc170600984)

[3.5 Data Diversity 20](#_Toc170600985)

[3.6 Version Control 23](#_Toc170600986)

[3.7 Iterative Process 23](#_Toc170600987)

[4. Analytic Approach 23](#_Toc170600988)

[5. Bibliography 24](#_Toc170600989)

# 1. Introduction

## What is the goal?

The plan is to develop a machine learning model to predict diabetes risk using comprehensive patient data.

## Why should such a model exist?

The motivation behind this project is the global rise in diabetes prevalence over the past few decades, which poses significant long-term health risks. Developing a predictive model for diabetes can serve as an early warning system, enabling timely intervention and management to mitigate complications and improve patient outcomes. This project addresses a critical health issue and aims to contribute to better preventative healthcare.

## Who would benefit from such a system?

Everybody who is interested in the safety of their health would benefit from such a system. The only people who could be excluded from the targeted audience are the people who already have it – either developed or inherited.

## When will the project be realized?

* Phase 1 – Research and planning – Week 17
  + Research about the project will be conducted in the first couple of weeks of the semester, to get a better understanding of recommendation systems and what would be the most suitable approach in our case.
* Phase 2 – Development – Week 17 & Week 18
  + During that time, the data will be edited and structured properly for the selected algorithm, which will also be chosen during this time, based on what currently matches our expectations.
* Phase 3 – Delivery and evaluation – Week 17 & Week 18
  + After the creation and testing of the project it will be evaluated to see if it matches up to our expectations.

## How will the project be done?

During the preprocessing, will examine the distinct kinds of patient data - patient ID, demographic details, lifestyle factors, medical history, clinical measurements, medications, symptoms, quality of life scores, environmental exposures, health behaviours, and diagnosis information. Afterwords, with the insights that we have gathered, the model will analyse these factors to identify patterns and risk indicators associated with diabetes.

# 2. Domain Understanding

To develop such a system, it is essential to first comprehend the significance of diabetes and why it is a critical factor and disease that necessitates our attention. Not only that but it is crucial for us to understand how the healthcare system has developed and fought against such sickness.

## 2.1 What is diabetes?

Diabetes is a chronic medical condition where the body cannot properly regulate blood sugar levels. There are two main types: Type 1 diabetes, where the immune system attacks insulin-producing cells in the pancreas, and Type 2 diabetes, which develops when the body becomes resistant to insulin, or the pancreas cannot produce enough insulin. Both types require careful management to prevent serious health complications. [(Diabetes Fonds, n.d.)](#_5._Bibliography)

There are many possible complications that can occur if any type of diabetes is left untreated:

* Type 1 diabetes complications – headaches, dry skin and mouth, stomach pain and without treatment – death [(Basina, 2022)](#_5._Bibliography)
* Type 2 diabetes complications - kidney damage, eye damage, increased risk of hear disease or stroke [(UnityPoint Health, n.d.)](#_5._Bibliography)

By taking this into consideration, it is important to understand the urgency of an available quick evaluation on the possibility of diabetes. With a program where one can insert some of their medical data, which then can be quickly assessed, they can take measures in time and contact a doctor for an appropriate appointment.

## 2.2 Spread of diabetes.

As of 2021 over 537 million adults (20-79) were registered to have diabetes – that was one in every 10 people, and this continues to grow. It is estimated that the rise of the illness will continue and by 2045 it will reach approximately 783 million people or 1 in 8, worldwide. Having these numbers gives helps us understand the gravity of the situation. It is a phenomenon which can happen to anyone. [(International Diabetes Federation, 2021)](#_5._Bibliography)

The documented death cases attributed to diabetes in 2011 were 4.6 million. This number has increased to 6.7 million by 2021, highlighting the exponential growth and severity of the issue. These statistics underscore the escalating prevalence of diabetes and the increasing number of individuals affected by this condition. [(International Diabetes Federation, 2021)](#_5._Bibliography)

In summary, there is a great amount of evidence, which supports the increasing prevalence of diabetes and its growing impact on global health, highlight the urgency of addressing this issue. The significant rise in diabetes-related deaths over the past decade further emphasizes the severity of the situation. These trends illustrate the critical need for enhanced global efforts in prevention, management, and treatment to mitigate the impact of diabetes and improve health outcomes.

## 2.3 History of diabetes and findings.

Some of the earliest documented treatments for diabetes are referenced in the Ebers Papyrus from around 1500 BC, identifying the condition by symptoms such as unquenchable thirst, weight loss, and frequent urination. Subsequently, in the 5th and 6th centuries BC, the use of ants was discovered as a method to diagnose diabetes. Initial therapeutic approaches included incorporating specific seeds into the diet, fasting regimens, the use of certain narcotics and opiates, and adherence to carbohydrate-restricted diets. [(Diabetes: Past treatments, new discoveries, 2023)](#_5._Bibliography)

One of the earliest individuals credited with identifying different types of diabetes, yet often overlooked in discussions of the disease, is Johann Peter Frank. Frank is undoubtedly one of the most influential figures in the early history of public health and community medicine. In addition to his seminal work, "A Complete System of Medical Policy," Frank authored a seven-volume textbook on internal medicine and made significant clinical discoveries, including the distinction between diabetes mellitus and diabetes insipidus. [(Last, n.d.)](#_5._Bibliography)

In 1889, Joseph von Mering and Oskar Minkowski discovered that the removal of the pancreas from dogs resulted in fatal diabetes, providing the first indication that the pancreas plays a crucial role in regulating glucose concentrations. In 1910, Edward Albert Sharpey-Schafer hypothesized that diabetes was caused by the deficiency of a single chemical produced by the pancreas.

In 1921, Frederick Banting and Charles Best discovered insulin by reversing diabetes induced in dogs using an extract from the pancreatic islet cells of healthy dogs. Along with James Collip and John Macleod, they purified the hormone insulin from bovine pancreases and were the first to use it to treat a patient with diabetes. The production of insulin and its therapeutic application rapidly spread worldwide. [(UMass Diabetes Center of Excellence, n.d.)](#_5._Bibliography)

This series of events represents a remarkable example of the swift translation of a basic science discovery into a significant benefit for patients. Once insulin injections became available, young individuals with insulin deficiency, who had previously faced almost certain and painful death within weeks to months, were able to survive for extended periods. [(The Past 200 Years in Diabetes, 2012)](#_5._Bibliography)

The historical progression of diabetes understanding and treatment illustrates a profound journey from ancient diagnostic methods and rudimentary treatments to groundbreaking scientific discoveries and life-saving therapies. From the early identification of diabetes symptoms in the Ebers Papyrus to Johann Peter Frank's critical differentiation of diabetes types, and culminating in the revolutionary discovery and therapeutic use of insulin by Banting, Best, Collip, and Macleod, each milestone has significantly enhanced the management of this chronic condition. These advances underscore the critical importance of scientific research in translating fundamental discoveries into practical medical applications, profoundly improving patient outcomes and transforming the prognosis for those with diabetes.

## 2.4 Treatment for diabetes.

The increasing incidence of diabetes does not imply a lack of preventive measures. Healthcare systems worldwide are collaborating to support patients in managing their illness and ensuring they receive the highest quality of treatment depending on the diabetes that a person has, regardless of their country of origin. However, this does come at a cost.

### 2.4.1 Treatments for type 1 Diabetes

Type 1 diabetes necessitates insulin administration because the pancreas fails to produce this hormone. Individuals with type 1 diabetes must administer insulin multiple times daily, typically coordinating doses with meals to regulate blood glucose levels effectively. [(Insulin, Medicines, & Other Diabetes Treatments, 2022)](#_5._Bibliography)

#### 2.4.1.1 Types of insulin that can be taken for type 1 diabetes.

* ***Rapid-acting Insulin*** - starts working within about 15 minutes of injection, with its peak effectiveness occurring around 60 minutes afterward. It remains active for approximately 4 hours. This insulin type is usually administered 15 to 20 minutes before meals to control post-meal blood sugar spikes. Examples of rapid-acting insulin include glulisine (Apidra), lispro (Humalog, Admelog, Lyumjev), and aspart (Novolog, FiAsp).
* ***Short-acting Insulin (Regular Insulin)*** - also known as regular insulin, begins to work approximately 30 minutes after injection. It reaches its peak effectiveness between 90 to 120 minutes and typically lasts for about 4 to 6 hours. Examples of short-acting insulin include Humulin R, Novolin R, and Afrezza. This type of insulin is often used to manage blood glucose levels during meals.
* **Intermediate-acting Insulin (NPH Insulin)** - also known as NPH insulin, typically starts working within 1 to 3 hours of injection. Its peak effectiveness is reached between 6 to 8 hours, and it can last for 12 to 24 hours. Examples of NPH insulin include Novolin N and Humulin N. This type of insulin provides a longer duration of action compared to short-acting and rapid-acting insulins, making it suitable for maintaining basal insulin levels between meals and overnight.
* **Long- and Ultra-long-acting Insulin** - Long-acting and ultra-long-acting insulins provide extended coverage, ranging from 14 to 40 hours. They are designed to provide a steady level of insulin over an extended period, reducing the need for multiple daily injections. Examples include glargine (Lantus, Toujeo Solostar, Basaglar), detemir (Levemir), and degludec (Tresiba). These insulins are crucial for maintaining basal insulin needs throughout the day and night, offering greater flexibility and stability in managing blood glucose levels.

#### 2.4.1.2 Options for delivering insulin.

* ***Injections*** - Insulin can be injected using a fine needle and syringe or an insulin pen. Insulin pens, resembling ink pens, are available in both disposable and refillable varieties. When opting for injections, a combination of different types of insulin is typically required for day and night use.
* ***Insulin Pump*** - An insulin pump is a small device worn externally, programmed to deliver specific amounts of insulin throughout the day and during meals. The device connects a reservoir of insulin to a catheter inserted under the skin of the abdomen. There is also a tubeless pump option, which involves wearing a pod containing the insulin on the body, with a small catheter inserted under the skin.

#### 2.4.1.3 Monitoring blood sugar

Depending on the type of insulin therapy required, it is necessary to check and record blood sugar levels at least four times daily. Careful monitoring ensures that blood sugar levels remain within the target range. Despite a strict insulin regimen and dietary schedule, blood sugar levels can fluctuate due to factors such as food intake, physical activity, illness, medications, stress, hormonal changes, and alcohol consumption.

* ***Continuous glucose monitoring (CGM)*** - devices track blood sugar levels and are particularly effective in preventing hypoglycemia and lowering A1C. These monitors attach to the body with a fine needle just under the skin, providing glucose readings every few minutes.
* ***Closed loop system*** - A closed loop system links a continuous glucose monitor to an insulin pump, automatically adjusting insulin delivery based on blood sugar readings. The Food and Drug Administration has approved several hybrid systems for type 1 diabetes, which require some user input, such as reporting carbohydrate intake and confirming blood sugar levels. Fully autonomous closed loop systems are not yet available but are currently in clinical trials.
* **Finger-Prick Blood Sugar Testing** - A finger-prick test allows the user to regularly check their blood sugar levels, either at home or while on the go. This test indicates whether the blood sugar levels are too low, too high, or within the normal range. The test strip from the finger-prick is inserted into a device called a glucometer, which reads and displays the blood sugar level on a screen. [(MyWay digital health, n.d.)](#_5._Bibliography)
* **HbA1c Blood Test** - The HbA1c blood test assesses whether the blood sugar levels have been too high, on target, or returned to normal over the past three months. It measures the amount of blood sugar attached to hemoglobin (Hb), the protein in red blood cells that carries oxygen throughout the body. Since red blood cells have an average lifespan of three months, the HbA1c test provides clinical teams with valuable information about the average blood sugar level over this period. [(MyWay digital health, n.d.)](#_5._Bibliography)

Managing diabetes with insulin often involves a combination of long-acting insulin to provide basal coverage and rapid-acting insulin to address mealtime spikes in blood sugar. This regimen mimics the body's natural insulin production more closely than older regimens that required fewer injections per day. Studies have shown that utilizing three or more insulin injections daily can significantly improve blood sugar control and help individuals with diabetes achieve better long-term health outcomes. [(Mayo Clinic, 2024)](#_5._Bibliography)

### 2.4.2 Treatments for type 2 Diabetes

Conversely, type 2 diabetes often allows for initial management through lifestyle adjustments such as adopting a balanced diet and engaging in regular physical activity. These measures can help control blood glucose levels, particularly for those managing excess weight. [(Insulin, Medicines, & Other Diabetes Treatments, 2022)](#_5._Bibliography)

#### Types of treatment for type 2 diabetes:

* ***Healthy Eating*** - There is no specific diabetes diet, but maintaining a balanced diet is essential. Focus on regular meal schedules, smaller portions, high-fiber foods, and healthy oils. Limit refined grains, starchy vegetables, sweets, and overall calorie intake. Consulting a registered dietitian can help plan balanced meals, develop healthy habits, and monitor carbohydrate intake to stabilize blood sugar levels.
* ***Physical Activity*** - Exercise is vital for weight management and blood sugar control. There are many kinds of exercise that can be implemented into your regiment, but it is recommended to review possible **aerobic training** for more than 30 minutes, for most days of the week. Another way of exercising is with some resistance training, where the patient should be aiming for 2 to 3 sessions during the week.
* ***Weight loss*** - it improves control of blood sugar, cholesterol, triglycerides, and blood pressure. Even a 5% reduction in body weight can yield significant health benefits, with greater weight loss offering increased advantages. In some cases, losing up to 15% of body weight may be recommended. Consulting your healthcare provider or dietitian is recommended to set appropriate weight-loss goals and receive guidance on lifestyle changes to achieve them.
* ***Diabetes Medications*** - If diet and exercise alone do not maintain your target blood sugar levels, your healthcare provider may prescribe medications to lower glucose levels or recommend insulin therapy. Medications for type 2 diabetes include the following options – Metformin, Sulfonylureas, Glinides, Thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists.

In conclusion, while diabetes management has advanced significantly, encompassing collaborative global efforts and a range of treatment options tailored to different types of diabetes, challenges persist, notably financial implications. Type 1 diabetes necessitates lifelong insulin therapy, involving significant daily management, while type 2 diabetes often begins with lifestyle adjustments but may progress to requiring medications or insulin.

## 2.5 Causes, risks and symptoms

To comprehensively discuss the causes of diabetes, it is essential to first understand the nature of the condition. Diabetes arises due to a deficiency of insulin in the body, which can occur either because the body simply does not produce sufficient insulin or because the body has developed resistance to it. Numerous factors contribute to the increasing prevalence of this health issue, but the exact reason for its development is still not fully understood.

### 2.5.1 Risks and causes

* ***Genetics and family history –*** In the majority of type 1 diabetes cases, individuals must inherit risk factors from both parents. It is hypothesized that these factors are more prevalent among white populations, as they exhibit the highest incidence of type 1 diabetes. Since most individuals with these risk factors do not develop diabetes, researchers are keen to identify the specific environmental triggers. One potential trigger is exposure to cold weather. While the causes of type 1 and type 2 diabetes differ, two critical factors are common to both: a genetic predisposition to the disease and an environmental trigger that activates it. [(About Diabetes, Genetics of Diabetes, n.d.)](#_5._Bibliography)
* ***Age*** - Age is a significant factor in the risk of developing type 2 diabetes, with the likelihood increasing notably after the age of 35. Type 1 diabetes, on the other hand, can manifest at any age but shows two distinct incidence peaks. The first peak occurs in children aged between 4 and 7 years, while the second peak is observed in children aged between 10 and 14 years.
* ***Inactivity and obesity*** – As the world is ever evolving, we frequently hear about the world hunger that needs to be solved, however in the case of diabetes, it is more of a problem with overeating, than the Inactivity and weight are significant risk factors. Being overweight or obese increases the risk of developing diabetes. Additionally, a sedentary lifestyle further exacerbates this risk. Regular physical activity aids in weight management, utilizes glucose for energy, and enhances cellular sensitivity to insulin.

### 2.5.2 Symptoms

The symptoms of both types of diabetes are quite similar. However, in type 1 diabetes, the symptoms can develop suddenly, whereas in type 2 diabetes, they may develop gradually over several years and often go unnoticed.

* Symptoms for type 1 diabetes include:
  + Excessive thirst
  + Frequent urination
  + Bed-wetting in children who previously did not wet the bed at night.
  + Intense hunger
  + Unintentional weight loss
  + Irritability or mood changes
  + Fatigue and weakness
  + Blurry vision
* Symptoms for type 2 diabetes include:
  + Increased thirst
  + Frequent urination
  + Increased hunger
  + Unintended weight loss
  + Fatigue
  + Blurred vision
  + Slow-healing sores
  + Frequent infections
  + Numbness or tingling in the hands or feet.
  + Areas of darkened skin, usually in the armpits and neck

In any case, in the case of type 2 diabetes it is not possible to identify it in a quick notice, however with regular health checkups it is easier to prevent in in the pre-diabetic stage. As for type 1 diabetes, if the sudden changes occur it is important to seek immediate medical attention.

## 2.6 Diagnosis

There are multiple diagnostic methods available for the diagnosis of diabetes, many of which are applicable to both types of diabetes. Additionally, specific tests exist that are particularly useful in the diagnosis of type 2 diabetes.

* ***Glycated Hemoglobin (A1C) Test -*** This blood test indicates your average blood sugar level over the past 2 to 3 months by measuring the amount of sugar attached to hemoglobin, the oxygen-carrying protein in red blood cells. Higher blood sugar levels result in more sugar attached to hemoglobin. An A1C level of 6.5% or higher on two separate tests confirms a diagnosis of diabetes.

If the A1C test is unavailable, or if conditions such as pregnancy or the presence of a hemoglobin variant (uncommon form of hemoglobin) may affect its accuracy, alternative tests include:

* ***Random Blood Sugar Test:*** A blood sample is taken at a random time, with confirmation from additional tests. Blood sugar values are expressed in milligrams per deciliter (mg/dL) or millimoles per liter (mmol/L). Regardless of when you last ate, a q
* ***Fasting Blood Sugar Test:*** A blood sample is taken after an overnight fast. A fasting blood sugar level below 100 mg/dL (5.6 mmol/L) is considered normal. Levels between 100 and 125 mg/dL (5.6 to 6.9 mmol/L) indicate prediabetes. A fasting blood sugar level of 126 mg/dL (7 mmol/L) or higher on two separate tests confirms a diagnosis of diabetes.

Below are additional diagnostic approaches available for the diagnosis of type 2 diabetes [(Mayo Clinic, 2023):](#_5._Bibliography)

* ***The oral glucose tolerance test –*** This is less frequently employed compared to other diagnostic methods, except in cases of pregnancy. It involves fasting for a specified duration followed by the consumption of a sugary solution at a healthcare provider's office. Blood sugar levels are subsequently monitored at intervals over a two-hour period. Results are interpreted as follows:
  + Less than 140 mg/dL (7.8 mmol/L) after two hours is indicative of a healthy response.
  + Levels ranging from 140 to 199 mg/dL (7.8 to 11.0 mmol/L) indicate prediabetes.
  + A reading of 200 mg/dL (11.1 mmol/L) or higher after two hours suggests diabetes.
* ***Screening*** – Screening for type 2 diabetes is recommended by the American Diabetes Association as a routine practice for all adults aged 35 years and older, as well as for specific demographic groups, including:
  + Individuals younger than 35 who are overweight or obese and exhibit one or more risk factors associated with diabetes.
  + Women with a history of gestational diabetes.
  + Individuals previously diagnosed with prediabetes.
  + Overweight or obese children with a family history of type 2 diabetes or other identified risk factors.

## 2.7 The healthcare system.

Neglecting to gain a comprehensive understanding of the healthcare system would be a significant oversight, given that the contributions and findings of its members are crucial for the treatment of diabetes and potential future advancements in combating the disease.

### 2.7.1 What is the healthcare system?

A health system, healthcare system, or health care system is an organization of people, institutions, and resources designed to deliver healthcare services to meet the health needs of target populations.

Globally, there is a wide variety of health systems, each with unique histories and organizational structures shaped by their respective nations. Nations must design and develop health systems according to their specific needs and resources, though common elements across nearly all health systems include primary healthcare and public health measures.

In some countries, health system planning is decentralized, with various market stakeholders assuming responsibilities. Conversely, in other regions, healthcare service provision is a collaborative effort involving governmental entities, labor unions, philanthropic organizations, religious institutions, and other organized bodies, each working to address the specific needs of their populations. It is important to note that healthcare planning is often an evolutionary process rather than a revolutionary one.

Like other social institutional structures, health systems are influenced by the history, culture, and economics of the states in which they evolve. These unique characteristics complicate international comparisons and prevent the establishment of a universal standard of performance. [(Wikipedia, n.d.)](#_5._Bibliography)

## 2.8 World Health Organization

The World Health Organization (WHO), as the directing and coordinating authority on international health within the United Nations system, adheres to the UN values of integrity, professionalism, and respect for diversity.

The values of the WHO workforce also reflect the principles of human rights, universality, and equity established in WHO’s Constitution, along with the ethical standards of the Organization. These values are inspired by WHO’s vision of a world in which all people attain the highest possible level of health and its mission to promote health, ensure global safety, and serve the vulnerable, with measurable impact at the country level. [(World Health Organization, n.d.)](#_5._Bibliography)

### 2.8.1 What is the role of WHO

The leading role of WHO is health systems governance, which refers to the processes, structures, and institutions responsible for overseeing and managing a country's healthcare system. It encompasses the relationships between various actors and stakeholders, including government agencies, healthcare providers, patients and their families, communities, civil society organizations, and private sector entities.

Health systems governance involves ensuring the existence of strategic policy frameworks, effective oversight, coalition-building, appropriate regulations and incentives, system design considerations, and accountability mechanisms.

The primary objective of effective health systems governance is to ensure that healthcare services are accessible, equitable, efficient, affordable, and of high quality for all individuals. Achieving this requires:

* ***Efficient and Equitable Allocation of Resources:*** Ensuring that healthcare resources are distributed in a manner that meets the needs of the entire population.
* ***Policy and Regulatory Frameworks:*** Developing and implementing policies and regulations to guide healthcare delivery
* ***Monitoring and Evaluation:*** Establishing mechanisms to monitor, evaluate, and review the performance of the healthcare system.

Effective governance is crucial for promoting equity and social justice in healthcare. It aims to make the healthcare system responsive to the needs of all societal members, regardless of socioeconomic status, ethnicity, culture, gender, or other factors. [(World Health Organization, n.d.)](#_5._Bibliography)

Ultimately, WHO's role in health systems governance is essential. It involves overseeing and managing healthcare systems through strategic policy frameworks, effective oversight, coalition-building, regulations, incentives, system design, and accountability. This ensures that healthcare services are accessible, equitable, efficient, affordable, and high-quality for all. Effective governance promotes equity and social justice, making healthcare responsive to the needs of all societal members.

### 2.8.2 Ethical guidelines for AI in healthcare

The World Health Organization has recognized artificial intelligence (AI) as a tool with the potential to significantly enhance the capabilities of health-care providers. AI can improve patient care, provide accurate diagnoses, optimize treatment plans, support pandemic preparedness and response, inform health policy decisions, and efficiently allocate resources within health systems. To fully realize this potential, health-care professionals and systems must have comprehensive information on the contexts in which AI can be safely and effectively utilized, the necessary conditions for its reliable and appropriate use, and the mechanisms for continuous auditing and performance assessment.

However, The World Health Organization's new report highlights the need for caution regarding the benefits of artificial intelligence in health care, emphasizing that these should not overshadow essential investments and strategies for achieving universal health coverage. The report identifies several challenges and risks associated with AI, including unethical data collection and use, algorithmic biases, and threats to patient safety, cybersecurity, and the environment.

It warns that unregulated AI use might prioritize the interests of technology companies or government surveillance over the rights of patients and communities. AI systems developed using data from high-income countries may not perform effectively in low- and middle-income settings, necessitating careful design to reflect diverse socio-economic and health-care contexts.

Ensuring ethical principles are integrated into AI technology requires collaboration among governments, providers, and designers, guided by existing laws and human rights obligations, as well as new laws and policies are stated as key factors, when talking about the possibility of implementing AI in healthcare.

#### 2.8.2.1 The Principles

To mitigate the risks and maximize the benefits inherent in the use of artificial intelligence (AI) for health, the World Health Organization (WHO) outlines the following principles for AI regulation and governance: [(World Health Organization , 2021 )](#_5._Bibliography)

* ***Protecting Human Autonomy:*** In health care, humans must retain control over health-care systems and medical decisions. Privacy and confidentiality should be safeguarded, and patients must provide valid informed consent through appropriate legal frameworks for data protection.
* ***Promoting Human Well-Being, Safety, and the Public Interest: AI*** technologies must meet regulatory standards for safety, accuracy, and efficacy for specific use cases or indications. Quality control measures and mechanisms for quality improvement in AI use must be established.
* ***Ensuring Transparency, Explainability, and Intelligibility:*** Transparency requires the publication or documentation of sufficient information before designing or deploying an AI technology. This information must be easily accessible and support meaningful public consultation and debate on the technology's design and appropriate use.
* ***Fostering Responsibility and Accountability:*** Stakeholders are responsible for ensuring AI technologies are used under suitable conditions and by adequately trained personnel. Effective mechanisms for questioning and redress must be available for individuals and groups adversely affected by algorithm-based decisions.
* ***Ensuring Inclusiveness and Equity:*** AI for health should be designed to promote equitable use and access across all demographics, regardless of age, sex, gender, income, race, ethnicity, sexual orientation, ability, or other characteristics protected under human rights codes.
* ***Promoting Responsiveness and Sustainability:*** AI applications should be continuously and transparently assessed during actual use to ensure they meet expectations and requirements. AI systems must be designed to minimize environmental impact and enhance energy efficiency. Governments and companies should address anticipated workplace disruptions, including training health-care workers to adapt to AI systems and managing potential job losses due to automation.

#### 2.8.2.2 World health organization’s opinion on diabetes and AI

***Predictive Capabilities of AI in Healthcare***

Artificial Intelligence has the potential to be utilized for predicting illnesses or major health events before they occur. For instance, AI technology could be adapted to assess the relative risk of disease, which could subsequently be employed for the prevention of lifestyle diseases such as cardiovascular disease and diabetes.

***Ethical Considerations in AI-Driven Predictive Health Care***

However, significant ethical concerns may arise regarding a specific AI technology or its use case, even if it provides accurate and valuable information. For example, an AI technology capable of predicting individuals likely to develop type 2 diabetes or HIV infection could benefit at-risk individuals or communities. However, it could also lead to unnecessary stigmatization, questioning or criminalization of individuals' choices and behaviors, over-medicalization of otherwise healthy individuals, unnecessary stress and anxiety, and exposure to aggressive marketing by pharmaceutical companies and other for-profit health-care services. Additionally, certain AI technologies, if not carefully deployed, could exacerbate health care disparities related to ethnicity, socioeconomic status, or gender. While a shift toward patient-based care may be empowering and beneficial for some, others may find the increased responsibility stressful, and it could potentially limit access to formal health-care services.

### 2.8.3 Investments and fundings

The estimated global direct health expenditure on diabetes was USD 760 billion in 2019 and is projected to increase to USD 825 billion by 2030 and USD 845 billion by 2045. There exists significant variability in annual health expenditures related to diabetes. The United States of America leads with an estimated expenditure of USD 294.6 billion, followed by China with USD 109.0 billion and Brazil with USD 52.3 billion. Among age groups, individuals aged 60-69 years incur the highest annual diabetes-related health expenditure at USD 177.7 billion, followed by those aged 50-59 years and 70-79 years, with expenditures of USD 173.0 billion and USD 171.5 billion, respectively. Women demonstrate slightly higher diabetes-related health expenditures compared to men (USD 382.6 billion versus USD 377.6 billion, respectively), a trend expected to persist through 2030 and 2045. [(National Library of Medicine, 2020)](#_5._Bibliography)

The American Diabetes Association (ADA) released the Economic Costs of Diabetes in the U.S. in 2022 (Economic Report), a comprehensive analysis examining the economic impact of diabetes in the United States. Conducted every five years, the Economic Report reveals that the total annual cost of diabetes in 2022 amounts to $412.9 billion. This figure includes $306.6 billion in direct medical costs and $106.3 billion in indirect costs. Individuals diagnosed with diabetes represented one-quarter of all healthcare expenditures in the United States. [(American Diabetes Association, 2023)](#_5._Bibliography)

## 2.9 Pros and cons

### 2.9.1 Pros

* ***Early Detection and Prevention*** – The development of such a system promises significant utility in the prevention of type 2 diabetes. This system will operate based on user-provided inputs to calculate the likelihood of the illness or prediabetes. Upon identifying potential prediabetic conditions, it is advisable for individuals to seek medical consultation for further evaluation. This proactive approach may prevent progression to type 2 diabetes, a condition for which a definitive cure remains elusive.

Additionally, there are instances where the healthcare system covers the cost of diabetes treatment, as seen in countries like the Netherlands and Bulgaria. Early identification of the condition can lead to significant cost savings, which could then be redirected into further research aimed at discovering a permanent cure for diabetes.

* ***Integration with Electronic Health Records (EHRs)*** – With the possibility of integrating EHRs into the system authorized users/doctors can see the more detailed information about the patient they are working with and more in-depth information about the patient’s past readings.
* ***Scalability and reach -*** The development of such a system holds the potential for implementation in underserved areas where individuals may struggle to recognize possible diabetes symptoms. By creating an accessible device or application, this technology can facilitate a better understanding of diabetes, making it easier to reach a broader audience who might be at risk of developing the condition. This approach ensures that even those in less fortunate circumstances can benefit from early detection and preventative measures, ultimately contributing to better health outcomes.

### 2.9.2 Cons

* ***Patient Resistance and Acceptance -*** Some patients may exhibit hesitancy or resistance toward trusting and such diagnostic tools, favoring traditional healthcare methods instead. Concerns about the reliability and accuracy of these advanced technological solutions can contribute to this reluctance. Patients may also express apprehensions regarding the potential depersonalization of healthcare, fearing that the human touch in medical interactions might diminish as technology takes a more central role.
* ***Accuracy and Reliability*** - Machine learning (ML) models, while powerful, may not always achieve optimal accuracy and reliability, potentially resulting in incorrect predictions or diagnoses. The performance of these models is heavily influenced by the quality and variability of the data used for training. Inconsistent or poor-quality data can lead to inaccurate outputs, undermining the effectiveness of ML applications in healthcare and putting the health of the patient at greater risk.
* ***Data Privacy and Security Risks*** - Handling sensitive health data in the context of machine learning (ML) systems poses significant privacy and security risks, including the potential for data breaches, unauthorized access, and misuse. These risks can compromise patient confidentiality and undermine trust in both the technology and the broader healthcare system.

### 2.8.3 Conclusion

Regarding the potential implementation and guidance for patients using such systems, it is recommended that the application be endorsed and explained by a physician. This involves providing recommendations on optimal usage times, frequency of application use, and systematic data input procedures. Instructions should be clearly delineated within the application interface to remind patients of the correct operational protocols. This approach aims to mitigate the potential for misuse.

In terms of informing the user, if the results following system usage indicate cause for concern, the system should display a message advising the user to schedule an appointment with a physician. Conversely, if the system determines that everything is within normal parameters, it should greet the user with an informative message confirming that all conditions are as expected.

In summary, the creation of such a system would be beneficial for individuals who are not accustomed to regular doctor appointments. In regions where regular medical checkups are not considered necessary, such as Bulgaria, the system could serve as an indicator prompting individuals to schedule a doctor's appointment. However, in areas where consistent health checkups are the norm, the system's utility may be limited. Additionally, it is important to note that this model cannot replace a professional diagnosis obtained through a doctor's appointment and comprehensive blood work.

# 3. Data sourcing

## 3.1 Objective

The goal of this project is to determine if a certain person, based on different kinds of data is sustainable to developing either kind of diabetes. This will be done with

## 3.2 Data requirements

To achieve such a system, we will need different kinds of mostly medical data, related to the person. Not only that be we would probably need some kind of lifestyle data since it is something that should be taken into consideration, when trying to predict something of this sort. General data would also be nice, something like the gender and age as an example to start with. The BMI also could also be useful since it can be affected by/affect diabetes. Sport activities may be somehow related, since there is research that shows that people who do more physical activity are less likely to develop the illness. Also, blood sugar levels should be of great help when it comes to determining the possibly of development.

* ***Age – Number***
* ***Gender – Male or Female – 0/1***
* ***BMI – Number***
* ***Smoking – yes/no – 1/0***
* ***Activity – yes/No – 1/0***
* ***Blood Sugat Levels – Number***

## 3.3 Data sources

Publicly available dataset:

<https://www.kaggle.com/datasets/rabieelkharoua/diabetes-health-dataset-analysis>

This dataset comprises extensive health information for 1,879 patients, each uniquely identified with IDs ranging from 6000 to 7878. It includes demographic details, lifestyle factors, medical history, clinical measurements, medication usage, symptoms, quality of life scores, environmental exposures, and health behaviors. Each patient is linked to a confidential primary care doctor, ensuring privacy and confidentiality.

## 3.4 Data Legality and Ethics

The data provenance for this dataset is confidential to protect the identities and privacy of the patients involved.

## 3.5 Data Diversity

***Patient Information – All the types of information we will be using.***

* Patient ID
* Demographic Details
* Lifestyle Factors
* Medical History
* Clinical Measurements
* Medications
* Symptoms and Quality of Life
* Environmental and Occupational Exposures
* Health Behaviors
* Diagnosis Information

***Patient Information***

Patient ID

* PatientID: A unique identifier assigned to each patient (6000 to 7878).

Demographic Details

* Age: The age of the patients ranges from 20 to 90 years.
* Gender: Gender of the patients, where 0 represents Male and 1 represents Female.
* Ethnicity: The ethnicity of the patients, coded as follows:
  + 0: Caucasian
  + 1: African American
  + 2: Asian
  + 3: Other
* SocioeconomicStatus: The socioeconomic status of the patients, coded as follows:
  + 0: Low
  + 1: Middle
  + 2: High
* EducationLevel: The education level of the patients, coded as follows:
  + 0: None
  + 1: High School
  + 2: Bachelor's
  + 3: Higher

***Lifestyle Factors***

* BMI: Body Mass Index of the patients, ranging from 15 to 40.
* Smoking: Smoking status, where 0 indicates No and 1 indicates Yes.
* AlcoholConsumption: Weekly alcohol consumption in units, ranging from 0 to 20.
* PhysicalActivity: Weekly physical activity in hours, ranging from 0 to 10.
* DietQuality: Diet quality score, ranging from 0 to 10.
* SleepQuality: Sleep quality score, ranging from 4 to 10.

***Medical History***

* FamilyHistoryDiabetes: Family history of diabetes, where 0 indicates No and 1 indicates Yes.
* GestationalDiabetes: History of gestational diabetes, where 0 indicates No and 1 indicates Yes.
* PolycysticOvarySyndrome: Presence of polycystic ovary syndrome, where 0 indicates No and 1 indicates Yes.
* PreviousPreDiabetes: History of previous pre-diabetes, where 0 indicates No and 1 indicates Yes.
* Hypertension: Presence of hypertension, where 0 indicates No and 1 indicates Yes.

***Clinical Measurements***

* SystolicBP: Systolic blood pressure, ranging from 90 to 180 mmHg.
* DiastolicBP: Diastolic blood pressure, ranging from 60 to 120 mmHg.
* FastingBloodSugar: Fasting blood sugar levels, ranging from 70 to 200 mg/dL.
* HbA1c: Hemoglobin A1c levels, ranging from 4.0% to 10.0%.
* SerumCreatinine: Serum creatinine levels, ranging from 0.5 to 5.0 mg/dL.
* BUNLevels: Blood Urea Nitrogen levels, ranging from 5 to 50 mg/dL.
* CholesterolTotal: Total cholesterol levels, ranging from 150 to 300 mg/dL.
* CholesterolLDL: Low-density lipoprotein cholesterol levels, ranging from 50 to 200 mg/dL.
* CholesterolHDL: High-density lipoprotein cholesterol levels, ranging from 20 to 100 mg/dL.
* CholesterolTriglycerides: Triglycerides levels, ranging from 50 to 400 mg/dL.

***Medications***

* AntihypertensiveMedications: Use of antihypertensive medications, where 0 indicates No and 1 indicates Yes.
* Statins: Use of statins, where 0 indicates No and 1 indicates Yes.
* AntidiabeticMedications: Use of antidiabetic medications, where 0 indicates No and 1 indicates Yes.

***Symptoms and Quality of Life***

* FrequentUrination: Presence of frequent urination, where 0 indicates No and 1 indicates Yes.
* ExcessiveThirst: Presence of excessive thirst, where 0 indicates No and 1 indicates Yes.
* UnexplainedWeightLoss: Presence of unexplained weight loss, where 0 indicates No and 1 indicates Yes.
* FatigueLevels: Fatigue levels, ranging from 0 to 10.
* BlurredVision: Presence of blurred vision, where 0 indicates No and 1 indicates Yes.
* SlowHealingSores: Presence of slow-healing sores, where 0 indicates No and 1 indicates Yes.
* TinglingHandsFeet: Presence of tingling in hands or feet, where 0 indicates No and 1 indicates Yes.
* QualityOfLifeScore: Quality of life score, ranging from 0 to 100.

***Environmental and Occupational Exposures***

* HeavyMetalsExposure: Exposure to heavy metals, where 0 indicates No and 1 indicates Yes.
* OccupationalExposureChemicals: Occupational exposure to harmful chemicals, where 0 indicates No and 1 indicates Yes.
* WaterQuality: Quality of water, where 0 indicates Good and 1 indicates Poor.

***Health Behaviors***

* MedicalCheckupsFrequency: Frequency of medical check-ups per year, ranging from 0 to 4.
* MedicationAdherence: Medication adherence score, ranging from 0 to 10.
* HealthLiteracy: Health literacy score, ranging from 0 to 10.

***Diagnosis Information (Target Variable)***

* Diagnosis: Diagnosis status for Diabetes, where 0 indicates No and 1 indicates Yes.

## 3.6 Version Control

A history of the project and the processing of the data will be kept on a Git repository in case of an incident and as a way for a version control.

The link to the git repository:

* <https://git.fhict.nl/I509460/diabetes_estemator.git>

## 3.7 Iterative Process

The model will be check for its performance after the iterations. Depending on the results, if they are found are satisfactory or not, changes will be made either to the model or the preparation of the data.

# 4. Analytic Approach

The approach for this project involves gaining a comprehensive understanding of the data through thorough exploratory data analysis. We will focus on features recognized as symptoms and causes of diabetes, including BMI, age, family history, and frequent urination, Fasting Blood Sugar, HbA1c and others.

Following this, we will apply a machine learning model suitable for classification. In this case, we will utilize Gradient Boosting Classifier due to its prominence in the field. Additionally, it has been suggested to explore the potential application of Deep Neural Network algorithms for this system through further research. Finally, we will use a confusion matrix to ensure accurate evaluation and distinction of the model's performance.

# 5. Bibliography

About Diabetes, Genetics of Diabetes. (n.d.). *About Diabetes, Genetics of Diabetes*. Retrieved from American Diabetes Association: https://diabetes.org/about-diabetes/genetics-diabetes

American Diabetes Association. (2023, November 1). *New American Diabetes Association Report Finds Annual Costs of Diabetes to be $412.9 Billion*. Retrieved from American Diabetes Association: https://diabetes.org/newsroom/press-releases/new-american-diabetes-association-report-finds-annual-costs-diabetes-be

Basina, M. (2022, April 11). *How Long Can Your Body Survive Without Insulin?* Retrieved from healthline: https://www.healthline.com/diabetesmine/ask-dmine-lifespan-sans-insulin

Diabetes Fonds. (n.d.). *Type 2 diabetes*. Retrieved from Diabetes Fonds: https://www.diabetesfonds.nl/over-diabetes/soorten-diabetes/diabetes-type-2

Diabetes Fonds. (n.d.). *What is type 1 diabetes?* Retrieved from Diabetes Fonds: https://www.diabetesfonds.nl/over-diabetes/soorten-diabetes/diabetes-type-1

*Diabetes: Past treatments, new discoveries*. (2023, May 4). Retrieved from MedicalNewsToday: https://www.medicalnewstoday.com/articles/317484#early-treatment

*Insulin, Medicines, & Other Diabetes Treatments*. (2022, March). Retrieved from National Institute of Diabetes and Digestive and Kidney Diseases: https://www.niddk.nih.gov/health-information/diabetes/overview/insulin-medicines-treatments

International Diabetes Federation. (2021). *IDF Diabetes Atlas*. Retrieved from IDF Diabetes Atlas: https://diabetesatlas.org/

Last, J. M. (n.d.). *Frank, Johann Peter*. Retrieved from Encyclopedia.com: https://www.encyclopedia.com/education/encyclopedias-almanacs-transcripts-and-maps/frank-johann-peter

Mayo Clinic. (2023, March 14). *Type 2 diabetes - Diagnosis and treatment*. Retrieved from Mayo Clinic: https://www.mayoclinic.org/diseases-conditions/type-2-diabetes/diagnosis-treatment/drc-20351199

Mayo Clinic. (2024, March 27). *Type 1 diabetes*. Retrieved from Mayo Clinic: https://www.mayoclinic.org/diseases-conditions/type-1-diabetes/diagnosis-treatment/drc-20353017

MyWay digital health. (n.d.). *Different ways to monitor blood sugar levels*. Retrieved from knowdiabetes: https://www.knowdiabetes.org.uk/know-more/your-blood/blood-glucose/different-ways-to-monitor-blood-sugar-levels/

MyWay digital health. (n.d.). *knowdiabetes*. Retrieved from Different ways to monitor blood sugar levels: https://www.knowdiabetes.org.uk/know-more/your-blood/blood-glucose/different-ways-to-monitor-blood-sugar-levels/

National Library of Medicine. (2020, Feb 13). *Global and regional estimates and projections of diabetes-related health expenditure: Results from the International Diabetes Federation Diabetes Atlas, 9th edition*. Retrieved from National Library of Medicine: https://pubmed.ncbi.nlm.nih.gov/32061820/

*The Past 200 Years in Diabetes*. (2012, October 4). Retrieved from The New England Journal of Medicine homepage: https://www.nejm.org/doi/10.1056/NEJMra1110560

UMass Diabetes Center of Excellence. (n.d.). *The First Human Insulin Injection to Treat Diabetes*. Retrieved from UMass Diabetes Center of Excellence: https://www.umassmed.edu/dcoe/diabetes-education/patient-resources/first-insulin-injection/

UnityPoint Health. (n.d.). *Untreated Diabetes: What to Know*. Retrieved from UnityPoint Health: https://www.unitypoint.org/news-and-articles/untreated-diabetes-what-to-know

Wikipedia. (n.d.). *Health system*. Retrieved from Wikipedia, The free encyclopedia: https://en.wikipedia.org/wiki/Health\_system

World Health Organization . (2021 , June 28 ). *WHO issues first global report on Artificial Intelligence (AI) in health and six guiding principles for its design and use*. Retrieved from World Health Organization : https://www.who.int/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use

World Health Organization. (n.d.). *Health system governance*. Retrieved from World Health Organization: https://www.who.int/health-topics/health-systems-governance#tab=tab\_1

World Health Organization. (n.d.). *Our values* . Retrieved from World Health Organization: https://www.who.int/about/values