KAMI LIMU

Cohort 8 Group Members and Roles

1. Brian Kinyanjui - Team lead / ML/ MLops
2. Stephanie Mukami - UI/UX designer
3. Justine Kebiba - Data scientist

ICT Track Mentor: Sidney Ochieng

iPREVENT

Problem Background

Diabetes Mellitus, commonly known as diabetes, is a group of endocrine diseases characterized by defects in insulin secretion [1]. It is one of the four major non-communicable diseases (NCDs), alongside cardiovascular diseases, cancers, and chronic respiratory diseases, contributing to a staggering 63% of NCD-related deaths worldwide [2]. Diabetes is classified into two types: Type 1, which involves insulin deficiency and can develop at any point in life, and Type 2 diabetes (T2DM), which leads to the body becoming resistant to insulin, preventing normal glucose regulation [3].

Globally, diabetes presents a formidable challenge, with its prevalence projected to rise by 46% by 2045 unless proactive measures are implemented [4]. In Kenya, diabetes is recognized as a significant public health issue, accounting for approximately 10,000 of the 284,000 deaths attributed to NCDs [5][6][7]. Like most NCDs, diabetes arises from a combination of genetic, physiological, behavioral, and environmental factors. For instance, individuals with a family history of diabetes are more susceptible, and exposure to certain toxins, cultural influences, and

unhealthy lifestyle choices, such as insufficient physical activity, significantly heighten the risk of developing diabetes.

A pressing concern is the high prevalence of undiagnosed T2DM, which is alarming on a global scale. Worldwide, an estimated 50% of individuals aged 20–79 with diabetes remain undiagnosed [8]. In Africa, this figure is even more troubling, with approximately 69% of cases undiagnosed, compared to 37% in high-income countries. This trend is echoed in Kenya, particularly in urban settings, where the prevalence of undiagnosed diabetes is nearly twice as high in urban areas (3.4%) compared to rural areas (1.9%) [9].

A national household survey conducted among 4,500 eligible individuals between April and June 2015 revealed a 51% prevalence of undiagnosed diabetes in urban settings [9]. Additionally, a cross-sectional study involving 50 patients at a primary health facility in Nairobi found that 52% were undiagnosed [10]. To further investigate these findings, we conducted a pilot survey in August 2024 among 53 adults aged 18-36 living in urban areas. Alarmingly, 71.7% of respondents reported that they had never been tested for diabetes. Those who did seek screening often did so only after experiencing symptoms associated with T2DM. The missed opportunities for early diagnosis are concerning, as individuals are frequently diagnosed only when serious complications, such as retinopathy and cardiovascular diseases, arise. This results in needless suffering, financial burdens, and further strains on an already overtaxed healthcare system [11].

Timely diagnosis of T2DM is crucial for reducing the number of cases and improving the overall well-being of the Kenyan adult population. Additionally, it aligns with Goal 3 of the Sustainable Development Goals (SDGs), which aims to reduce premature mortality from non-communicable diseases through prevention and treatment [15].

Market Opportunity

1. Current Alternative Solutions

a. Traditional Healthcare Clinics and Hospitals

- Patients often rely on clinics and hospitals for diabetes management, including routine check-ups, lab tests, and dietary consultations. However, these services can be limited in accessibility and affordability, particularly for low-income populations. Long waiting times and a lack of personalized care can result in poor diabetes management and treatment adherence.

b. Digital Health Apps

- **One Drop**: This app allows users to track their glucose levels, food intake, physical activity, and medication. It also provides personalized coaching and insights based on user data. However, it lacks local context and specific dietary advice relevant to the Kenyan population.
- **MySugr**: A popular diabetes management app that enables users to log their glucose levels, meals, and other health metrics. While effective in some regions, it does not fully cater to the specific needs and cultural context of Kenyan users, limiting its usability.

Market Opportunity for iPrevent: T2DM Solutions in Kenya

1. Improvements Needed in Current T2DM Solutions

While traditional healthcare clinics and digital health apps like One Drop and MySugr offer valuable support for diabetes management, significant improvements are needed in the following areas:

- Predictive Analytics: Current software solutions primarily focus on managing diabetes rather than preventing it. Traditional healthcare clinics provide reactive care, while digital apps like One Drop and MySugr often rely on users to input data without offering predictive risk assessments. iPrevent, with its machine learning capabilities, specifically the Random Forest algorithm, fills this gap by assessing user risk factors and providing early warnings about potential T2DM development.
- Personalization of Recommendations: Existing digital health apps tend to provide generic advice that may not cater to individual user needs. For instance, while MySugr allows users to log their health metrics, it lacks the tailored lifestyle recommendations that iPrevent offers based on specific user data, such as age, weight, physical activity, and family history. iPrevent's ability to deliver personalized dietary and exercise suggestions enhances its effectiveness in helping users mitigate their risk.
- Cultural Relevance: Many solutions, including One Drop and MySugr, do not adequately address local dietary practices or cultural contexts. iPrevent can capitalize on this by ensuring that its lifestyle recommendations resonate with Kenyan users, promoting adherence and effectiveness in managing their health.

2. Leveraging the Gaps for iPrevent

- Proactive Risk Assessment: Unlike existing solutions that focus mainly on management, iPrevent offers a proactive approach by predicting early risk factors for T2DM. This unique feature allows users to take preventive actions before complications arise, distinguishing it from traditional clinics and digital apps.
- Tailored Lifestyle Recommendations: The personalized advice generated by iPrevent based on individual assessments positions it advantageously against apps like One Drop and MySugr, which often provide generalized information. This tailored approach fosters greater user engagement and motivation to adopt healthier habits.

- Focus on Prevention: With the increasing global emphasis on preventive healthcare, iPrevent can leverage the trend toward early intervention. As users become more health-conscious, the app's focus on predicting and mitigating risk will resonate strongly with its target audience.
- Collaboration with Healthcare Providers: By partnering with health professionals, iPrevent can enhance its credibility and effectiveness, offering expert-backed recommendations that improve user trust and satisfaction, a feature not prominently emphasized in existing digital health apps.

3. Market Size and Economic Contribution

The economic burden of managing T2DM in Kenya is substantial, with total costs amounting to approximately US$ 633 million (KES 74,324 million) in 2021. This figure represents about 60% of the national health budget, indicating a critical need for effective preventive solutions. Management of T2D complications alone accounted for US$ 387 million (KES 42,465 million), highlighting the pressing challenge posed by conditions such as nephropathy, which incurs the highest costs at US$ 332 million. As T2D prevalence continues to rise, projections suggest that by 2045, annual costs could reach up to US$ 1.6 billion (KES 177 billion). The target audience for iPrevent includes adults aged 30 and above, particularly in urban settings where lifestyle-related risk factors are prevalent. By leveraging predictive analytics and tailored lifestyle recommendations, iPrevent addresses the urgent need for prevention and early intervention, ultimately contributing to reduced healthcare expenditures and improved quality of life for users. Additionally, shifting focus toward early screening and lifestyle modification can alleviate the economic strain on the healthcare system, making it a vital component of Kenya's public health strategy.

Solution Idea

1. Target User

The iPrevent Diabetes WebApp targets individuals aged 25 to 55 in urban areas of Kenya who are at risk of developing T2DM. This demographic was chosen based on health data indicating a rising prevalence of diabetes in this age group. Insights from the pilot study that was conducted with potential users highlighted their interest in digital health solutions for diabetes management and revealed common challenges such as lack of awareness about diabetes risk.

This age group is strategically targeted for early intervention as they are at a critical stage where health decisions significantly impact long-term wellness. They are also more likely to engage with digital health technologies, making them ideal candidates for proactive health management solutions like iPrevent Diabetes.

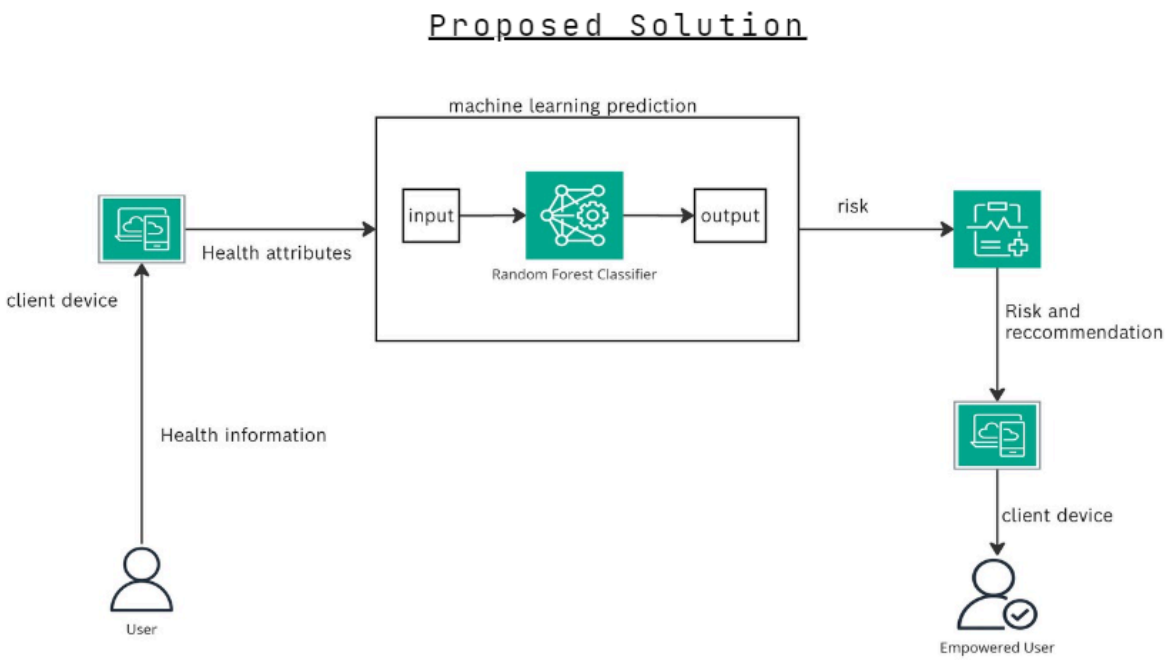
2. Solution Prototype

The iPrevent Diabetes WebApp is designed to predict early risk factors for T2DMand provide lifestyle recommendations tailored to each user's needs. The core offering of the solution is the ability to assess a user's risk using machine learning technology, specifically a Random Forest algorithm. This algorithm analyzes user-provided data, such as age, weight, physical activity level, and family history of diabetes, to deliver a comprehensive risk analysis. Based on the results, the WebApp generates personalized advice, including dietary recommendations, exercise routines, and lifestyle changes that can help reduce the risk of developing T2DM.

Process Overview

The iPrevent Diabetes WebApp follows a clear and structured process that begins with data collection and concludes with actionable insights to help users reduce their risk of developing T2DM.

1. Data Collection: Users are prompted to input personal health-related information such as their age, weight, family medical history, and lifestyle habits (including diet and physical activity). This stage is crucial as the app needs comprehensive data to provide accurate predictions.
2. Risk Assessment: The app's AI-powered engine, utilizing the Random Forest algorithm, processes the collected data to compute a personalized risk score. This assessment predicts the likelihood of the user developing T2DM based on their input.
3. Personalized Recommendations: Based on the calculated risk score, the WebApp provides users with tailored lifestyle recommendations. These suggestions may include diet modifications, specific exercise routines, and general health tips aimed at lowering the user's diabetes risk.



The iPrevent Diabetes WebApp directly addresses the problem of delayed diagnosis and prevention of T2DMby offering users an accessible tool for early risk detection. Unlike traditional diagnostic methods that rely on clinical visits, iPrevent empowers users by providing predictive insights from the comfort of their homes. By identifying key risk factors early, users can make informed decisions about lifestyle changes, reducing their likelihood of developing the disease. The WebApp also acts as an educational platform, promoting long-term health improvements.

Assumptions Made

The following key assumptions are made regarding the deployment and usage of iPrevent:

1. User Motivation: It is assumed that users are motivated to take control of their health and will regularly engage with the WebApp to monitor their risk and follow the recommended lifestyle changes.
2. Access to Technology: The solution assumes that the target user base has access to smartphones or computers, as well as stable internet connections, to interact with the WebApp effectively.
3. Trust in AI Solutions: It is assumed that users trust AI-driven health tools to provide accurate and reliable risk assessments and are comfortable following AI-generated health advice.
4. Willingness to Share Information: The solution presupposes that users are willing to provide personal health data in exchange for personalized insights into their diabetes risk, with the assurance that their privacy will be respected and safeguarded.

Value proposition

The iPrevent Diabetes Web App provides individuals aged 25 to 55 at risk of T2DM with personalized insights and resources to promote early prevention and healthier living.

Designed Solution

Technologies Used

Summarize here the technologies that you used, justifying the choices.

Machine learning (ML) Technologies

The language used for all ML and ML operation aspects was Python. This language was chosen for its robust libraries and packages for ML while offering seamless deployment and integration.

Python libraries used in this project include:

1. Pandas: Used for dataframe and series manipulation to handle data. Chosen due its to how well it works with other python libraries
2. Matplotlib: Used for visualization of data during Exploratory Data Analysis (EDA). Chosen over seaborn for the wider range of plots it offers
3. Scikit-learn: Used to implement machine learning concepts such as preprocessing, data splitting, accuracy calculation and modeling. Chosen over DaskML and CuML which, despite their ability to run on GPU, do not offer most algorithms such as scikit-learn
4. Optuna and Hyperopt: Used for Bayesian hyperparameter optimization on models
5. Imblearn: Used for oversampling data using SMOTE
6. Joblib: Used to pickle and load the trained model into an executable

The model used for this project was random forest boosted using ADABoostClassifier. This model was chosen over XGBoost and Gradient Boosting Machine (GBM) due to its ability to model complex data with huge dimensionality. The model was tuned using Optuna and Hyperopt.

Flask was used to expose the model as an API and it was tested using Postman and Curl.

Web Application Technologies used were

1. Next.js (Backend and Frontend) - Provided full-stack framework for server-side and client-side rendering. With built-in API routes for backend functionality.

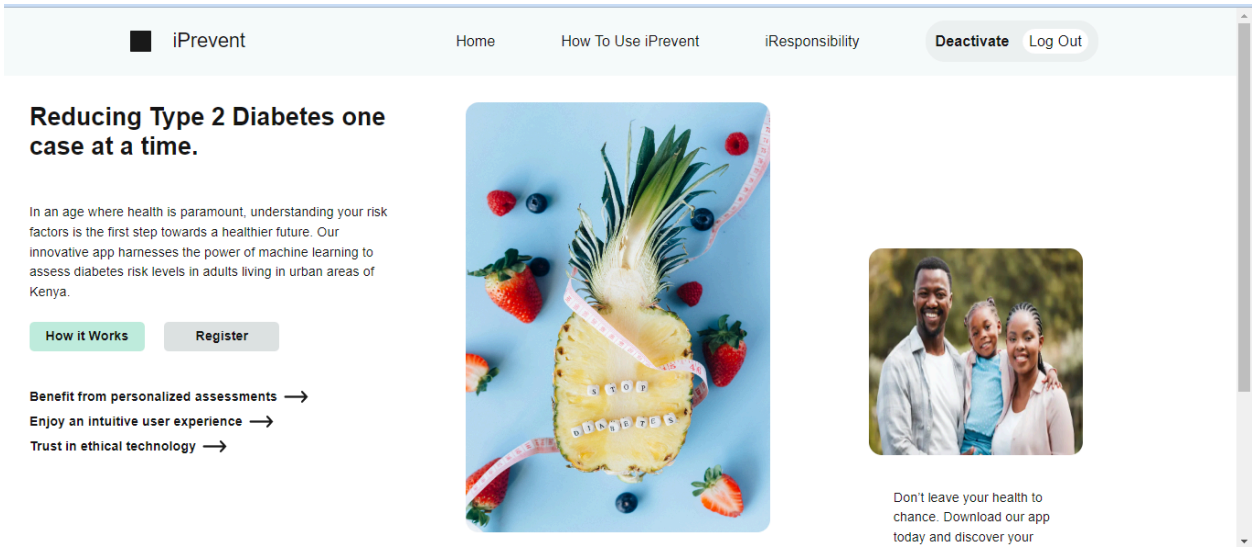
2. NextAuth (Authentication) - Simplifies authentication integration with Next.js and provides secure session management.
3. Tailwind CSS (Styling): - utilized as it offers utility-first CSS framework for fast and customizable styling.
4. Neon DB (PostgreSQL as a Service): - used as it is a managed PostgreSQL database with scalable, high availability.

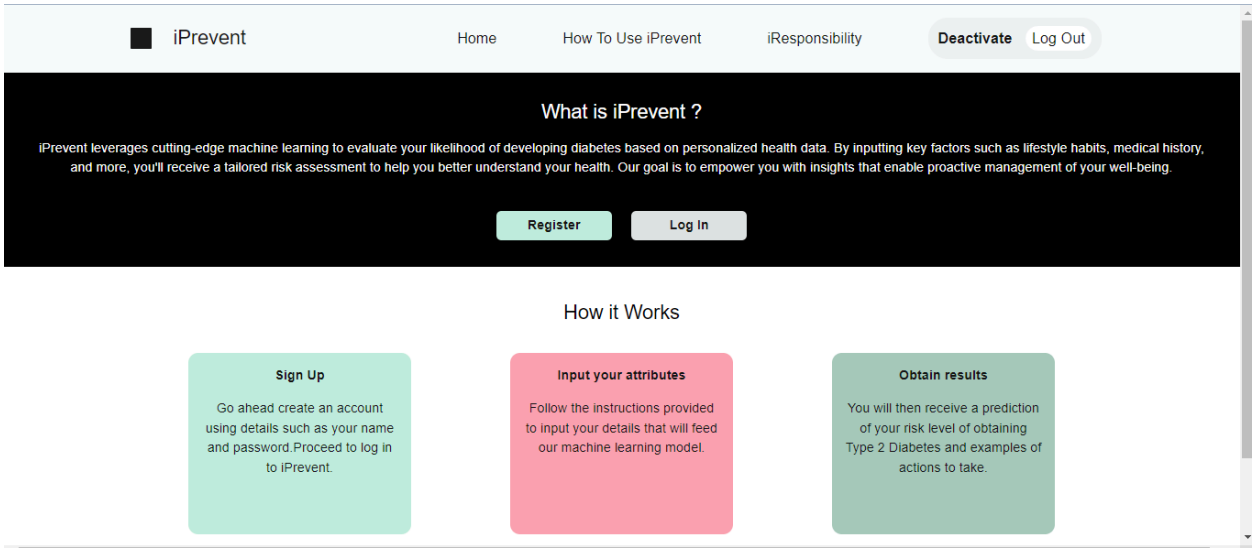
Screenshots of Main Modules

Here, showcase the main parts of your projects with screenshots after you complete building your solution and then explain what each module does.

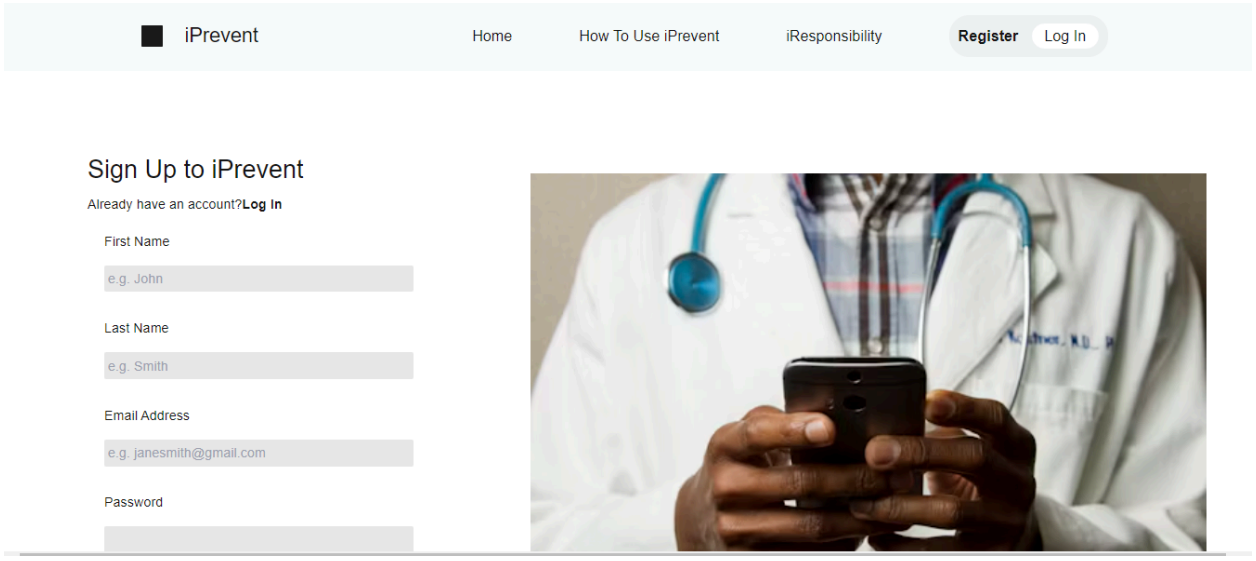
Client Side

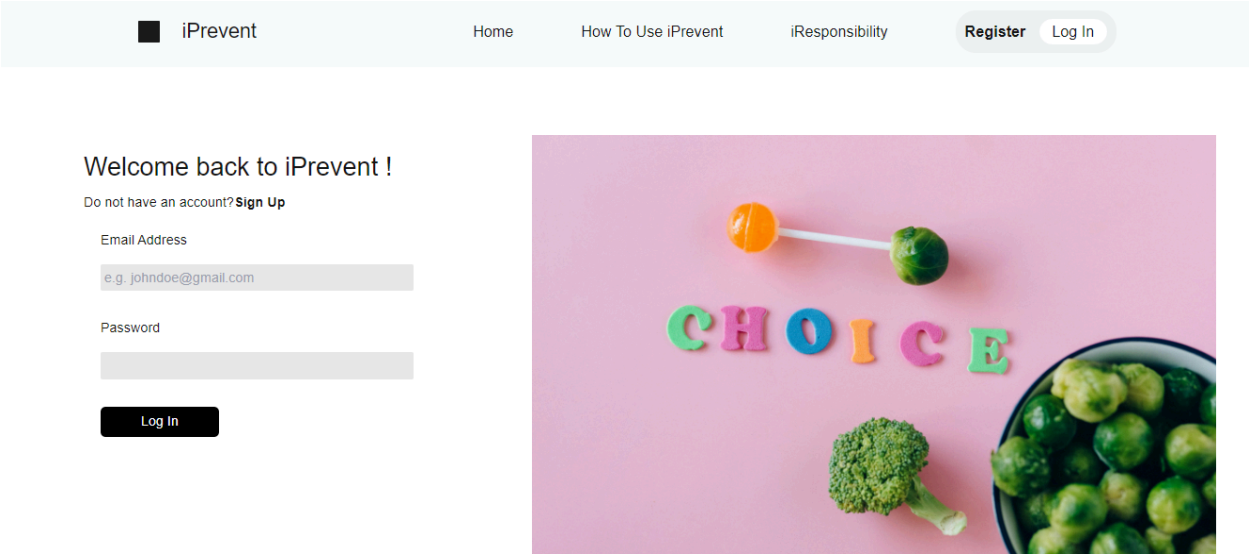
Home and About Page - These sections give an explainer on what iPrevent is



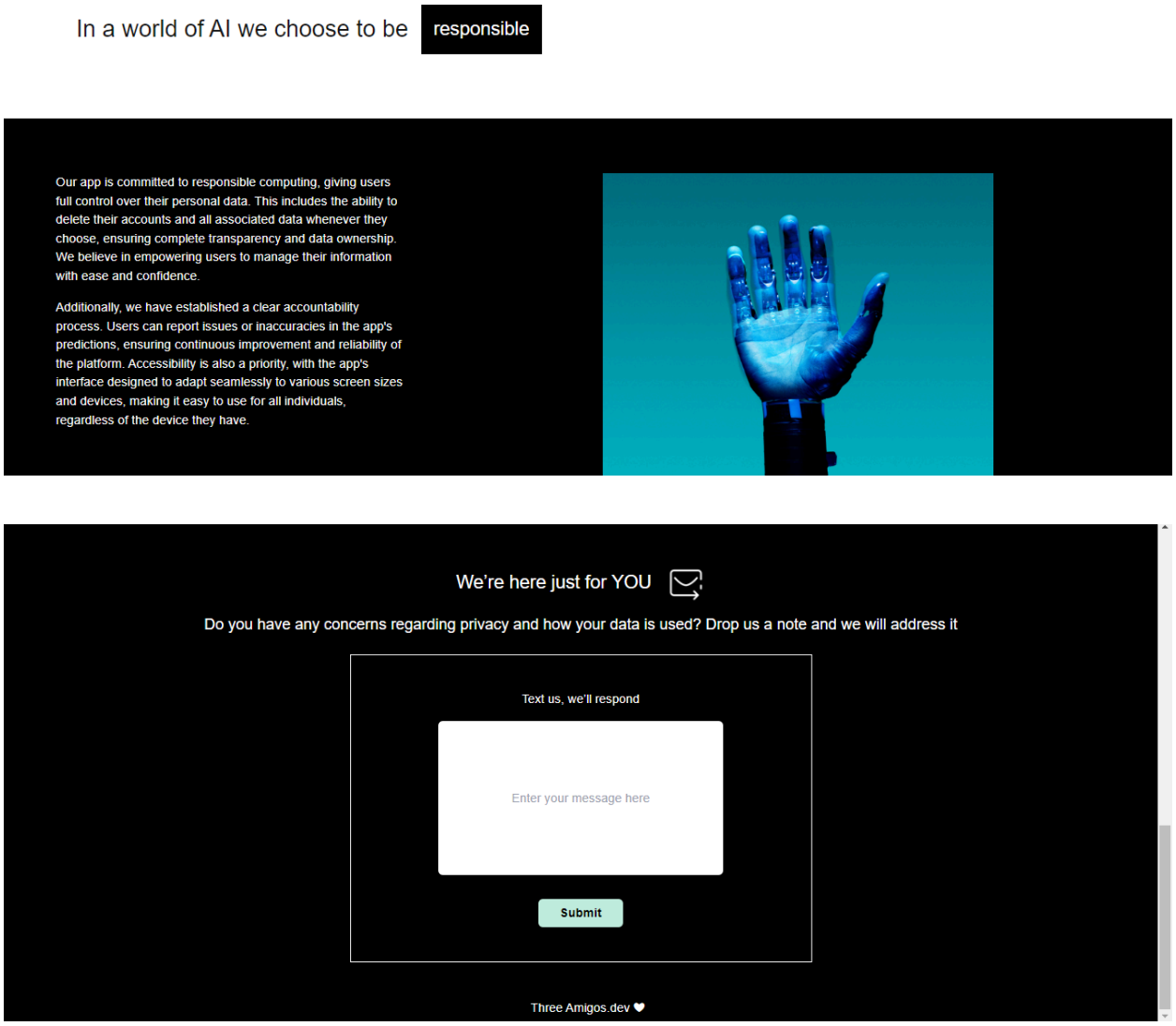


Authentication Page - This page allows a user to create an account and log in

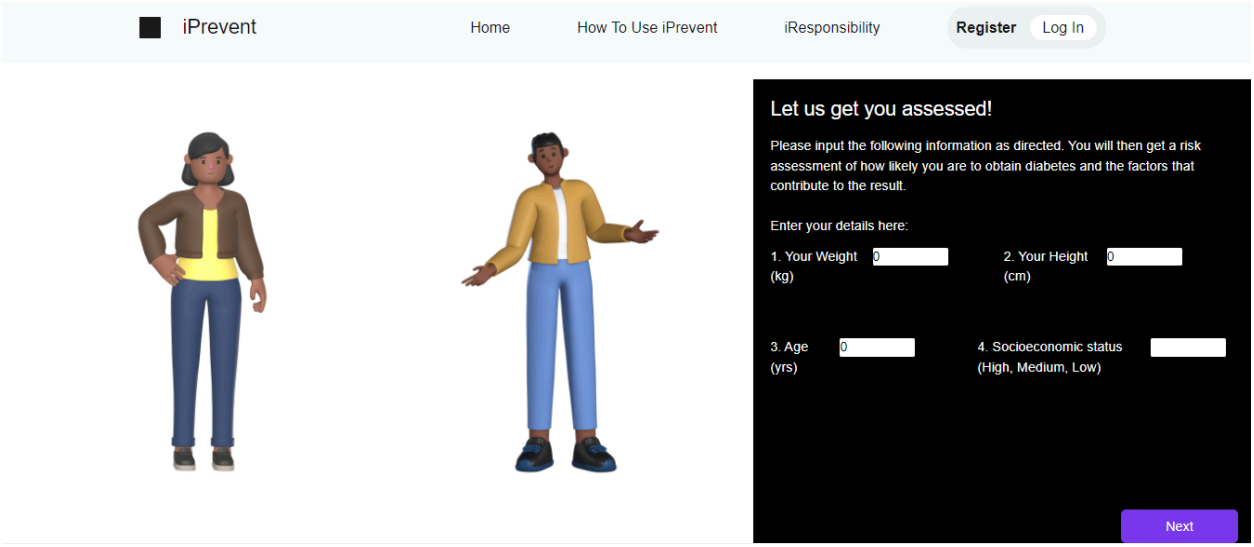


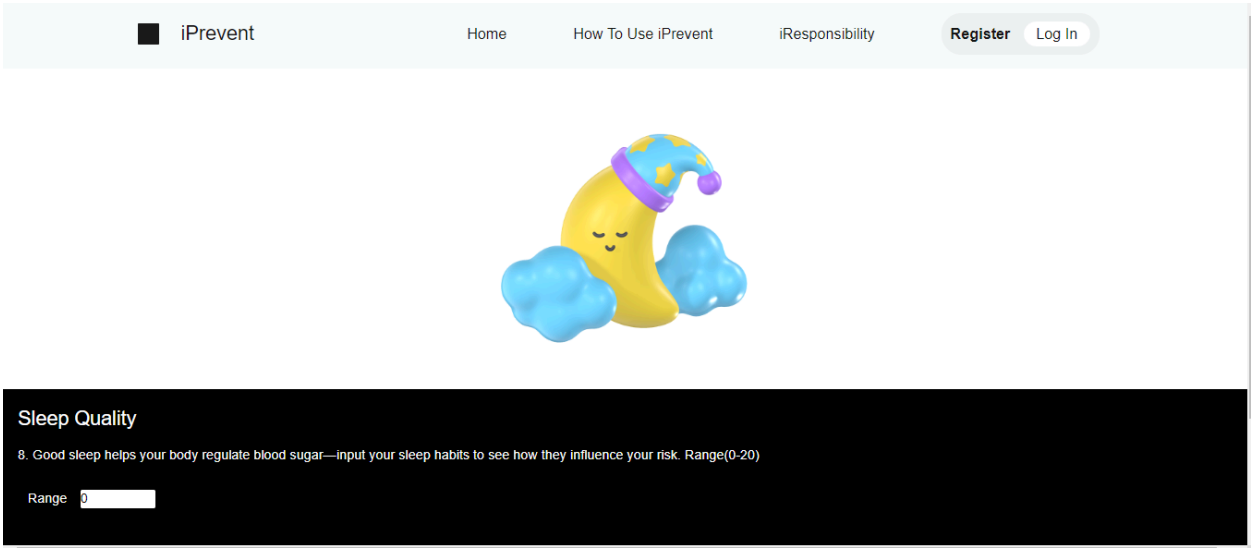


Responsible Computing Page - This section highlights the measures taken to make our application as ethically responsible as possible.



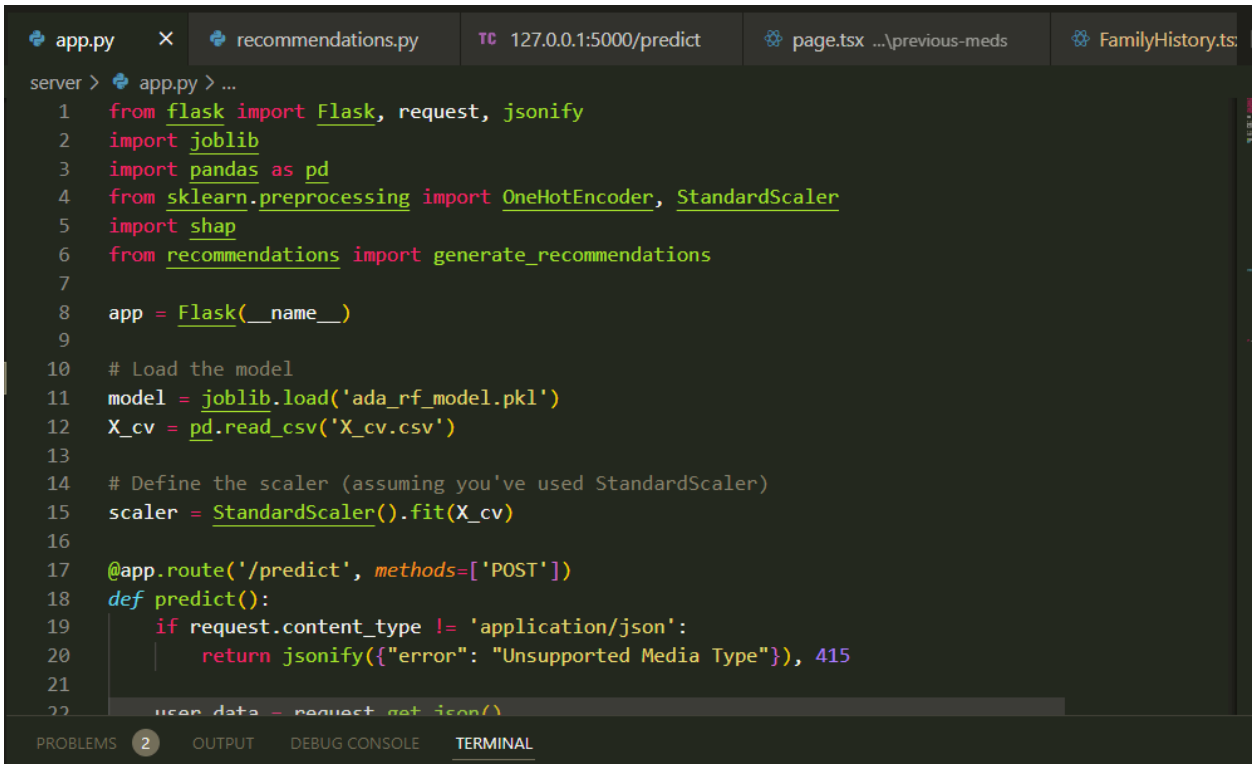
Data Input Page - examples of pages where user inputs various attributes e.g age and lifestyle factors


iPrevent
Home
How To Use iPrevent
iResponsibility
Register
Log In
Let us get you assessed!
Please input the following information as directed. You will then get a risk assessment of how likely you are to obtain diabetes and the factors that contribute to the result.
Enter your details here:
1. Your Weight (kg)
0
2. Your Height (cm)
0
3. Age (yrs)
0
4. Socioeconomic status (High, Medium, Low)
Next


iPrevent
Home
How To Use iPrevent
iResponsibility
Register
Log In
Sleep Quality
8. Good sleep helps your body regulate blood sugar—input your sleep habits to see how they influence your risk. Range(0-20)
Range
0

Server Side

Machine learning model encapsulated in Flask Web Api - This module shows the pickled model and routes to access it from the client side



```
# Define the scaler (assuming you've used StandardScaler)
scaler = StandardScaler().fit(X_cv)

@app.route('/predict', methods=['POST'])
def predict():
    if request.content_type != 'application/json':
        return jsonify({"error": "Unsupported Media Type"}), 415

    user_data = request.get_json()
    user_df = pd.DataFrame([user_data])

    # Preprocess user data
    user_df = preprocessing.encode_categorical_values(user_df)
    user_df = preprocessing.normalize_data(user_df)
    user_df = preprocessing.scale(user_df)

    # Make prediction
    prediction = int(model.predict(user_df)[0])
    prediction_proba = model.predict_proba(user_df)[0].tolist()

    # Calculate SHAP values
    explainer = shap.KernelExplainer(model.predict_proba, scaler.transform(X_cv))
```

Examples of recommendations to guide user after assessment

```python
def generate_recommendations(top_3, prediction):

    recommendations_dict = {

        "GestationalDiabetes": {
            "High": (
                "You seem to be experiencing gestational diabetes. Here's what you can do: Follow a meal plan that includes smaller, "
                "frequent meals balanced in carbohydrates, proteins, and fats. Engaging in regular physical activity, such as walking, "
                "can help manage your blood sugar levels effectively. It's crucial to monitor your blood sugar as advised by your "
                "healthcare provider and attend all prenatal appointments. If insulin is prescribed, make sure you feel comfortable "
                "administering it."
            ),
            "Low": (
                "You may not have gestational diabetes, but it's still important to maintain a balanced diet rich in whole grains, "
                "lean proteins, fruits, and vegetables. Regular physical activity is vital, and attending prenatal check-ups will help "
                "monitor both your health and that of your baby."
            )
        },
        "PreviousPreDiabetes": {
            "High": (
                "Given your history of pre-diabetes, it's essential to focus on lifestyle changes to prevent progression to diabetes. "
                "Consider adopting a Mediterranean-style diet that features healthy fats, lean proteins, and plenty of vegetables. "
                "You might also benefit from participating in a structured lifestyle change program with regular meetings with a lifestyle "
                "coach. Be sure to monitor your blood glucose levels as recommended by your healthcare provider."
            ),
            "Low": (
                "Even without a history of pre-diabetes, continue with healthy eating habits, regular exercise, and annual check-ups to "
                "monitor your blood glucose levels and other health markers."
            )
        },
        "Hypertension": {
            "High": (
                "It seems you have high blood pressure. Here's what you can do: Adopt a DASH (Dietary Approaches to Stop Hypertension) "
                "diet rich in fruits, vegetables, whole grains, and low-fat dairy. Reducing sodium intake and avoiding high-sodium "
                "processed foods is crucial. Incorporate regular aerobic exercises like walking, cycling, or swimming. Limiting alcohol "
```

Link to the solution

Github Link - [here]

Business Model

The iPrevent Diabetes WebApp is designed with a hybrid approach, incorporating both for-profit and nonprofit elements to ensure financial sustainability while maintaining accessibility for users. The goal is to provide a valuable health tool while generating revenue to sustain and grow the platform.

1. For-Profit Model

The WebApp will generate revenue through the following avenues:

- **Freemium Model**: iPrevent will offer a basic version of the WebApp for free, allowing users to assess their T2DM risk and receive general recommendations. However, to access advanced features such as detailed reports, personalized meal plans, and in-depth health tracking, users can subscribe to a **premium version**. The subscription tiers would include:
    - Basic Plan (Free): Access to general risk assessments and limited lifestyle recommendations.
    - Premium Plan (Monthly/Annual Subscription): Advanced insights, personalized meal and exercise plans, continuous health monitoring, and access to expert consultations.
    - Corporate Partnerships and Sponsorships: The WebApp will partner with healthcare providers, fitness brands, insurance companies, and pharmaceutical companies. These partners may sponsor or advertise their services/products on the platform. For example, insurance companies could provide subsidized health services to users or offer coverage discounts to those actively using the app for health improvement.

2. Non-Profit Model

As part of its social mission to reduce the incidence of T2DM, iPrevent will incorporate a non-profit arm to raise funds for offering the app free to at-risk, underserved populations. This will be done through:

- Grants and Donations: The WebApp will apply for grants from public health organizations, governmental agencies, and NGOs focusing on diabetes prevention, health

innovation, or digital health solutions. Donations will also be sought from philanthropic organizations committed to improving global health.

- Public Health Partnerships: Collaborations with public health departments, international health agencies, and non-profit organizations will help bring iPrevent to communities with high diabetes risk. The platform could be provided as a free resource for diabetes prevention programs in low-income areas or developing countries.

3. Ensuring Financial Sustainability

To ensure financial sustainability, the iPrevent Diabetes WebApp will rely on:

- Diversified Revenue Streams: By combining the freemium model with partnerships and sponsorships, iPrevent will have multiple sources of income. The focus on corporate partnerships, particularly with health and wellness sectors, will ensure consistent revenue flow from trusted brands.
- Scalable Platform: iPrevent will operate as a web-based solution, minimizing overhead costs associated with app maintenance and updates. The scalability of the platform allows for easy expansion into new markets without incurring significant development costs, while leveraging AI technology will ensure ongoing relevance and innovation.
- Cost-Efficiency: Operating as a primarily digital service, iPrevent has low operating costs. Marketing and development costs will be kept manageable through a focus on online campaigns, social media marketing, and strategic partnerships.

Responsible Computing

Responsible Computing Criteria (RCC) for iPrevent Diabetes App

1. Accountability: For user accountability, the app features a user-friendly feedback form accessible on the iResponsibility page. This form allows users to report issues, inaccuracies, or suggestions related to the app's predictions. This will empower users to actively participate in improving the app, ensuring that their experiences and concerns are addressed accordingly.

2. Accessibility: The app's interface has been designed to adapt seamlessly to various screen sizes and device types, including smartphones and tablets. This commitment to accessibility ensures that all users, regardless of their device, can navigate the app effortlessly. By prioritizing an inclusive design, we aim to enhance user satisfaction and broaden our user base.

3. Privacy and Security:

- Data Encryption: User credentials are encrypted both during transmission and while stored.
- Clear Privacy Policy: A privacy policy is made available on the iResponsibility page. The document clearly articulates how user data is collected, utilized, and safeguarded, helping users understand their rights and the protective measures we implement. This clarity is essential in building trust with the user base.
- User Control Over Data: To enhance user autonomy, the app also features a straightforward delete account functionality, allowing users to easily delete their accounts and all associated data. This capability empowers users to take control of their personal information and reinforces our commitment to user privacy.

Traction

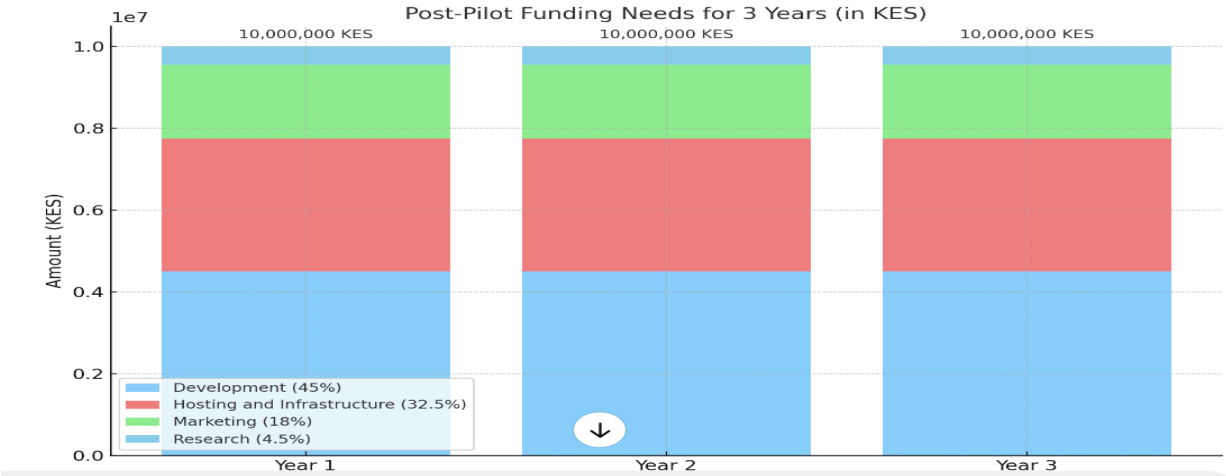
Review from a medical student : Brian Wari

- iPrevent stands to be a great tool to raise awareness among young adults on Type 2 Diabetes. Many patients come when it is too late.

Funding/Support Need

The graph below illustrates the projected funding requirements for the three years following the pilot stage, with a total estimated budget of 9.95 million KES per year. The funding is divided into four key areas: Development, Hosting and Infrastructure, Marketing, and Research, each playing a vital role in the project's ongoing success.

- Development accounts for the largest portion of the budget, comprising 45% of the total annual funds. This reflects the continued need for software updates, feature enhancements, and general system maintenance. Each year, approximately 4.5 million KES will be required to support these efforts.
- Hosting and Infrastructure forms the second largest expense, representing 32.5% of the total budget. This allocation ensures that the necessary technology, such as servers and cloud storage, can scale to meet growing user demand. The yearly cost for hosting and infrastructure is estimated at 3.25 million KES.
- Marketing, at 18% of the total budget, will require 1.8 million KES per year. This will cover advertising campaigns, brand promotion, and customer acquisition efforts, which are crucial for expanding the project's market reach.
- Finally, Research, while the smallest portion at 4.5%, remains essential for the project’s evolution. This includes market research, user feedback analysis, and exploration of new features, with an annual cost of 450,000 KES.



The Team

Stephanie Mukami
UX & Web Developer

Brian Kinyanjui
ML & MLOps Engineer

Justine Kebiba
Data Scientist

Stephanie Mukami – UX and Web Developer

Stephanie Mukami, the team's UX and Web Developer, holds an IBM certification in design thinking. Her expertise ensures the iPrevent Diabetes Web App delivers a user-friendly, intuitive

interface. Stephanie is responsible for optimizing the app's design for seamless interaction across various devices, ensuring accessibility for all users. Her focus on user-centered design principles will enhance the overall experience, making it easier for users to engage with the app and its features effectively.

Justine Kebiba – Data Scientist

Justine Kebiba, serving as the team's Data Scientist, brings extensive experience in data analysis and interpretation. As a former Secretary of the JKUAT Data Science Club, he is well-versed in handling large datasets and deriving actionable insights. His role will focus on interpreting user data to provide accurate health risk assessments, contributing to the app's predictive analytics. Justine's expertise ensures that the data-driven components of the web app function effectively and deliver valuable insights to users.

Brian Kinyanjui – ML and MLOps Engineer

Brian, the team's ML and MLOps Engineer, will lead the development and deployment of the machine learning models that power the app's predictive capabilities. With experience as a committee member for the upcoming Nairobi School of AI, Brian brings a deep understanding of machine learning technologies and operational best practices. This expertise will be applied to ensure the app delivers accurate, real-time health risk predictions, offering users personalized insights to support diabetes prevention.

References

1. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015 - PMC (nih.gov)

2. (Determination of microbial metagenomic markers of Type 2 Diabetes Mellitus (T2DM) in patients visiting South C Health Centre in Nairobi Kenya | Journal of Agriculture, Science and Technology (ajol.info))
3. IDF Diabetes Atlas 2021 | IDF Diabetes Atlas
4. https://apps.who.int/iris/handle/10665/274512.
5. World Health Organization. (2018). Noncommunicable diseases country profiles 2018. World Health Organization.
6. Ministry of Health, The Kenya Non-Communicable Diseases & Injuries Poverty Commission Report, 2018
7. Diabetes, like most NCDs, results from a combination of genetic, physiological, behavioural and environmental factors
8. https://static1.squarespace.com/static/55d4de6de4b011a1673a40a6/t/5b637739562fa77c7bbf430a/1533245242346/Kenya+Report+layout+23-07-18_JUSTIFIED.pdf.
9. [3, 5 [6] Non-communicable diseases surveillance: overview of magnitude and determinants in Kenya from STEPwise approach survey of 2015 | BMC Public Health | Full Text (biomedcentral.com).
10. 5]. IDF DIABETES ATLAS [Internet] - PubMed (nih.gov)
11. Prevalence and factors associated with pre-diabetes and diabetes mellitus in Kenya: results from a national survey - PMC (nih.gov).
12. What drives distress? Rethinking the roles of emotion and diagnosis among people with diabetes in Nairobi, Kenya - PubMed (nih.gov).
13. https://doi.org/10.2174/1876524601407010001
14. The financial burden from non-communicable diseases in low- and middle-income countries: a literature review - PMC (nih.gov).
15. (http://apps.who.int/gb/ebwha/pdf_files/WHA71/A71_R1-en.pdf, accessed 16 August 2018).
16. https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-024-11700-x