

 SANTÉ PUBLIQUE

Détection Précoce du Diabète de Type 2

Un outil de triage intelligent pour révolutionner le dépistage et sauver des vies



Made with **GAMMA**



Le Défi du Diabète : Un Tueur Silencieux

Le Problème

Le diabète de type 2 progresse sans symptômes pendant des années. Les patients ignorent leur condition jusqu'à l'apparition de complications irréversibles.

- Diagnostic souvent tardif
- Progression asymptomatique
- Conséquences graves évitables

Impact Majeur

Les complications touchent profondément la vie des patients et pèsent lourdement sur les systèmes de santé.

- Amputations et cécité
- Accidents vasculaires cérébraux
- Coûts de santé exponentiels

Notre Solution : Intelligence Artificielle au Service du Dépistage



Machine Learning

Algorithmes d'apprentissage automatique pour identifier les patients à haut risque



Triage Intelligent

Priorisation automatique des dépistages sanguins selon le niveau de risque



Détection Précoce

Intervention avant l'apparition des complications graves





Analyse des Données : Base Solide pour la Prédiction

Dataset Pima Indians

Notre modèle s'appuie sur une base de données clinique éprouvée contenant 768 patientes.

8 indicateurs clés :

- Taux de glucose sanguin
- Indice de masse corporelle
- Âge et grossesses
- Pression artérielle

768

Patientes

8

Variables

Échantillon clinique validé

Indicateurs biométriques

0.47

Corrélation

Glucose - Diagnostic

Traitements des Données : Rigueur Scientifique

01

Nettoyage des Données

Détection et correction des valeurs incohérentes (glucose, IMC, pression artérielle à zéro)

02

Imputation Statistique

Remplacement des valeurs manquantes par la médiane pour préserver la distribution naturelle

03

Équilibrage des Classes

Pondération automatique pour compenser le déséquilibre entre cas sains et diabétiques

04

Normalisation

Standardisation des variables pour optimiser les performances du modèle

Performance du Modèle : Résultats Probants

Nous avons sélectionné le modèle **Random Forest** pour son équilibre optimal entre sensibilité et spécificité, garantissant un dépistage efficace tout en minimisant les faux positifs. Un Recall supérieur à 0.75 était notre critère clé.

0.837

ROC-AUC

Capacité de discrimination

0.757

Recall

Détection des vrais positifs

0.704

F1-Score

Précision équilibrée

Comparaison

	ROC-AUC	Recall	F1-Score
log_reg	0.845	0.720	0.682
random_forest	0.837	0.757	0.704
xgboost	0.833	0.752	0.682

Interprétation des Métriques Clés

Pour évaluer la pertinence de notre modèle de détection précoce du diabète, nous nous concentrons sur des métriques spécifiques.

Recall (Sensibilité) : La Priorité en Santé

En matière de santé, il est primordial de **ne rater aucun patient malade** (réduire les Faux Négatifs), même si cela implique de tester occasionnellement une personne saine (Faux Positif).

Notre objectif est de maximiser la capacité du modèle à identifier tous les patients réellement diabétiques pour déclencher un dépistage sanguin.

AUC-ROC : Robustesse et Capacité Discriminatoire

L'AUC-ROC mesure la capacité globale de notre modèle à distinguer entre les patients sains et les patients diabétiques, **indépendamment du seuil de décision.**

Un score élevé prouve que le modèle est "robuste" et qu'il performe significativement mieux qu'un classement aléatoire.

F1-Score : Précision Équilibrée sur Données Déséquilibrées

Le F1-Score est la moyenne harmonique de la Précision et du Recall. Il est particulièrement pertinent car notre jeu de données est déséquilibré (plus de personnes saines que de diabétiques).

Il fournit une **image plus honnête de la performance** globale du modèle que la simple exactitude (Accuracy), qui peut être trompeuse sur des données déséquilibrées.

Explicabilité : Au-delà de la Boîte Noire

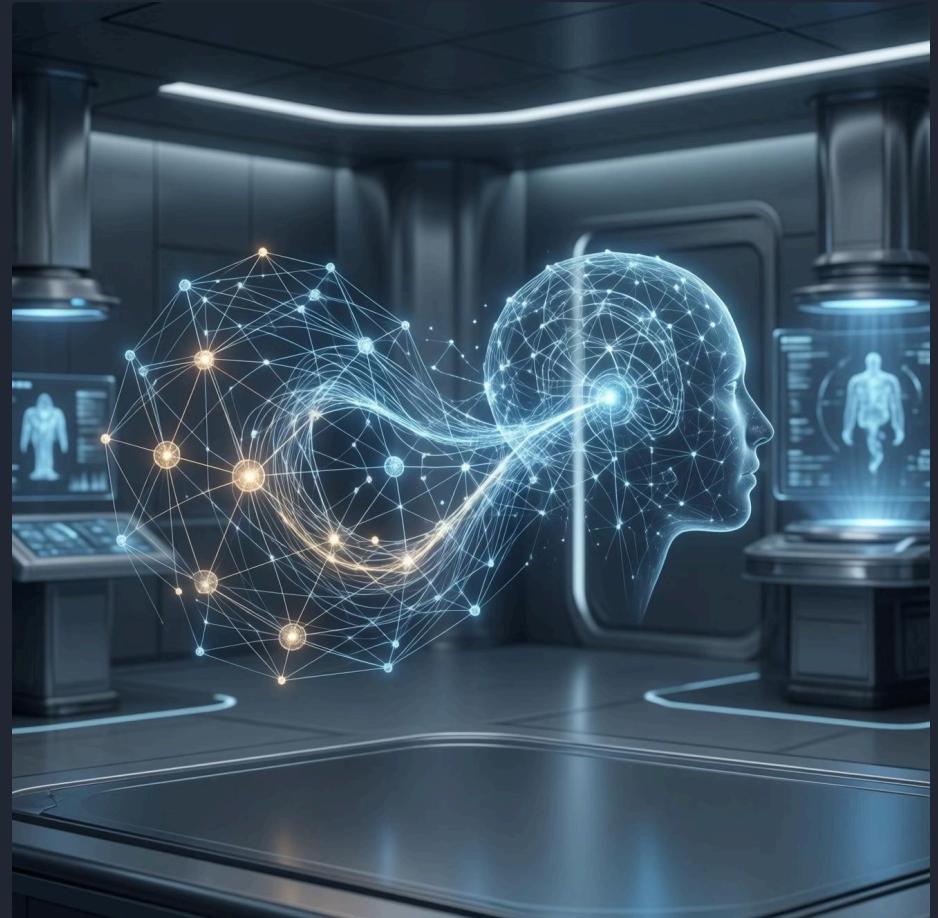
Transparence Totale avec SHAP

Chaque prédiction est justifiée et compréhensible grâce à l'analyse SHAP (SHapley Additive exPlanations).

Facteurs d'influence identifiés :

- Glucose sanguin (impact majeur)
- Indice de masse corporelle
- Âge du patient
- Historique de grossesses

Cette transparence permet aux cliniciens de comprendre et valider les recommandations du système.



Dashboard KPI : Interface Clinique Intuitive



Saisie Patient

Formulaire interactif pour entrer rapidement les données biométriques



Score de Risque

Probabilité transformée en score 0-100 avec niveau d'alerte (Faible/Moyen/Élevé)

Développé avec Streamlit pour une expérience utilisateur fluide et professionnelle.



Visualisation SHAP

Explication visuelle des facteurs de risque pour chaque patient



Conformité Légale et Déploiement Sécurisé

Notre solution est conçue en parfaite adéquation avec le cadre réglementaire européen et français, garantissant une utilisation éthique et responsable de l'IA en santé.



Positionnement Légal

Un outil d'aide au dépistage, jamais un substitut au diagnostic médical. Le jugement du clinicien reste primordial.



Protection des Données (RGPD)

Données de santé sensibles : finalité claire, minimisation, pseudonymisation et chiffrement rigoureux pour la confidentialité.



AI Act (IA à Haut Risque)

Documentation détaillée, traçabilité, gestion des biais et supervision humaine obligatoire pour chaque prédition.

Chiffrage Azure : Solution Optimisée

Votre devis final reflète une architecture professionnelle sécurisée située en France Central pour la conformité HDS.

Service Azure	Configuration	Coût Mensuel Estimé
Azure Machine Learning	Instance F4s v2 (Usage partiel 8h)	1.37 €
App Service (Streamlit)	Plan B1 (Basic Tier)	11.17 €
Azure SQL Database	Serverless (Mode économie)	5.64 €

TOTAL GÉNÉRAL : 18,18 € / mois



Perspectives d'Évolution



Surveillance Continue

Monitoring du data drift pour maintenir la performance du modèle dans le temps

Enrichissement Temporel

Intégration de l'historique glycémique pour des prédictions plus fines

Généralisation

Ré-entraînement sur populations diversifiées (hommes, origines variées)

Prochaines étapes : Validation clinique en conditions réelles et extension à d'autres pathologies chroniques.



Conclusion