

# Ocena Ryzyka Raka Szyjki Macicy

---

## Opis projektu

Aplikacja do oceny ryzyka raka szyjki macicy to narzędzie prognostyczne wykorzystujące algorytm XGBoost do przewidywania potencjalnego ryzyka rozwoju raka szyjki macicy na podstawie danych klinicznych i demograficznych pacjentki.

## Spis treści

- [Wprowadzenie](#)
- [Instalacja](#)
- [Użytkowanie](#)
- [Dokumentacja techniczna](#)
- [Model predykcyjny](#)
- [Uwagi](#)
- [Kontakt](#)

## Wprowadzenie

Aplikacja umożliwia ocenę ryzyka raka szyjki macicy na podstawie danych wprowadzonych przez użytkownika, takich jak:

- Wiek
- Historia seksualna (wiek rozpoczęcia aktywności, liczba partnerów)
- Historia ciąż
- Palenie tytoniu
- Stosowanie antykoncepcji hormonalnej
- Stosowanie wkładki wewnątrzmacicznej (IUD)
- Historia chorób przenoszonych drogą płciową (STD)
- Wcześniejsze diagnozy

## Instalacja

### Wymagania systemowe

- Python 3.7+
- Flask
- XGBoost
- NumPy
- Pandas
- scikit-learn

### Kroki instalacji

1. Instalacja zależności:

```
pip install -r requirements.txt
```

## 2. Uruchomienie aplikacji:

W cmd przejdź do folderu z projektem i wpisz:  
python app.py  
Nie zamykaj terminala!

## Użytkowanie

### 1. Otwórz przeglądarkę internetową i przejdź pod adres:

```
http://localhost:5000
```

### 2. Wypełnij formularz danymi pacjentki:

- Nie wszystkie pola są wymagane
- Wartości numeryczne muszą być nieujemne
- Pola wyboru (tak/nie) muszą zostać zaznaczone

### 3. Kliknij przycisk "Oblicz Ryzyko", aby otrzymać prognozę:

- Wynik zostanie wyświetlony jako procent ryzyka
- Próg klasyfikacji wynosi 10% (wartości  $\geq 10\%$  są klasyfikowane jako wysokie ryzyko)

## Dokumentacja techniczna

### Struktura projektu

```
Cancer_Predict_App/  
├─ app.py                # Serwer Flask  
├─ cervical_cancer.csv   # Dane treningowe  
├─ cervical_cancer_model.json # Zapisany model XGBoost  
├─ scaler.pkl            # Zapisany skaler StandardScaler  
├─ cervical_cancer_prediction_using_xg_boost_algorithm_4-1.py # Skrypt trenujący model  
└─ templates/  
    ├─ home.html         # Strona główna  
    └─ index.html        # Formularz oceny ryzyka
```

### Komponenty aplikacji

#### Backend (app.py)

- Serwer Flask obsługujący requesty HTTP
- Endpoint `/predict` przetwarzający dane z formularza
- Ładowanie modelu XGBoost i skalera
- Przetwarzanie danych wejściowych do formatu wymaganego przez model

## Frontend (index.html)

- Responsywny formularz HTML/CSS/JavaScript
- Walidacja danych po stronie klienta
- Dynamiczna aktualizacja pól formularza
- Wizualizacja wyników za pomocą paska postępu

## Przepływ danych

1. Użytkownik wprowadza dane do formularza
2. JavaScript waliduje dane i wysyła request do `/predict`
3. Backend przetwarza dane:
  - Konwersja typów danych
  - Skalowanie danych przy użyciu StandardScaler
4. Model XGBoost generuje predykcję
5. Backend zwraca JSON z wynikami
6. Frontend wyświetla rezultaty

## Model predykcyjny

### Dane treningowe

Model został wytrenowany na zbiorze danych "cervical\_cancer.csv", zawierającym informacje kliniczne od kobiet poddanych badaniom przesiewowym w kierunku raka szyjki macicy.

### Algorytm

- **XGBoost** (Extreme Gradient Boosting)
- Parametry: `learning_rate=0.1`, `max_depth=50`, `n_estimators=100`
- Dokładność na zbiorze testowym: [wartość]

### Preprocessing danych

- Zastąpienie brakujących wartości średnią
- Standaryzacja danych przy użyciu StandardScaler
- Konwersja wartości kategorycznych na liczbowe

### Ocena modelu

- Model wykorzystuje próg decyzyjny 0.1 (10%) do klasyfikacji ryzyka
- Ewaluacja na podstawie: dokładności, precyzji, czułości i F1-score

## Uwagi

### Ograniczenia modelu

- Model jest narzędziem pomocniczym i nie zastępuje profesjonalnej diagnozy medycznej
- Dokładność modelu zależy od jakości i kompletności wprowadzonych danych
- Możliwe są fałszywie dodatnie i fałszywie ujemne wyniki

## Kontakt

- W razie pytań proszę o kontakt pod adresem
- [jakub.aurzecki@o2.pl](mailto:jakub.aurzecki@o2.pl) (Nr albumu 44869) lub [dawid.bartoszek04@wp.pl](mailto:dawid.bartoszek04@wp.pl) (Nr albumu 44871)