

Modelowanie i Analiza Systemów Informatycznych

*Aplikacja internetowa do przewidywania przypadłości zdrowotnych
na podstawie mikromacierzy DNA*

Autor:

Daniel Czyż

Rzeszów 2023

Spis treści

1. Opis świata rzeczywistego.....	3
2. Modelowanie systemu i jego otoczenia	5
3. Opis techniczny projektu.....	16
4. Prezentacja warstwy użytkowej projektu	17
5. Podsumowanie	20

1. Opis świata rzeczywistego

1.1. Opis zasobów ludzkich

Osoby pełniące rolę administratora zarządzają systemem w aplikacji „Aplikacja internetowa do przewidywania przypadłości zdrowotnych na podstawie mikromacierzy DNA”. Zarządzanie systemem obejmuje zarządzanie zasobami bazy danych, zarządzanie modelami klasyfikacyjnymi, zarządzanie kontami oraz zarządzanie metodami klasyfikacyjnymi, selekcji cech, preprocessingu danych. Zarządzanie zasobami bazy danych polega na dodawaniu / modyfikowaniu / usuwaniu danych składających się na zbiory mikromacierzy DNA oraz zbiory cech klinicznych powiązanych z konkretnymi mikromacierzami DNA. Dane te są podstawą do trenowania i testowania modeli klasyfikacyjnych. Zarządzanie modelami klasyfikacyjnymi polega na tworzeniu / modyfikowaniu / usuwaniu modeli klasyfikacyjnych (obejmuje proces tworzenia modeli o wysokiej jakości tj. trenowanie i testowanie modeli oraz dostrajanie parametrów i opcji metod biorących udział przy tworzeniu modelu tak aby finalna jego skuteczność była zmaksymalizowana i spełniała standardy dobrego modelu klasyfikującego), które klasyfikują w kontekście konkretnej przypadłości zdrowotnej. Modele te są widoczne i mogą być wykorzystane z poziomu zalogowanych użytkowników w systemie. Zarządzanie kontami polega na dodawaniu / modyfikowaniu / usuwaniu kont oraz informacji odnośnie istniejących kont w systemie. W obręb zarządzania kontami wchodzi również możliwość tworzenia kont z uprawnieniami administratora. Zarządzanie metodami klasyfikacyjnymi, selekcji cech, preprocessingu danych polega na dodawaniu / modyfikowaniu / usuwaniu metod, jakie są dostępne w systemie, które później mogą być wykorzystane do tworzenia modeli klasyfikacyjnych.

Gość może dokonać podglądu jakie modele klasyfikacyjne oferuje aplikacja. Gość może również założyć konto w systemie, a po zalogowaniu staje się użytkownikiem.

Użytkownik może wczytać do systemu swoją mikromacierz DNA i dokonać klasyfikacji wybranym modelem w kontekście konkretnej przypadłości zdrowotnej. Otrzymane wyniki klasyfikacji użytkownik może pobrać na urządzenie, a także są one zapisywane w bazie danych i dostępne w każdej chwili dokonując przegląd swojej historii klasyfikacji z poziomu panelu użytkownika.

1.2. Przepisy i strategia firmy

System określa maksymalną ilość klasyfikacji, jako 1/miesiąc dla użytkownika bez wykupionej subskrypcji. Po wykupieniu subskrypcji przez użytkownika ograniczenie to przestaje być aktywne. Zmiana opcji subskrypcji jest dostępna z poziomu panelu użytkownika.

Aby ograniczyć możliwość tworzenia wielu kont przez jedną osobę w celu dokończenia rejestracji użytkownika wymagane będzie potwierdzenie numeru telefonu za pomocą kodu SMS.

1.3. Dane techniczne

Korzystanie z aplikacji i jej wszystkich funkcjonalności powinno być dostępne z poziomu przeglądarki internetowej. Ponieważ zadania dot. klasyfikacji wymagają dużej złożoności obliczeniowej, a przechowywanie danych mikromacierzy DNA wymaga stosunkowo dużej ilości pamięci dostawcy usług muszą być dobrani z dużą ostrożnością, aby aplikacja mogła działać możliwie szybko i sprawnie. Ze względu na przechowywanie danych o zarejestrowanych użytkownikach i ich danych kontaktowych (e-mail, telefon) należy zadbać o ich wysokie bezpieczeństwo (przed utratą i próbą kradzieży danych).

1.4. Wymagania funkcjonalne i нефункционалне

14.1. Wymagania funkcjonalne

- Rejestracja użytkowników w systemie (należy podać: imię, nazwisko, data urodzenia, e-mail, hasło, telefon, kod SMS).
- Zarządzanie kontami w systemie z poziomu administratora, w tym dodawanie do systemu nowych kont z uprawnieniami administratora (należy podać: imię, nazwisko, data urodzenia, e-mail, hasło, telefon, kod SMS).
- Logowanie do systemu użytkowników/administratorów (należy podać: e-mail, hasło).
- Wczytywanie danych (użytkownik: swoją mikromacierz DNA).
- Przeglądanie panelu użytkownika gdzie między innymi można sprawdzić historię klasyfikacji oraz zmienić opcję subskrypcji. Panel jest dostępny po zalogowaniu się do systemu przez użytkownika.
- Przeglądanie dostępnych modeli klasyfikacyjnych (badających konkretną przypadłość), możliwe z poziomu gościa/ zalogowanego użytkownika.
- Dokonanie klasyfikacji z poziomu zalogowanego użytkownika wybranym modelem wykorzystując wczytaną mikromacierz DNA.
- Możliwość pobrania wyniku klasyfikacji przez użytkownika na urządzenie.
- Zarządzanie zasobami bazy danych (zbiory mikromacierzy DNA oraz cech klinicznych, stanowiących podstawę do trenowania i testowania modeli klasyfikacyjnych).
- Zarządzanie metodami klasyfikacyjnymi (dodawanie/usuwanie/modyfikowanie dostępnych metod klasyfikacji, selekcji cech, preprocessingu danych).

Definicja aktorów:

Aktor	Opis	Przypadek użycia
Gość	Gość może dokonać podglądu, jakie modele klasyfikacyjne badające mikromacierz DNA pod kątem danej przypadłości zdrowotnej są dostępne w aplikacji. Gość może założyć konto w systemie, a po zalogowaniu staje się użytkownikiem.	-PU Logowanie / Rejestracja -PU Przeglądanie dostępnych modeli klasyfikacyjnych
Użytkownik	Użytkownik korzystając z przeglądarki internetowej może wczytać po zalogowaniu się do systemu swoją mikromacierz DNA i dokonać jej analizy wybranym modelem klasyfikacyjnym, który sprawdzi czy użytkownik jest zagrożony określoną przypadłością zdrowotną.	-PU Logowanie / Rejestracja -PU Przeglądanie dostępnych modeli klasyfikacyjnych -PU Wybranie modelu klasyfikacyjnego powiązane przez <<include>> z PU Wczytanie swojej mikromacierzy DNA powiązane przez <<extends>> z PU Odczytanie wyniku klasyfikacji powiązane przez <<extends>> z PU Pobranie wyników na urządzenie -PU Przeglądanie panelu użytkownika
Administrator	Administrator może zarządzać systemem, co obejmuje modyfikowanie/dodawanie/usuwanie kont, metod klasyfikacyjnych, metod selekcji danych, metod preprocessingu, danych mikromacierzowych DNA oraz modeli klasyfikacyjnych, (z których mogą korzystać zalogowani użytkownicy).	-PU Zarządzanie systemem powiązane przez <<extends>> z PU Zarządzanie kontami oraz powiązane przez <<extends>> z PU zarządzanie zasobami bazy danych oraz powiązane przez <<extends>> z PU Zarządzanie metodami klasyfikacyjnymi, selekcji cech, preprocessingu danych oraz powiązane przez <<extends>> z PU Zarządzanie modelami klasyfikacyjnymi

Definicja scenariuszy przypadków użycia:**PU Logowanie / Rejestracja****Opis**

Cel: Utworzenie konta w systemie lub zalogowanie się do systemu na istniejące konto, aby w pełni korzystać z aplikacji.

WS: Inicjalizacja przez uruchomienie aplikacji (otwarcie odpowiedniej strony WWW z poziomu przeglądarki internetowej)

WK: Rejestracja nowego konta lub zalogowanie się do systemu na istniejące konto podając prawidłowe dane w formularzu rejestracji / logowania.

Przebieg:

Przy rejestracji należy podać: imię, nazwisko, data urodzenia, e-mail, hasło, telefon oraz kod SMS przesłany na podany numer telefonu. W formularzu rejestracyjnym można wybrać opcje darmowego konta lub konto z subskrypcją. Dane jakie trzeba podać przy logowaniu to e-mail i hasło. Przy rejestracji należy sprawdzić czy dane są prawidłowe i czy konto o podanych danych już istnieje. Po podaniu błędnych danych lub gdy konto będzie już istniało w systemie wyświetli się stosowny komunikat. PU kończy się w momencie sukcesywnego zalogowania się na istniejące konto lub po utworzeniu nowego.

PU Wybranie modelu klasyfikacyjnego

Opis

Cel: Celem jest wybranie gotowego modelu klasyfikacyjnego z listy dostępnych w systemie, dzięki któremu użytkownik pozyska informację o wyniku klasyfikacji w kontekście konkretnej przypadłości zdrowotnej, na podstawie mikromacierzy DNA przesłanej przez użytkownika.

WS: Inicjalizacja przez wybranie konkretnego modelu klasyfikacyjnego z listy dostępnych modeli w systemie.

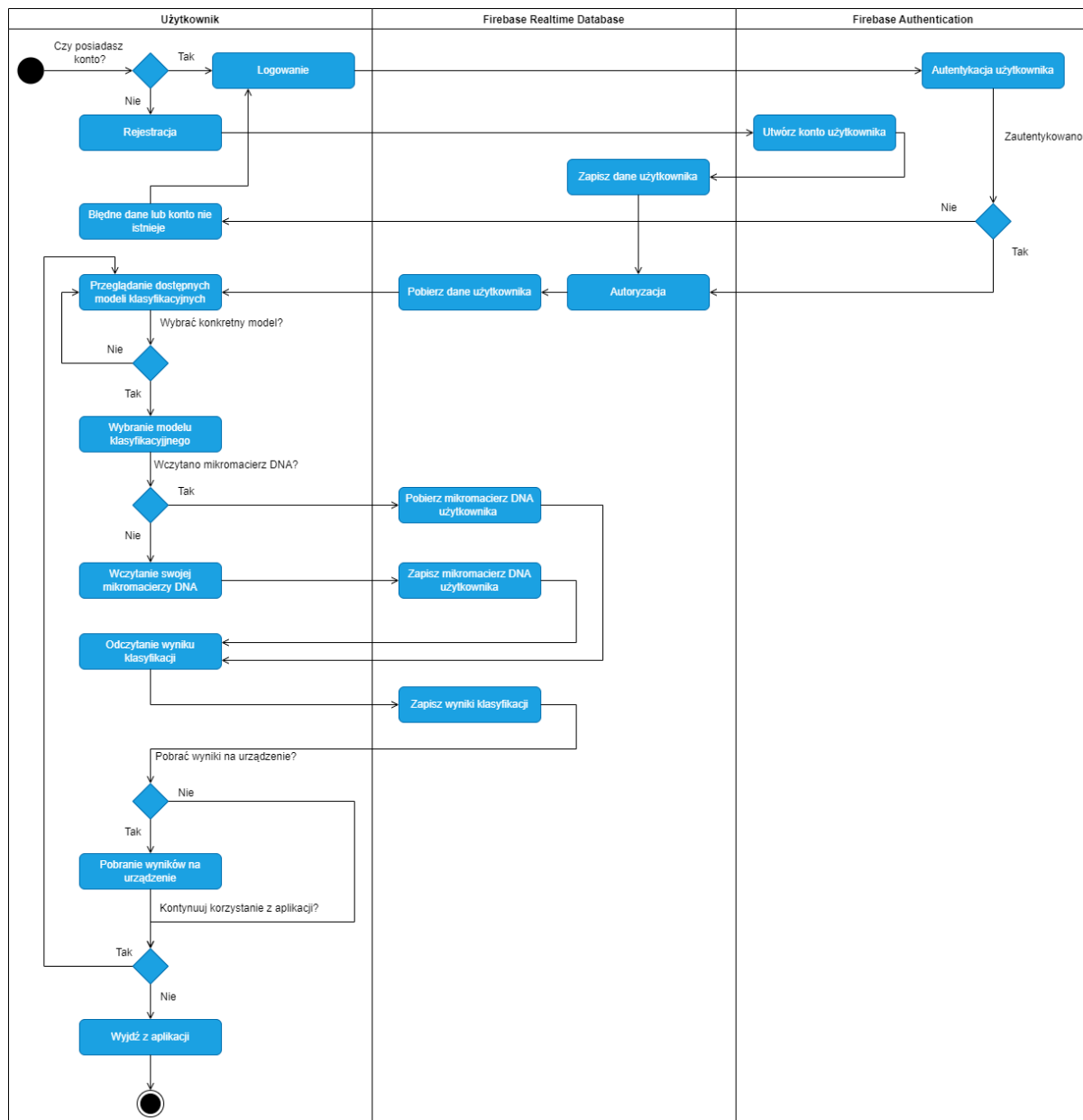
WK: Pozyskanie wyniku analizy mikromacierzy DNA użytkownika przez wybrany model klasyfikujący.

Przebieg:

Przeglądając listę dostępnych modeli klasyfikacyjnych użytkownik wybiera model, który bada przypadłość zdrowotną, która interesuje użytkownika. Za pomocą przesłanej mikromacierzy DNA przez użytkownika, model dokonuje klasyfikacji czy użytkownik ten jest lub będzie zagrożony daną przypadłością zdrowotną. PU kończy się wraz z zapisem do bazy danych i wyświetleniem wyniku klasyfikacji.

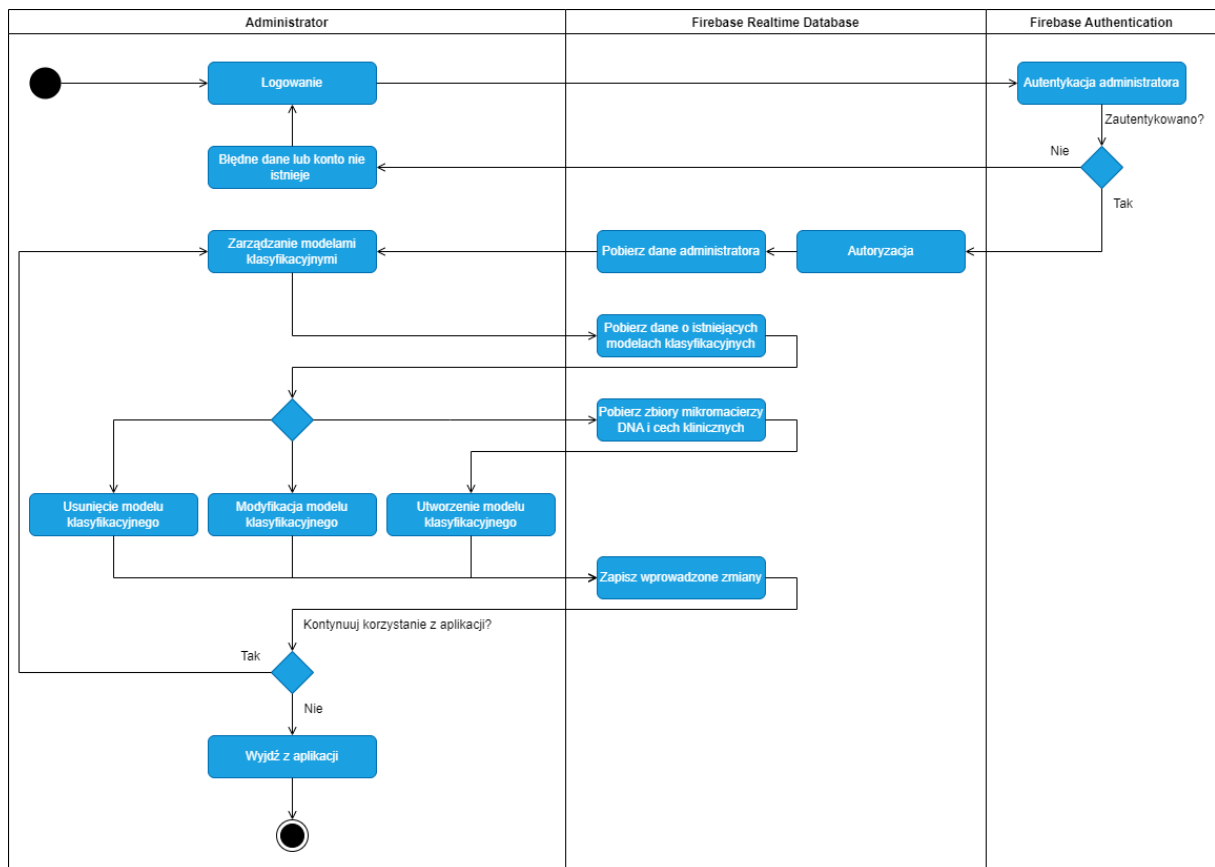
2.2. Diagramy aktywności

Poniższy obraz przedstawia pierwszy diagram aktywności wizualizujący procesy zachodzące w systemie, gdzie celem jest poprawne dokonanie przez użytkownika czynności „Wybranie modelu klasyfikacyjnego” wraz z „Odczytanie wyniku klasyfikacji”.



Rysunek 2. Diagram aktywności wizualizujący procesy w trakcie realizowania PU Wybranie modelu klasyfikacyjnego, wraz z PU Odczytanie wyniku klasyfikacji

Poniższy obraz przedstawia drugi diagram aktywności wizualizujący procesy zachodzące w systemie, gdzie celem jest poprawne dokonanie przez administratora czynności „Zarządzanie modelami klasyfikacyjnymi”.

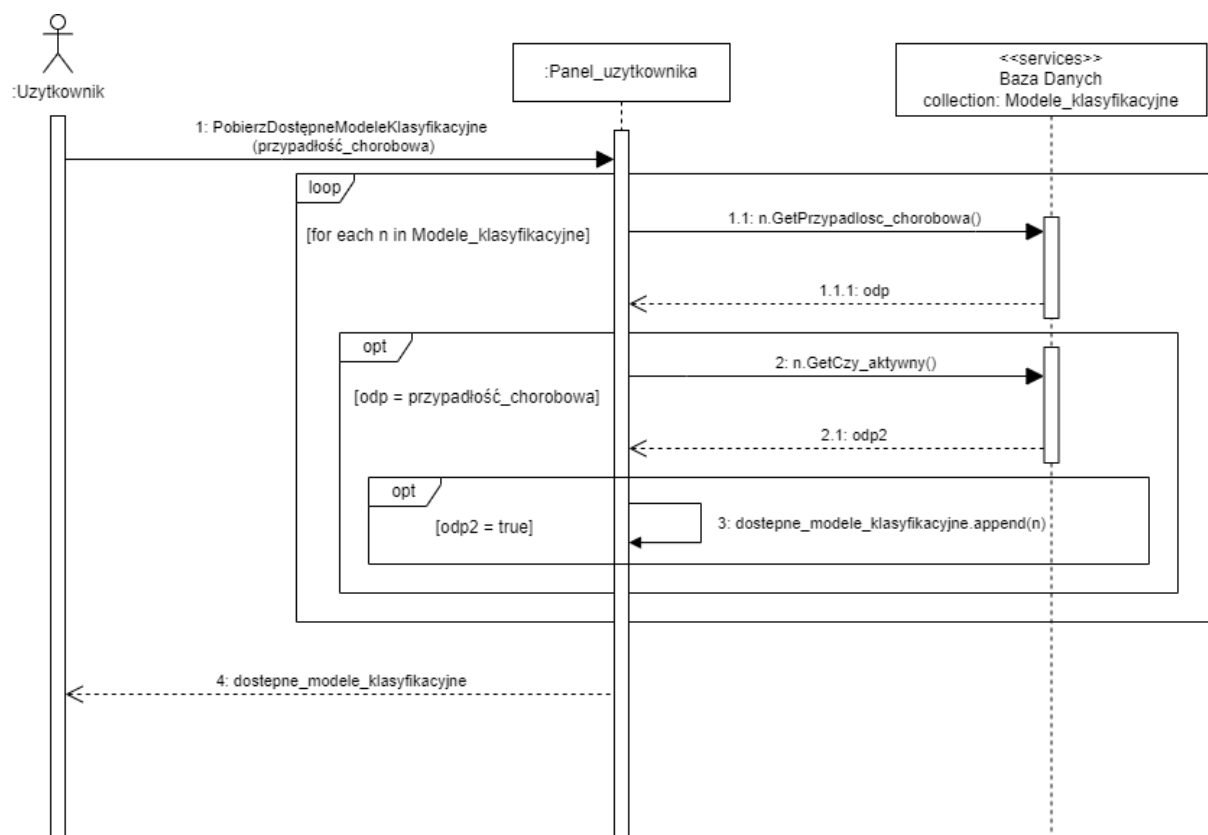


Rysunek 3. Diagram aktywności wizualizujący procesy w trakcie realizowania PU Zarządzanie modelami klasyfikacyjnymi.

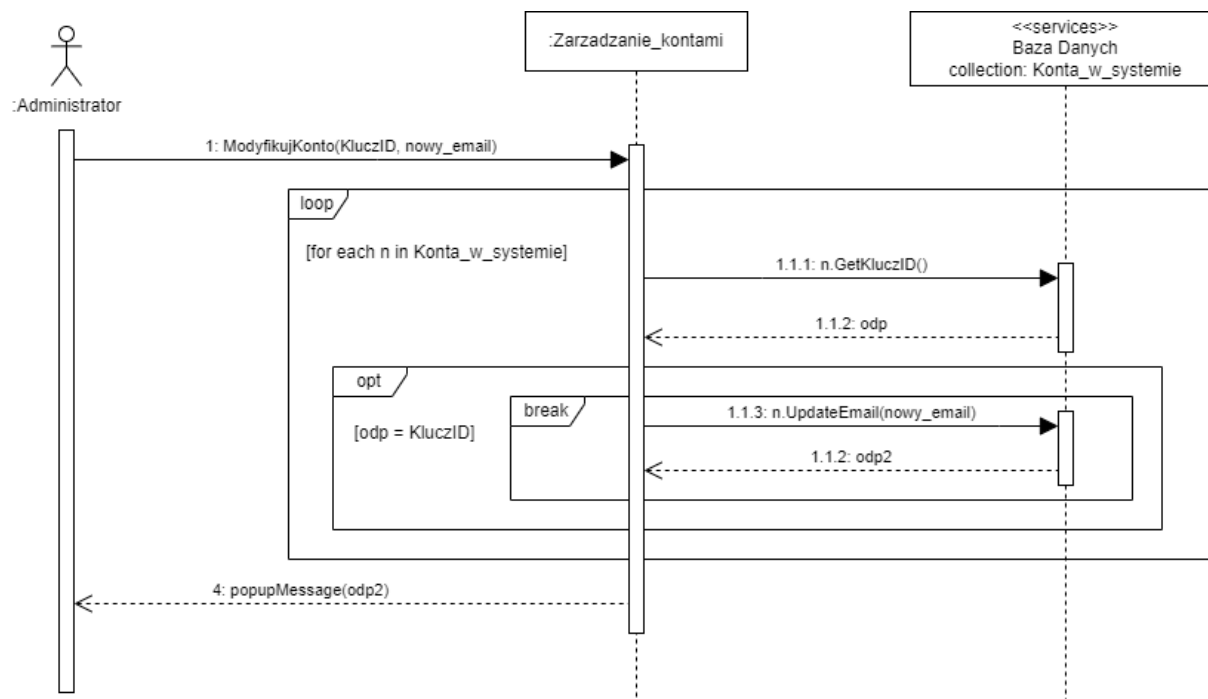
2.3. Diagram klas

Poniższy obraz przedstawia diagram klas opisujący strukturę projektowanego programu (a także zależności między nimi), co znajduje odzwierciedlenie w kodzie. Wszystkie pola w klasach domyślnie posiadają metody dostępu tzw. gettery i setery.

zamieszczonych diagramów obrazuje interakcje pomiędzy obiektami zachodzące w procesie modyfikacji istniejącego konta w systemie przez administratora.



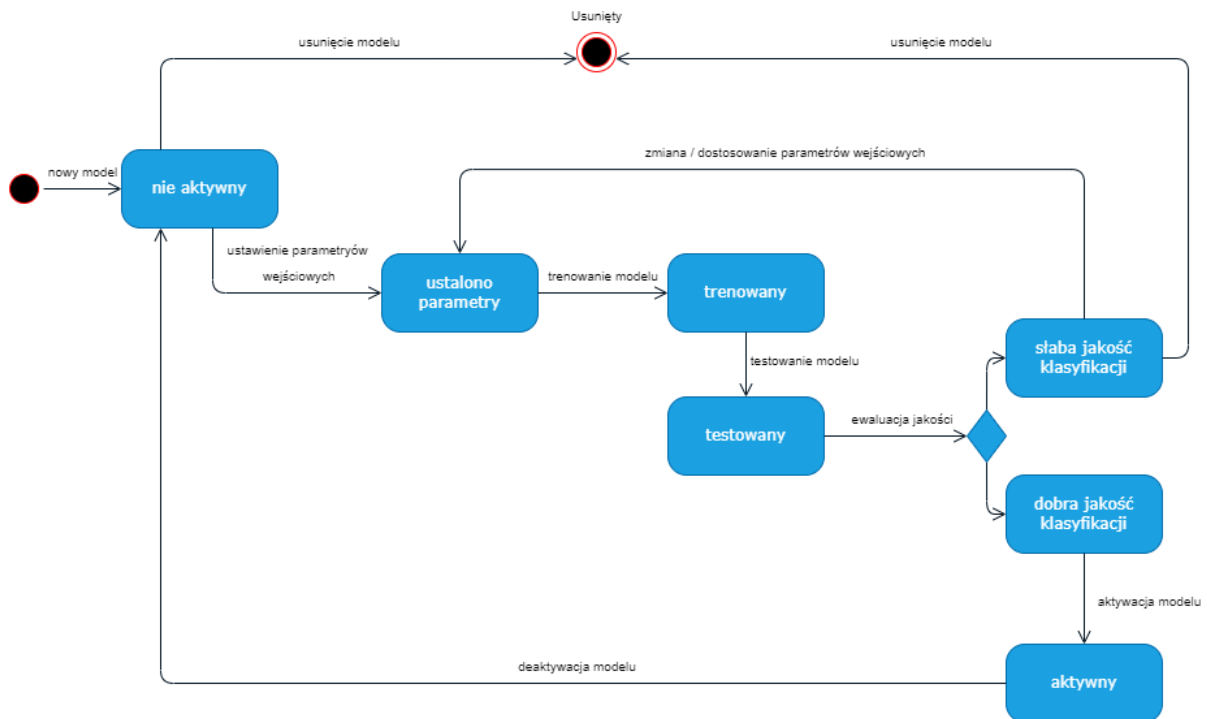
Rysunek 5. Diagram sekwencji od strony użytkownika – pobranie dostępnych modeli klasyfikacyjnych w systemie.



Rysunek 6. Diagram sekwencji od strony administratora – modyfikacja istniejącego konta w systemie.

2.5. Diagram stanów

Diagram maszyny stanowej jest jednym z diagramów aktywności UML. Diagram ten przedstawia zbiór stanów przyjmowanych przez obiekty w odpowiedzi na zdarzenie zachodzące w czasie jego życia, a także reakcje obiektu na te zdarzenia. Przedstawiony diagram obrazuje zbiór stanów dla obiektu modelu klasyfikacyjnego w systemie.

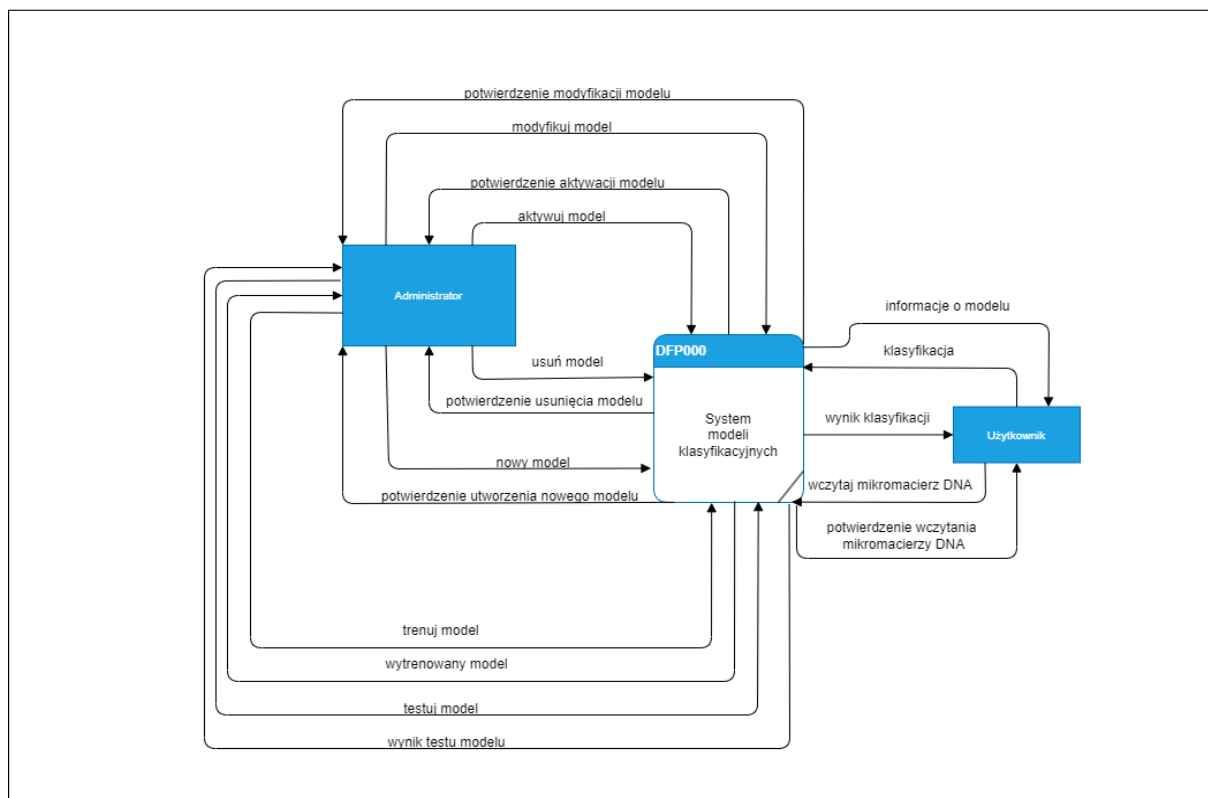


Rysunek 7. Diagram stanów dla obiektu modelu klasyfikacyjnego w systemie.

2.6. Diagram przepływu danych

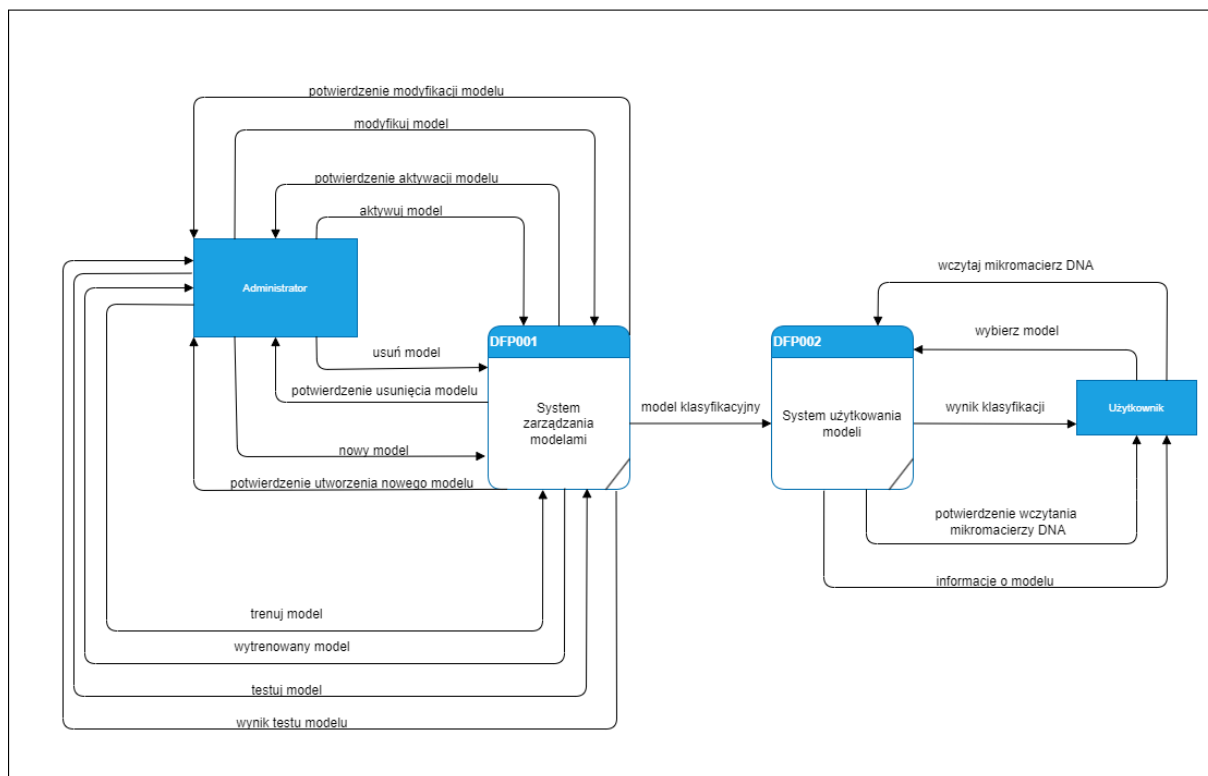
Diagram przepływu danych (DPD) odwzorowuje przepływ informacji dla dowolnego procesu lub systemu. Przedstawione diagramy opisują na coraz głębszych poziomach przepływ informacji podczas korzystania z systemu modeli klasyfikacyjnych w aplikacji przez administratora jak i użytkownika.

Diagram kontekstowy



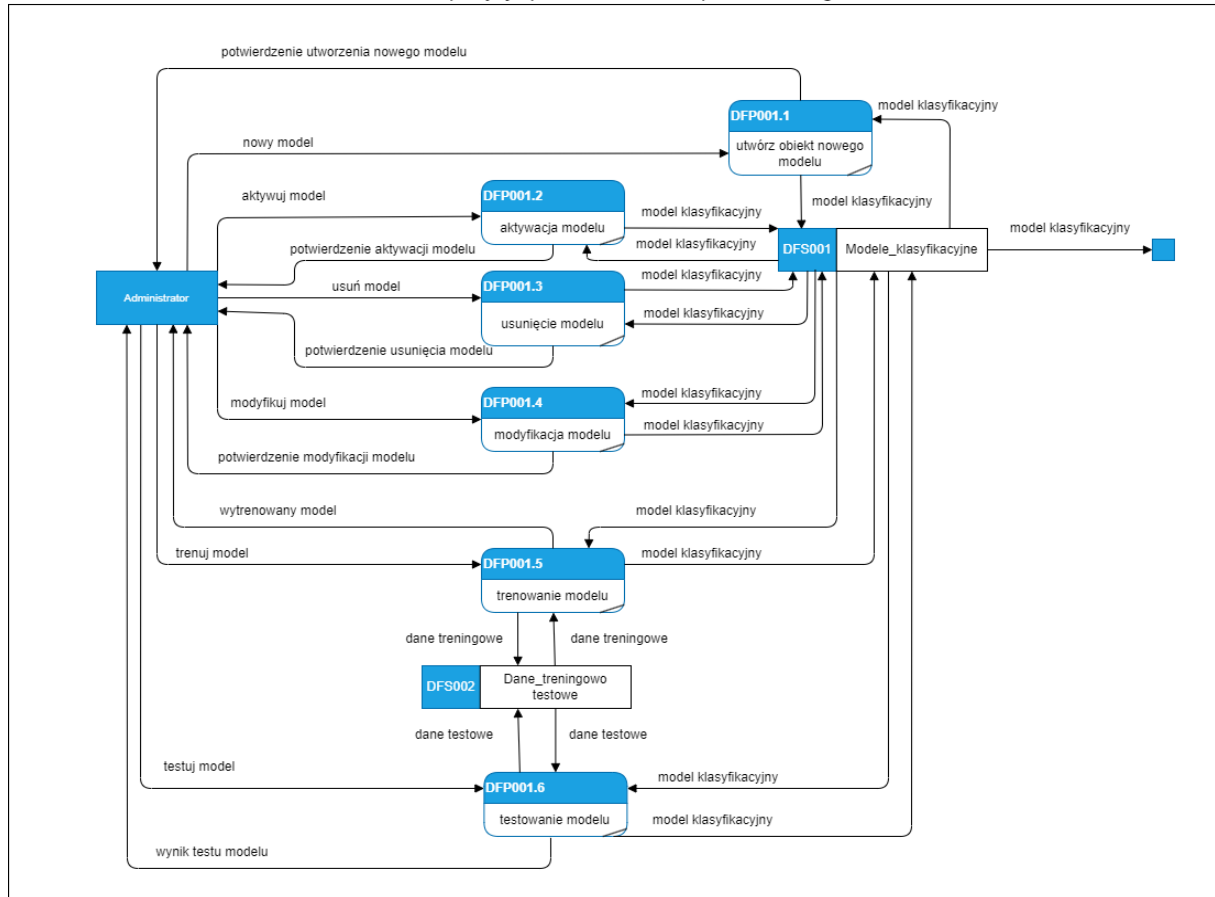
Rysunek 8. Diagram kontekstowy DPD - system modeli klasyfikacyjnych (DFP000).

Dekompozycja procesu DFP000 poziomu 0-ego



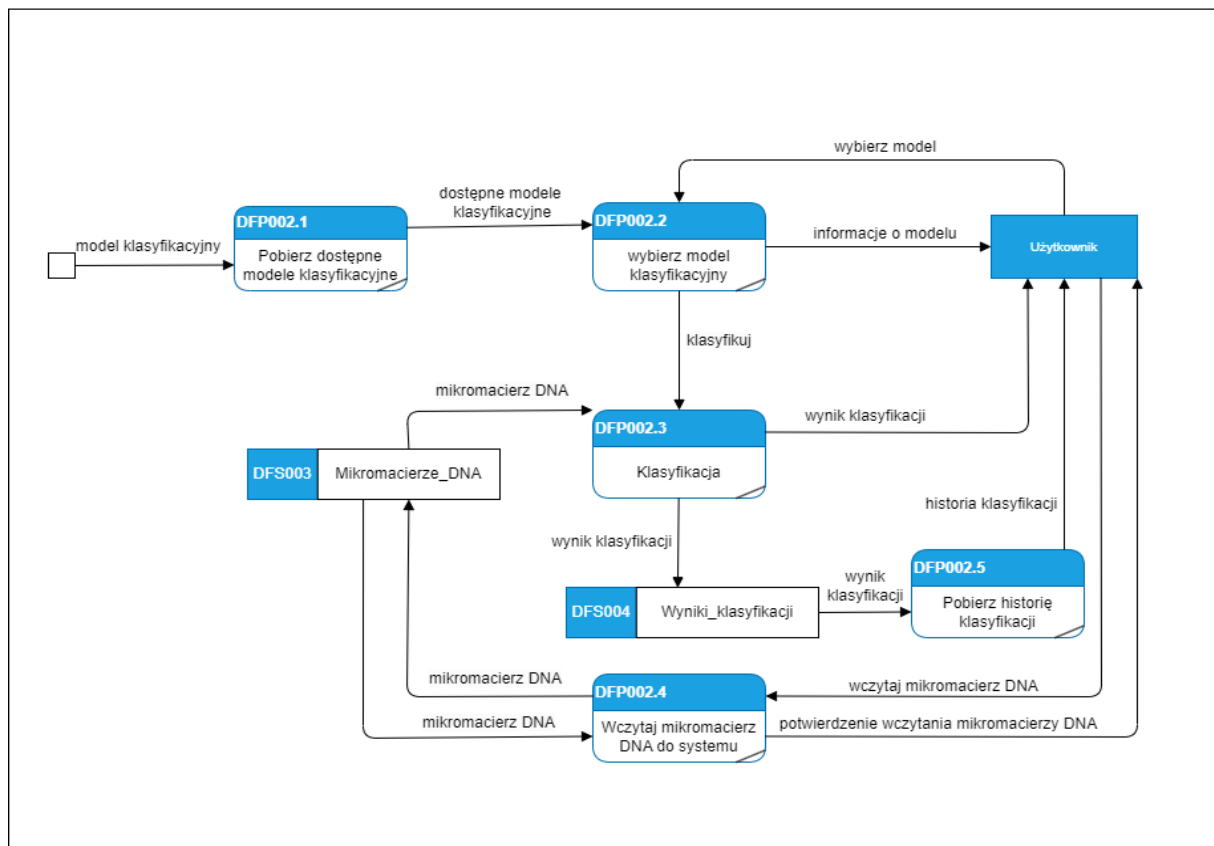
Rysunek 9. Diagram obrazujący dekompozycję poziomu zerowego – system modeli klasyfikacyjnych (DFP000).

Dekompozycja procesu DFP001 poziomu 1-ego



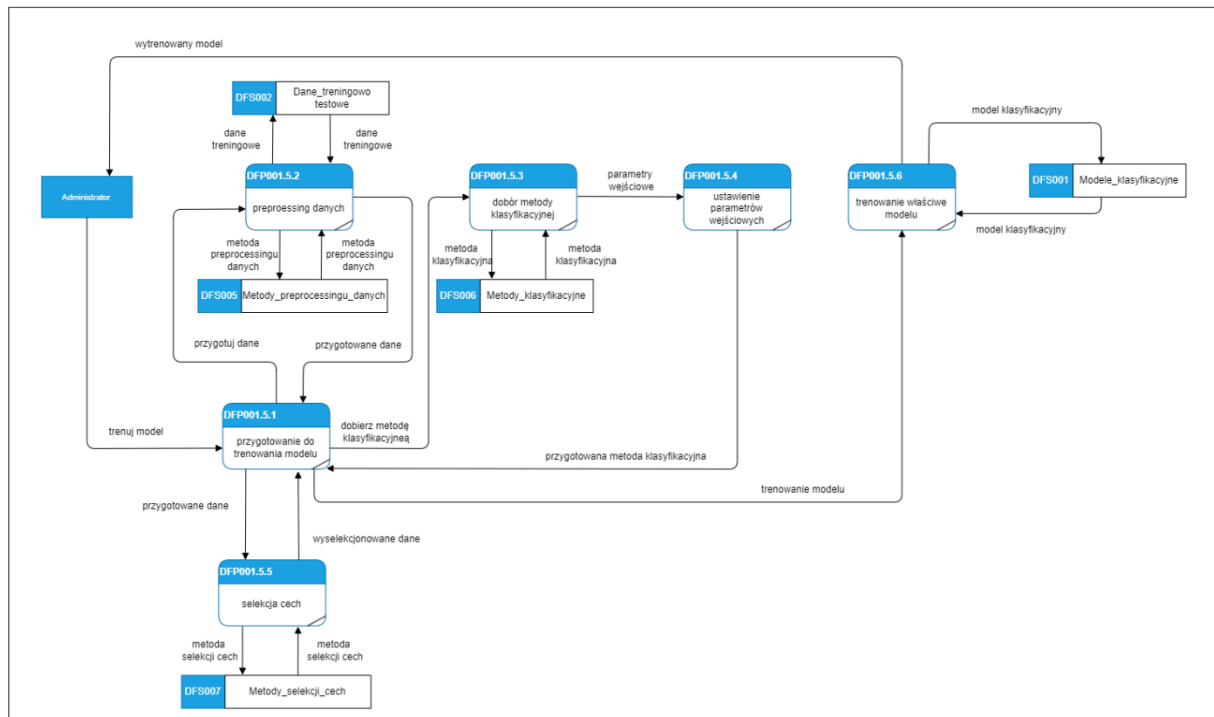
Rysunek 10. Diagram obrazujący dekompozycję poziomu pierwszego – system zarządzania modelami (DFP001).

Dekompozycja procesu DFP002 poziomu 1-ego



Rysunek 11. Diagram obrazujący dekompozycję poziomu pierwszego – system użytkowania modeli (DFP002).

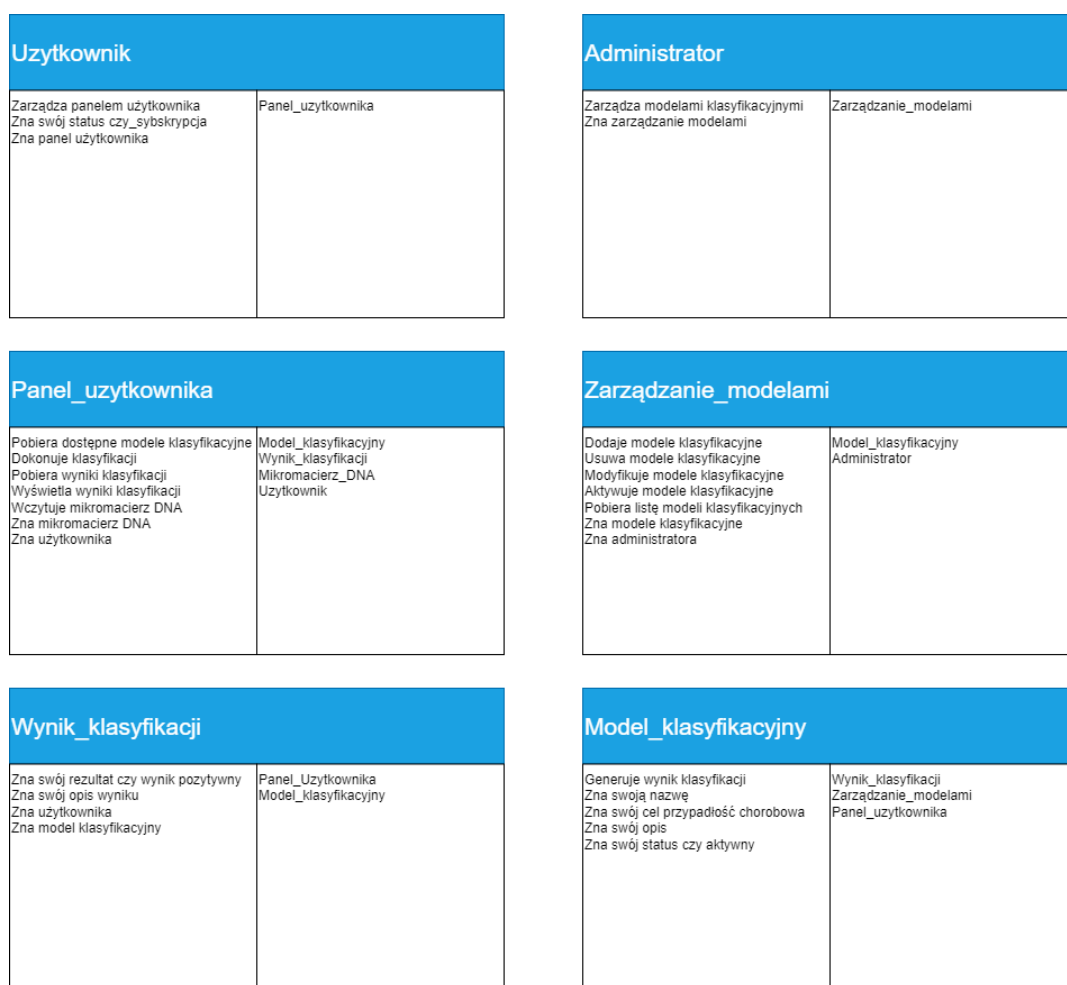
Dekompozycja procesu DFP001.5 poziomu 2-ego



Rysunek 12. Diagram obrazujący dekompozycję poziomu drugiego – trenowanie modeli (DFP001.5).

2.7. Karta CRC - Class Responsibility Collaboration (CRC)

Karty CRC zostały pierwotnie wprowadzone jako technika nauczania pojęć obiektowych, ale z powodzeniem są również wykorzystywane jako pełnoprawna technika modelowania. Model CRC jest zbiorem standardowych kart indeksowych, które zostały podzielone na trzy sekcje (nazwa klasy / obiektu, obowiązki, czyli to coś, co klasa wie lub robi oraz współpracownicy, tj. inna klasa, z którą wchodzi w interakcję, aby wypełnić swoje obowiązki).



Rysunek 13. Karta CRC utworzona w kontekście PU Wybranie modelu klasyfikacyjnego.

3. Opis techniczny projektu

Informacje o wykorzystanych technologiach i narzędziach:

- środowisko programistyczne: PyCharm Professional 2023.1.2,
- wykorzystany język: Python 3.9, JavaScript ECMAScript 2022,
- wykorzystane biblioteki: Scikit-learn 1.2.2, React.js 18.2.0,

- wykorzystane usługi: hosting - Netlify, baza danych - Firebase Realtime DB,

Wymagania sprzętowe:

- dwurdzeniowy procesor o częstotliwości taktowania co najmniej 2 GHz,
- minimum 4 GB RAM,
- najnowsze wersje popularnych przeglądarek, takich jak Google Chrome, Mozilla Firefox, Safari, Opera lub Microsoft Edge,
- kompatybilność z różnymi systemami operacyjnymi, takimi jak Windows, macOS i Linux,
- szybkie i stabilne połączenie internetowe z przepustowością co najmniej 10 Mbps,
- brak wymaganej rozdzielczości ekranu (optymalizacja dla różnych rozdzielczości ekranów, w tym dla urządzeń mobilnych).

4. Prezentacja warstwy użytkowej projektu

GUI aplikacji będzie budowane w oparciu o czysto neutralną paletę kolorów. Głównymi motywatorami było uzyskanie takich efektów jak:

- Elegancja i nowoczesność.
- Czystość i prostota.
- Bezczasowość designu.
- Wywołanie poczucia rzetelności i jakości.
- Efekt uspokajający.



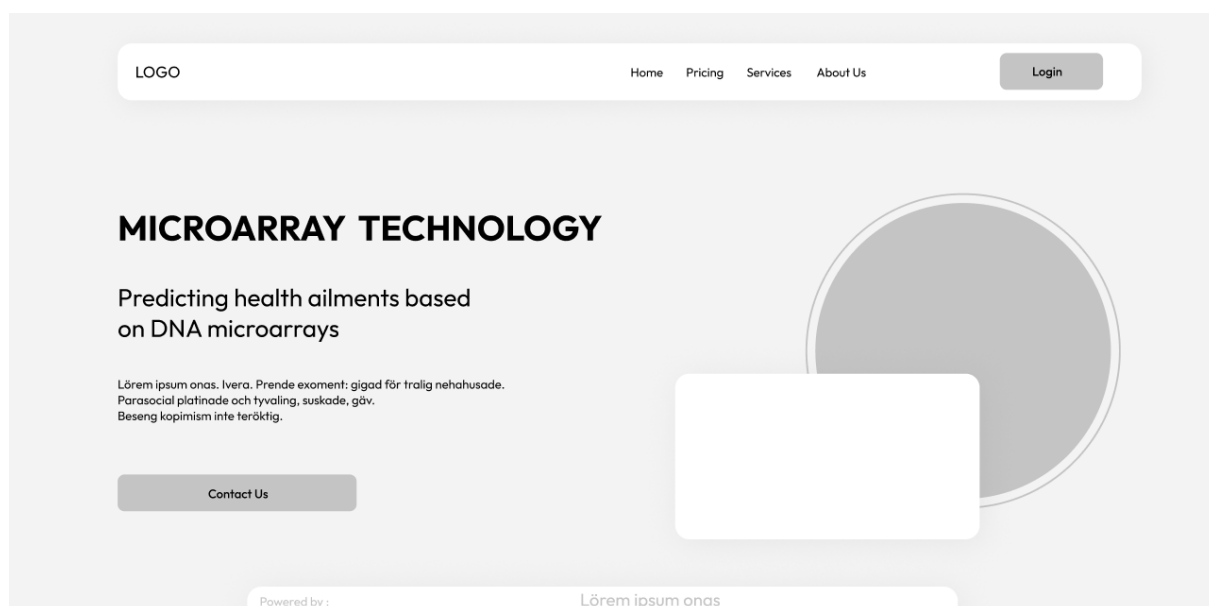
Rysunek 14. Wykorzystana paleta kolorów przy budowie GUI tworzonej aplikacji.

Przykłady okien warstwy prezentacji tworzonej aplikacji:

Na rysunku 15 przedstawiono główne okno aplikacji. Użytkownik po uruchomieniu aplikacji otrzymuje ekran startowy, na którym może:

- Zalogować się lub zarejestrować klikając przycisk „Login”.
- Przejść w zakładkę „Pricing” z informacjami o opcji subskrypcji.
- Przejść w zakładkę „Services” z opisem dostępnych usług.
- Przejść w zakładkę „About Us” z ogólnymi informacjami o aplikacji.

- Kliknąć przycisk „Contact Us” w celu wyświetlenia informacji kontaktowych oraz gotowym formularzem do przesłania wiadomości e-mail.

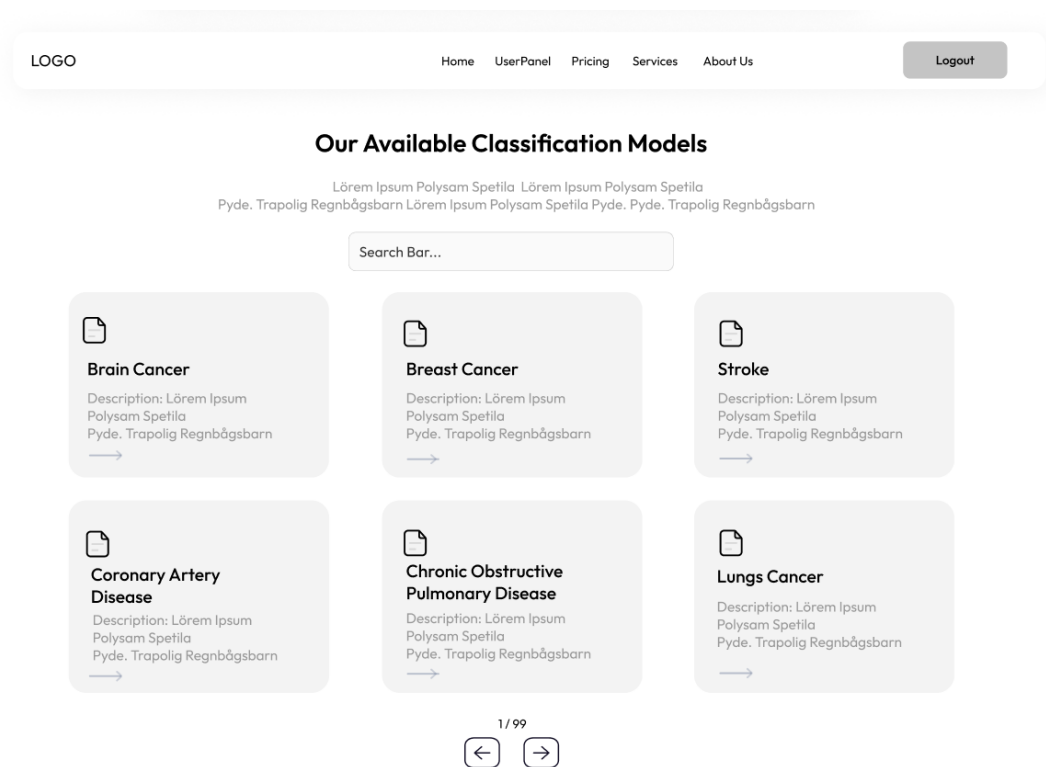


Rysunek 15. Strona startowa tworzonego projektu.

Na rysunku 15 przedstawiono okno pozwalające na przeglądanie dostępnych modeli klasyfikacyjnych w systemie przez użytkownika. Z poziomu tego okna użytkownik może:

- Przeglądać dostępne modele klasyfikacyjne wraz z ich opisem i nazwami przypadłości, pod kątem których dokonywana będzie klasyfikacja.
- Kliknąć w pożądany obiekt modelu klasyfikującego w celu dokonania klasyfikacji.
- Możemy wyszukać konkretne modele klasyfikacyjne za pomocą wyszukiwarki.

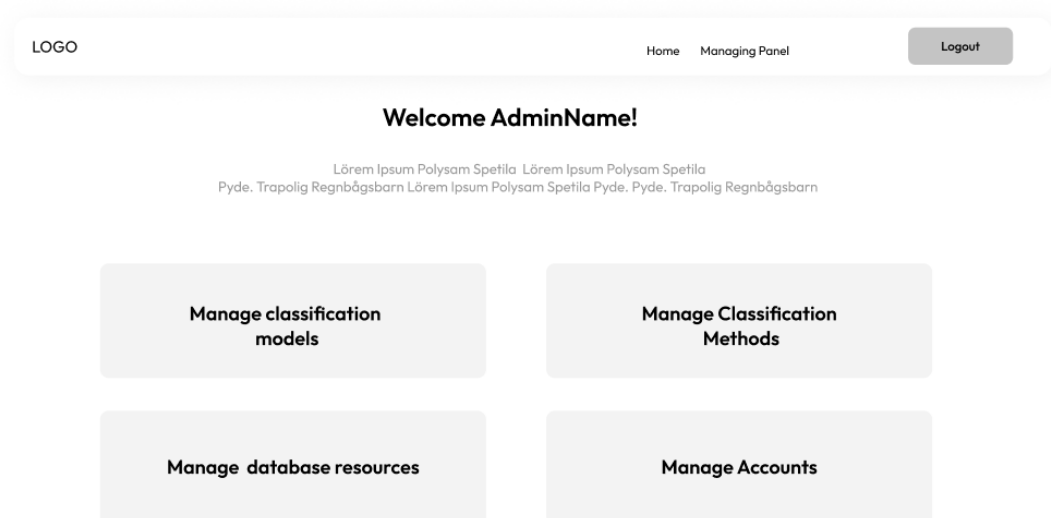
Po zalogowaniu się do systemu mamy teraz możliwość wylogowania się z poziomu paska menu. Pojawia się również możliwość przejścia do panelu użytkownika klikając w „UserPanel” z poziomu paska menu.



Rysunek 16. Widok z przeglądem dostępnych modeli klasyfikacyjnych z poziomu użytkownika.

Na rysunku 16 przedstawiony jest wygląd okna zarządzania systemem. Dostępne jest ono po zalogowaniu się przez administratora systemu. Ma on wtedy takie możliwości jak:

- Przejść do zarządzania modelami klasyfikacyjnymi.
- Przejść do zarządzania metodami biorącymi udział w tworzeniu modeli klasyfikacyjnych.
- Przejść do zarządzania zasobami bazy danych.
- Przejść do zarządzania kontami w systemie.
- Opcja wylogowania się oraz przejścia na stronę startową.



Rysunek 17. Widok z przeglądem dostępnych modeli klasyfikacyjnych z poziomu użytkownika.

5. Podsumowanie

Projekt aplikacji pt. "Aplikacja internetowa do przewidywania przypadłości zdrowotnych na podstawie mikromacierzy DNA" to innowacyjny pomysł implementacji aplikacji umożliwiającej eksplorowanie przekazanej przez użytkownika mikromacierzy DNA pod kątem różnego rodzaju chorób uwarunkowanych genetycznie. Znajomość o zwiększonej możliwości wystąpienia konkretnej choroby pozwoliłaby takiej osobie przygotować się i rozpocząć leczenie lub profilaktykę w odpowiednim czasie.

W projekcie opisano zadania i możliwości poszczególnych uczestników systemu, oraz przybliżono strategię firmy. Opisano aspekty techniczne oraz wymagania funkcjonalne i niefunkcjonalne. W następnym kroku rozpoczęto modelowanie systemu prezentując kolejno diagramy UML, tj. diagram przypadków użycia wraz z definicją scenariuszy przypadków użycia, diagram aktywności, diagram klas, diagram sekwencji, diagram stanów oraz przedstawiono przykład karty CRC. W następnym kroku wprowadzono opis techniczny projektu obejmujący informacje o wykorzystanych technologiach i narzędziach. W ostatnim kroku dokonano prezentacji warstwy użytkowej projektu na przykładach, które ukazują docelowy design i zastosowaną paletę barw przy budowie części front-endowej.