PROJEKT ZESPOŁOWY 02fz-pz

Aplikacja do analizowania dokumentów CV

**Projekt aplikacji do analizy CV z wykorzystaniem Flask**

Spis treści

[1. Wstęp 2](#_Toc185123264)

[2. Cel projektu 2](#_Toc185123265)

[3. Główne funkcjonalności 3](#_Toc185123266)

[Rejestracja 3](#_Toc185123267)

[Logowanie 6](#_Toc185123268)

[Przesyłanie dokumentu CV 10](#_Toc185123269)

[Analiza dokumentu CV 12](#_Toc185123270)

[Ekstrakcja treści CV 12](#_Toc185123271)

[Wyszukiwanie kluczowych informacji 13](#_Toc185123272)

[Ocena dopasowana do stanowiska 14](#_Toc185123273)

[Prezentacja wyników 15](#_Toc185123274)

[Zarządzanie stanowiskami 16](#_Toc185123275)

[Dodawanie stanowisk 17](#_Toc185123276)

[Edytowanie stanowisk 19](#_Toc185123277)

[Usuwanie stanowisk 24](#_Toc185123278)

[Ranking kandydatów 25](#_Toc185123279)

[Wybór stanowiska pracy i ograniczenie liczby wyników 26](#_Toc185123280)

[Prezentacja wyników 27](#_Toc185123281)

[Dodatkowe akcje dla kandydatów 29](#_Toc185123282)

[3. Technologie i biblioteki 30](#_Toc185123283)

[Języki programowania 30](#_Toc185123284)

[Pozostałe języki 30](#_Toc185123285)

[Frameworki i biblioteki backendowe 30](#_Toc185123286)

[Frameworki i biblioteki frontendowe 30](#_Toc185123287)

[Biblioteki NLP i analizy danych 30](#_Toc185123288)

[Bazy danych 30](#_Toc185123289)

[Infrastruktura i narzędzia wdrożeniowe 31](#_Toc185123290)

[Zarządzanie użytkownikami i bezpieczeństwo 31](#_Toc185123291)

[Funkcje pomocnicze 31](#_Toc185123292)

[Systemy kontroli wersji i organizacja pracy 31](#_Toc185123293)

[4. Struktura katalogów 32](#_Toc185123294)

[5. Możliwe rozszerzenia 32](#_Toc185123295)

[7. Podsumowanie 33](#_Toc185123296)

# Wstęp

Prezentowana aplikacja została stworzona, aby wspierać procesy rekrutacyjne poprzez automatyczne przetwarzanie i analizę dokumentów aplikacyjnych. Jej głównym zadaniem jest ułatwienie selekcji kandydatów poprzez identyfikację kluczowych informacji i ocenę ich dopasowania do zdefiniowanych wymagań stanowiskowych. Dzięki wykorzystaniu nowoczesnych technologii, takich jak przetwarzanie języka naturalnego (NLP) i zaawansowana analiza danych, system oferuje intuicyjne zarządzanie procesem rekrutacji. Aplikacja umożliwia precyzyjne przypisywanie parametrów do stanowisk pracy, co pozwala na dokładne porównanie kwalifikacji kandydatów. Zastosowanie tej technologii pozwala firmom zoptymalizować pracę zespołów HR, skracając czas potrzebny na selekcję aplikacji i wspierając proces podejmowania obiektywnych decyzji.

# Cel projektu

Podstawowym celem projektu jest stworzenie narzędzia, które w znaczący sposób automatyzuje procesy rekrutacyjne, eliminując konieczność ręcznej analizy CV. System koncentruje się na wykorzystaniu algorytmów przetwarzania języka naturalnego do wydobywania i oceny danych kluczowych, takich jak kwalifikacje, doświadczenie czy umiejętności kandydatów. Funkcjonalności aplikacji obejmują definiowanie stanowisk pracy wraz z wymaganiami oraz tworzenie mechanizmów, które klasyfikują kandydatów na podstawie stopnia spełnienia tych wymagań. Ważnym aspektem systemu jest jego zdolność do automatycznego generowania rankingów kandydatów, co przyspiesza i usprawnia proces decyzyjny. Realizacja tego celu wspiera organizacje w zmniejszeniu kosztów rekrutacji, poprawie jakości wyboru kandydatów oraz bardziej efektywnym zarządzaniu zasobami ludzkimi.

# Główne funkcjonalności

## Rejestracja

Moduł rejestracji polegający na założeniu konta użytkownika, aby miał dostęp do swoich indywidualnych danych kandydatów i stanowisk został zaprojektowany w sposób umożliwiający prosty proces zakładania konta użytkownika, bez dodatkowych uwierzytelnień. Rejestracja wymaga podania podstawowych danych, takich jak:

1. Nazwa użytkownika
2. Adres e-mail
3. Hasło

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Aby zapewnić bezpieczeństwo danych użytkowników, hasła są szyfrowane za pomocą odpowiedniej funkcji przed zapisaniem ich w bazie danych. System wykorzystuje metodę set\_password z importu biblioteki werkzeug.security w modelu użytkownika, która implementuje szyfrowanie haseł.

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

W przypadku błędów, takich jak próba rejestracji z już istniejącą nazwą użytkownika, system wyświetla komunikat za pomocą funkcji flash, a użytkownik jest przekierowywany z powrotem na stronę rejestracji.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Po pomyślnym utworzeniu konta, nowy użytkownik jest zapisywany w bazie danych, a sesja kończy się przekierowaniem na stronę logowania z odpowiednim komunikatem potwierdzającym.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Użytkownik po zarejestrowaniu jest widoczny w bazie danych w tabeli user z zaimplementowanymi danymi na kolumnie id, email, password\_hash (szyfrowanie haseł).

Obraz zawierający tekst, numer, linia, Czcionka

Opis wygenerowany automatycznie

## Logowanie

Moduł logowania został zaprojektowany, aby umożliwić użytkownikom dostęp do ich indywidualnych kont w systemie, gdzie mogą zarządzać danymi kandydatów i stanowisk. Logowanie wymaga podania dwóch podstawowych danych uwierzytelniających:

1. Nazwa użytkownika
2. Hasło

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

System weryfikuje podane dane na podstawie zapisanych w bazie informacji. Hasło użytkownika jest sprawdzane za pomocą funkcji check\_password, która wykorzystuje algorytm porównujący podane hasło z jego zaszyfrowaną wersją przechowywaną w bazie danych.

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

W przypadku błędnych danych logowania, takich jak niepoprawna nazwa użytkownika lub hasło, użytkownik otrzymuje odpowiedni komunikat o błędzie za pomocą funkcji flash, a system odświeża stronę logowania. Mechanizm ten zapewnia jasną informację zwrotną bez konieczności przerywania pracy aplikacji.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Po prawidłowym uwierzytelnieniu użytkownik jest logowany do systemu, a jego identyfikator (user\_id) zostaje zapisany w sesji za pomocą mechanizmu Flask Sessions. Umożliwia to autoryzowany dostęp do chronionych zasobów aplikacji oraz personalizację danych. Po zalogowaniu system przekierowuje użytkownika na stronę główną aplikacji.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

## Przesyłanie dokumentu CV

Moduł przesyłania dokumentów CV pozwala użytkownikom na dodanie plików CV kandydatów do systemu w celu przeprowadzenia ich analizy i dopasowania do wybranych stanowisk pracy.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

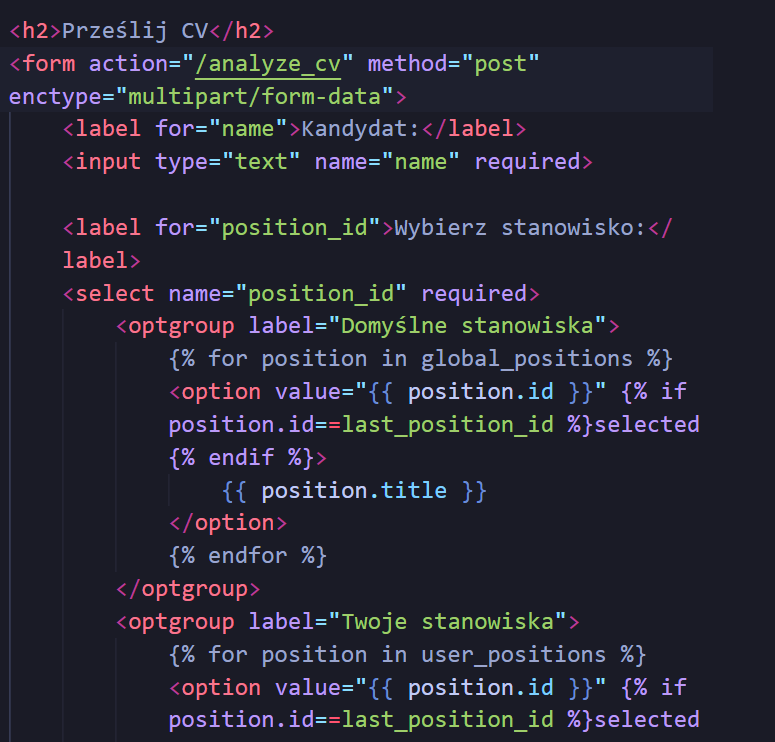
Formularz przesyłania danych składa się z kilku prostych sekcji:

1. Pole "Kandydat" umożliwia wpisanie dowolnej nazwy, co pozwala na przypisanie dokumentu do odpowiedniej osoby.
2. Lista rozwijana "Wybierz stanowisko" pozwala użytkownikowi wybrać stanowisko, dla którego CV będzie analizowane. Lista jest podzielona na kategorie:
   1. Domyślne stanowiska dostępne globalnie w systemie.
   2. Stanowiska użytkownika, które zostały wcześniej zdefiniowane przez zalogowanego użytkownika.
3. Pole „Plik CV” pozwala na przesłanie pliku w formacie .pdf

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, numer

Opis wygenerowany automatycznie

Po wypełnieniu formularza i wybraniu pliku użytkownik może przesłać dane do analizy, klikając przycisk „Analizuj”. Dane zostają wysłane na serwer w celu przetworzenia i oceny dopasowania CV do wybranego stanowiska.



W przypadku potrzeby powrotu do poprzedniego ekranu użytkownik może skorzystać z przycisku „Wróć”.

## Analiza dokumentu CV

Proces analizy rozpoczyna się od odebrania danych przesłanych w formularzu:

* Dowolna nazwa kandydata
* Identyfikator stanowiska pracy wybranego z listy
* Plik CV w formacie PDF

Obraz zawierający tekst, zrzut ekranu, Czcionka, design

Opis wygenerowany automatycznie

Plik CV jest zapisywany w bazie danych pod unikalną nazwą, która łączy imię kandydata z nazwą przesłanego pliku.

Ekstrakcja treści CV

Dokument CV jest przetwarzany w celu wyodrębnienia zawartości tekstowej. Dla plików PDF, używana jest biblioteka pdf2image do konwersji stron PDF na obrazy, a następnie pytesseract do odczytu tekstu z tych obrazów. W przypadku błędów podczas ekstrakcji system informuje użytkownika o problemie za pomocą komunikatów flash.

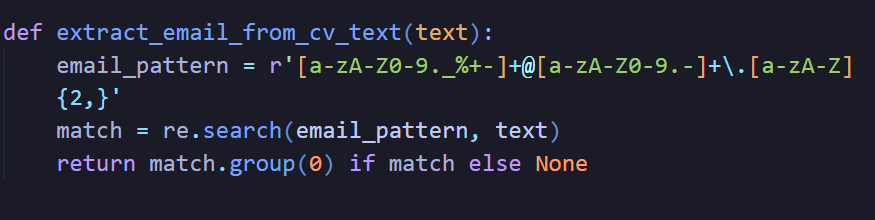
Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

### Wyszukiwanie kluczowych informacji

System przeprowadza analizę zawartości tekstowej w celu wyodrębnienia istotnych danych, takich jak:

* Imię i nazwisko kandydata – identyfikowane na podstawie analizy tekstu w pierwszych liniach dokumentu.
* Adres e-mail – wyodrębniany za pomocą wzorca regularnego dopasowującego popularne formaty adresów e-mail.
* Numer telefonu – znajdowany na podstawie wyrażeń regularnych identyfikujących formaty numerów telefonicznych.



Ocena dopasowana do stanowiska  
  
Wyszukiwane są wystąpienia słów kluczowych zdefiniowanych dla wybranego stanowiska pracy. Każde słowo kluczowe posiada określoną wagę, która wpływa na końcową punktację. Mechanizm działa następująco:

1. Tekst CV jest normalizowany poprzez usunięcie znaków diakrytycznych oraz zamianę na małe litery.
2. Słowa kluczowe są porównywane z zawartością CV, a każda zgodność zwiększa wynik punktowy zgodnie z wagą przypisaną do danego słowa kluczowego.
3. Wynik dla każdego słowa kluczowego obejmuje liczbę jego wystąpień, wagę oraz łączną liczbę zdobytych punktów.

Obraz zawierający tekst, numer, zrzut ekranu, paragon

Opis wygenerowany automatycznie

### Prezentacja wyników

Wyniki analizy prezentowane są w nowym oknie z takimi informacjami jak:

* Wprowadzona nazwa kandydata przez użytkownika
* Łączna liczba punktów po analizie dokumentu CV

Obraz zawierający tekst, zrzut ekranu, Czcionka, design

Opis wygenerowany automatycznie

Wyniki analizy po kliknięciu przycisku „Pokaż szczegóły” są prezentowane użytkownikowi w przejrzystej formie tabelarycznej, zawierającej:

* Słowa kluczowe.
* Liczbę ich wystąpień.
* Wagę przypisaną do słowa kluczowego.
* Liczbę punktów zdobytych za każde słowo.

Obraz zawierający tekst, zrzut ekranu, numer, Czcionka

Opis wygenerowany automatycznie

Wyniki analizy są zapisywane w bazie danych w tabeli kandydata, która przechowuje m.in.:

1. Pełną treść CV.

* Wyodrębnione informacje (imię i nazwisko, e-mail, numer telefonu).
* Całkowity wynik punktowy oraz szczegółowe wyniki dla każdego słowa kluczowego.
* Ścieżkę do przesłanego pliku CV.

Dane są przechowywane w bazie danych, z takiego powodu, że mogą być wykorzystane w aplikacji do dalszego przetwarzania.

## Zarządzanie stanowiskami

Moduł zarządzania stanowiskami umożliwia użytkownikom dodawanie, edytowanie i usuwanie stanowisk pracy, co stanowi kluczowy element w procesie personalizacji analizy dokumentów CV. Dzięki tej funkcji użytkownik może definiować wymagania dla poszczególnych stanowisk, co wpływa na skuteczność analizy i dopasowania kandydatów.

### Dodawanie stanowisk

Dodanie nowego stanowiska jest intuicyjne i odbywa się za pośrednictwem dedykowanego formularza. Użytkownik musi podać nazwę stanowiska oraz zdefiniować słowa kluczowe, które mają być uwzględniane podczas analizy CV, oddzielając je od siebie przecinkami.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Funkcja dodawania stanowiska korzysta z formularza HTML oraz mechanizmu obsługi żądań POST w aplikacji, co umożliwia zapisanie danych w tabeli Position w bazie danych. Każde stanowisko może być przypisane do konkretnego użytkownika ze względu kolumnę user\_id w tej tabeli jako klucz obcy, tworząc relację z tabelą User, dzięki czemu dane są dostępne wyłącznie dla jego konta.

Obraz zawierający tekst, zrzut ekranu, linia, Czcionka

Opis wygenerowany automatycznie

Obraz zawierający tekst, Czcionka, linia, numer

Opis wygenerowany automatycznie

Po dodaniu stanowiska słowa kluczowe zostają zaimplementowane do tabeli Keyword w bazie danych i mają przypisane position\_id tworząc kolejną relację, dzięki czemu słowa kluczowe są przypisane tylko do tego stanowiska.

Obraz zawierający tekst, zrzut ekranu, numer, wyświetlacz

Opis wygenerowany automatycznie

### Edytowanie stanowisk

Edytowanie istniejących stanowisk pozwala na aktualizację nazwy oraz powiązanych słów kluczowych, pobranych wcześniej z bazy danych przez konkretny endpoint. Mechanizm edycji stanowiska jest realizowany poprzez dynamiczne załadowanie danych stanowiska do formularza edycyjnego, co pozwala użytkownikowi na wygodne aktualizowanie zawartości.

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

Funkcja ta jest szczególnie przydatna w przypadku zmiany wymagań stanowiskowych lub rozszerzenia ich zakresu. Użytkownik może z łatwością wprowadzać modyfikacje za pomocą interfejsu użytkownika, gdzie każda zmiana jest odzwierciedlana w bazie danych w czasie rzeczywistym.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

W oknie edycji stanowiska najważniejszym blokiem jest formularz HTML wspierany przez narzędzie Jinja2 zintegrowane z frameworkiem Flask. Po zatwierdzeniu zmian, dane są zapisywane w bazie za pomocą odpowiedniej funkcji obsługującej żądanie POST.

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

Na samej górze formularza wyświetla się etykieta „Tytuł stanowiska” z dotychczasowo pobraną nazwą stanowiska, którą możemy edytować, a pod nią zostaje pobrana lista wszystkich słów kluczowych i wag, dotycząca danego stanowiska, na której również użytkownik może „manipulować” danymi m.in.:

1. zmienić nazwę
2. zmienić wagę (po utworzeniu stanowiska wszystkie słowa kluczowe mają domyślną wagę 1), zakres wartości wagi to 1-5
3. usunąć słowo kluczowe, przyciskiem Usuń znajdującym się pod danym elementem

Obraz zawierający tekst, zrzut ekranu, design

Opis wygenerowany automatycznie

W dolnej sekcji formularza, użytkownik ma możliwość dodania słowa kluczowego i przypisanie do niego wag. Kliknięcie przycisku „Dodaj” spowoduje wywołanie funkcji dodającej słowo kluczowe do formularza.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Za dodawanie i usuwanie słów kluczowych odpowiada skrypt z użyciem języka JavaScript.

Funkcja removeKeyword(button, keywordId), odpowiada za:

1. usuwanie wybranego elementu z listy słów kluczowych <li>
2. dodawanie ukrytego pola do formularza, aby oznaczyć które słowa kluczowe zostały usunięte. Dzięki temu backend (Flask) wie, jakie słowa kluczowe usunąć z bazy danych.

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

Funkcja addKeyword(), odpowiada za:

1. dodawanie nowego słowa kluczowego i jego wagi do listy
2. tworzenie nowego elementu <li> w liście słów kluczowych z odpowiednimi polami tekstowymi i liczbowymi.

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

Funkcja removeKeyword(button) odpowiada za usuwanie wybranego słowa kluczowego z listy.

Obraz zawierający tekst, Czcionka, zrzut ekranu

Opis wygenerowany automatycznie

Po zakończeniu edycji stanowiska, użytkownik może zapisać zmiany i wrócić do okna swoich stanowisk.

Usuwanie stanowisk

Dla stanowisk, które nie są już potrzebne, dostępna jest opcja ich usunięcia. Funkcja ta zapobiega przechowywaniu nieaktualnych danych w systemie, co zwiększa jego przejrzystość i wydajność. Usunięcie stanowiska skutkuje także usunięciem powiązanych z nim słów kluczowych, zapewniając integralność danych.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

## Ranking kandydatów

Moduł rankingu kandydatów umożliwia użytkownikowi przeglądanie i analizowanie wyników dopasowania kandydatów do wybranego stanowiska pracy. W przejrzystym układzie użytkownik może z łatwością wybrać stanowisko, określić liczbę wyświetlanych wyników oraz zapoznać się z kluczowymi informacjami na temat każdego kandydata. Dzięki zastosowaniu dynamicznego systemu filtracji i sortowania, użytkownik jako potencjalny rekruter, może łatwo uzyskać listę najlepszych kandydatów na wybrane stanowisko, co znacząco usprawnia proces rekrutacji.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

### Wybór stanowiska pracy i ograniczenie liczby wyników

Formularz umożliwia wybranie jednego z dostępnych stanowisk pracy. Lista stanowisk jest dynamicznie generowana z bazy danych i obejmuje zarówno globalne stanowiska, jak i te dodane przez użytkownika.

Wybrane stanowisko określa kryteria oceny kandydatów oraz słowa kluczowe uwzględniane w analizie.

Obraz zawierający tekst, elektronika, zrzut ekranu, oprogramowanie

Opis wygenerowany automatycznie

Użytkownik może ustawić maksymalną liczbę wyników wyświetlanych w rankingu (zakres: 1–50). Ta funkcja pozwala na skupienie się na najlepszych kandydatach, co ułatwia podejmowanie decyzji.

Obraz zawierający tekst, zrzut ekranu, Czcionka, linia

Opis wygenerowany automatycznie

### Prezentacja wyników

Ranking jest przedstawiony w formie tabeli, w której wyświetlane są następujące informacje o kandydacie:

1. Pozycja w rankingu (#).
2. Nazwa pliku CV.
3. Imię i nazwisko (lub inne zidentyfikowane dane z dokumentu CV).
4. Adres e-mail.
5. Numer telefonu.
6. Liczba uzyskanych punktów w analizie.

Kandydaci są uporządkowani malejąco według liczby punktów.

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Endpoint pobiera informacje o kandydacie z bazy danych z tabeli Candidate, którzy są tam przechowywani po zakończonej pomyślnie analizie dokumentu CV.

Obraz zawierający tekst, numer, zrzut ekranu

Opis wygenerowany automatycznie

Powyższy zrzut ekranu, przedstawia przykładowe dane testowe, stworzone na rzecz modułu rankingu kandydatów. Każdy zalogowany użytkownik ma swoją bazę kandydatów, z powodu relacyjnego modelu bazy danych, ponieważ w tabeli Candidate znajduję się kolumna user\_id.

### Dodatkowe akcje dla kandydatów

Użytkownik posiada możliwość zarządzania kandydatami z poziomu rankingu:

1. Klikając ikonę „oko” (fa-eye) z biblioteki Font Awesome, użytkownik może otworzyć CV w osobnym oknie przeglądarki
2. Klikając ikonę „pobierz” (fa-download), użytkownik ma możliwość zapisania pliku CV na swoim urządzeniu.
3. Klikając w ikonę „czerwonego kwadratu”, kandydat z zostanie usunięty z listy w rankingu

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Obraz zawierający zrzut ekranu, tekst, oprogramowanie, Oprogramowanie multimedialne

Opis wygenerowany automatycznie

# Baza danych

## Struktura bazy danych

Baza danych aplikacji została zaprojektowana w oparciu o relacyjny model danych, aby efektywnie przechowywać i zarządzać danymi związanymi z procesem analizy CV oraz zarządzaniem użytkownikami i stanowiskami. Schemat bazy danych obejmuje pięć głównych tabel: users, positions, keywords, candidates oraz alembic\_version (służącą do zarządzania migracjami).

## Tabele

**Users (Użytkownicy)**

* Przechowuje dane logowania i dane identyfikacyjne użytkowników.
* Pola: id (klucz główny), username, email (unikalne), password\_hash.
* Relacje:
  + Jeden użytkownik może posiadać wiele stanowisk (positions).
  + Jeden użytkownik może mieć przypisanych wiele kandydatów (candidates).

**Positions (Stanowiska)**

* Zawiera dane dotyczące stanowisk pracy, na które kandydaci są oceniani.
* Pola: id, title (nazwa stanowiska), is\_default (czy stanowisko jest globalne), user\_id.
* Relacje:
  + Każde stanowisko może mieć przypisane wiele słów kluczowych (keywords).
  + Stanowisko jest powiązane z jednym użytkownikiem, który je utworzył.

**Keywords (Słowa Kluczowe**)

* Przechowuje słowa kluczowe przypisane do stanowisk oraz ich wagi.
* Pola: id, word (słowo kluczowe), position\_id (relacja do stanowiska), weight (waga).
* Relacje:
  + Powiązane z tabelą positions.

**Candidates (Kandydaci)**

* Przechowuje dane przesłanych CV oraz wyniki analizy.
* Pola: id, name, cv\_text, position\_id, points (punkty uzyskane w analizie), first\_words, email\_cv, phone\_number, user\_id, path (ścieżka do przesłanego pliku).
* Relacje:
  + Kandydat jest powiązany z określonym stanowiskiem (positions).
  + Kandydat jest przypisany do użytkownika, który przesłał jego CV.

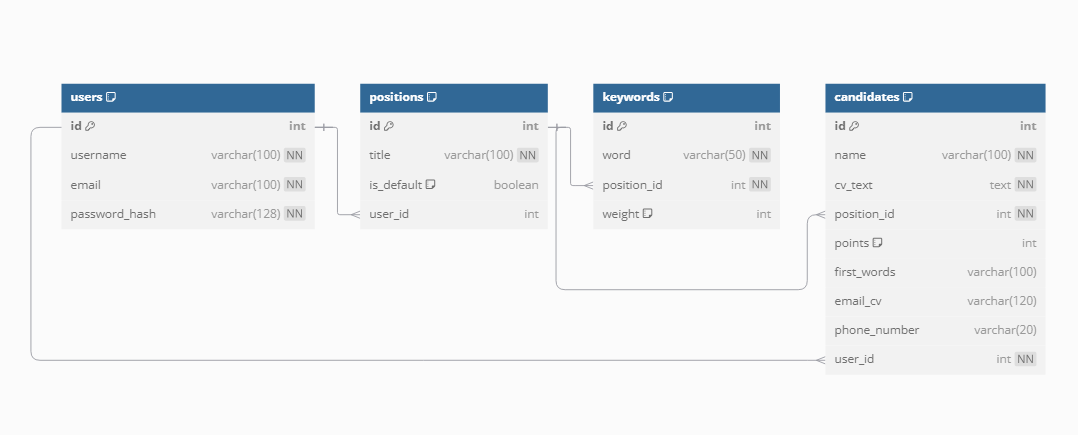
**Alembic\_version**

* Przechowuje dane dotyczące wersji migracji bazy danych, co pozwala na zarządzanie zmianami w schemacie bazy.

## Diagram ERD

Struktura bazy danych została przedstawiona w formie diagramu ERD (ang. Entity-Relationship Diagram). Diagram ukazuje relacje między tabelami:

* users ma relację jeden-do-wielu z positions oraz candidates.
* positions ma relację jeden-do-wielu z keywords oraz z candidates.



# Technologie i biblioteki

## Języki programowania

1. Python – główny język backendu.
2. JavaScript – do dynamicznych interakcji (funkcje w edit\_position.html)

## Pozostałe języki

1. HTML – do tworzenia szablonów stron
2. CSS – do stylizacji i projektowania interfejsu użytkownika

## Frameworki i biblioteki backendowe

1. Flask - framework webowy do obsługi backendu i integracji z frontendem.
2. Flask-SQLAlchemy - ORM do zarządzania bazą danych.
3. Flask-Migrate - narzędzie do migracji schematów bazy danych.
4. Jinja2 - silnik szablonów w Flask, używany do dynamicznego generowania HTML w szablonie i szkielecie interfejsu aplikacji tworzonym na języku znaczników
5. gunicorn - serwer WSGI do produkcyjnego wdrożenia aplikacji.

## Frameworki i biblioteki frontendowe

1. Bootstrap - stylizacja i responsywność interfejsu.
2. Font Awesome - ikony używane w elementach interfejsu (np. przyciski).

## Biblioteki NLP i analizy danych

1. Pytesseract – narzędzie OCR do ekstrakcji tekstu z obrazów PDF
2. Pdf2image – konwersja plików PDF na obraz\

## Bazy danych

1. SQLite – lokalna baza danych używana w środowisku deweloperskim.

## Infrastruktura i narzędzia wdrożeniowe

1. Docker - konteneryzacja aplikacji:
   1. Obraz oparty na python:3.8-slim.
   2. requirements.txt do instalacji zależności.
2. Heroku - platforma chmurowa do hostowania aplikacji:
   1. Plik heroku.yml do konfiguracji wdrożenia.
   2. Plik Procfile.txt do uruchamiania aplikacji z użyciem Gunicorn.
3. Alembic - narzędzie do migracji bazy danych (konfiguracja w alembic.ini).

## Zarządzanie użytkownikami i bezpieczeństwo

1. Werkzeug.security – generowanie i weryfikacja hashów haseł użytkowników

## Funkcje pomocnicze

1. re - do wyszukiwania wzorców w tekście (np. e-maile, numery telefonów).
2. unicodedata - normalizacja tekstu w celu usunięcia znaków diakrytycznych.
3. os - operacje na plikach i ścieżkach.
4. jsonify - zwracanie danych w formacie JSON w odpowiedzi HTTP.

## Systemy kontroli wersji i organizacja pracy

1. GitHub - repozytorium kodu.
2. .gitignore - konfiguracja ignorowanych plików (np. środowiska wirtualne, bazy danych, klucze .env).

# Struktura katalogów

analyzer\_cv/

├── app.py

├── models.py

├── templates/

│ ├── base.html

│ ├── home.html

│ ├── login.html

│ ├── register.html

│ ├── upload.html

│ ├── ranking.html

│ ├── results.html

│ ├── add\_position.html

│ ├── edit\_position.html

│ ├── view\_positions.html

├── static/

│ ├── stylesHome.css

│ ├── stylesLogin.css

│ ├── stylesPosition.css

│ ├── stylesResult.css

├── uploads/

# Możliwe rozszerzenia

1. Frontend:
   1. Interaktywne dashboardy i nowoczesny design (np. React/Vue.js).
   2. Tryb ciemny/jasny.
   3. Zaawansowane filtry wyszukiwania kandydatów.
2. Analiza CV
   1. Obsługa dodatkowych formatów (np. .doc, .docx).
   2. Lepsza analiza semantyczna z wykorzystaniem zaawansowanych modeli NLP.
   3. Punktowanie doświadczenia i kwalifikacji na podstawie dat i certyfikatów.
3. Ranking kandydatów:
   1. Zaawansowane kryteria rankingowe z wizualizacjami wyników (np. wykresy).
   2. Automatyczne rekomendacje stanowisk na podstawie analizy treści CV.
   3. Eksport raportów rankingowych do PDF lub Excel.
4. Funkcjonalności użytkownika:
   1. Różne role użytkowników (np. administrator, rekruter, kandydat).
   2. Historia przesłanych CV i analiz.
   3. Powiadomienia e-mail o wynikach analizy.

# 7. Podsumowanie

Stworzona aplikacja do analizy CV stanowi rozwiązanie dla współczesnych wyzwań związanych z procesami rekrutacyjnymi. Dzięki zastosowaniu zaawansowanych technologii, takich jak przetwarzanie języka naturalnego (NLP) oraz odpowiednio zaprojektowane algorytmy dopasowujące, narzędzie to automatyzuje najbardziej czasochłonne aspekty rekrutacji. Główne funkcje, takie jak ekstrakcja kluczowych informacji z CV, dopasowanie kandydatów do zdefiniowanych stanowisk oraz generowanie rankingów, sprawiają, że aplikacja wspiera zarówno rekruterów, jak i firmy, w szybkim i efektywnym pozyskiwaniu kandydatów.