

# Dokumentacja specyfikacji wymagań

## 1. Wprowadzenie

Dokument zawiera specyfikacje wymagań dot. kodu napisanego w języku R, badającego sentymenty na podstawie zawartości plików .txt z kodowaniem UTF-8. Skrypt ma umożliwić wczytanie dwóch plików tekstowych i umożliwić sprawdzenie różnic w sentymentach pomiędzy nimi.

## 2. Cele systemu

- Wczytanie dwóch plików wejściowych w formacie .txt z kodowaniem UTF-8
- Przetworzenie i oczyszczenie tekstów
- Pozbycie się stopwords
- Analiza sentymentów na podstawie słowników afin, bing, loughrun i nrc
- Wizualizacja wyników w postaci wykresów

## 3. Wymagania funkcjonalne

- **Wczytywanie danych**
  - Skrypt obsługuje wczytywanie danych z plików w formacie .txt i kodowaniu UTF-8
- **Przetwarzanie i oczyszczanie tekstu**
  - Skrypt pomija znaki interpunkcyjne i liczby
  - Skrypt nie zwraca uwagi na wielkość liter i ujednolica formy skrócone zawierające apostrofy z formami nieskróconymi.
- **Analiza sentymentu**
  - Skrypt umożliwia wczytanie słowników afinn.csv, bing.csv, loughrun.csv i nrc.csv
- **Wizualizacja danych**
  - Program tworzy wykresy ggplot2
  -

## 4. Wymagania нефункционалне

- **Wydajność**
  - Analiza pliku o długości 1000 zdań nie zajmuje dłużej niż 20 sekund
- **Bezpieczeństwo**
  -
- **Niezawodność**
  - Skrypt poprawnie obsługuje brakujące wartości
- **Użyteczność**
- **Kompatybilność**

## 5. Interfejsy użytkownika i wymagania dotyczące danych

- **Wejście**
  - 2 pliki w formacie .txt
  - ew. pliki słowników
- **Wyjście**
  - Chmura słów
  - Liczba słów dla każdego słownika
  - Skumulowany sentyment dla każdego ze słowników

- Wykresy zmiany sentymentu w czasie

## **6. Słownictwo dokumentacji**

## **7. Przypadki użycia**

- Użytkownik:
  - wczytuje plik .txt
  - uruchamia analizę
  - generuje raport
- Skrypt:
  - przetwarza tekst
  - oczyszcza tekst
  - generuje chmurę słów
  - wyświetla liczbę słów dla słowników
  - generuje skumulowany sentyment dla każdego słownika
  - generuje wykres zmiany sentymentu w czasie
- Testowe przypadki użycia:
  - Testy z plikami z różnymi poziomami sentymentu
  - Test z plikiem zaw. znaki specjalne

## **8. Scenariusze użytkownika**

- a. Analiza opinii w czasie
    - Jako: Socjolog
    - Chcę: Przeanalizować w czasie wypowiedzi wyborców na temat prezydenta
    - Aby: Zobaczyć jak zmieniła się ich opinia w trakcie urzędowania
- Kryteria akceptacji:
- Użytkownik może wczytać wypowiedzi wyborców z mediów społecznościowych
  - Skrypt przeprowadza analizę sentymentów