Analiza sentymentu i emocji w tekstach religijnych: Biblia i Koran.

Niniejszy dokument przedstawia porównawczą analizę sentymentu oraz emocji zawartych w dwóch najważniejszych księgach religijnych: Biblii (wersja King James Version) oraz Koranie (tłumaczenie Pickthalla). Analiza przeprowadzona została w języku R, z zastosowaniem technik przetwarzania języka naturalnego (NLP). W raporcie zawarto wizualizacje, statystyki, najczęściej występujące słowa, bigramy, oraz porównania emocjonalnego profilu obu tekstów.

Opis danych

W analizie tekstów religijnych wykorzystano łącznie:

* 110 978 zdań z dwóch źródeł: Biblii (bible\_df – 51 992) i Koranu (quran\_df – 28 868),
* 1 165 203 bigramów (par wyrazów) przed czyszczeniem, z czego 82 435 to bigramy oczyszczone z cyfr i stop words (bigrams\_clean),
* 44 019 unikalnych bigramów zliczonych (bigrams\_count),
* 376 936 słów przypisanych do emocji (emotion\_words) na podstawie leksykonu NRC (nrc – 13 872 słowa).

Dane zostały oczyszczone z liczb, znaków specjalnych, wersów bez tekstu oraz poddane lematyzacji i usunięciu stop words. Analizę wykonano w języku R z użyciem pakietów: tidytext, textstem, dplyr, ggplot2.

Top 10 najczęstszych słów (procentowo)

Obraz zawierający tekst, zrzut ekranu, diagram, numer

Zawartość wygenerowana przez AI może być niepoprawna.

Wykres przedstawia porównanie dziesięciu najczęściej występujących słów w Biblii i Koranie, znormalizowanych jako procent wszystkich słów w każdej księdze.

* **"God"** jest najczęściej używanym słowem w Koranie, a także bardzo częstym w Biblii.
* **"Lord"**, **"thy"**, **"thou"** i **"ye"** dominują w Biblii, wskazując na formalny i archaiczny styl.
* **"Mohammed"** i **"verily"** pojawiają się wyłącznie w Koranie.
* Wspólne słowa, takie jak **"god"**, **"day"**, **"you"**, **"king"** i **"israel"**, występują w obu tekstach, ale z różną częstotliwością.

Analiza wskazuje na różnice w słownictwie, strukturze językowej oraz teologicznym ukierunkowaniu obu ksiąg.

Chmura słów – Koran

Obraz zawierający tekst, Czcionka, zrzut ekranu, design

Zawartość wygenerowana przez AI może być niepoprawna.

W przypadku Koranu (tłumaczenie Pickthalla), także pojawiają się archaizmy typu **"thou"**, **"hath"**, **"ye"**, ale widoczne są również charakterystyczne dla tekstu religijnego odniesienia do proroków (**"mohammed"**, **"prophet"**), miejsc świętych (**"mecca"**) oraz słownictwo związane z moralnością: **"truth"**, **"punishment"**, **"reward"**, **"verily"** (zaprawdę).

Chmura słów – Biblia

**Obraz zawierający tekst, Czcionka, zrzut ekranu, projekt graficzny

Zawartość wygenerowana przez AI może być niepoprawna.**

Na chmurze słów z Biblii (KJV) dominują archaiczne zaimki osobowe i formy gramatyczne, takie jak **"thou"** (ty), **"thy"** (twój), **"hath"** (ma) czy **"ye"** (wy). Wskazuje to na styl językowy charakterystyczny dla XVII wieku. Często występują też słowa: **"lord"**, **"god"**, **"israel"**, **"king"**, **"son"** – odzwierciedlające centralne tematy duchowe i społeczne.

Top 10 najczęstszych bigramów w Biblii i Koranie

Obraz zawierający tekst, zrzut ekranu, numer, diagram

Zawartość wygenerowana przez AI może być niepoprawna.

Wykres przedstawia dziesięć najczęściej występujących bigramów (czyli par wyrazów występujących obok siebie) w Biblii i Koranie. Bigramy zostały wyodrębnione na podstawie analizy tekstu obu świętych ksiąg.

* W **Biblii** dominują konstrukcje archaiczne, charakterystyczne dla języka staroangielskiego, np. *"thou hast"*, *"thou shalt"*, *"shalt thou"*.
* Często pojawiają się kombinacje odnoszące się do Boga: *"lord god"*, *"thy god"*, *"god hath"*, *"jesus christ"*.
* W **Koranie** oprócz odniesień religijnych (*"god hath"*, *"fear god"*, *"merciful god"*) pojawiają się także frazy związane z określeniami społecznymi lub komentatorskimi, np. *"al beidâwî"* (nazwa znanego komentatora Koranu) czy *"true believers"*.
* Obie księgi zawierają frazę *"thou shalt"*, co sugeruje podobne struktury nakazowe (np. przykazania, zasady).

Analiza bigramów ujawnia charakterystyczne wyrażenia dla każdej z ksiąg – w Biblii dominuje archaiczny styl i bezpośrednie odniesienia do Boga i przykazań, natomiast Koran zawiera również odniesienia do komentatorów oraz terminologii doktrynalnej. Wspólne elementy, takie jak *"thou shalt"*, pokazują podobieństwa w formie przekazu religijnego.

Rozkład sentymentu zdań

Obraz zawierający tekst, zrzut ekranu, diagram, Wykres

Zawartość wygenerowana przez AI może być niepoprawna.

Wykres 1 przedstawia histogram rozkładu wartości sentymentu. Obie księgi wykazują silne skupienie wyników wokół wartości neutralnych (0), ale Biblia ma więcej wartości skrajnych – zarówno pozytywnych, jak i negatywnych.

Wariacja sentymentów zdań

Obraz zawierający tekst, zrzut ekranu, diagram, Wykres

Zawartość wygenerowana przez AI może być niepoprawna.

Wykres 2 pokazuje wartości sentymentu każdego zdania w kolejności ich występowania. Pomaga to zobaczyć zmienność emocjonalną tekstów – obie księgi prezentują podobną rozpiętość, ale Koran wykazuje nieco większą stabilność.

Średni sentyment zdań

Obraz zawierający tekst, zrzut ekranu, diagram, Czcionka

Zawartość wygenerowana przez AI może być niepoprawna.

Wykres 3 to porównanie średnich wartości sentymentu dla obu ksiąg. Koran wykazuje wyższą średnią wartość sentymentu, co oznacza, że jego język jest nieco bardziej pozytywny w ogólnym tonie w porównaniu do Biblii.

Emocje w Biblii i Koranie (procentowo)

**Obraz zawierający tekst, zrzut ekranu, diagram, Wykres

Zawartość wygenerowana przez AI może być niepoprawna.**

Powyższy wykres przedstawia procentowy udział słów powiązanych z określonymi emocjami w tekstach **Biblii** i **Koranu**, zgodnie z klasyfikacją słownika NRC. Każda kolumna oznacza procent słów nacechowanych daną emocją spośród wszystkich słów z emocjami w danej księdze.

* **Najczęściej występującą emocją** w obu tekstach jest **trust** (zaufanie), z wyraźną przewagą w Koranie.
* Drugą dominującą emocją w obu księgach jest **joy** (radość), również częstsza w Koranie.
* **Fear** (strach) i **anger** (złość) mają zauważalny udział, przy czym **Biblia** zawiera nieco więcej słów powiązanych z tymi emocjami niż Koran.
* **Sadness** (smutek) i **disgust** (odraza) pojawiają się rzadziej, ale w porównywalnym stopniu w obu księgach.
* Emocje **anticipation** (oczekiwanie), **surprise** (zaskoczenie) i **negative**/ **positive** (ogólne nacechowanie emocjonalne) są bardziej zróżnicowane, przy czym Koran wykazuje większy udział emocji pozytywnych.

Analiza emocji pokazuje, że choć obie księgi zawierają zarówno emocje pozytywne, jak i negatywne, **Koran ma większy udział słów nacechowanych pozytywnie**, szczególnie w kategoriach **zaufania i radości**, co współgra z wynikami wcześniejszej analizy sentymentu.

Podsumowanie analizy

Analiza porównawcza tekstów Biblii i Koranu wykazała, że Koran zawiera nieco wyższy średni poziom pozytywnego sentymentu niż Biblia. W obu tekstach dominują podobne emocje, takie jak zaufanie i radość, choć ich natężenie się różni. Najczęstsze słowa i bigramy pokazują obecność charakterystycznych fraz religijnych i nakazowych, z wyraźną obecnością archaizmów w Biblii. Ogólnie rzecz biorąc, obie księgi wykazują zbliżony styl pod względem struktury emocjonalnej i tematycznej, mimo różnic kulturowych i językowych.