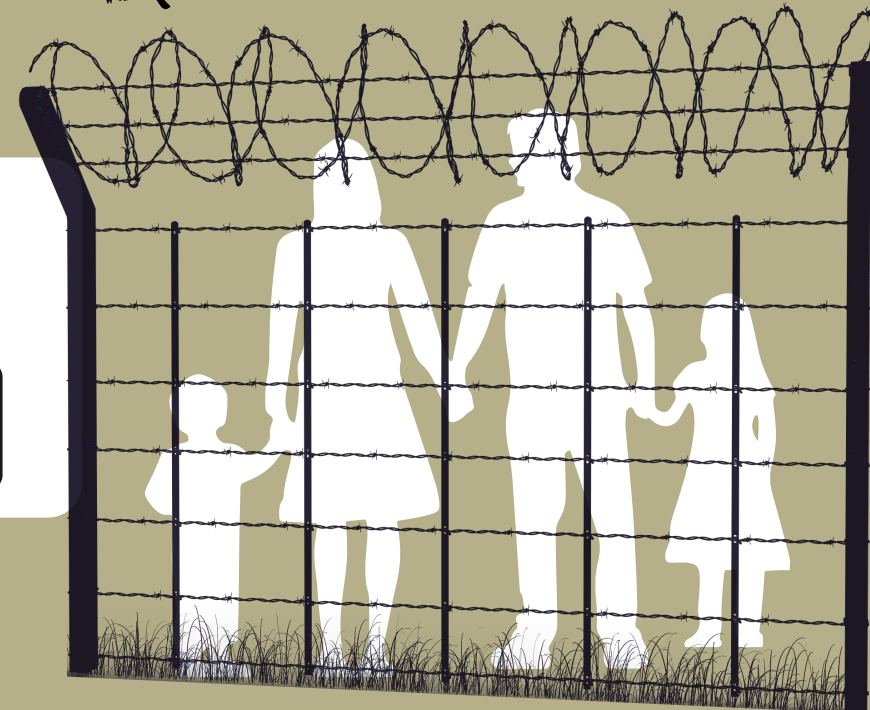


KONFLIKT NA GRANICY POLSKO-BIAŁORUSKIEJ

Zbadaliśmy stanowisko użytkowników serwisu Twitter wobec decyzji podejmowanych przez polski rząd w sprawie konfliktu na granicy polsko-białoruskiej. Jest to kontrowersyjny temat mający realny wpływ na wizerunek naszego kraju na arenie międzynarodowej.



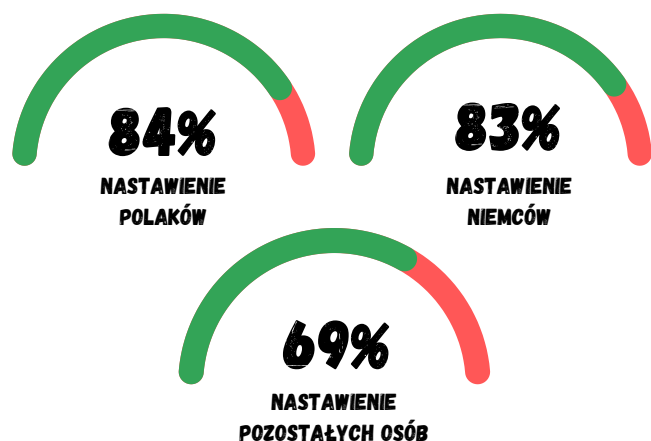
AUTORZY:

JAKUB PYKA

KAROL SZYMOŃCZYK

MICHAŁ WAWRZYNIAK

NASTAWIENIE



Postanowiliśmy podzielić zebrane przez nas tweety na trzy grupy: **popierające**, **neutralne** oraz **krytyczne** wobec działań na granicy podejmowanych przez polski rząd.

Ostatecznie skupiliśmy się głównie na badaniu skrajnych stanowisk. Analiza pokazała, że większość postów zawiera treści wskazujące na **poparcie** rządu, co oznacza większą część kraju nastawioną negatywnie w stosunku do imigrantów. Warto zwrócić uwagę na podobne nastawienie polskich i niemieckich użytkowników odstające od nastawienia pozostałych osób. Może to wynikać z faktu, że badana sytuacja dotyczy Polski oraz Niemiec bezpośrednio.

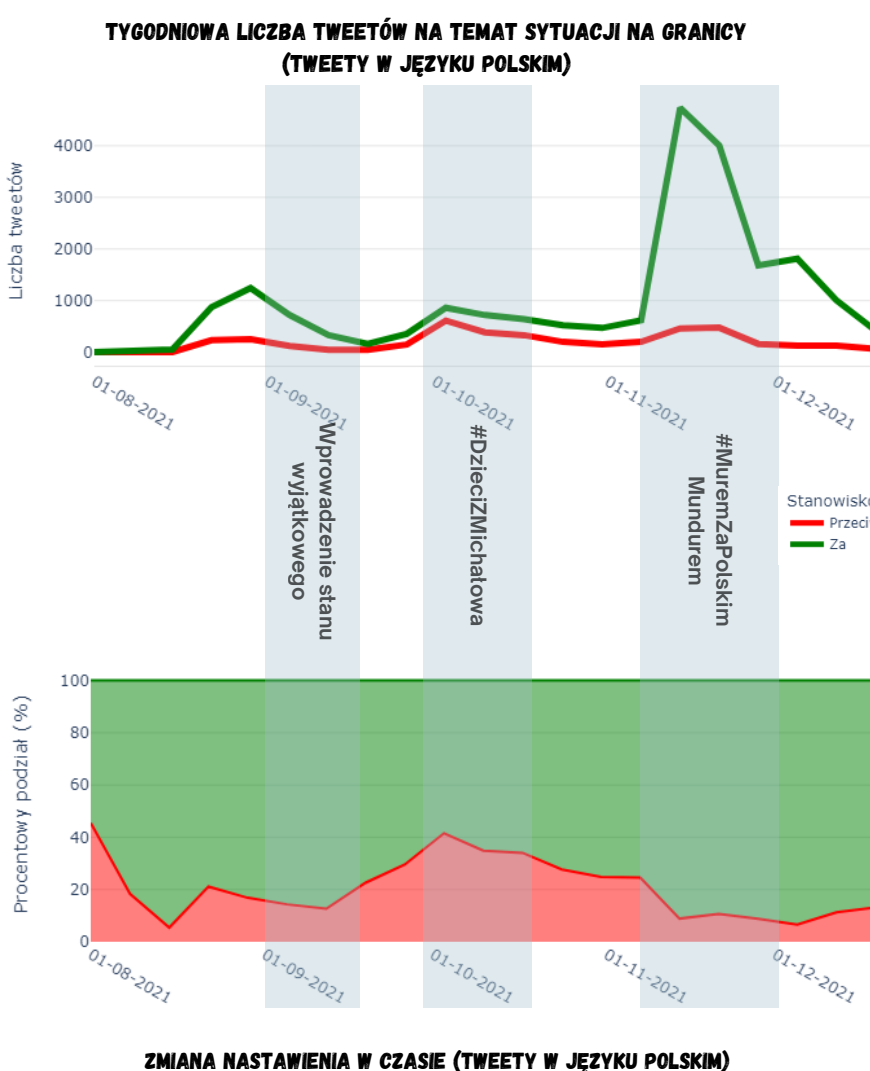
ANALIZA WYNIKÓW

Początkowo po wybuchu konfliktu temat nie był zbyt popularny, a liczba zwolenników i przeciwników była porównywalna. Następnie media rozgłosiły temat ukazując uchodźców jako celowy zabieg łukaszenki (przełom sierpnia i września).

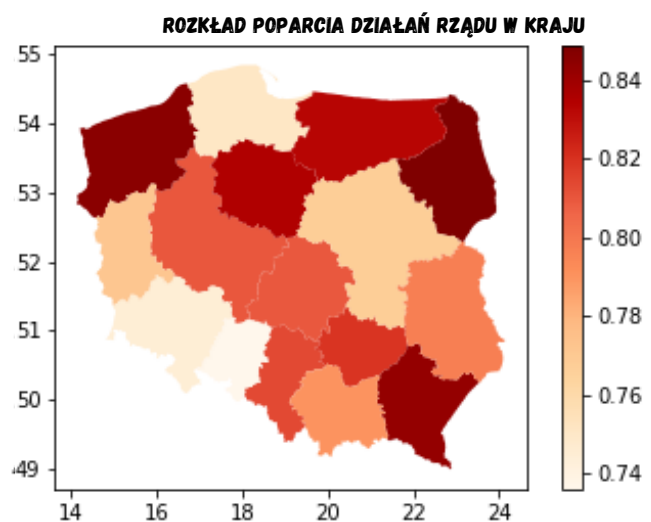
Po wprowadzeniu stanu wyjątkowego na początku września, poparcie dla działań rządu zaczęło spadać, a sytuacja na granicy zaczęła mocniej angażować ludzi, co widać po liczbie publikowanych w tym okresie tweetów.

Pod koniec września zaczęło występować coraz więcej incydentów związanych z uchodźcami, które powodowały dalszy spadek poparcia. Jednym z nich jest historia dzieci z Michałowa, która została nagłośniona dzięki hashtagom **#DzieciZMichałowa** oraz **#GdzieSaDzieci**.

W odpowiedzi na rosnące niezadowolenie zaczęły tworzyć się zorganizowane kampanie mające na celu poprawę wizerunku i morale żołnierzy. Przykładem jest wprowadzenie akcji **#MuremZaPolskimMundurem**, która zyskała popularność na początku listopada, co również wpłynęło na wzrost liczby publikowanych postów na temat sytuacji na granicy polsko-białoruskiej.



POPARCIE DLA RZĄDU



Kolorem ciemnoczerwonym oznaczone zostały regiony najbardziej przychylnie działaniom polskiej władzy. Bezwzględne poparcie w tej skali wynosi 1.

Można zauważyć, że największą aprobatą polski rząd cieszy się w województwie leżącym przy granicy polsko-białoruskiej - **województwie podlaskim**, którego mieszkańców sytuacja bezpośrednio dotyczy. Z kolei najmniejsze poparcie działań na granicy wyrażali w postach użytkownicy zamieszkujący województwo opolskie.

PRZEBIEG PRACY

ZBIERANIE DANYCH

Zebraaliśmy ok. **100 000 tweetów** z okresu 01.02.2021 - 13.12.2021 napisanych przez niemal **34 000 użytkowników**, którzy wypowiadali się na temat sytuacji na granicy polsko-białoruskiej. Finalny zbiór zawierał ok. 48 000 wpisów w języku polskim, 28 000 - w języku angielskim oraz 7 000 - w języku niemieckim, pochodzących zarówno z dużych serwisów informacyjnych, jak i zwykłych użytkowników Twittera. Następnie na podstawie profili użytkowników oraz narzędzia OpenStreetMap zdekodowaliśmy ich miejsce zamieszkania, co pozwoliło na analizę problemu z uwzględnieniem województwa. Dane były zbierane za pomocą pomocą scraper'a wykorzystującego bibliotekę stweet oraz przetrzymywane w MongoDB.



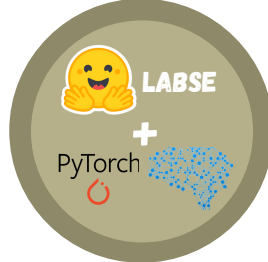
ANOTACJA DANYCH

Spośród polskich tweetów stworzyliśmy losową próbkę **2500 wpisów**, zachowującą rozkład liczby publikowanych postów w przestrzeni miesiąca. Zostały one sklasyfikowane przez **3 anatorów** względem stanowiska: **przeciwko**, **brak** lub **popieram**. Zgodność anotacji określona przez miarę Kappa Cohena wyniosła 0.7.



BUDOWA MODELU

Ze względu na wielojęzyczny charakter zebranych przez nas tweetów postanowiliśmy wykorzystać pre-trenowany model **LabSE**, pozwalający wyliczyć wektory osadzeń całych zdań we wspólnej przestrzeni dla 109 najpopularniejszych języków (w tym polskiego, angielskiego i niemieckiego). Model został pobrany z serwisu Hugging Face. Z racji na ograniczoną długość wpisów narzuconą przez Twitter, posty zostały potraktowane w całości jako zdania.

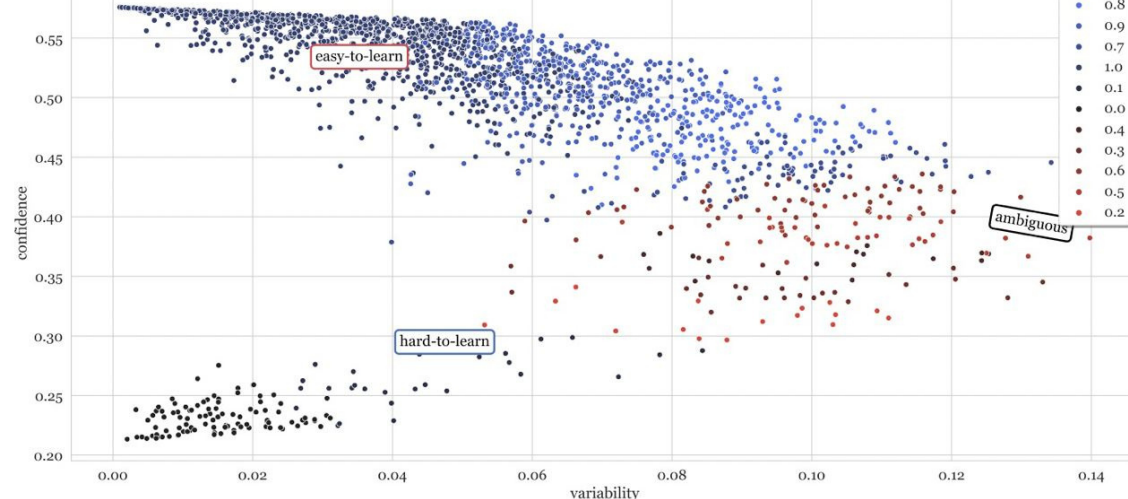


ANALIZA I WIZUALIZACJA

Wytrenowany model wykorzystaliśmy do wykrycia stanowiska dla pozostałych wpisów ze zbioru. Tak uzyskane dane poddane zostały analizie. Następnie przystąpiliśmy do wizualizacji uzyskanych wyników.



MAPA DANYCH MODELU WYTRENOWANEGO NA ZAANOTOWANYCH PRÓBKACH



ZNALEZISKA

