

Możliwości powiązania danych geolokacyjnych i analizy sentymentu w analizie zachowań użytkowników w wybranych portalach społecznościowych

Dariusz Mydlarz

Promotor: dr inż. Anna Zygmunt

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
Wydział Informatyki, Elektroniki i Telekomunikacji – Katedra Informatyki

Kraków, 10 grudnia 2014 roku

Agenda

- 1 Motywacja i cele
- 2 Sposób realizacji i przebieg prac
- 3 Uzyskane rezultaty

Motywacja i cele

Cel – zbadanie możliwości połączenia trzech dziedzin:

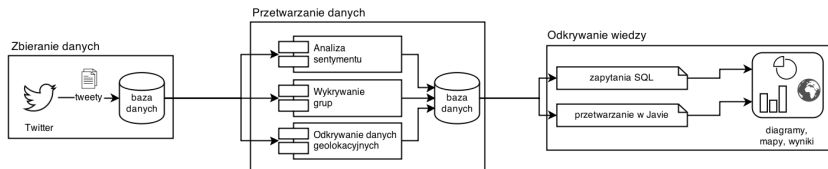
- analizy sieci społecznościowych
- analizy sentymentu
- geolokacji

Motywacje:

- dziedziny nie łączone wcześniej
- ogromna popularność serwisów społecznościowych
- wykorzystanie szerokiej wiedzy zdobytej na studiach

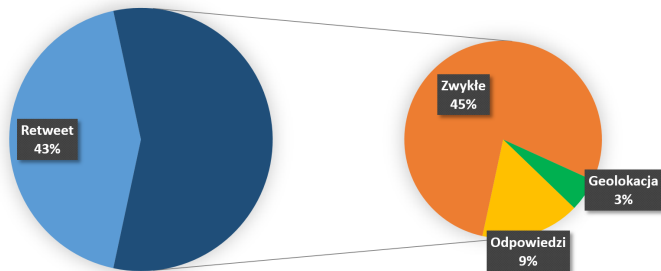
Sposób realizacji i przebieg prac

- 1 zbieranie danych z Twittera
- 2 wybór i budowa narzędzi do ekstrakcji wiedzy
- 3 odkrywanie wiedzy

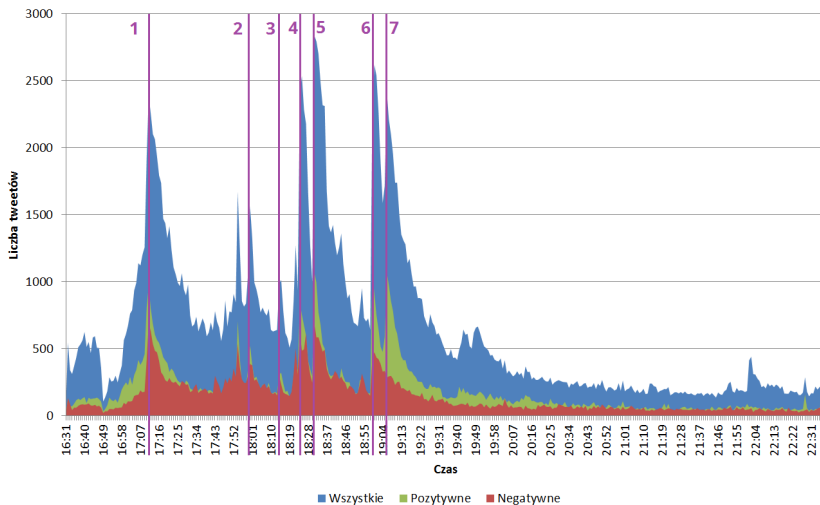


Uzyskane rezultaty

- 7 mln tweetów
- 49% pozytywnych, 47% negatywnych

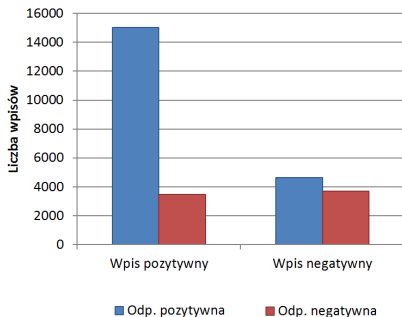


Aktywność użytkowników i rozkład sentymentu

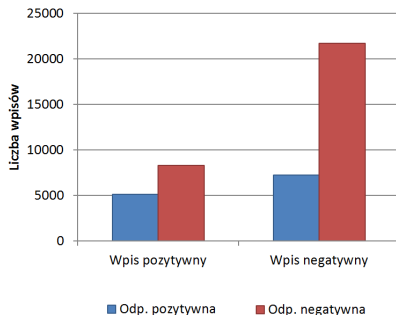


Interakcja między użytkownikami

Sentyment odpowiedzi. Odpowiada zwolennik na wpis zwolennika

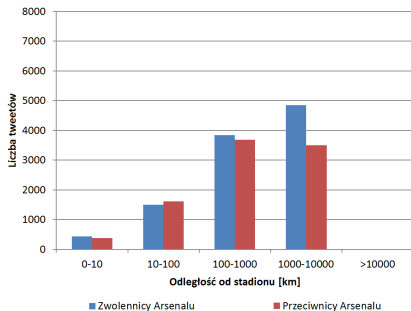


Sentyment odpowiedzi. Odpowiada przeciwnik na wpis przeciwnika

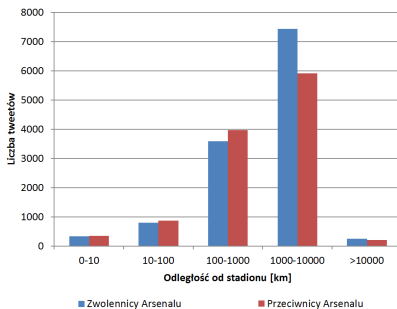


Aktywność kibiców, a odległość od stadionu

Mecze Arsenalu u siebie. Odległość komentujących od stadionu



Mecze Arsenalu na wyjeździe. Odległość komentujących od stadionu



Inne eksperymenty

- ❶ struktura grup między meczami
 - zauważalna stała grupa kibiców
 - im bardziej popularne mecze, tym mniejsze podobieństwo użytkowników w nich uczestniczących
- ❷ zmiana sentymentu w meczach
 - gdy drużyna wygrywa – wydźwięk bardziej pozytywny, gdy przegrywa – bardziej negatywny
- ❸ struktura komunikacji
 - przecwinicy najczęściej odpowiadają sobie nazwajem
 - zwolennicy najczęściej podają swoje wpisy dalej (retweetują)
- ❹ im kibice fizycznie bliżej siebie, tym częściej się ze sobą komunikują

Podsumowanie eksperymentów

- użytkownicy sieci społecznościowych zachowują się w taki sam sposób jak w życiu codziennym
- najnowsze technologie komputerowe pozwalają na badanie dużych społeczności w sposób transparentny
- badanie sieci społecznych można zautomatyzować i zastosować również do innych dziedzin

Podsumowanie pracy

Osiągnięte cele:

- połączono geolokację, sentyment i sieci społeczne
- zbudowano narzędzie do przetwarzania dużej ilości danych z Twittera pod kątem analizy sieci społecznych
- odkryto ciekawe zależności wśród badanej społeczności

Możliwe kierunki rozwoju:

- rozbudowanie narzędzia do analizy sentymentu
- odkrywanie tematów rozmów
- bliższe przyjrzenie się grupom