

# Rekomendacje artykułów opisujących produkty w serwisach e-commerce

Łukasz Dragan

Informatyka spec. Metody sztucznej inteligencji, MiNI PW

31.10.2017

- 1 Opis problemu
- 2 Systemy rekomendacji
- 3 Techniki przetwarzania języka naturalnego
- 4 Metody ewaluacji
- 5 Testy
- 6 Podsumowanie
- 7 Wybrane źródła

Czy metody semantycznej analizy tekstu mogą być alternatywą dla dotychczas używanej przez *Allegro* metody generowania rekomendacji artykułów tekstowych?

Praca

## Praca: 43 360 ofert pracy



🔍 Szukaj

Określ odległość  0 km

### Filtruj wyniki




#### Miejsce pracy

- ☐ cała Polska (42350)
- ☐ dolnośląskie (4499)
- ☐ kujawsko-pomorskie (1587)
- ☐ lubelskie (1133)
- ☐ lubuskie (1058)
- ☐ łódzkie (2787)
- ☐ małopolskie (4030)
- ☐ mazowieckie (10245)
- ☐ opolskie (1071)
- ☐ pomorskie (2881)
- ☐ podkarpackie (1191)
- ☐ podlaskie (911)
- ☐ świętokrzyskie (886)
- ☐ śląskie (3775)
- ☐ warmińsko-mazurskie (950)
- ☐ wielkopolskie (3765)
- ☐ zachodniopomorskie (1581)
- ☐ zagranica (1010)

Zastosuj

### Oferty rekomendowane dla Ciebie

Na podstawie Twojej aktywności wybraliśmy oferty dopasowane do Twoich oczekiwań

	<b>Stażysta IT&amp;ccw w Roche Diagnostics</b> Roche Diagnostics 📍 Warszawa, mazowieckie	2017-10-21	★
	<b>Stażysta w dziale Analiz</b> Innogy Polska S.A. 🏢 o firmie 📍 Warszawa, mazowieckie	2017-10-20	★
	<b>Stażysta w Zespole Modeli Ryzyka</b> ALIOR BANK 🏢 o firmie 📍 Warszawa, mazowieckie	2017-10-15	★
...	<b>Stażysta w programie NN Pro</b> Nationale-Nederlanden Usługi Finansowe SA 📍 Warszawa, mazowieckie	2017-10-02	★
...	<b>Stażysta w Spółce IT</b> PGE Systemy S.A.		★

Wniosek / Złóż CV

Łukasz Dragan

FILMY
 SERIALE
 GRY
 REPERTUARIUM KIN
 PROGRAM TV
 MAGAZYN
 WFF
 MÓJ FILMWEB

## Mój Filmweb

Aktywność znajomych, których obserwujesz:

Martha oglądała film **Persona**  
 13 godzin temu

**Persona (1966)**

gatunki: Dramat, Psychologiczny  
 reżyser: Ingmar Bergman  
 obsada: Bibi Andersson, Liv Ullmann

Aktorka traci głos podczas przedstawienia teatralnego. Okazuje się, że demonstruje ona w ten sposób bezradność wobec otaczającego ją świata.

kuba1004 ocenił rolę aktorki Marion Cotillard w filmie To tylko koniec świata na 8

### Mój asystent

wszystkie w guście w guście newsy recenzje ustawienia

Filmweb znalazł film w Twoim guście: **Ślódmy kontynent**  
 83%

Filmweb znalazł film w Twoim guście: **Podróż na Księżyc**  
 84%

Filmweb znalazł film w Twoim guście: **Głową w mur**  
 81%

Filmweb znalazł film w Twoim guście: **Boska Florence**  
 75%

Filmweb znalazł film w Twoim guście: **Zagraj to jeszcze raz, Sam**  
 80%

allegro

czego szukasz?

wszystkie działy



koszyk jest pusty

Elektronika

Moda  
i urodaDom  
i zdrowie

Dziecko

Kultura  
i rozrywkaSport  
i wypoczynek

Motoryzacja

Kolekcje  
i sztuka

Firma

Strefa  
okazji

Allegro - Poradniki - Dom i zdrowie - Jaka farba dla alergika?

## Jaka farba dla alergika?



autor: Ewelina Wojtunik, data publikacji: 23-04-2015

Za chwilę wiosna, a wraz z nią potrzeba porządków i odświeżenia ścian. Jak co roku będziemy sprzątać, wietrzyć i wymieniać zimę z kątów mieszkania. Zaraz po tym zaczną się pierwsze remonty.

**Ewelina Wojtunik**

Zawodowo związana z Social Media, pisała m.in. do Aktivist.pl. Prywatnie pasjonatka projektowania wnętrz, zdrowego stylu życia i roślin doniczkowych. Podróże i kuchnie świata są dla niej inspiracją. W wolnym czasie spełnia się jako mama i uczy języków.

**może Cię również  
zainteresować**



wszystkim chemikalia i detergenty. znajdujące się w nich alergeny mogą być powodem problemów zdrowotnych, a także nasilać objawy nadwrażliwości takie jak łzawienie oczu, zapalenie skóry czy kaszel astmatyczny.

### Szkodliwe związki lotne

W styczniu 2010 roku Unia Europejska wprowadziła normę, która reguluje zawartość szkodliwych lotnych związków organicznych tak zwanych LZO (VOCs, ang. volatile organic compounds) w trafiających do sprzedaży farbach i lakierach. Warto wiedzieć, że lotne związki lubią pozostawać aktywne pomimo wyschnięcia farby i starannego wentrowienia mieszkania. Co więcej, mogą uwalniać się ze ścian całymi latami, nasilając objawy alergiczne i pogarszając samopoczucie mieszkańców. Im mniej ich w składzie, tym lepiej dla nas.

Pamiętajmy więc, że kupowana przez nas farba powinna posiadać **atest hipoalergiczny** – najlepiej specjalny certyfikat potwierdzający bezpieczeństwo dla osób cierpiących z powodu nadwrażliwości na alergeny. Opatrzony certyfikatem farby gwarantują nawet trzydziestokrotnie niższą szkodliwość! Dlatego kupując je, zwróćmy uwagę na obecność stosownego oznaczenia na opakowaniu, dzięki czemu zyskamy pewność, że nie zawierają żadnych substancji uczulających i pozostają w pełni bezpieczne dla zdrowia naszego i naszych bliskich. Oprócz farb szkodliwe związki lotne mogą pojawiać się także w klejach, wykładzinach dywanowych, **tapetach ściennych**, a nawet materiałach do wykończenia podłóg.



EKO ŚNIEŻKA BIAŁA FARBA  
EMULSJA 10L  
HIPOALERGICZNA  
kup teraz 43,97 zł



EKO ŚNIEŻKA BIAŁA FARBA  
EMULSJA 10L  
HIPOALERGICZNA  
kup teraz 46,90 zł



ŚNIEŻKA EKO Farba Emulsja  
Hipoalergiczna 10l  
kup teraz 50,10 zł



Emulsja Hipo  
Śnieżka EKO  
kup teraz 9,9l



### Wnętrzarski hit – ściany ombre

Ombre stało się hitem w wizażu i modzie już kilka sezonów temu! Chętnie rozjaśniamy końcówki włosów, cieniujemy kolory na paznokciach, a także nosimy ubrania w przenikających się tonach. Czy tę technikę mo...



### Jak przemaalować ciemną ścianę?

Planujesz remont mieszkania, a jednym z jego etapów będzie przemaalowanie ciemnej ściany? A może po prostu znudził ci się niemiódny już kolor? Jeśli zastanawiasz się, jak prawidłowo przemaalować ścianę, spraw...

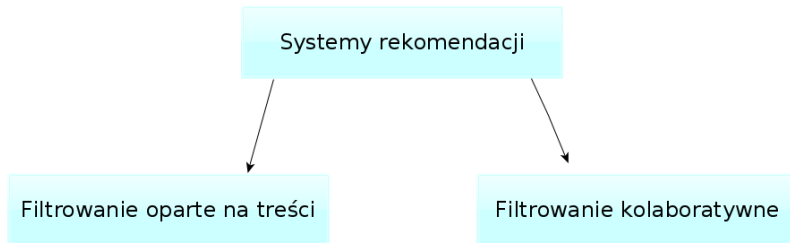




„Elasticsearch is a distributed, JSON-based search and analytics engine designed for horizontal scalability, maximum reliability, and easy management.”



W ujęciu ogólnym systemy wyszukiwania mają na celu sugerowanie tego, co użytkownik chciałby otrzymać. Natomiast systemy rekomendacji mają sugerować przedmioty potrzebne użytkownikowi nawet, jeżeli potrzeby te nie zostały bezpośrednio wyrażone.



## Praca magisterska

1.168785 0.060346 0.502299 0.291747 -0.365562  
0.257444 -0.329024 0.758068 0.139132 -0.066573  
1.171894 0.067840 -0.002970 -0.360585 -0.144586  
0.105688 -0.528267 0.377016 0.220084 -0.13236  
0.232592 0.338373 0.106514 0.096009 -0.068181  
-0.698880 0.040483 -0.820396 0.110031 -0.493751  
-0.339397 0.278281 -0.000135 -0.121884 0.107060  
-0.001215 -0.348834 0.399166 0.391983 0.197091  
-0.837996 -0.081890 -0.534775 0.589362 0.278594  
-0.724953 0.143085 -0.300889 -0.051467 0.133181  
0.110936 -0.159592 -0.338680 0.324832 -0.227569  
-0.257161 -0.403050 -0.355761 -0.11366 0.127871  
-0.045948 0.256404 -0.413172 -0.565309 0.252026  
-0.178040 0.353451 -0.043467 0.437229 -0.364093  
0.620433 0.491961 -0.044899 0.075592 -0.035806  
0.552777 0.539595 -0.307839 -0.488252 0.494307  
-0.506171 0.517397 0.100668 -0.247984 0.322363

# Zarys podejścia

## Wnętrze zask. hit – ściany cmiere



Copyright: Shutterstock.com

Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.

Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.

Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.

Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.

Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.

Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.



Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.

Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.

Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.



Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.

Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.

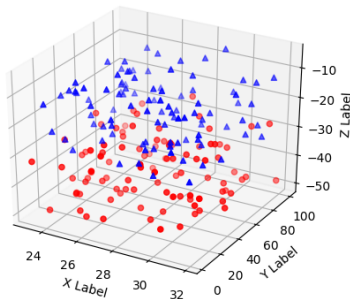
Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.

Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.

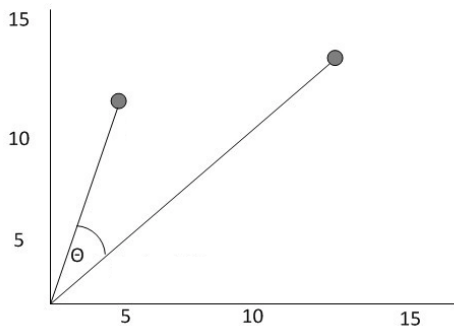
Całkowicie czyste i jasne ściany w kolorze białym, które są idealnym tłem dla kolorowych elementów wnętrza. Wnętrze zask. hit – ściany cmiere.



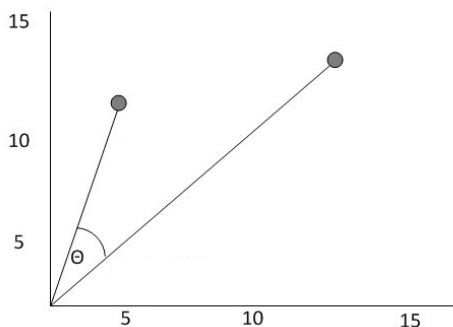
```
1.168785 0.060346 0.502299 0.291747 -0.365562
0.257444 -0.329024 0.758068 0.139132 -0.066573
1.171894 0.076840 -0.002970 -0.360585 -0.144586
0.105688 -0.528267 0.377016 0.220084 -0.132361
-0.232592 0.338373 0.106514 0.096009 -0.068181
-0.698880 0.040483 -0.820396 0.110031 -0.493751
-0.339397 0.278281 -0.000135 -0.121884 0.107060
-0.001215 -0.348834 0.399166 0.391983 0.197091
-0.837996 -0.081890 -0.534775 0.589362 0.278594
-0.724953 0.143085 -0.308889 -0.051467 0.133181
0.110936 -0.159592 -0.338680 0.324832 -0.227569
-0.257161 -0.403050 -0.355761 0.111366 0.127810
-0.045948 0.256404 -0.413172 -0.565309 0.252026
-0.178040 0.353451 -0.043467 0.437229 -0.364093
0.620433 0.491961 -0.044899 0.075592 -0.035806
0.552777 0.539595 -0.307839 -0.488252 0.494307
-0.506171 0.517397 0.010668 -0.247984 0.322363
```



# Dystans między wektorami



# Dystans między wektorami



$$\text{sim} = \cos(\theta) = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}, \quad (1)$$

gdzie  $A_i$  i  $B_i$  są składowymi wektorów  $A$  i  $B$

# Bag-of-words

(1) John likes to watch movies. Mary likes movies too.

(2) John also likes to watch football games.

# Bag-of-words

(1) John likes to watch movies. Mary likes movies too.

(2) John also likes to watch football games.

```
[  
  "John",  
  "likes",  
  "to",  
  "watch",  
  "movies",  
  "Mary",  
  "too",  
  "also",  
  "football",  
  "games"  
]
```



# Bag-of-words

(1) John likes to watch movies. Mary likes movies too.

(2) John also likes to watch football games.

```
[  
  "John",  
  "likes",  
  "to",  
  "watch",  
  "movies",  
  "Mary",  
  "too",  
  "also",  
  "football",  
  "games"  
]
```

(1) [1, 2, 1, 1, 2, 1, 1, 0, 0, 0]

(2) [1, 1, 1, 1, 0, 0, 0, 1, 1, 1]

Wartość *TF-IDF* słowa  $w_i$  w dokumencie  $d_j$ :

$$tfidf_{ij} = tf_{ij} * idf_i, \quad tf_{ij} = \frac{n_{ij}}{\sum_k n_{kj}}, \quad idf_i = \log \frac{|D|}{|d : w_i \in d|} \quad (2)$$

- $tf_{ij}$ : liczba wystąpień słowa  $w_i$  w dokumencie  $d_j$  podzielona przez liczbę słów dokumentu  $d_j$ ,
- $idf_i$ : liczba dokumentów w korpusie podzielona przez liczbę dokumentów zawierających przynajmniej jedno wystąpienie słowa  $w_i$ .

Zalety:

- prostota

Zalety:

- prostota

Wady:

- duża wymiarowość wektorów

Zalety:

- prostota

Wady:

- duża wymiarowość wektorów
- wektory niemalże ortogonalne

Distributional hypothesis — „słowa występujące w tym samym kontekście niosą ze sobą podobne znaczenie.”

# Latent semantic indexing (1988)

- Redukcja wymiarowości macierzy wystąpień słów w dokumentach

	$d_1$	$d_2$	$d_3$	$d_4$	$d_5$	$d_6$
statek	1	0	1	0	0	0
łódź	0	1	0	0	0	0
ocean	1	1	0	0	0	0
podróż	1	0	0	1	1	0
wycieczka	0	0	0	1	0	1

- Hiperparametr: docelowa wymiarowość

- Rozkład według wartości osobliwych:

$$A = U\Sigma V^T, \quad (3)$$

$U$  i  $V$  to macierze ortogonalne

$\Sigma$  to macierz diagonalna, taka, że  $\Sigma = \text{diag}(\sigma_i)$ , gdzie  $\sigma_i$ , to nieujemne wartości szczególne macierzy  $A$ .



- Rozkład według wartości osobliwych:

$$A = U\Sigma V^T, \quad (3)$$

$U$  i  $V$  to macierze ortogonalne

$\Sigma$  to macierz diagonalna, taka, że  $\Sigma = \text{diag}(\sigma_i)$ , gdzie  $\sigma_i$ , to nieujemne wartości szczególne macierzy  $A$ .

- $\{(\text{statek}), (\text{łódź}), (\text{ocean})\} \rightarrow$   
 $\{(1.3452 * \text{statek} + 0.2828 * \text{łódź}), (\text{ocean})\}$

- Automatyczne wykrywanie tematów zawartych w dokumentach
- Dokumenty jako mieszanki tematów
- Tematy jako rozkłady prawdopodobieństwa na zbiorze słów
- Hiperparametr: docelowa liczba tematów

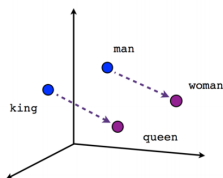
Algorytm — próbkowanie Gibbsa:

- 1 Przejdź przez każdy dokument i losowo (zgodnie z rozkładem Dirichleta) przypisz każde słowo dokumentu do jednego z  $T$  tematów.
- 2 Dla każdego dokumentu  $d$ , dla każdego słowa  $w$  należącego do  $d$ , dla każdego tematu  $t$  oblicz:  $p(t|d)$  oraz oblicz  $p(w|t)$  Przypisz słowu  $w$  nowy temat poprzez losowanie z prawdopodobieństwem  $p(t_i|d) * p(w|t)$  dla każdego tematu  $t_i$ .

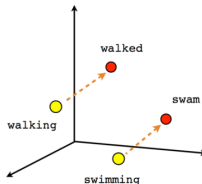
- Osadzanie słów w przestrzeni wektorowej
- Uczenie nienadzorowane
- Niska wymiarowość wektorów
- Reprezentacja słów wraz z zależnościami pomiędzy nimi

# Word embeddings

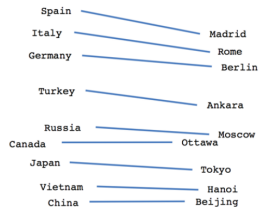
- Osadzanie słów w przestrzeni wektorowej
- Uczenie nienadzorowane
- Niska wymiarowość wektorów
- Reprezentacja słów wraz z zależnościami pomiędzy nimi



Male-Female

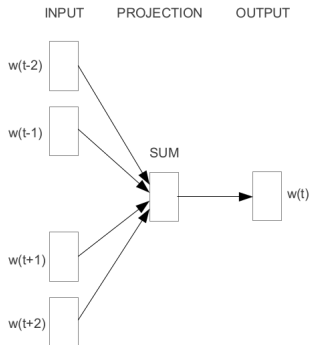


Verb tense

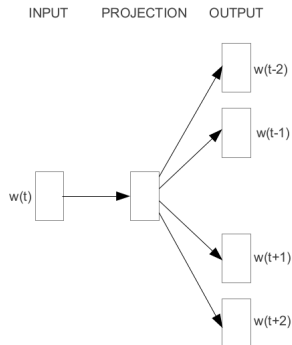


Country-Capital

## płytkiej sieci neuronowej typu feed-forward



**CBOW**



**Skip-gram**

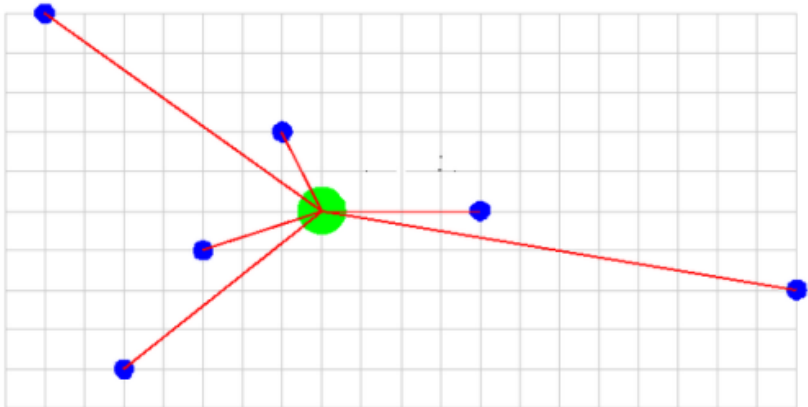


- Globalna macierz współwystąpień słów. Ile razy słowo  $w_i$  występuje w kontekście słowa  $w_j$
- ❶ Zgromadź współwystąpienia słów w formie macierzy  $X$ . Każdy element  $X_{ij}$  takiej macierzy reprezentuje jak często słowo  $i$  występuje w pobliżu słowa  $j$ . Zazwyczaj macierz buduje się poprzez skanowanie bazowego korpusu oknem o ustalonej szerokości, w obrębie którego centralne słowo leży w kontekście słów je otaczających. Dodatkowo można tu wprowadzić wagi dla słów malejące wraz ze wzrostem dystansu od słowa centralnego.
- ❷ Zdefiniuj ograniczenie dla każdej pary słów:  
 $w_i^T w_j + b_i + b_j = \log(X_{ij})$ , gdzie  $w_i$  oznacza wektor głównego słowa,  $w_j$  słowa leżącego w pobliżu  $i$ ,  $b_i$  i  $b_j$  to skalary.
- ❸ Zdefiniuj funkcję kosztu 4:

$$J = \sum_{i=1}^V \sum_{j=1}^V f(X_{ij})(w_i^T w_j + b_i + b_j - \log X_{ij})^2, \quad (4)$$

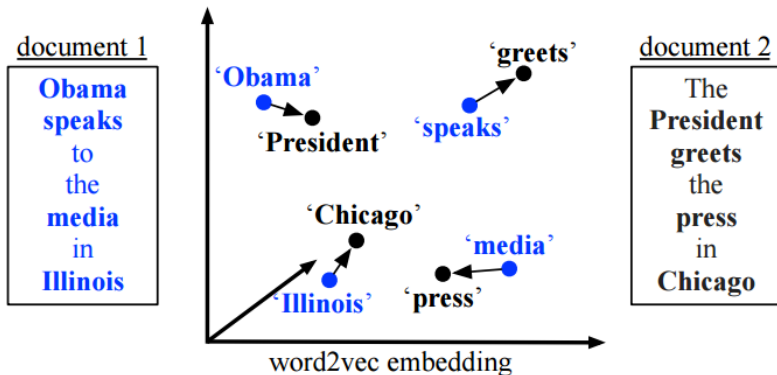


# Centroid



# Word Mover's Distance

Dystans pomiędzy dokumentami  $A$  i  $B$  to minimalny skumulowany dystans jaki słowa dokumentu  $A$  muszą „przebyć”, aby osiągnąć słowa dokumentu  $B$



- 20000 artykułów tekstowych w formacie *JSON*
- język polski
- słowa specyficzne dla różnych branż
- struktura artykułu:
  - treść: tytuł, nagłówek, tekst
  - metadane: id, kategoria, słowa kluczowe

## ❶ Oczyszczanie tekstu ze znaczników

- 1 Oczyszczanie tekstu ze znaczników
- 2 Usunięcie słów stopu

a, aby, ach, acz, aczkolwiek, aj, albo, ale, ależ, ani, aż, bardziej, bardzo, bo, bowiem, by, byli, bynajmniej, być, był, była, było, były, będzie, będą, cali, cała, cały, ci, cię, ciebie, co, cokolwiek, coś, czasami, czasem, czemu, czy, czyli, daleko, dla, dlaczego, dlatego, do, dobrze, dokąd, dość, dużo, dwa, dwaj, dwie, dwoje, dziś, dzisiaj, gdy, gdyby, gdyż, gdzie, gdziekolwiek, gdzieś, go, i...

- ❶ Oczyszczanie tekstu ze znaczników
- ❷ Usunięcie słów stopu

- ❶ Oczyszczanie tekstu ze znaczników
- ❷ Usunięcie słów stopu
- ❸ Zamiana na małe litery



- ❶ Oczyszczanie tekstu ze znaczników
- ❷ Usunięcie słów stopu
- ❸ Zamiana na małe litery
- ❹ Tokenizacja i lematyzacja

# Preprocessing - przykład

Każda mama cieszy się, gdy jej maluszek z apetytem zjada przygotowany przez nią posiłek.



"mama",  
"cieszyć",  
"maluszek",  
"apetyt",  
"zjadać",  
"przygotować",  
"posiłek"

# Liczba wspólnych kategorii

# Miara jakości wyszukiwania: Normalized Discounted Cumulative Gain

- Discounted Cumulative Gain:

$$\text{DCG}_p = \sum_{i=1}^p \frac{rel_i}{\log_2(i+1)}, \quad (6)$$

gdzie  $p$  to liczba elementów rankingu,  $i$  to miejsce przedmiotu w rankingu, a  $rel$  to poziom relewantności elementu.

# Miara jakości wyszukiwania: Normalized Discounted Cumulative Gain

- Discounted Cumulative Gain:

$$DCG_p = \sum_{i=1}^p \frac{rel_i}{\log_2(i+1)}, \quad (6)$$

gdzie  $p$  to liczba elementów rankingu,  $i$  to miejsce przedmiotu w rankingu, a  $rel$  to poziom relewantności elementu.

- Normalized Discounted Cumulative Gain:

$$nDCG_p = \frac{DCG_p}{IDCG_p}. \quad (7)$$

# Liczba wspólnych słów kluczowych

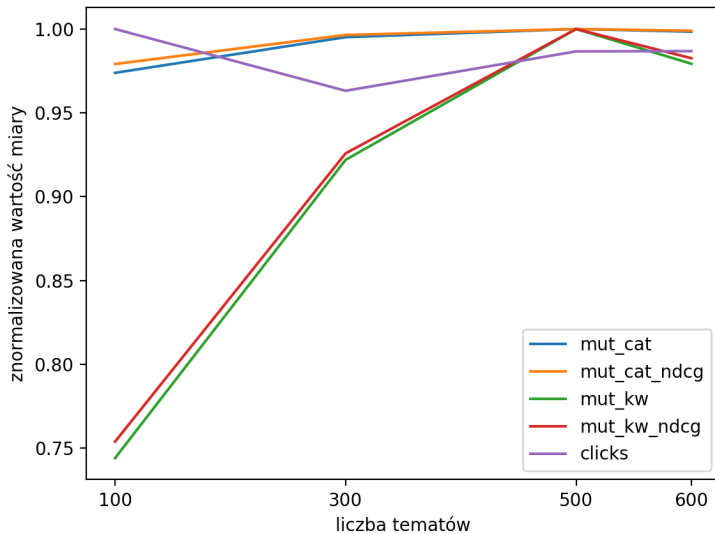
*clicks* — ocena na podstawie historycznej aktywności użytkowników mierzona na podstawie liczby kliknięć w odnośniki.



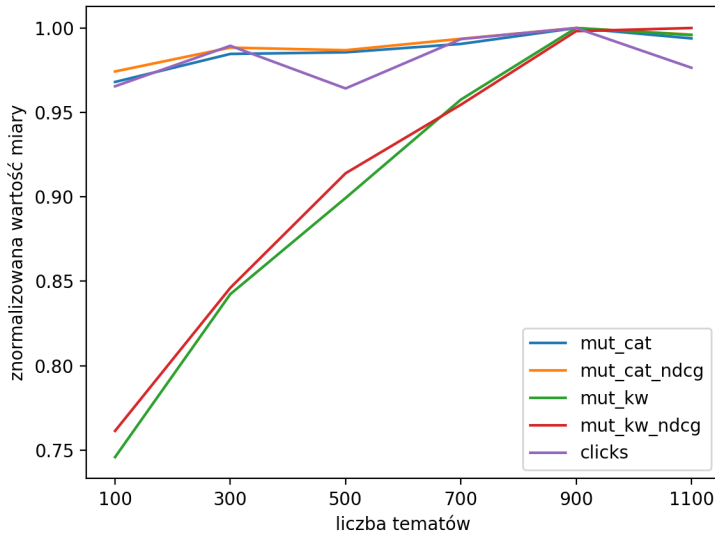


- 1  $mut\_kw[_ndcg]$  — relewantność wyszukanych artykułów liczona na podstawie liczby wspólnych słów kluczowych z artykułem bazowym. Również stosuję dwa warianty: średnia relewantność wyszukanych artykułów oraz miara  $nDCG$ .
- 2  $users$  — ocena na podstawie eksperckiej oceny użytkowników. W badaniu wykorzystałem 5 użytkowników operujących każdy na tym samym zbiorze par testowych. Pary zostały wygenerowane (zgodnie z wcześniejszym opisem metody) na podstawie 50 artykułów bazowych wylosowanych spośród wszystkich artykułów udostępnionych mi przez *Allegro*.

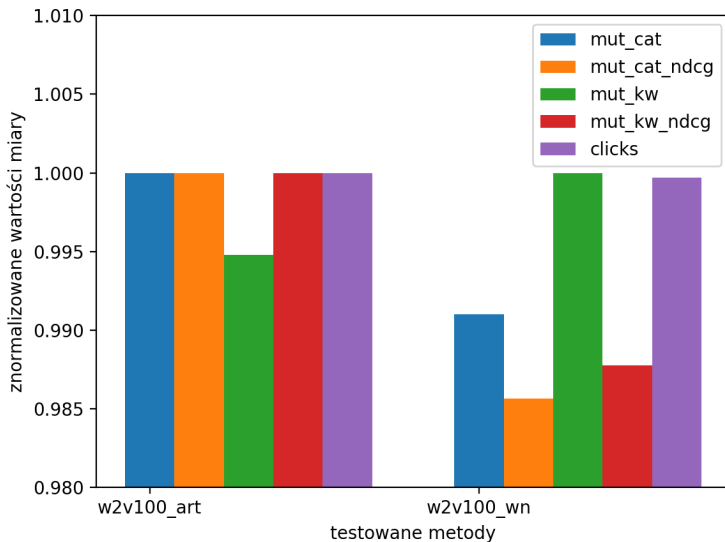
# LSI w zależności od liczby tematów



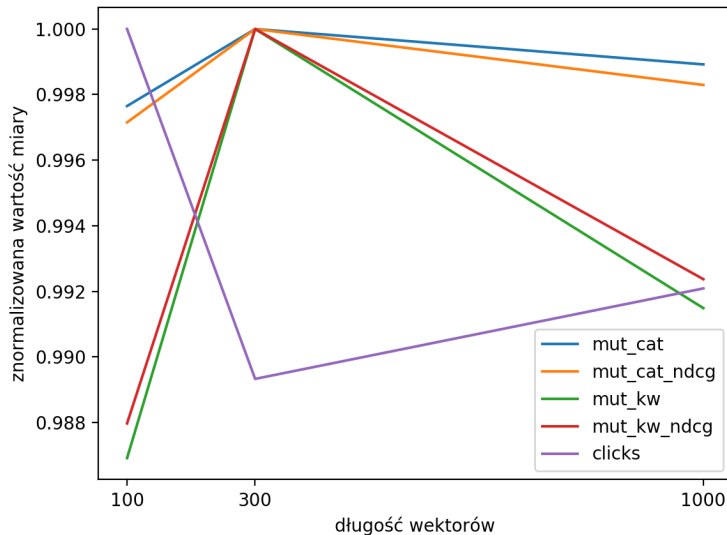
# LDA w zależności od liczby tematów



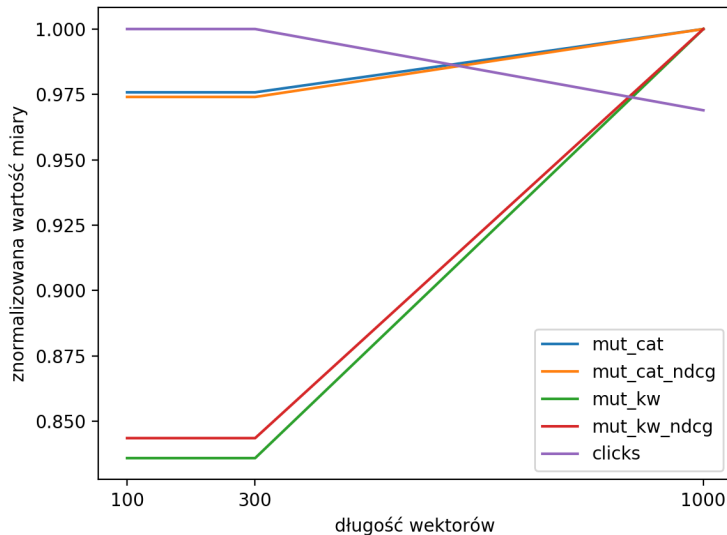
# Word2vec w zależności od korpusu



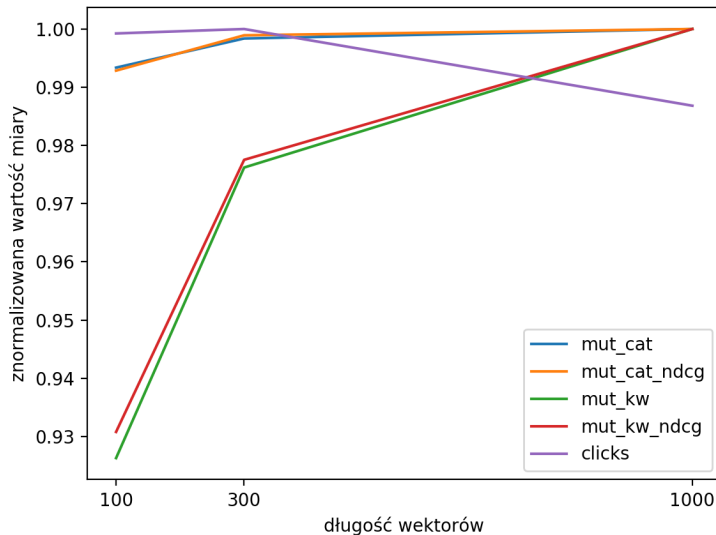
# Word2vec w zależności od długości wektorów



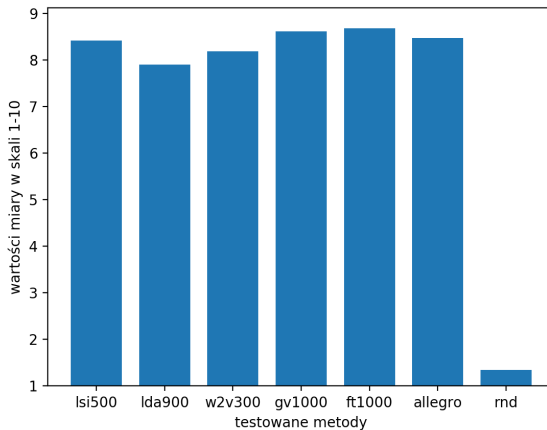
# GloVe w zależności od długości wektorów



# FastText w zależności od długości wektorów

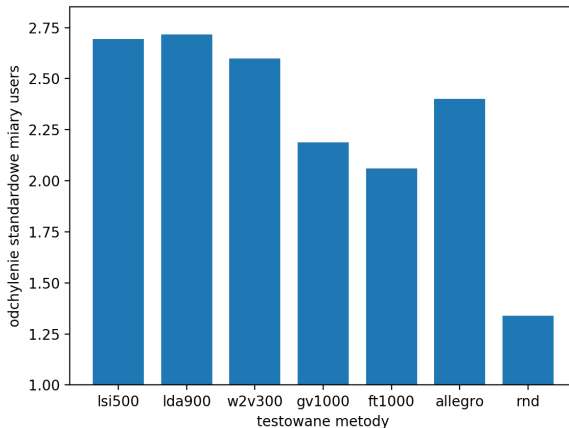


# Wyniki ewaluacji eksperckiej dla wybranych metod





# Porównanie odchyleń standardowych ocen eksperckich dla wybranych metod



- Brak istotnych statystycznie różnic między wynikami wszystkich metod






- Brak istotnych statystycznie różnic między wynikami wszystkich metod
- Im dłuższe wektory *word embeddings* tym lepsze rezultaty

- Brak istotnych statystycznie różnic między wynikami wszystkich metod
- Im dłuższe wektory *word embeddings* tym lepsze rezultaty
- Większa liczba tematów nie implikuje lepszych rezultatów

- Brak istotnych statystycznie różnic między wynikami wszystkich metod
- Im dłuższe wektory *word embeddings* tym lepsze rezultaty
- Większa liczba tematów nie implikuje lepszych rezultatów
- ...



testowane metody nie odbiegają jakością od dotychczasowej.  
python elasticsearch nlp trudno jest zmierzyc efekty

-  D. M. Blei, A. Y. Ng, M. I. Jordan, *Latent Dirichlet Allocation*, Journal of Machine Learning Research, tom 3 num. 4–5, 2003
-  S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, R. Harshman, *Indexing by latent semantic analysis*, Journal of the American Society for Information Science, tom 41, num. 6, 1990
-  A. Joulin, E. Grave, P. Bojanowski T. Mikolov, *Bag of Tricks for Efficient Text Classification*, Facebook AI Research, 2016
-  T. Mikolov, K. Chen, G. Corrado, J. Dean, *Efficient Estimation of Word Representations in Vector Space*, International Conference on Machine Learning (ICML), 2013
-  J. Pennington, R. Socher, C. D. Manning, *GloVe: Global Vectors for Word Representation*, Computer Science Department, Stanford University, Stanford, CA 94305, 2014