

Latex

Лазар Вукадиновић

21. април 2024.

1 Математика

1.1 Фуријеова анализа

Свака периоди?на функција¹, по теорији Фуријеових редова, може се представити помо?у бескона?но много ортогоналних функција. Како овај на?ин представљања функције пружа могу?ност сасвим друга?ије анализе у односу на анализу у временском домену, поставља се питање да ли је могу?е исту идеју проирити на функције које нису периоди?не. Ако се непериоди?на функција посматра као периоди?на са бескона?но великим периодом, Фуријеова трансформација проирује овај концепт разлагања функција и на непериоди?не функције.

1.1.1 Формуле

1. Периодицне функције $f(x) = f(x + 2\pi)$
2. Представљање: $a \cos \omega x + b \sin \omega x$ $a, b \in \mathbf{R}$
3. $\prod_{i=N}^N = f(i)$

2 Моделовање тема

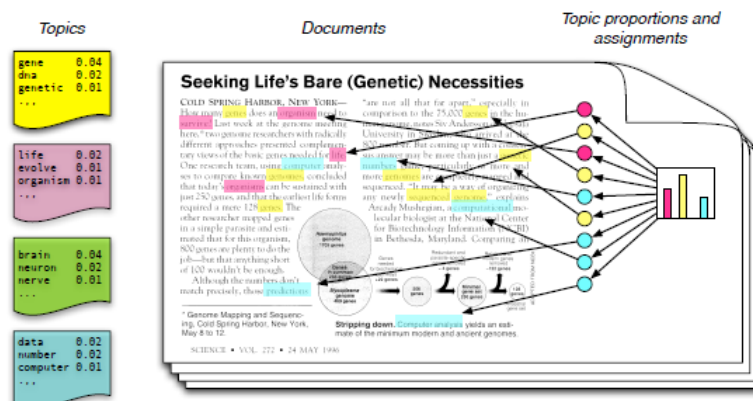
Latent Dirichlet Allocation, надаље LDA, је најједноставнији приступ проблему моделовања тема и његова примена је предмет овог рада. Основна карактеристика LDA алгоритма је могу?ност **издвајања** тема које су присутне у некој колекцији докумената без било каквог додатног знања. Дакле, применом LDA-а могу?е је открити теме "о којима говорижадати скуп докумената а да се притом никакво додатно експертско знање не укљу?ује. Полазна претпоставка LDA-а је да сваки документ у колекцији докумената "говори о"вие

тема. Оправданост ове претпоставке би?е илустрована на неколико примера. Добро је познат роман Бранка ?опи?а "Орлови рано лете". Уколико би неко ко није про?итао ови књигу елео да зна "о ?ему се ради" у њој, највероватније би добио одговор да је у питању књига која се бави доивљајима групе де?ака на по?етку Другог свецког рата. Иако је то најири оквир романа, у њему су присутне и теме о љубави, дружењу, пријатељству, рату, пустоловинама итд. Према томе, роман, опте гледано, обухвата вие тема, али се са неколико њих интензивно бави.

2.1 Графицки пример документа

Генерално, процес закљу?ивања тематике документа може се илустровати следе?им примером. На Слици 1 представљен је ?ланак Сеекинг Лифе'с Баре(Генетиц) Нецеситиес који говори о употреби анализе података за одре?ивање броја гена који организам треба да поседује да би преивео (у еволутивном смислу). Може се уо?ити да су три најзаступљеније области у овом тексту - анализа података, еволутивна биологија и генетика. На слици су ру?но означене неке ре?и које припадају овим областима. Ре?и које се могу сврстати у област *анализе података* означене су плавом бојом, ре?и које припадају *генетици* означене су утом бојом, док су ре?и које се односе на *еволутивну биологију* означене розе бојом. Уколико би се ова процедура применила на сваку ре? текста, јасно би се уо?ило колико је која тема заступљена у овом тексту. Математи?ки, *присуство* теме у тексту се означава односом броја ре?и "обојених"јатом бојом и укупног броја ре?и у тексту.

¹погледати како се дефинисе функција



Слика 1: Пример цланкова

Наравно, постоје реџи које се могу сврстати у више од једне теме. Такве реџи би биле обојене са две или више боја, али због прегледности слике, такви случајеви су изостављени.



2.2 Приказ рада алгоритама

Рад алгоритама приказаног на Слици 1 је следећи

1. Први корак:

Пример документа				
математика	физика	хемија	биологија	историја

2. Други корак:

Пример документа				
$\sum_1^N f(reci)$	$\sum_1^N f(reci)$	$\sum_1^N f(reci)$	$\sum_1^N f(reci)$	$\sum_1^N f(reci)$
математика	физика	хемија	биологија	историја

3. Финални корак:

Реци	Теме		
	1	2	3
математика	n_1^1	n_2^1	n_3^1
физика	n_1^2	n_2^2	n_3^2
хемија	n_1^3	n_2^3	n_3^3
биологија	n_1^4	n_2^4	n_3^4
историја	n_1^5	n_2^5	n_3^5

Литература

- [1] Mell, Peter, and Tim Grance. The NIST definition of cloud computing. (2011).
- [2] Buyya, Rajkumar, et al. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility.”Future Generation computer systems 25.6 (2009): 599-616.

Садржај

1	Математика	2
1.1	Фуријеова анализа	2
1.1.1	Формуле	2
2	Моделовање тема	2
2.1	Графицки пример документа	2
2.2	Приказ рада алгоритама	4