

Latex

Име и презиме

3. март 2018

# 1 Математика

## 1.1 Фуријеова анализа

Свака периодична функција<sup>1</sup>, по теорији Фуријеових редова, може се представити помоћу бесконачно много ортогоналних функција. Како овај начин представљања функције пружа могућност сасвим другачије анализе у односу на анализу у временском домену, поставља се питање да ли је могуће исту идеју проширити на функције које нису периодичне. Ако се непериодична функција посматра као периодична са бесконачно великим периодом, Фуријеова трансформација проширује овај концепт разлагања функција и на непериодичне функције.

### 1.1.1 Формуле

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

# 2 Моделовање тема

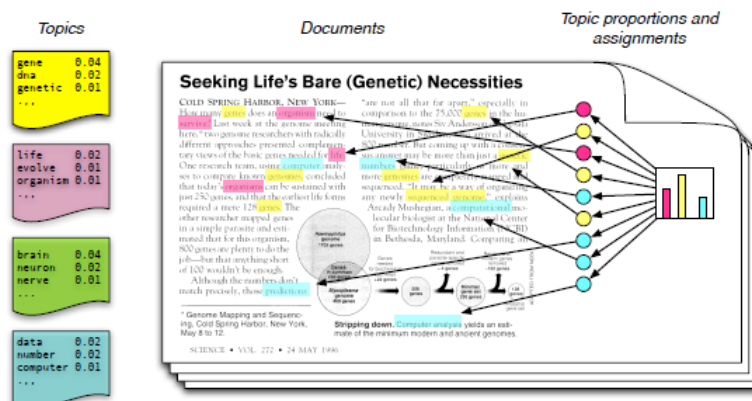
*Latent Dirichlet Allocation*, надаље LDA, је најједноставнији приступ проблему моделовања тема, и његова примена је предмет овог рада. Основна карактеристика LDA алгорита је могућност **издвајања** тема које су присутне у некој колекцији докумената без било каквог додатног знања. Дакле, применом LDA-а могуће је открити теме "о којима говорижадати скуп докумената а да се притом никакво додатно експертско знање не укључује. Полазна претпоставка LDA-а је да сваки документ у колекцији докумената "говори о" више тема. Оправданост ове претпоставке биће илу-

стрована на неколико примера. Добро је познат роман Бранка Ћопића "Орлови рано лете". Уколико би неко ко није прочитао ови књигу желео да зна "о чему се ради" у њој, највероватније би добио одговор да је у питању књига која се бави доживљајима групе дечака на почетку Другог светског рата. Иако је то најшири оквир романа, у њему су присутне и теме о љубави, дружењу, пријатељству, рату, пустоловинама итд. Према томе, роман, опште гледано, обухвата више тема, али се са неколико њих интензивно бави.

## 2.1 Графички пример документа

Генерално, процес закључивања тематике документа може се илустровати следећим примером. На Слици 1 представљен је чланак *Seeking Life's Bare(Genetic) Necessities* који говори о употреби анализе података за одређивање броја гена који организам треба да поседује да би преживео (у еволутивном смислу). Може се уочити да су три најзаступљеније области у овом тексту - анализа података, еволутивна биологија и генетика. На слици су ручно означене неке речи које припадају овим областима. Речи које се могу сврстати у област *анализе података* означене су плавом бојом, речи које припадају *генетици* означене су жутом бојом, док су речи које се односе на *еволутивну биологију* означене розе бојом. Уколико би се ова процедура применила на сваку речтекста, јасно би се уочило колико је која тема заступљена у овом тексту. Математички, *присуство* теме у тексту се означава односом броја речи "обојених"јатом бојом и укупног броја речи у тексту.

<sup>1</sup>погледати како се дефинише функција



Слика 1: Пример чланка

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



## 2.2 Приказ рада алгоритма

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 :

Речи	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

<b>1</b>	<b>Математика</b>	<b>2</b>
1.1	Фуријеова анализа . . . . .	2
1.1.1	Формуле . . . . .	2
<b>2</b>	<b>Моделовање тема</b>	<b>2</b>
2.1	Графички пример документа . . . . .	2
2.2	Приказ рада алгоритма . . . . .	4