

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение
высшего профессионального образования

Национальный исследовательский университет
«Высшая школа экономики»
Факультет информатики, математики и компьютерных наук

Программа подготовки магистров по направлению
01.04.02 Прикладная математика и информатика

ПРОЕКТ ПО НАУЧНО-ИССЛЕДОВАТЕЛЬСКОМУ СЕМИНАРУ

На тему «Поиск и автоматический анализ научных публикаций.»

Руководитель проекта: Крылов Владимир Владимирович

Список участников проекта:

Ларюшина Юлия
Шашкин Павел
Желонкин Дмитрий
Соколов Артем
Кузнецов Владимир
Куренков Евгений
Мозохин Дмитрий
Груздев Алексей
Вороня Ксения

Нижний Новгород, 2016

Оглавление:

1 Сбор данных для исходной задачи, создание модели поиска тематически близких документов	3
2 Исследование различных методов построения summary и выявление наиболее качественного саммаризатора.....	6
3 Построение summary с использованием модуля Gensim и извлечение keywords из статьи.....	10
4 Исследование возможностей по созданию тематической модели на основе научных работ.....	14

1 Сбор данных для исходной задачи, создание модели поиска тематически близких документов

Задача группы	Сбор данных для задачи, создание модели для поиска тематически близких документов, создание интерфейса для поиска тематических близких документов к заданному, визуализация и оценка полученных результатов.
Исполнители	<ul style="list-style-type: none"> ◦ Ларюшина Юлия ◦ Шашкин Павел ◦ Желонкин Дмитрий
Используемые библиотеки и программные средства, данные	<ul style="list-style-type: none"> ◦ Язык R, python2 ◦ Пакеты R: magrittr, dplyr, quanteda, tidyr, stm, lubridate, stmCorrViz, LDAvis, aRxiv, pdftools, parallel, rplos ◦ Библиотеки python: nltk, gensim, numpy, pandas, scipy ◦ Библиотека для мультимодального тематического моделирования bigARTM ◦ Источники данных: https://www.plos.org/ https://arxiv.org/
Решенные подзадачи	<p>1. Сбор документов для обучения моделей из открытых источников.</p> <p>Для тренировки модели были собраны данные по медицинской тематике из открытых источников.</p> <ul style="list-style-type: none"> • Для сбора данных из источника arXiv создан скрипт: https://github.com/artezio-kseniav/text_summarization/blob/master/data_collection/corpora_getter.R <p>Данные (статьи + метаданные) могут быть собраны в формате pdf и преобразованы в текст.</p> <ul style="list-style-type: none"> • Для сбора данных из источника plos создан скрипт:

https://github.com/artezio-kseniav/text_summarization/blob/master/data_collection/search_plos_example.R

Данные изначально представлены в формате xml и разбиты на необходимые структурные единицы (article, body, abstract + metadata).

2. Обработка корпуса для построения модели.

3. Построение мультимодальной тематической модели с использованием bigARTM.

4. Создание интерфейса для взаимодействия с мультимодальной тематической моделью (поиска близких документов).

Подробная инструкция по использованию и работе доступна по ссылке: https://github.com/artezio-kseniav/text_summarization/tree/master/python-topic-modeling

5. Построение структурированной тематической модели для поиска близких документов.

Для обработки корпуса и построения модели был создан функционал: https://github.com/artezio-kseniav/text_summarization/blob/master/topics/plos_stm.R

1. Для начала средствами библиотек dplyr, tidyr и lubridate данные, собранные с plos, приводятся к виду id + дата_публикации + секция_документа + текст.
2. Средствами quanteda осуществляется построение матрицы частот для тренировки модели.
3. Производится обучение модели stm, в которой распределение тем зависит не только от секции документа, но и от даты публикации.
4. Полученные темы (topics) визуализируются средствами LDAvis (производится многомерное шкалирование векторов статей) и stmCorrViz (производится иерархическая кластеризация матрицы корреляции тем).

6. Визуализация полученных результатов для формирования представления о собранном корпусе,

	<p>экспертной оценки выделенных тем и формирования представления о предметной области:</p> <p>https://github.com/artezio-kseniav/text_summarization/blob/master/topics/corrviz.html</p> <p>https://github.com/artezio-kseniav/text_summarization/blob/master/topics/simil.html</p>
Нерешенные подзадачи	<ol style="list-style-type: none"> 1. Создание функционала для взаимодействия с документом в формате pdf 2. Создание функционала для обновления внутренней базы статей 3. Объединение модуля с модулями других подгрупп
Связь с задачами других групп	<p>Созданная модель осуществляет сбор поискового корпуса, его обработку и поиск близких документов с точки зрения распределения тем в нём и в документах корпуса. После идентификации близких документов предполагается передача данных на вход модулю, осуществляющему формирование аннотаций (summary) для конечного пользователя.</p>

2 Исследование различных методов построения summary и выявление наиболее качественного саммаризатора

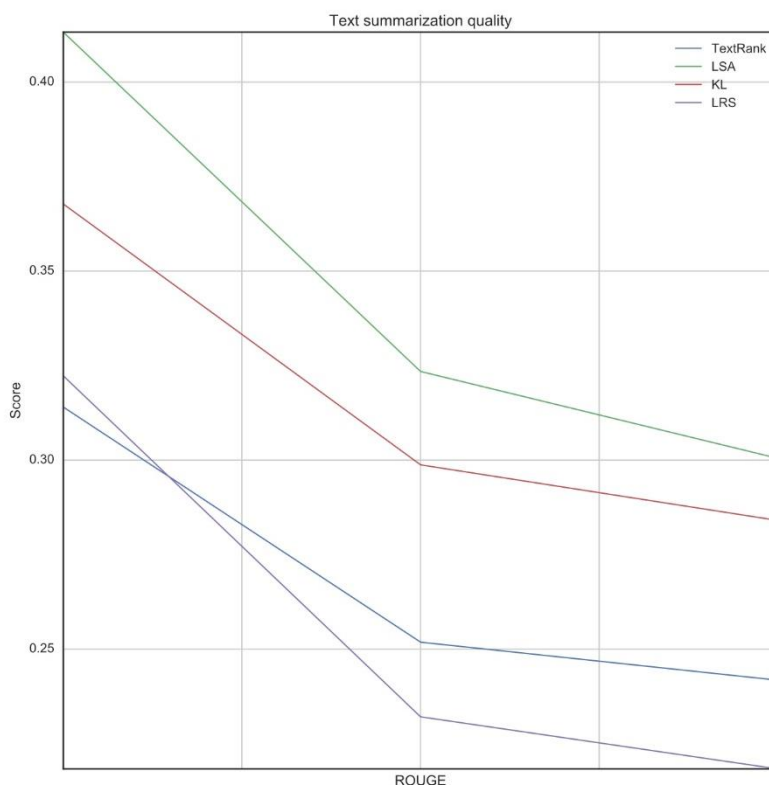
Задача группы	<ul style="list-style-type: none"> • Исследовать различные методы построения summary • Собрать тестовый набор данных • Выявить наиболее качественный саммаризатор
Исполнители	<ul style="list-style-type: none"> • Груздев Алексей • Мозохин Дмитрий
Используемые библиотеки и программные средства, данные	<ul style="list-style-type: none"> • python3 • Библиотеки python: nltk, sumy, seaborn • Источники данных: https://www.plos.org/
Решенные подзадачи	<p>1. Сбор документов для обучения моделей из открытых источников.</p> <p>2. Наиболее яркий представитель методов построения summary основанных на Deep Learning - Sequence-to-Sequence with Attention Model for Text Summarization by Google. Более подробно она описана в Google Research Blog'e. Данная модель была натренирована на GigaWord dataset'e, который является закрытым для доступа из вне - даже для ведения исследований в университете, без предварительной договоренности между университетом и владельцами датасета. Процесс тренировки модели с использованием Deep Learning библиотеки Tensorflow выложен в общий доступ, но в данной модели отсутствуют веса, поскольку они были получены на закрытом датасете, поэтому без получения данного датасета, невозможно использовать Sequence-to-Sequence модель для построения саммаризатора. Поэтому далее были рассмотрены более классические методы построения summary:</p>

- a) TextRank – метод, основанный на представлении исходного текста в качестве графа, где вершине соответствует некоторый токен(например, предложение). Весам ребер приписывается значение семантической близости смежных вершин и после чего применяется Google Page Rank алгоритм, с помощью которого определяется набор "наиболее значимых" токенов, выступающих в качестве summary.
- b) LSA - то метод обработки информации на естественном языке, анализирующий взаимосвязь между коллекцией документов и терминами в них встречающимися, сопоставляющий некоторые факторы (тематики) всем документам и терминам. В основе метода латентно-семантического анализа лежат принципы факторного анализа, в частности выявление латентных связей изучаемых явлений или объектов. При классификации/кластеризации документов этот метод используется для извлечения контекстно-зависимых значений лексических единиц при помощи статистической обработки больших корпусов текстов.
- c) Kullback–Leibler summarizator - метод основанный на эвристическом добавлении предложений к имеющемуся саммари; в качестве критерия добавления/оценки текущего качества summary используется мера расстояние Кульбака — Лейблера которая считается для summary и исходного текста. Таким образом, суть метода состоит в нахождении набора предложений, которые наилучшим образом аппроксимирует исходный текст по КЛ-расстоянию.
- d) LexRank - метод, схожий с TextRank, но конструктивной особенностью которого является его применимость к целому набору исходных текстов с одной темой.

3. Проведен сравнительный анализ указанных методов на основе собранных текстов.

В качестве метрики использовались значения ROUGE-1, ROUGE-2, ROUGE-3, усредненные для всей коллекции документов.

Заметим, что большее значение метрики соответствует лучшему качеству построенного summary.



Из графиков видно, что для данного набора текстов и методов, при ограничении длины summary в 15 предложений, метод LSA в среднем формирует более качественные краткие изложения исходных текстов.

Нерешенные подзадачи

1. Улучшение TextRank метода путем изменения метрики для подсчета семантической близости токенов (см. <https://arxiv.org/pdf/1602.03606v1.pdf>)
2. Использование Sequence-to-Sequence модели и других DL-based моделей
3. Объединение модуля с модулями других подгрупп

Source code	https://github.com/artezio-kseniav/text_summarization/tree/master/summary_v1
Вывод	<p>Для построения качественного саммаризатора необходимо использование DL-based моделей, обучение которых требует больших вычислительных мощностей и временных ресурсов. Однако, применение наивных методов позволяет построить модели, которые обладают необходимым качеством, чтобы выступать в качестве baseline-решения.</p>

3 Построение summary с использованием модуля Gensim и извлечение keywords из статьи

Задача группы	<ul style="list-style-type: none"> • Построить качественное summary с использованием модуля Gensim • Извлечь keywords из статьи
Исполнители	<ul style="list-style-type: none"> • Кузнецов Владимир • Вороная Ксения • Куренков Евгений
Используемые библиотеки и программные средства, данные	<ul style="list-style-type: none"> • python3 • Anaconda - the leading open data science platform powered by Python. • Jupyter Notebook - a web application that allows you to create and share documents that contain live code, equations, visualizations and explanatory text. • Библиотеки python: genism, nltk, lxml • Источники данных: Extract articles from on of the trusted websites http://www.psychiatrictimes.com/
Решенные подзадачи	<p>1. Произвели парсинг одного из предложенных вебсайтов с целью извлечения топ-10 самых последних статей по шизофрении: http://www.psychiatrictimes.com/ Добавили эти статьи в небольшой корпус.</p> <p>2. Произвели предобработку текста для извлечения keywords, а именно удалили стоп-слова и сделали лемматизацию.</p> <p>3. Извлечение keywords <u>Приводим пример для статьи:</u> Adjunctive Topiramate in People With Schizophrenia The authors are: Brian Miller, MD, PhD, MPH</p>

The date of publication is: September 29, 2016

```
In [41]: print (' ***** Extracted Keywords ***** ')
keywords = keywords(modified_article)
print(keywords)

***** Extracted Keywords *****
topiramate
trials
patients
antipsychotics
effect
effects
antipsychotic weight
patient schizophrenia
randomized trial
```

4. Построение summary (пример для той же статьи см. под таблицей)

Gensim использует алгоритм TextRank с метрикой BM25. TextRank каждому предложению присваивает метрику называемую прочностью соединения, которая ставится в соответствие количеству слов в предложении. Прочность соединения вычисляется по BM25 алгоритму и находится по следующей формуле

$$\text{score}(D, Q) = \sum_{i=1}^n \text{IDF}(q_i) \cdot \frac{f(q_i, D) \cdot (k_1 + 1)}{f(q_i, D) + k_1 \cdot (1 - b + b \cdot \frac{|D|}{\text{avgdI}})},$$

Где Q — предложение, состоящее из слов $q_1..q_n$.

D — документ (в нашем случае текст статьи).

$f(q_i, D)$ — частота слова q_i в документе

|D| - длина документа (количество слов в нем).

Avgdl — средняя длина документа.

k_1 и b — свободные коэффициенты (приблизительно равны 2.0 и 0.75 соответственно).

$\text{IDF}(q_i)$ — обратная документная частота для слова q_i .

Таким образом, имеем граф, узлы которого соответствуют прочностям соединения для каждого предложения и чем выше данная оценка, тем более ценным считается предложение.

	<p>В функции gensim.summarize() присутствует параметр ratio, который устанавливает степень сжатия статьи. Чем выше данный параметр, тем более подробное изложение получается и тем больше времени необходимо для выполнения.</p> <p>Также в программе присутствует функция keywords для подсчета и вывода наиболее популярных слов в файле. На вход программе передается документ, каждая строка которого содержит тело статьи, для которой необходимо сделать саммари. На выходе получается два файла: summary — каждая строка которого содержит краткое изложение статьи и keywords - каждая строка которого содержит наиболее значимые слова в документе.</p> <p>Для оценки качества саммаризации используется метрика ROUGE-n (Recall-Oriented Understudy for Gisting Evaluation, http://anthology.aclweb.org/W/W04/W04-1013.pdf), где n- это количество последовательных слов, используемых для оценки. Данная возможность не была имплементирована в данной программе и планируется к реализации в дальнейшей разработке.</p>
Нерешенные подзадачи	<ol style="list-style-type: none"> 1. Улучшение процесса извлечения key-words 2. Объединение модуля с модулями других подгрупп
Source code	<p>https://github.com/artezio-kseniav/text_summarization/tree/master/summary_version2</p> <p>https://github.com/artezio-kseniav/text_summarization/tree/master/text_summarization</p>

```

# ratio (default = 0.2) - to specify what fraction of sentences in the original text should be returned as
# output
# word_count - to specify the maximum amount of words in the summary
# "split" option if need in a list of strings instead of a single string

print(' ***** Summary ***** ')
summary = summarize(articles[0], word_count=100)
print(summary)

```

***** Summary *****

In a systematic search of PubMed/MEDLINE, the researchers looked for all published studies of antipsychotic augmentation with topiramate in patients with schizophrenia-spectrum disorders (both randomized, placebo-controlled trials or open-label trials with an untreated control group). The primary outcome was change in total score on either the Positive and Negative Syndrome Scale (PANSS) or the Brief Psychiatric Rating Scale (BPRS).

There was a trend for more paresthesias with topiramate use (relative risk = 2.0), but otherwise no difference in adverse effects reported in at least 3 trials. The authors found evidence that adjunctive topiramate was associated with significantly greater reductions in psychopathology (particularly in clozapine-treated patients) and body weight. Other than an increase in paresthesias, there were no differences in adverse effects or all-cause discontinuation between topiramate and placebo.

4. Исследование возможностей по созданию тематической модели на основе научных работ

Данное исследование провел Артем Соколов.

Изначально было запланировано что поиск и анализ публикаций будет основываться на тематическом моделировании. Это сравнительно новое направление в машинном обучении подразумевает что поиск будет выполняться на основе анализа документа в целом, т.е. с учетом текста, ссылок, авторов. В нашей группе мы разделились так что 4 человека отвечали непосредственно за исследование возможностей по созданию тематической модели на основе научных работ. 2 человека искали и исследовали возможности языка R и два человека делали то же самое для языка python с целью дальнейшего сравнения.

Я принимал участие в работе второй группы. Мы нашли очень перспективную библиотеку для создания и работе с тематическими моделями BigARTM, написанной на языке c++, но имеющей, также, API для языка python (<http://bigartm.org/>). Коллегами была найдена и выложена большая выборка научных статей на медицинские темы. Мы обработали перевели эти документы в формат vowpal wabbit, поддерживаемый BigARTM и далее, разделили наши усилия.

Коллега занимался исследованием возможностей модели в направлении мультимодальности (т.е. учитывания не только текста, но и ссылок и авторов для нахождения релевантных), а я занимался построением иерархической тематической модели. Такая модель помимо матриц документы-темы и темы-слова содержит матрицу темы-подтемы для каждой соседних уровней иерархии. (www.machinelearning.ru/). Как я выяснил, BigARTM имеет интерфейс для построения таких моделей, но не в основном, стабильном варианте. Я написал код на питоне для того что бы построить трёхуровневую модель на основе текстов документов. Подбирались параметры и

регуляризаторы для того что бы сделать матрицы разреженными (см. <http://bigartm.org/> для дополнительной информации о принципах работы BigARTM). Количество тем и подтем подбиралось экспериментально. Какой-то абсолютной числовой характеристики, по которой можно было бы судить о чистоте тем найдено не было. Я приведу здесь примеры для двух уровней, при этом я уверен что, полученные результаты еще возможно значительно улучшить.

topic_0:	cell	activation	receptor	induced	culture	incubated	signaling	inhibitor	treated	play	staining	medium	hour	rtt	inhibit
topic_1:	three	sequence	dna	strain	por	genome	resistance	bacterial	sequencing	isolates	bacteria	bp	plasmid	primer	fragm
topic_2:	specie	plant	tree	water	environmental	site	leaf	abundance	soil	root	fish	diversity	community	temperature	mosq
topic_3:	observed	similar	described	experiment	line	previously	specific	type	role	presence	analyzed	either	indicated	reduced	detect
topic_4:	protein	site	binding	domain	structure	interaction	peptide	residue	amino	complex	molecular	chain	blind	molecule	ligand
topic_5:	woman	child	hospital	mortality	death	malaria	birth	pregnancy	respiratory	delivery	maternal	admission	parent	infant	mother
topic_6:	population	case	year	individual	high	per	hiv	estimate	country	among	prevalence	proportion	cost	area	testing
topic_7:	expression	gene	pathway	ma	ma	target	transcript	involved	promoter	regulation	probe	transcript	functional	development	revers
topic_8:	group	control	effect	significant	difference	test	condition	response	subject	four	presented	greater	whether	ability	man
topic_9:	activity	well	concentration	ml	determined	min	acid	growth	high	mm	assay	material	decreased	reaction	amount
topic_10:	treatment	day	drug	week	following	month	therapy	baseline	period	dose	duration	received	treated	mg	hour
topic_11:	cancer	tumor	survival	breast	metastasis	biopsy	egfr	carcinoma	invasion	chemotherapy	prostate	colon	gastric	cervic	
topic_12:	infection	virus	infected	antibody	viral	cd4	immune	response	vaccine	antigen	parasite	influenza	cytokine	vaccin	
topic_13:	within	image	area	surface	volume	left	right	bone	lesion	cm	imaging	side	white	ct	sectio
topic_14:	male	female	body	weight	diabetes	muscle	glucose	insulin	food	bmi	adult	metabolic	mass	diet	error
topic_15:	study	journal	found	reported	risk	low	included	al	et	outcome	potential	evidence	identified	report	
topic_16:	mouse	tissue	animal	rat	lung	liver	injury	injection	inflammatory	inflammation	kidney	heart	cardiac	vascular	sectio
topic_17:	model	time	two	value	rate	mean	shown	figure	change	first	system	increase	show	due	functi
topic_18:	region	mutation	genetic	snp	association	genotype	allele	variant	locus	chromosome	polymorphism	exon	deletion	haplotype	phenot
topic_19:	patient	disease	however	compared	factor	associated	higher	increased	significantly	clinical	movement	pd	neural	noise	ampli
topic_20:	brain	neuron	m	stimulus	channel	functional	ad	signal	cortex	motor	disorder	social	men	pain	medic
topic_21:	age	participant	score	status	association	symptom	sex	physical	variable	gender	medical	part	intervention	general	suppl
topic_22:	also	may	could	use	health	important	would	research	care	likely	medical	part	intervention	general	suppl
topic_23:	using	analysis	data	result	used	number	one	different	method	based	information	obtained	size	standard	set
topic_24:	level	sample	total	performed	human	positive	blood	serum	negative	normal	collected	marker	protocol	plasma	detect

Таблица тем Уровень 1

topic_0:	cancer	tumor	survival	breast	metastasis	carcinoma	chemotherapy	prostate	invasion	egfr	gastric	biopsy	er	colon
topic_1:	patient	disease	clinical	diagnosis	acute	chronic	conjoint	characteristic	severe	severity	diagnosed	failure	pulmonary	median
topic_2:	time	value	mean	change	first	respectively	point	function	parameter	pattern	phase	range	therefore	second
topic_3:	two	type	one	however	first	present	well	table	small	all	et	even	since	found
topic_4:	task	subject	performance	participat	condition	visual	training	trial	experiment	memory	learning	target	stimulus	session
topic_5:	region	mutation	snp	genetic	genotype	allele	association	variant	focus	chromosome	polymorphism	frequency	haplotype	phenotype
topic_6:	intervention	information	quality	participat	practice	review	trial	decision	clinical	item	survey	survey	knowledge	guideline
topic_7:	receptor	signaling	differentiation	proliferat	migrat	endothelial	epithelial	tgf	adhesion	fibroblast	actin	veg	extracellular	primary
topic_8:	brain	neuron	channel	stimulus	signal	functional	functional	ad	cortex	region	pd	neural	nerve	amplitude
topic_9:	parameter	model	simulation	state	dynamic	probability	fig	input	r	constant	equation	space	scenario	property
topic_10:	child	woman	hospital	mortality	death	malaria	birth	pregnancy	maternal	delivery	respiratory	infant	admission	parent
topic_11:	within	image	area	volume	imaging	left	lesion	surface	cm	right	white	measurement	field	
topic_12:	change	increased	significantly	total	increase	condition	day	growth	bss	decreased	reduced	relative	significant	early
topic_13:	cd4	cytokine	immune	cd8	cd8	lymphocyte	dc	ifit	donor	spleen	antigen	subset	recipient	nk
topic_14:	expression	gene	ma	pathway	mma	transcription	target	promoter	involved	expressed	regulation	development	transcript	probe
topic_15:	infection	virus	infected	antibody	viral	response	vaccine	parasite	influenza	vaccinat	antigen	host	igg	hiv1
topic_16:	research	question	paper	issue	medicine	member	researcher	community	work	respondent	science	user	knowledge	people
topic_17:	activation	induced	incubated	treated	culture	macrophage	hour	inhibitor	play	vivo	medium	vitro	inhibition	staining
topic_18:	health	care	program	medical	service	facility	public	physician	family	medicine	access	worker	healthcare	home
topic_19:	mouse	tissue	lung	liver	animal	inflammatory	inflammation	section	organ	per	fibrosis	kidney	collagen	neutrophil
topic_20:	bone	eye	layer	surface	segment	slide	head	section	thickness	posterior	retinal	lateral	anterior	joint
topic_21:	test	case	clinical	positive	disease	laboratory	diagnostic	symptom	normal	reported	performed	negative	based	n
topic_22:	analysis	using	number	result	obtained	size	one	calculated	based	distribution	method	le	single	new
topic_23:	gene	identified	family	genome	sequence	read	candidate	set	database	predicted	member	position	cluster	novel
topic_24:	region	population	site	within	individual	size	pattern	distribution	variation	number	location	area	density	distance
topic_25:	study	risk	association	subject	woman	sex	age	men	cohort	factor	smoking	population	associated	adjusted
topic_26:	specie	plant	water	tree	environmental	leaf	site	soil	abundance	community	fish	temperature	root	mosquito
topic_27:	male	female	weight	body	muscle	glucose	diabetes	insulin	food	diet	bmi	mass	metabolic	adult
topic_28:	animal	rat	injection	injury	intestinal	group	saline	plg	injected	vehicle	sem	sham	experimental	fed
topic_29:	group	control	effect	difference	significant	test	response	condition	subject	four	presented	greater	whether	ability
topic_30:	binding	structure	interaction	residue	complex	ligand	site	molecule	structural	energy	chain	loop	molecular	substrate
topic_31:	data	used	different	one	method	possible	based	information	standard	available	large	comparison	many	developed
topic_32:	sequence	dna	three	por	genome	primer	sequencing	bp	plasmid	fragment	clone	nucleotide	amplification	copy
topic_33:	model	two	rate	figure	shown	system	show	increase	thus	due	average	small	process	
topic_34:	network	set	method	cluster	feature	algorithm	distance	class	map	information	correlation	pair	matrix	prediction
topic_35:	concentration	mm	temperature	nm	ca	correlation	estimate	index	table	measure	regression	interval	estimated	coefficient
topic_36:	model	score	value	variable	minute	protocol	exposure	zone	min	surgery	procedure	flow	use	assessed
topic_37:	time	week	hour	fig	minute	protocol	exposure	zone	min	surgery	procedure	flow	use	assessed
topic_38:	protein	domain	site	peptide	amino	binding	subunit	motif	purified	protease	recombinant	yeast	blind	core
topic_39:	strain	resistance	bacterial	isolates	bacteria	culture	pool	antibiotic	pathogen	colony	host	resistant	aureus	antimicrob

Таблица части тем Уровень 2

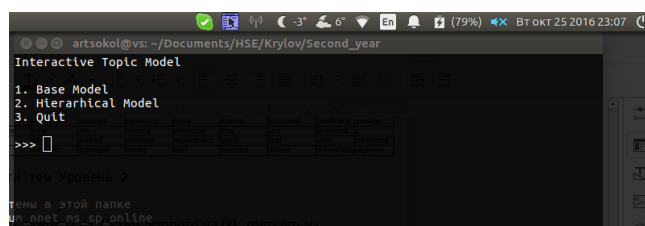
Всего документы первого уровня были разбиты на 25 тем, второго на 60, третьего на 120.

API библиотеки также позволяет получить все три матрицы (документы-темы, темы-слова, темы-подтемы), но ввиду их громоздкости здесь я их приводить не стану.

Натренированная модель была сохранена, для написания приложения по работе с ней.

Тут необходимо сказать что так как это была тестовая ветка библиотеки, существовало (и существует) множество не решенных проблем. Часть из них чинилась сообществом непосредственно во время моей работы с библиотекой и я просто ее пересобирал. Сохранение же иерархической модели не работало и чинил эту функциональность я самостоятельно, после чего пересобирал библиотеку и заново перестраивал модель.

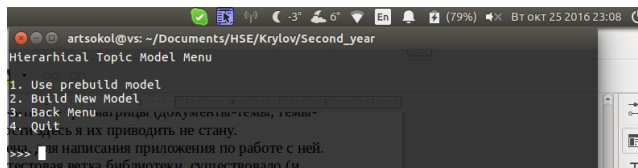
Я написал небольшое консольное приложение (с незаконченной функциональностью) на языке python и с BigARTM API идля возможности работы с моделью. Планировалось объединить наши исследования с коллегой. К сожалению, это не было реализовано до конца. Сейчас программа предоставляет небольшое меню, большая часть пунктов которого не реализованы.



Верхний уровень меню

Base model — уровни меню для работы с обычной, не иерархической моделью(не реализована).

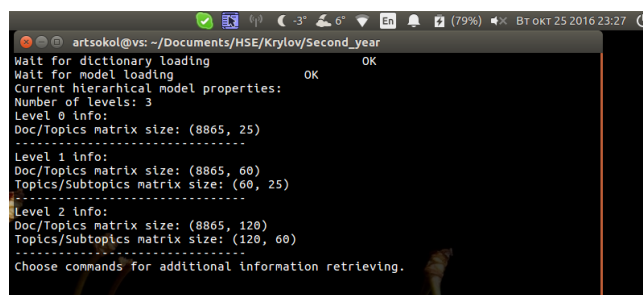
Hierarhical Model — работа с иерархической моделью.



Второй уровень меню

Use prebuild model — загрузка построенной модели с возможностью дальнейшего получения информации.

Build New Model — задание параметров для построения и сохранения новой модели.



Краткие данные о загруженной модели

Следующим этапом планировалось предоставить терминал пользователю для запросов информации, матриц, коэффициентов модели в целом и отдельных уровней.