

Исследование осложнений в нейрохирургии с помощью технологий искусственного интеллекта

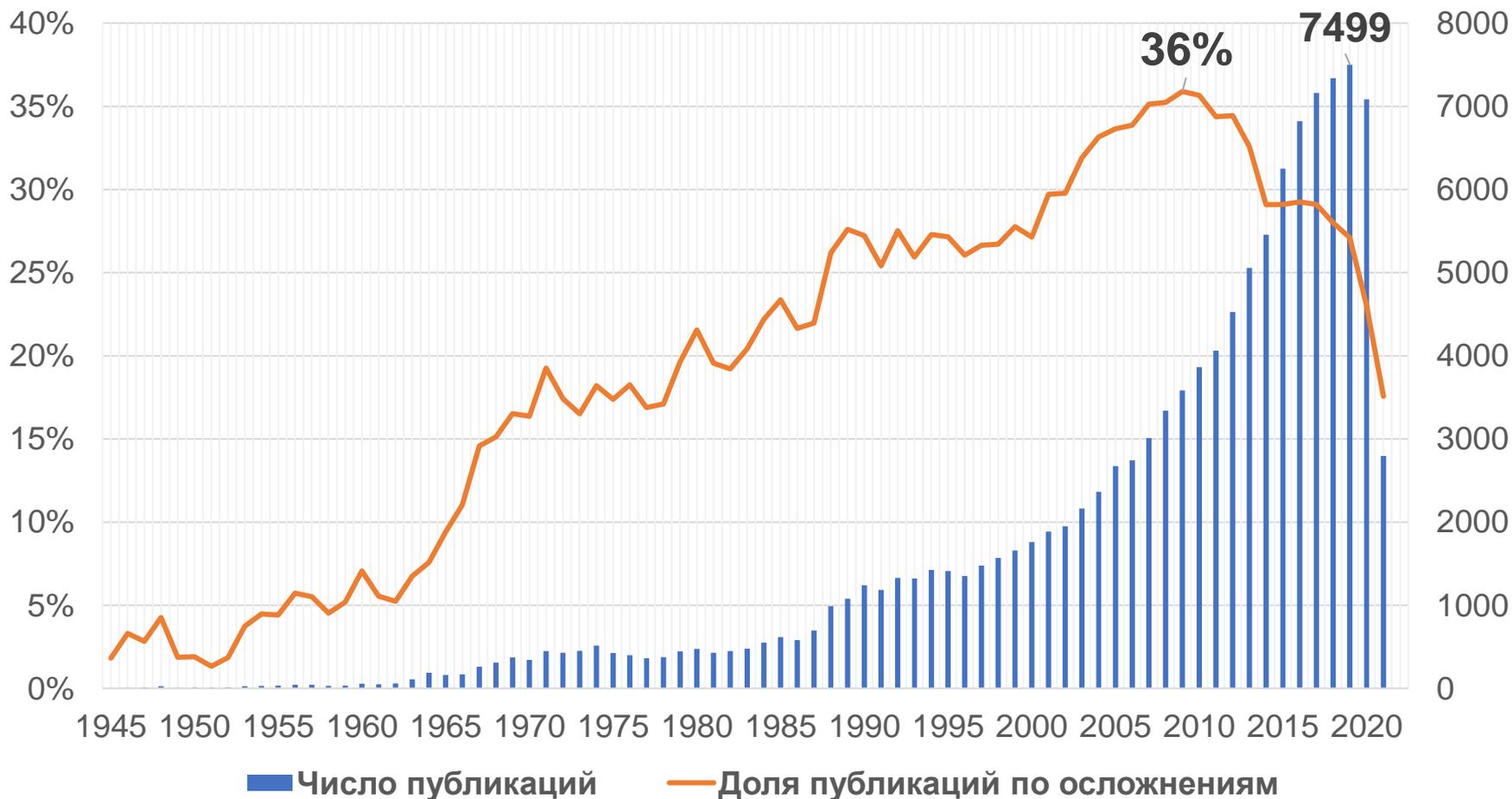
**Г.В. Данилов, А.А. Потапов, М.А. Шифрин, Т.В. Цуканова, Ю.В. Струнина,
А.В. Косырькова, М.А. Шульц, С.А. Мельченко, О.И. Шарипов,
Е.С. Макашова, Я.А. Латышев, М.Д. Варюхина, Р.А. Суфианов**

Лаборатория биомедицинской информатики и искусственного интеллекта

ФГАУ «НМИЦ нейрохирургии им. ак. Н.Н. Бурденко» Минздрава России

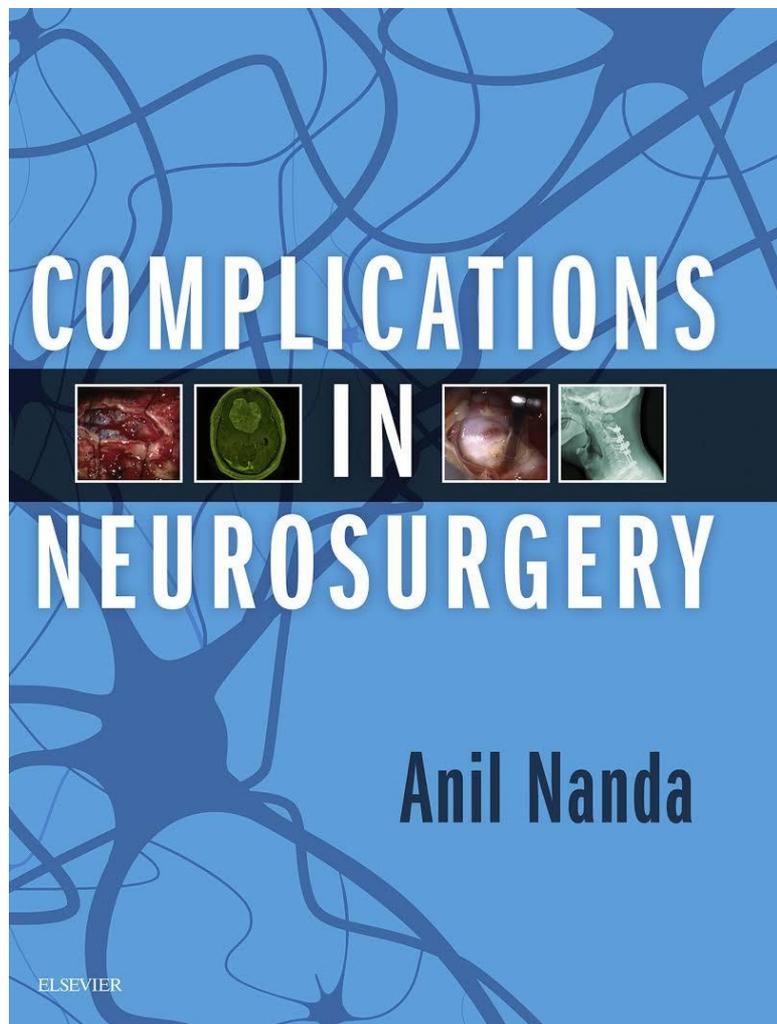
Москва, 2021

Число и доля публикаций, посвященных осложнениям в нейрохирургии по данным Национальной библиотеки медицинских институтов здоровья США (PUBMED)

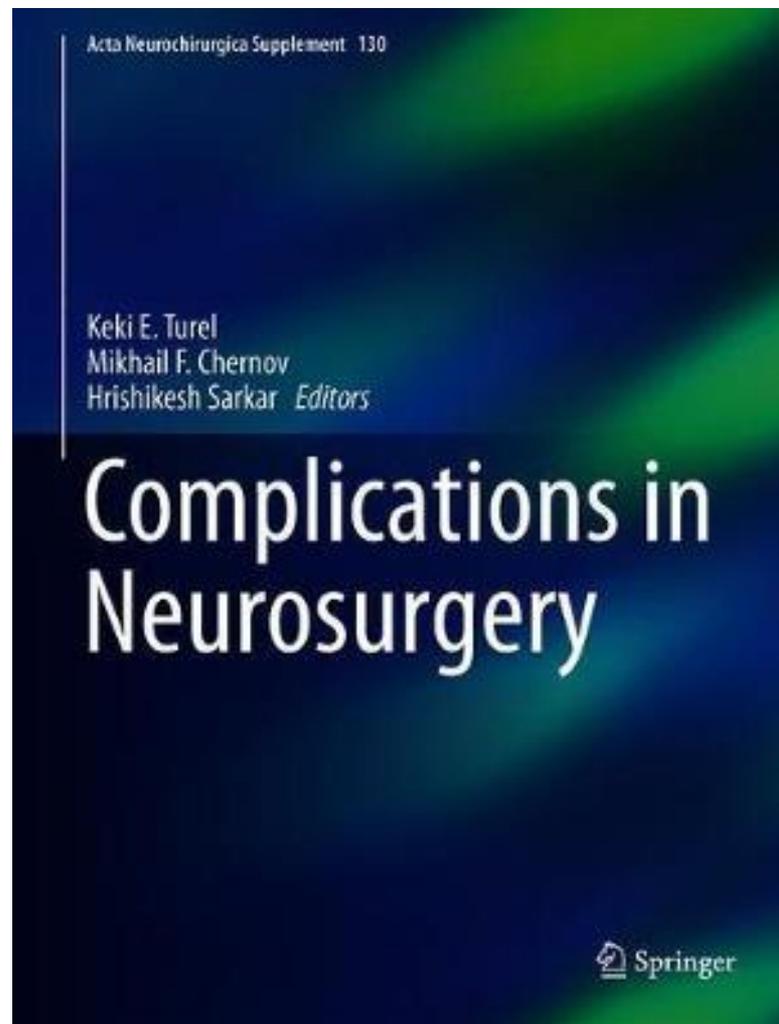


Запрос "neurosurgery AND complications", последнее обращение 15.06.2021

Книги, посвященные осложнениям в нейрохирургии



A.Nanda, 2019



K.Turel, 2020

Сложности изучения осложнений в нейрохирургии:

Отсутствие универсального, общепринятого определения «осложнения» в нейрохирургии



Отсутствие общепринятых систем классификации



Отсутствие единых подходов к учету и мониторингу



Несопоставимость отчетов по безопасности



Невозможность масштабного исследования осложнений

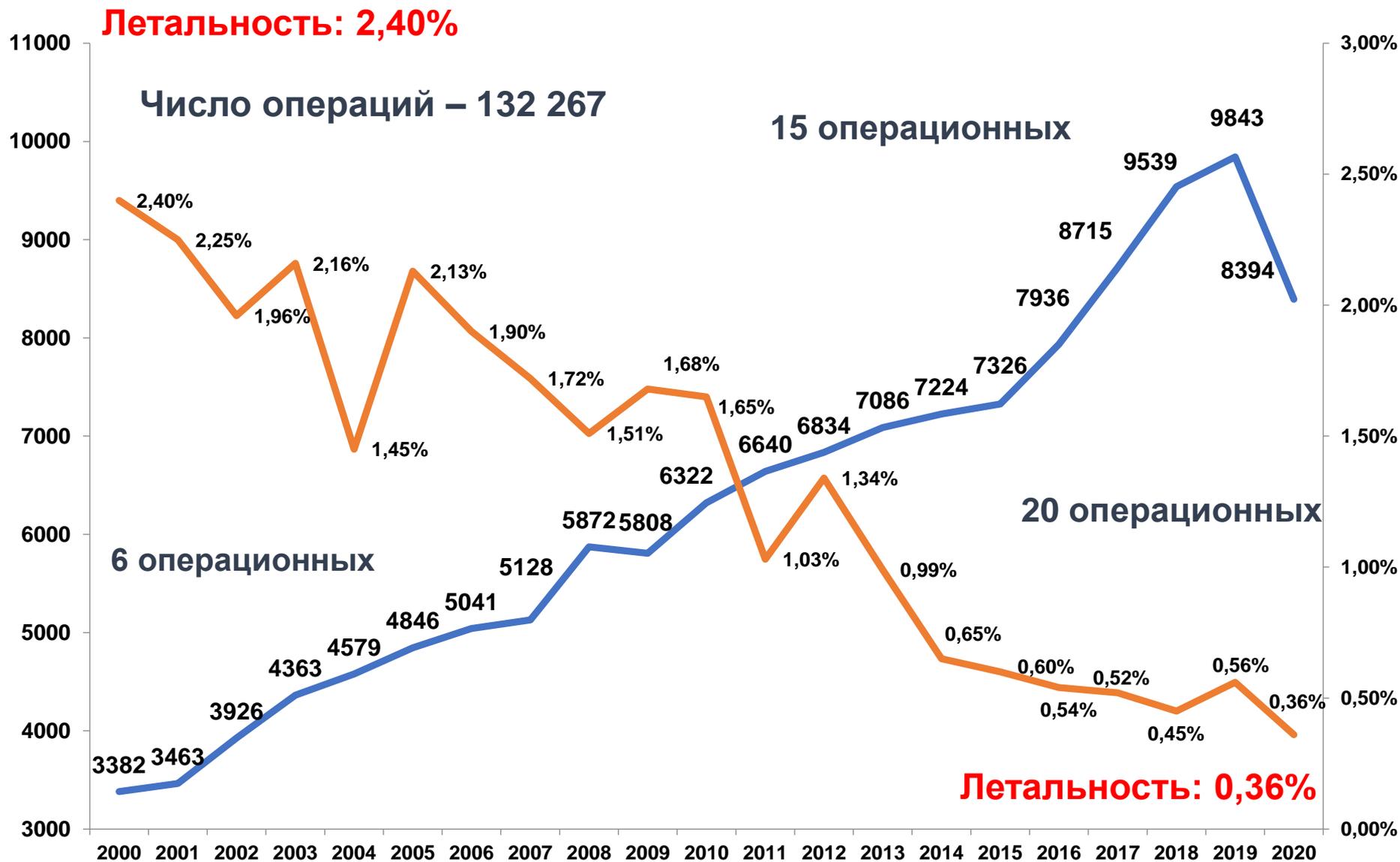
Системное исследование осложнений в нейрохирургии может быть направлено на:

1. Определение понятия «осложнение»
2. Идентификацию спектра возможных нежелательных явлений
3. Учет исходов осложненных случаев
4. Поиск факторов риска и неблагоприятного течения осложнений

Системный взгляд на осложнения важен для:

1. Оценки рисков медицинских воздействий
2. Предотвращения осложнений
3. Адекватного реагирования в ответ развитие осложнений

Динамика числа операций и послеоперационной летальности в НМИЦ нейрохирургии с 2000 г. по 2020 г.



ЭЛЕКТРОННЫЕ МЕДИЦИНСКИЕ ЗАПИСИ ФГАУ «НМИЦ нейрохирургии им. ак. Н.Н. Бурденко» Минздрава России

- Период период сбора данных: **2000 – 2017**
- Число случаев: **90 688**
- Текстовые данные извлечены из:



- **78** таблиц базы данных
- **588** текстовых полей
- **13 060 326** текстовых записей
- **229 019 413** слов

- **40 121** уникальных лемм, встречающихся чаще 5 раз



Прогноз длительности послеоперационного периода в нейрохирургии по текстам 101 654 протоколов операций

ICT for Health Science Research

A. Shabo (Shvo) et al. (Eds.)

© 2019 The European Federation for Medical Informatics (EFMI) and IOS Press.

This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).

doi:10.3233/978-1-61499-959-1-125

125

Prediction of Postoperative Hospital Stay with Deep Learning Based on 101 654 Operative Reports in Neurosurgery

Gleb DANILOV^{a,1}, Konstantin KOTIK^b, Michael SHIFRIN^a, Uliya STRUNINA^a, Tatyana PRONKINA^a and Alexander POTAPOV^a

^aNational Medical Research Center for Neurosurgery named after N.N. Burdenko, Moscow, Russian Federation

^bLomonosov Moscow State University, Moscow, Russian Federation

382

Digital Personalized Health and Medicine

L.B. Pape-Haugaard et al. (Eds.)

© 2020 European Federation for Medical Informatics (EFMI) and IOS Press.

This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).

doi:10.3233/SHTI200187

Predicting Postoperative Hospital Stay in Neurosurgery with Recurrent Neural Networks Based on Operative Reports

Gleb DANILOV^{a,1}, Konstantin KOTIK^a, Michael SHIFRIN^a, Uliya STRUNINA^a, Tatyana PRONKINA^a and Alexander POTAPOV^a

^aNational Medical Research Center for Neurosurgery named after N.N. Burdenko, Moscow, Russian Federation



**Послеоперационный период:
0-199 дней**

**Средняя ошибка прогноза (MAE):
3.0 дня**

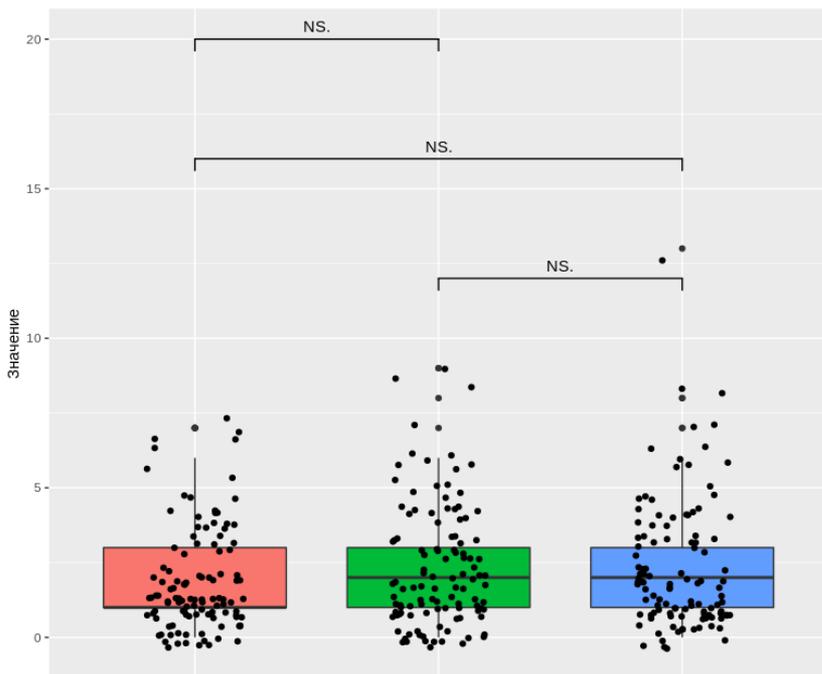
**Доля прогнозов с ошибкой
MAE 0-3 дня: 78.5%**



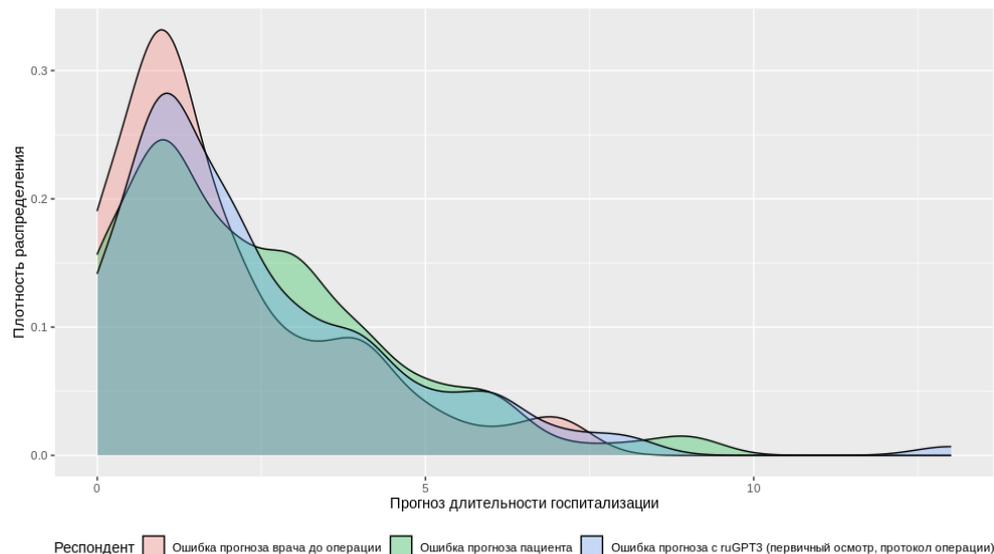
Исследование поддержано грантом РФФИ № 18-29-01052

Прогнозирование длительности госпитализации в нейрохирургии с помощью языковой модели ruGPT3

«Дообучение» ruGPT3 на 90 688 случаях



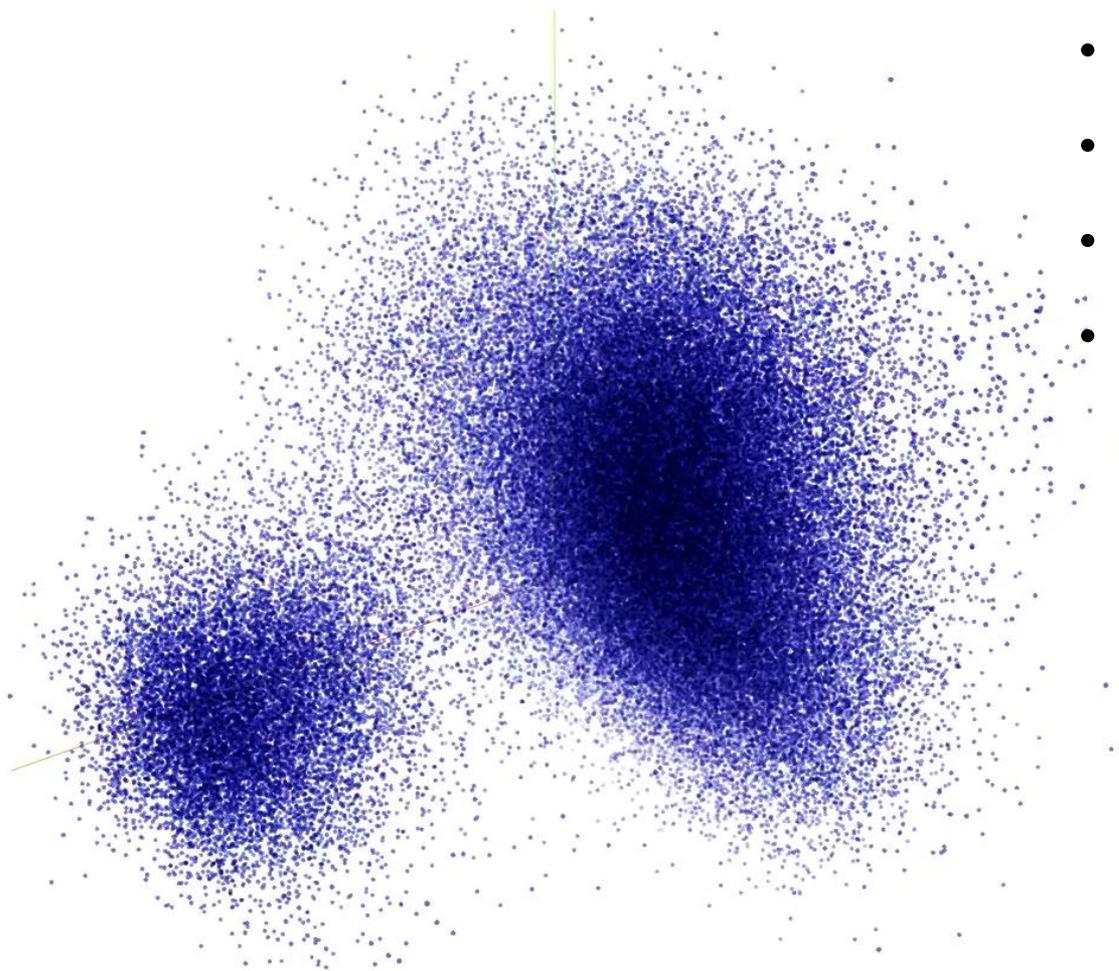
- Ошибка прогноза врача до операции
- Ошибка прогноза пациента
- Ошибка прогноза с ruGPT3 (первичный осмотр, протокол операции)



Параметр	Ошибка прогноза врача до операции	Ошибка прогноза пациента	Ошибка модели ruGPT3	p
Медиана [IQR]	1.00 [1.00,3.00]	2.00 [1.00,3.00]	2.00 [1.00,3.00]	<0.001

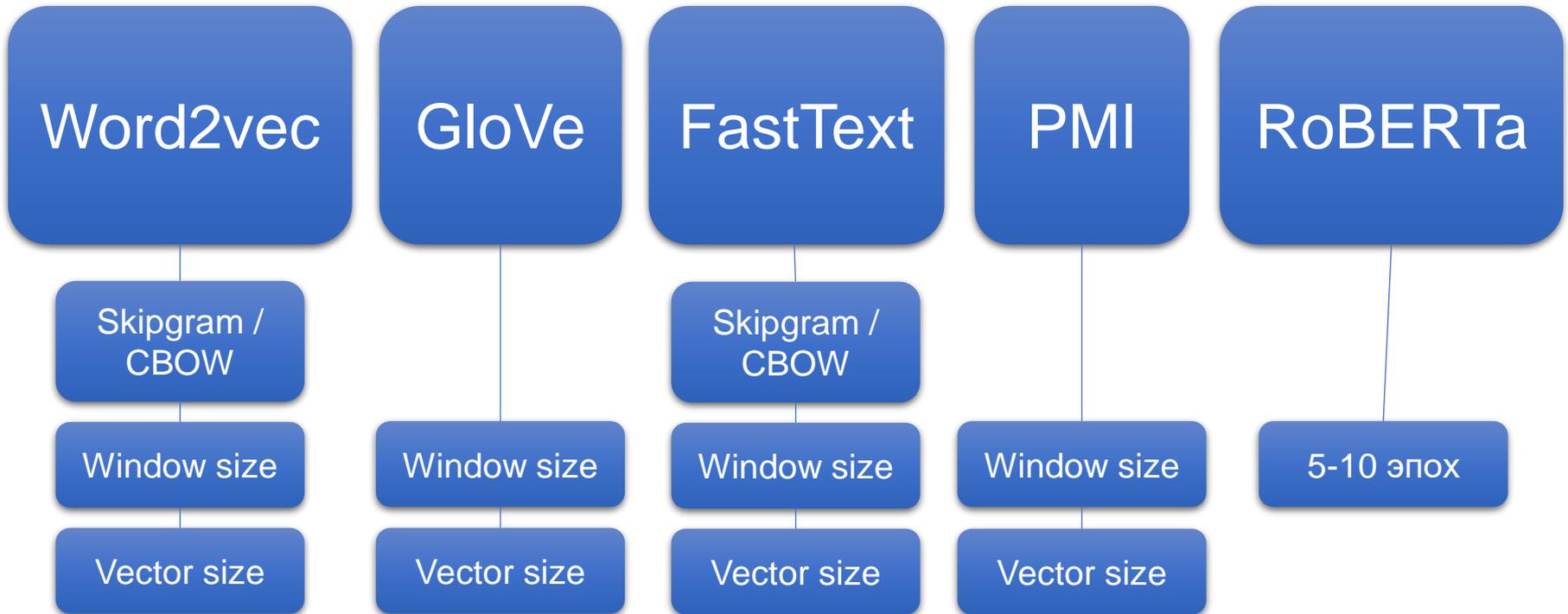


«ВСЕЛЕННАЯ» НЕЙРОХИРУРГИЧЕСКОЙ ЛЕКСИКИ



- 2000 – 2017 гг.
- 90 688 случаев
- ~13 млн. текстов
- 40 153 терминов

Сравнение векторных представлений в эксперименте с эталоном



84 вариантов векторных представлений



40 эталонных
кластеров для 258
существительных

GloVe

PMI

Word2vec

FastText

RoBERTa



СРАВНЕНИЕ РЕЗУЛЬТАТОВ КЛАСТЕРИЗАЦИИ ВЕКТОРОВ С ЭТАЛОНОМ

N = 5 853

2552 существительных	Эталонные кластеры	Форма	Система	Орган	МКБ 10	RoBERTa
		FastText	GloVe	Word2vec	PMI	
Слово 1	1	1	32	2	2	1
Слово 2	1	1	1	1	3	2
Слово 3	2	1	1	8	2	7
...
Слово 258	40	40		29	17	37

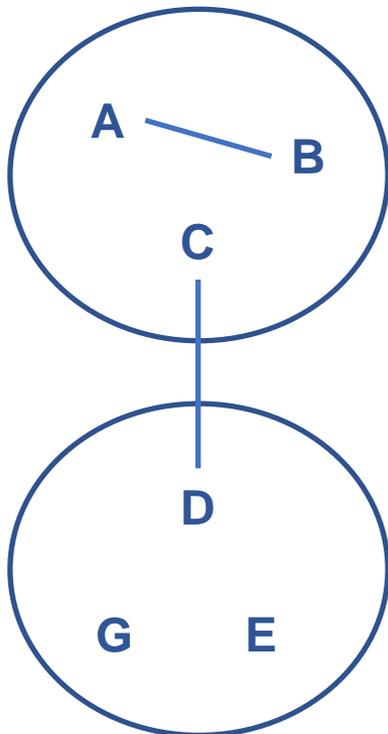
1359
прилагательных

505
глаголов

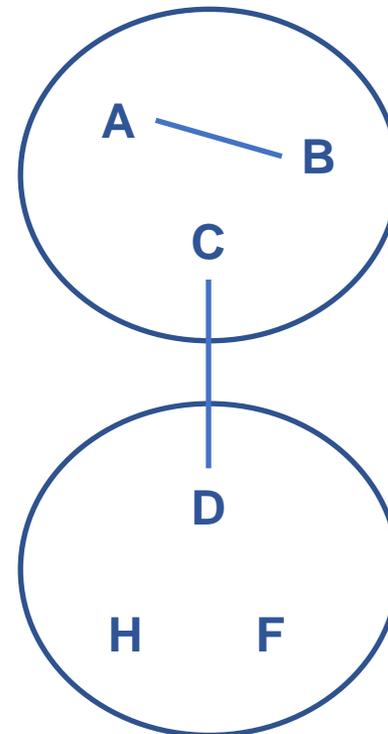
Кластеризация по методу k-средних

Сравнение результатов с помощью скорректированного индекса Рэнда

Кластеры векторов

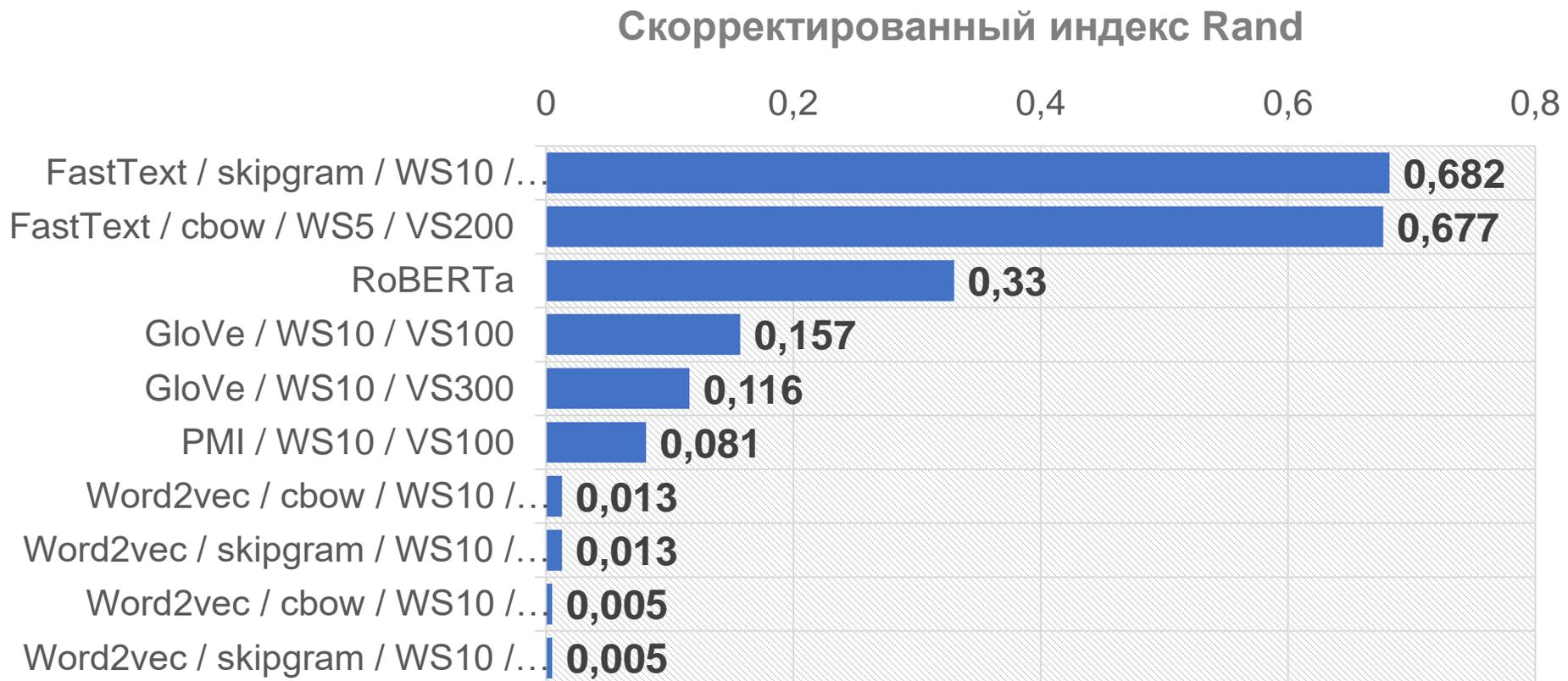


Эталонные кластеры



$k = 40$

Сравнение векторных представлений в задаче кластеризации



DATA

Points: 1000 | Dimension: 300 | Selected 7 points

Supervise ... No ignored label

Edit by Tag selection as

Load Publish Download Label

Spheroize data

Checkpoint: word2vec_skipgram_FS_WS10_VS300_co

Metadata: word2vec_skipgram_FS_WS10_VS300_lat_comp.tsv

UMAP T-SNE PCA CUSTOM

Dimension 2D 3D

Perplexity 8

Learning rate 10

Supervise 0

Stop Resume Perturb



Show All Data Isolate 7 points Clear selection

Search: ОСЛОЖН by

- менингоэнцефалит
- вентрикулит
- менингит
- менингит
- энцефалит
- вялотекущий
- нейроинфекция

BOOKMARKS (0)

ИНФЕКЦИОННЫЕ ОСЛОЖНЕНИЯ

DATA

Points: 1000 | Dimension: 300 | Selected 5 points

5 tensors found
Iris

Supervise ... No ignored label

Edit by Tag selection as

Load Publish Download Label

Spheroize data

Checkpoint: word2vec_skipgram_FS_WS10_VS300_co

Metadata: word2vec_skipgram_FS_WS10_VS300_lat_comp.tsv

UMAP T-SNE PCA CUSTOM

Dimension 2D 3D

Perplexity 8

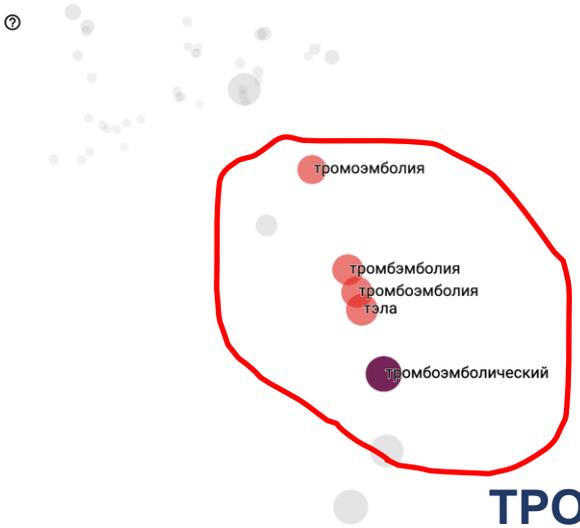
Learning rate 10

Supervise 0

Stop Resume Perturb

Iteration: 345

[How to use t-SNE effectively.](#)



Show All Data Isolate 5 points Clear selection

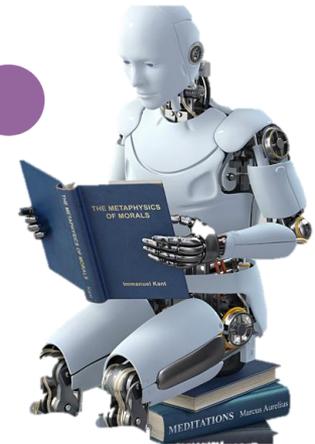
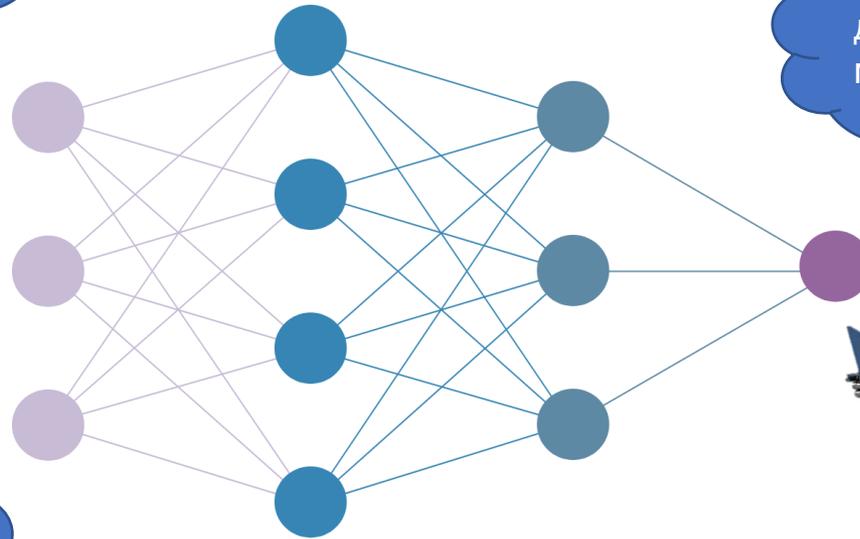
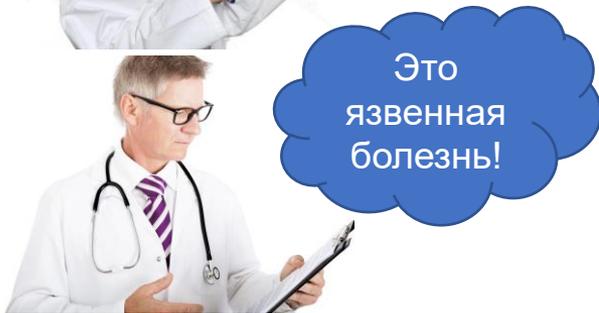
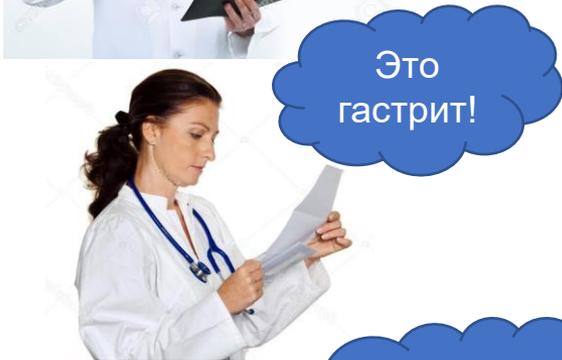
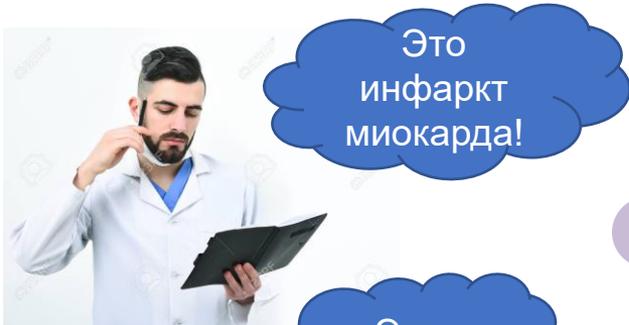
Search: ТЭЛА by

- тромбозомболия
- тромбозомболия
- тромбозомболический
- тромбозомболия
- тала

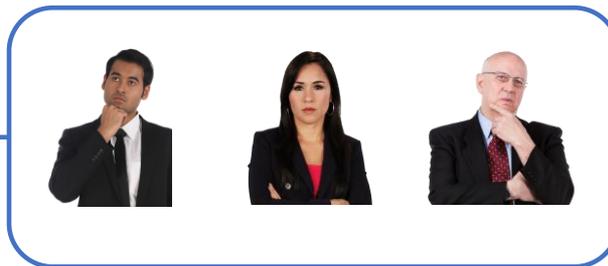
BOOKMARKS (0)

ТРОМБОЭМБОЛИЧЕСКИЕ ОСЛОЖНЕНИЯ

В КАКОЙ СТЕПЕНИ ЭКСПЕРТ, ПОДГОТОВИВШИЙ ДАННЫЕ, ВЛИЯЕТ НА РЕЗУЛЬТАТЫ МАШИННОГО ОБУЧЕНИЯ?



ИЗВЛЕЧЕНИЕ ИНФОРМАЦИИ ИЗ ТЕКСТОВ НЕСКОЛЬКИМИ СПЕЦИАЛИСТАМИ



**ПОЛНОЕ
СОГЛАСИЕ**
(n = 420)

ТЭЛА
(n = 78)

**ТЭЛА ПОД
ВОПРОСОМ**
(n = 18)

НЕТ ТЭЛА
(n = 324)

**УМЕРЕННОЕ
СОГЛАСИЕ:**
Коэффициент Light's
карра = 0.568 (p = 0)

**НЕПОЛНОЕ
СОГЛАСИЕ**
(n = 201)

Выделение информации из неструктурированных медицинских текстов

Любое отклонение от ожидаемого течения послеоперационного периода является осложнением: у пациента **появился двигательный дефицит**, что считается осложнением; **впервые развившийся эпилептический приступ** также можно считать осложнением; **выявлены повторные эпилептические приступы** – осложнения, связанные с неадекватной дозой препарата; если **диагностирована инфекция органов дыхания** – это осложнение

появился двигательный дефицит

развившийся эпилептический приступ

→ **выявлены повторные эпилептические приступы**

диагностирована инфекция раны

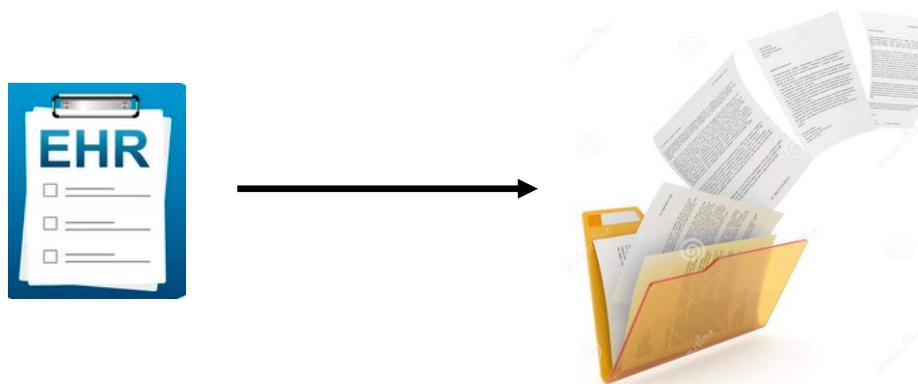


- **Вновь развившийся двигательный дефицит**
- **Впервые развившийся эпилептический приступ**
- **Повторные эпилептические приступы**
- **Инфекция области хирургического вмешательства**



ШАГ_1

> Выбор_источников_информации_из_ЭИБ



> Формирование_корпуса_текстов

which it had reported as advertising revenue. It will now book the sale of its stake in AOL Europe as a loss on the value of that stake. Dollar gains on Greenspan's speech
The dollar has hit its highest level against the euro in almost three months after the Federal Reserve head said the US trade deficit is set to stabilize.
Alan Greenspan highlighted the US government's willingness to curb spending and rising household savings as factors which may help to reduce it. In late trading
New York, the dollar reached \$1.2871 against the euro, from \$1.2974 on Thursday. Market concerns about the deficit has hit the greenback in recent months. On
Friday, Federal Reserve chairman Mr Greenspan's speech in London ahead of the meeting of G7 finance ministers sent the dollar higher after it had earlier tumbled
on the back of worse-than-expected US jobs data. "I think the chairman's taking a much more sanguine view on the current account deficit than he's taken for some
time," said Robert Sinche, head of currency strategy at Bank of America in New York. "He's taking a longer-term view, laying out a set of conditions under which
the current account deficit can improve this year and next." Worries about the deficit concerns about China do, however, remain. China's currency remains pegged to
the dollar and the US currency's sharp falls in recent months have therefore made Chinese export prices highly competitive. But calls for a shift in Beijing's
policy have fallen on deaf ears, despite recent comments in a major Chinese newspaper that the "time is ripe" for a loosening of the peg. The G7 meeting is thought
unlikely to produce any meaningful movement in Chinese policy. In the meantime, the US Federal Reserve's decision on 2 February to boost interest rates by a

ШАГ_2

> Подготовка текстов

- Приведение к нижнему регистру
- Удаление всех символов, кроме букв и пробелов

> Разбиение текста на слова

любое отклонение от
ожидаемого течения
послеоперационного
периода - осложнение



[1] любое
[2] отклонение
[3] ожидаемый
[4] течение
[5] послеоперационный
[6] осложнение

- Удаление «стоп-слов» и бессмысленных слов (букв, артефактов и т.д.)
- Исключение слов, встречающихся менее N раз (редко) в текстах

ШАГ_3

> Приведение_слов_к_начальной_форме

Уникальные словоформы		Число
1	состояние	596,807
2	состоянии	54,199
3	состояния	9,930
4	состояний	2,655
5	состояние	2,305
6	состоянию	1,310
7	состояние	1,249
...
170	юсостояние	6

ШАГ_4

> Формирование_лексики_осложнений

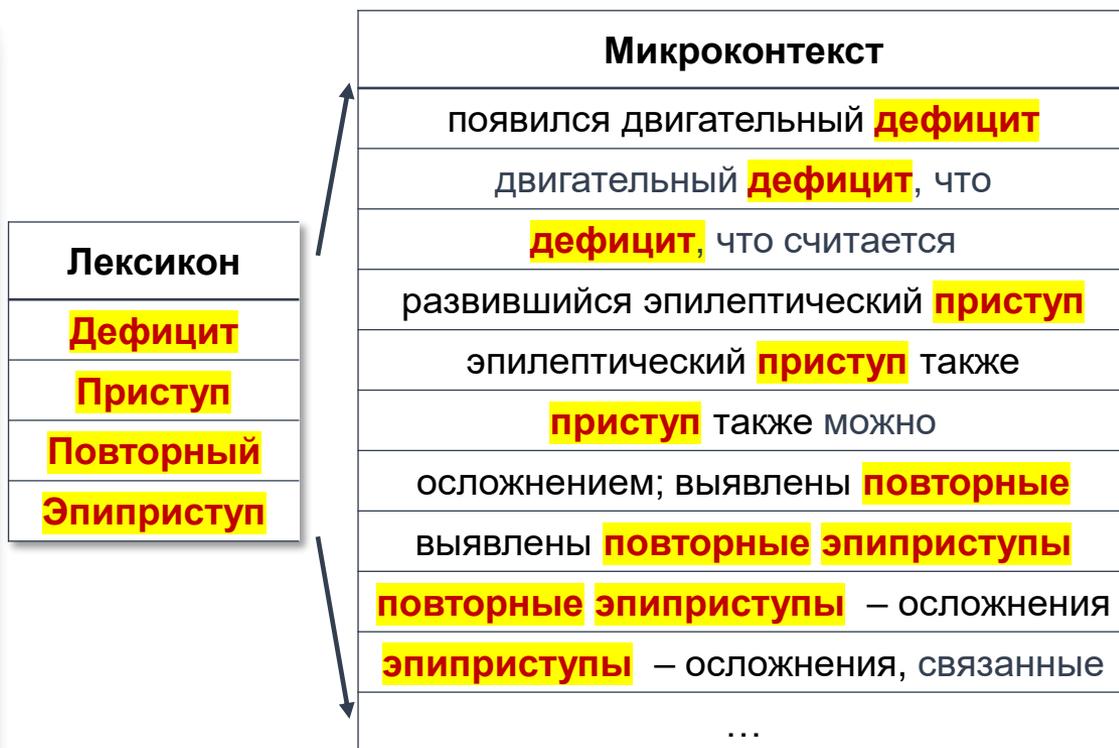
Любое отклонение от ожидаемого течения послеоперационного периода является осложнением: у пациента появился двигательный **дефицит**, что считается осложнением; впервые развившийся эпилептический **приступ** также можно считать осложнением; выявлены **повторные эпилептические приступы** – осложнения, связанные с неадекватной дозой препарата; если диагностирована инфекция органов дыхания – это осложнение

Слово	Осложнение
Дефицит	Да
Ожидаемый	
Впервые	
Инфекция	Да
Течение	
Любой	
Пациент	
Приступ	Да
Доза	
Развиться	
Препарат	
Повторный	Да

ШАГ_5

> Поиск_лексики_осложнений_в_словосочетаниях

Любое отклонение от ожидаемого течения послеоперационного периода является осложнением: у пациента появился двигательный **дефицит**, что считается осложнением; впервые развившийся эпилептический **приступ** также можно считать осложнением; выявлены **повторные эпилептические приступы** – осложнения, связанные с неадекватной дозой препарата; если диагностирована инфекция органов дыхания – это осложнение



ШАГ_6

> Разметка_словосочетаний_по_осложнениям

Микроконтекст	Произошло осложнение
появился двигательный дефицит	Да
двигательный дефицит , что	?
дефицит , что считается	?
развившийся эпилептический приступ	Да
эпилептический приступ также	?
приступ также можно	?
осложнением; выявлены повторные	?
выявлены повторные эпилептические приступы	Да
повторные эпилептические приступы – осложнения	?
эпилептические приступы – осложнения, связанные	?
...	...

ШАГ_7

> Разметка_документов_по_осложнениям

Микроконтекст
появился двигательный дефицит
развившийся эпилептический приступ
выявлены повторные эпилептические приступы



> Разметка_историй_болезни_по_осложнениям

Микроконтекст
появился двигательный дефицит
развившийся эпилептический приступ
выявлены повторные эпилептические приступы



← **! Осложнение !**

ICIMTH 2019

5-7 July 2019
Athens, Greece

17th International Conference on Informatics,
Management and Technology in Healthcare

"Health Informatics Vision: From Data via Information to Knowledge"

The Final Programme is available



194

Health Informatics Vision: From Data via Information to Knowledge
J. Mantas et al. (Eds.)
IOS Press, 2019

© 2019 The authors and IOS Press. All rights reserved.
doi:10.3233/SHTI190051

An Information Extraction Algorithm for Detecting Adverse Events in Neurosurgery Using Documents Written in a Natural Rich-in-Morphology Language

Gleb DANILOV^{a,1}, Michael SHIFRIN^a, Uliya STRUNINA^a, Tatyana PRONKINA^a
and Alexander POTAPOV^a

^aNational Medical Research Center for Neurosurgery named after N.N. Burdenko,
Moscow, Russian Federation

Abstract. Rich-in-morphology language, such as Russian, present a challenge for extraction of professional medical information. In this paper, we report on our solution to identify adverse events (complications) in neurosurgery based on natural language processing and professional medical judgment. The algorithm we proposed is easily implemented and feasible in a broad spectrum of clinical studies.

Keywords. Electronic Health Records, Neurosurgery, Natural Language Processing, Adverse Events

Извлечение информации из неструктурированных медицинских текстов

Патент RU 2751993 C1

НЕЙРОЛЕКСИКА: приложение для работы с терминами | Формирование лексикона | Разметка словосочетаний

Настройки
Время показа фразы, сек: 10
 Листать автоматически

Приложение готово к работе

№: 924 | Фраза: больной признаки тела мелкий ветвей | Заключение о нежелательном явлении: NA

Фраза указывает на нежелательное явление?
 Однзначно Не исключено Нет | Выполнено 28% | Встречается 1 раз

Загрузка файла:

Список фраз загружен
Всего фраз: 3331

Экспорт результата:
Формат файлов: ".xlsx"

Результат отбора:
Показать: 10 записей | Поиск: | Оценка

Фраза	Оценка
800 v рекомендуется тромб асс мг	0
192 th позвонков пва эмболами от	0
799 г признаки тромбоземболи мелкий ветвей	0.5
798 орбитали прямой эмболом не определяется	0
797 mthg заподозрена тала большая экстренно	1
350 mch mchc тромбо щипы mrv	0
796 l слева пва эмболами микрон	0
795 l позвонка эмболами пва проведена	0
794 ll желудка тромбом принято решение	0
793 ll желудка тромбом необходимо проведения	0

Записи с 1 до 10 из 923 записей
Предыдущая 1 2 3 4 5 ... 93 Следующая

Источники фразы:

Исследования поддержаны грантом РФФИ 18-29-22085



Способы полуавтоматического извлечение информации из неструктурированных медицинских текстов



2020

The Importance of Health Informatics in Public Health during a Pandemic
J. Mantas et al. (Eds.)
© 2020 The authors and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/SHIT200492

55

Semiautomated Approach for Muscle Weakness Detection in Clinical Texts

Gleb DANILOV^{a,1}, Michael SHIFRIN^a, Yuliya STRUNINA^a,
Konstantin KOTIK^a, Tatyana TSUKANOVA^a, Tatiana PRONKINA^a,
Timur ISHANKULOV^a, Elizaveta MAKASHOVA^a, Alexandra KOSYRKOVA^a,
Semen MELCHENKO^a, Timur ZAGIDULLIN^a and Alexander POTAPOV^a
^aLaboratory of Biomedical Informatics and Artificial Intelligence,
National Medical Research Center for Neurosurgery named after N.N. Burdenko,
Moscow, Russian Federation

Digital Personalized Health and Medicine
L.B. Pope-Hauggaard et al. (Eds.)

© 2020 European Federation for Medical Informatics (EFMI) and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/SHIT200143

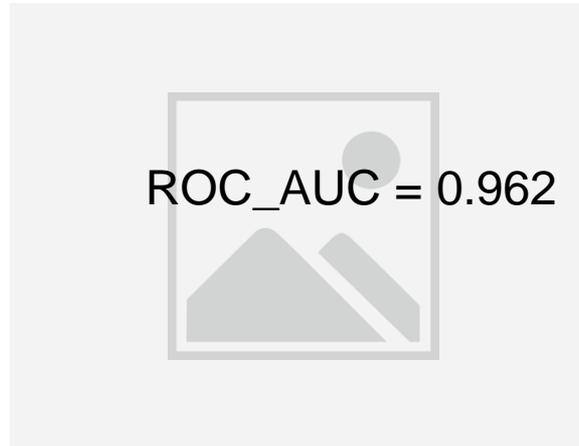
163

Detection of Muscle Weakness in Medical Texts Using Natural Language Processing

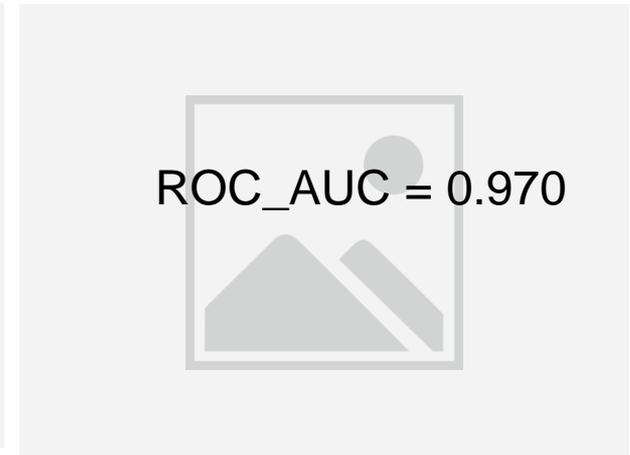
Gleb DANILOV^{a,1}, Michael SHIFRIN^a, Yuliya STRUNINA^a, Konstantin KOTIK^a,
Tatyana TSUKANOVA^a, Tatiana PRONKINA^a,
Timur ISHANKULOV^a, Elizaveta MAKASHOVA^a, Alexandra KOSYRKOVA^a and
Alexander POTAPOV^a

^aLaboratory of Biomedical Informatics and Artificial Intelligence, National
Medical Research Center for Neurosurgery named after N.N. Burdenko,
Moscow, Russian Federation

¹I.M. Sechenov First Moscow State Medical University, Moscow, Russian Federation



IEA2



IEA2 (nouns)



Cut-off by word occurrence

IEA1 (nouns)



Cut-off by word occurrence

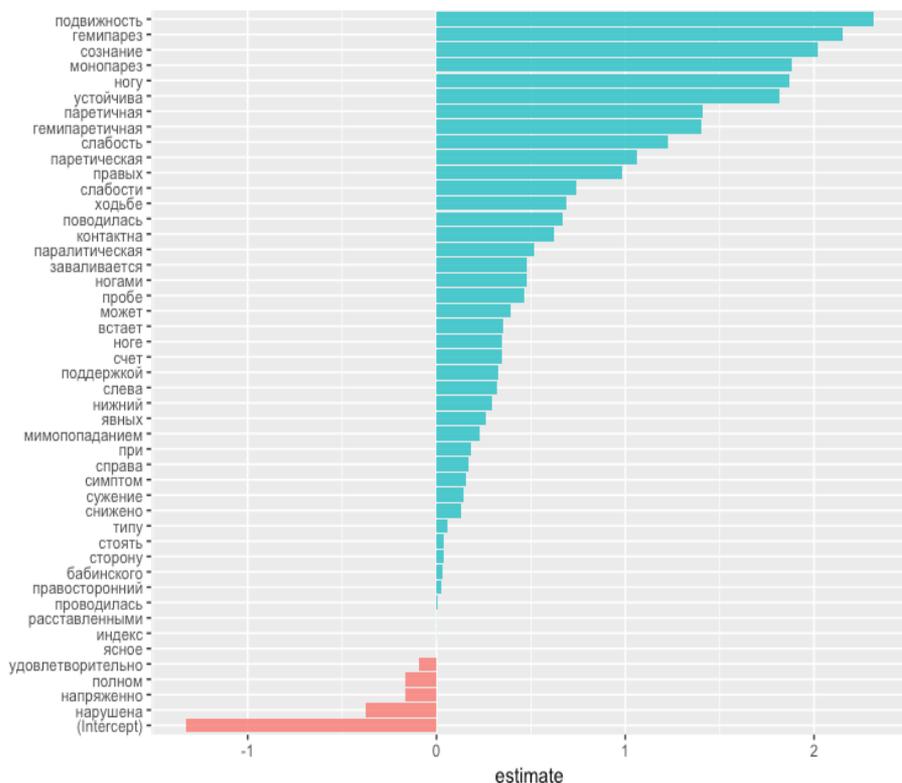
IEA1 (nouns)





ИДЕНТИФИКАЦИЯ ПАРЕЗОВ ПО ТЕКСТАМ ИСТОРИИ БОЛЕЗНИ

Сочетание машинного обучения и эксперт-зависимой модели (n = 429)



Слова, определяющие
вероятность идентификации
пареза

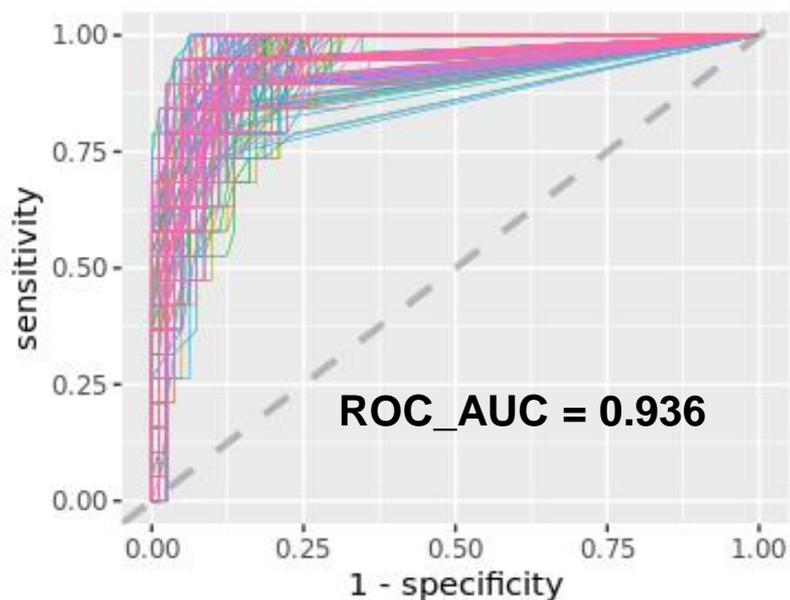


ROC (кривая качества
идентификации)

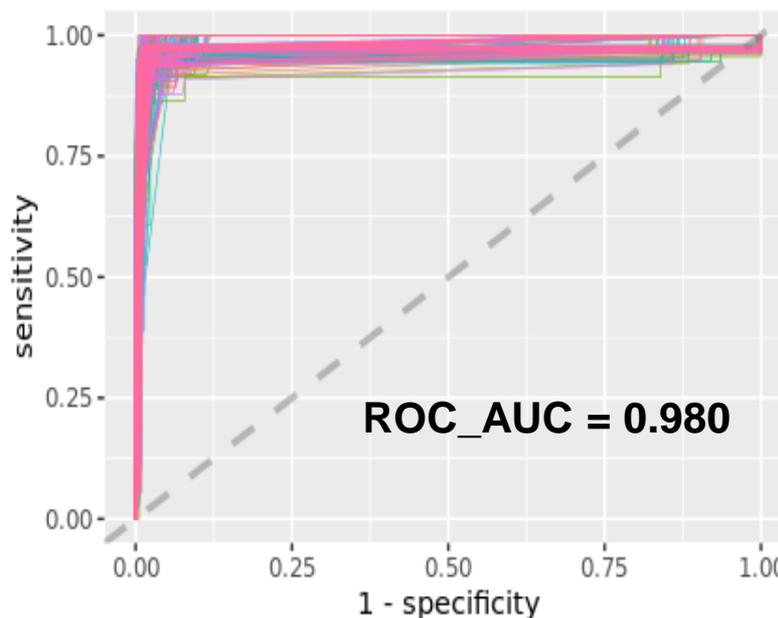
ИДЕНТИФИКАЦИЯ ТЭЛА ПО ТЕКСТАМ ИСТОРИИ БОЛЕЗНИ (n = 621)



Уровень согласия в разметке текстов 3 экспертов
(Light's kappa) = 0,568



ROC-кривые для моделей
классификации больших текстов
(n = 402)



ROC-кривые для моделей
классификации предложений
(n = 1565)

Словарь нейрохирургической лексики



НЕЙРОЛЕКСИКА: приложение для разметки терминов Разметка слов для словаря

Настройки
Время показа слова, сек
1 2 3 4 5 6 7 8 9 10
 Листать автоматически

Приложение готово к работе

№ Слово Лемма Исправлено пользователем

1176 **аллегия** **аллегия** **аллергия**

Применить ко всем ФИО Исключить

Как слово Как лемма Старт Стоп **Встречается 39 раз**

[Искать в Google](#) [Искать в Yandex](#)

Начальная форма слова (предложена алгоритмом):

аллергия

Debugging panel

	Уникальные словоформы	Число	Понятие
1	состояние	596,807	состояние
2	состоянии	54,199	состояние
3	состояния	9,930	состояние
4	состояний	2,655	состояние
5	состояние	2,305	состояние
6	состоянию	1,310	состояние
7	состояние	1,249	состояние
8	состоянеи	1,020	состояние
9	состояние	973	состояние
10	состоянием	874	состояние
...
170	юсостояние	6	состояние

ОБЛАСТЬ ПРИМЕНЕНИЯ



Исследование поддержано грантом РФФИ №18-29-22085



ЗАКЛЮЧЕНИЕ

1. Методы искусственного интеллекта **полипотентны** в задачах **изучения** осложнений в нейрохирургии
2. Методы искусственного интеллекта **перспективны** в задачах **автоматизированного мониторинга** осложнений в нейрохирургии
3. Методы искусственного интеллекта **эффективны** при **подготовке данных** к машинному обучению

БЛАГОДАРНОСТЬ КОЛЛЕГАМ

А.А. Потапов

М.А. Шифрин

Т.В. Цуканова

Т.Е. Пронкина

Ю.В. Струнина

К.В. Котик

Т.А. Ишанкулов

Ю.Н. Орлов

Е.А. Ильюшин

О.Н. Ершова

А.Г. Назаренко

М.А. Шульц

О.И. Шарипов

Е.С. Макашова

М.А. Варюхина

А.В. Косырькова

Я.А. Латышев

Н.В. Ласунин

А.И. Буров

С.А. Мельченко

С.А. Горяйнов

Р.А. Суфианов

А.Е. Быканов

Е.В. Шевченко

