

Laboratory work #4

Report

The error, which was described previously in the what's up chat was caused because firstly I tried to contain all tensors in pytorch tensors format.

It was solved by converting them to numpy array format.

```
array_of_tensors.append((predictions[0][0][i].detach().numpy()))
```

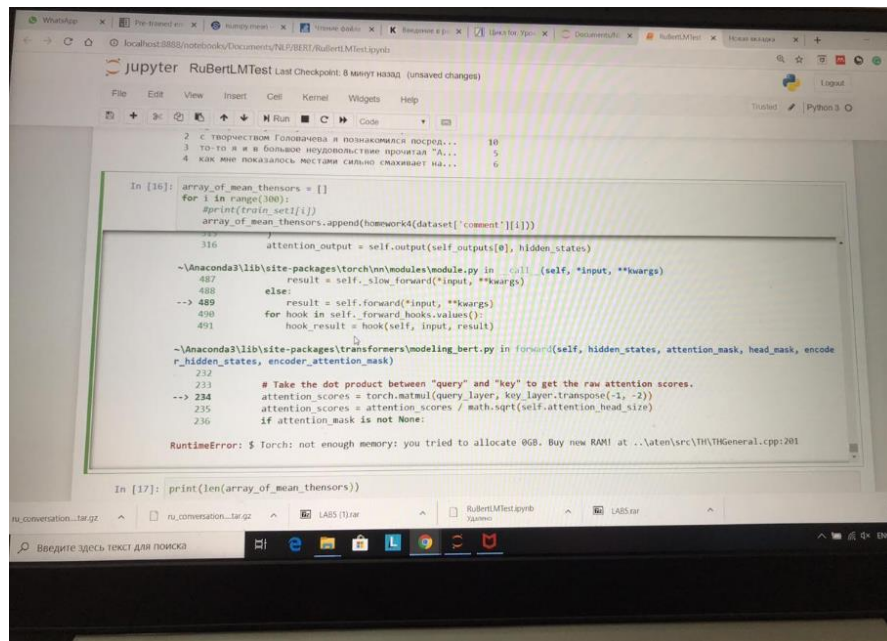


Figure 1 – Torch memory error

For comparing result I created 3 versions.

1. The final version is "lab4_Bert". The means are calculated as a means of tensors of all words by axis = 0.
2. Report is given in the notebook and in the end of the file (Figure 5)

Random forest

Accuracy score 67% from MAE and 55% from MSE. The Classification Report is given in the notebook.

Because of the limitization of computer power of the laptop I can implement only one type of classification algorithm on all rows. When I tried to implement on more on the same notebook a catch the Memory error (Figure 1).

```
MemoryError                                Traceback (most recent call last)
<ipython-input-34-3367162b2fad> in <module>
----> 1 clf.fit(X_train,y_train)

~\Anaconda3\lib\site-packages\sklearn\tree\tree.py in fit(self, X, y, sample_weight, check_input, X_idx_sorted)
    814         sample_weight=sample_weight,
    815         check_input=check_input,
--> 816         X_idx_sorted=X_idx_sorted)
    817     return self
    818

~\Anaconda3\lib\site-packages\sklearn\tree\tree.py in fit(self, X, y, sample_weight, check_input, X_idx_sorted)
    128     random_state = check_random_state(self.random_state)
    129     if check_input:
--> 130         X = check_array(X, dtype=DTYPE, accept_sparse="csc")
    131         y = check_array(y, ensure_2d=False, dtype=None)
    132         if issparse(X):

~\Anaconda3\lib\site-packages\sklearn\utils\validation.py in check_array(array, accept_sparse, accept_large_sparse, dtype, order,
r, copy, force_all_finite, ensure_2d, allow_nd, ensure_min_samples, ensure_min_features, warn_on_dtype, estimator)
    494     try:
    495         warnings.simplefilter('error', ComplexWarning)
--> 496         array = np.asarray(array, dtype=dtype, order=order)
    497     except ComplexWarning:
    498         raise ValueError("Complex data not supported\n")

~\Anaconda3\lib\site-packages\numpy\core\_asarray.py in asarray(a, dtype, order)
```

Figure 2 – Memory error

The file “result_of_testing.txt” contains the prediction data. Some examples of correct and incorrect result of the model are given below.

According the accuracy matrix (Figure 3) we can see that system gives better results with positive comments (7>) and gives better marks than was given in classification.

4	3	2	4	8	6	8	14	20	7
0	1	3	5	3	5	7	6	6	13
4	3	4	4	9	6	12	11	25	27
8	3	10	8	17	9	16	33	35	31
4	6	6	15	31	20	44	45	56	49
1	2	7	10	17	14	17	34	45	39
6	6	12	10	30	17	38	73	71	81
19	14	15	20	35	40	58	112	150	178
14	16	27	40	45	41	91	188	264	271
14	19	21	27	42	36	90	150	308	377

Figure 3- The confusion matrix

The main reason, why I add this report to the lab work is that in some cases I am not sure in a grade, which was given in the dataset. The next few comments are the examples (from results of training):

Comment:

Книга современная, открывающая глаза на образ мышления и жизни современной молодежи. Изначально вызывает эмоции: гадко, мерзко, цинично, но в действительности – это и жизнь, это реальность и осознание написанного заставляет о многом задуматься.

y_target: 1
y_predicted: 8

As for me (agree with the model it is an example of a positive comment)

Comment:

Надо смотреть всю трилогию в целом. Однозначно.

y_target: 7
y_predicted: 4

I think, it is difficult to evaluate an emotional color in this case

Comment:

Прочитала модную и рекомендованную мне здешним роботом книжку Вишневского :) первая мысль – по мотивам моей хистори аськи можно было бы и более впечатляющее что-то написать :D понимаю, что людям, не особо знакомым с инетом, читать про вирт.любовь по кайфу... а по мне так банальщина.. те же "Мосты округа Мэдиссон" по сути своей :)

y_target: 7
y_predicted: 3

The world “банальщина” gave negative meaning (neutral at least) to this comment.

Comment:

Довольно забавно читать книгу, зная что это не первая и не последняя Единственная в жизни автора :)

y_target: 3
y_predicted: 1

I don't find meaning of this comment (Honestly, *ничего не поняла*).

Comment:

Довольно интересны книги этого автора, пока читаю

y_target: 9
y_predicted: 5

Comment:

идеальный импрессионизм, как по нотам. сложно читается, но главное достоинство этой книги – вовсе не в ней самой, не в идеях и персонажах, а в идеально выверенной форме. вульф чуть с ума не сошла, пятьсот раз его переписывая, чтобы достичь идеальной архитектоники текста – это заслуживает хотя бы уважения.

y_target: 6
y_predicted: 10

As for me words (идеальный импрессионизм, идеально выверенной форме, идеальной архитектоники текста) mostly identify positive comment.

Comment:

Высокобюджетный, пышно-костюмный и авантюрно-патриотичн ый исторический колосс. В 1709 году король Луи XIV высылает из Франции двух завязых

игроков и дуэлянтов. Один из них попадает в Россию, другой в Швецию -- во время битвы под Полтавой они, естественно, оказываются противниками. Но одними только масштабными баталиями приключения героев не ограничиваются -- будет и любовь, и интриги, и даже мистика. В ролях Дмитрий Миллер, Алексей Чадов, Александр Бухаров, Ксения Князева, Евгений Меньшов, Николай Чиндяйкин, Валерий Маликов, Андрей Сухов и Дарья Семенова.

<http://www.vremya.ru/2007/31/10/172010.html>

y_target: 6

y_predicted: 10

Comment:

Фильм не оставил равнодушным.Вышел из кинотеатра и @хуел!

y_target: 9

y_predicted: 3

Очень двусмысленно.

Comment:

История о красивой богатой женщине, которой безумно скучно жить. Ни восхищенные взгляды, ни наряды, ни ее дети, ни комфортные апартаменты не могут успокоить нетерпение ее сердца. Она жаждет приключений и получает их на свою голову... она считала себя хозяйкой положения, но в результате запутывается в своей собственной паутине. Один неверный шаг - и уже никогда она не сможет вернуть назад свою прежнюю спокойную богатую жизнь...

y_target: 9

y_predicted: 5

The same as previously, there is only description of the story without any emotional color.

In the bigger cases (really dataset consist mostly from positive or neutral comments) model gives accurate result:

Comment:

Это прекраснейшее произведение русской литературы. Лучше не бывает!

y_target: 10

y_predicted: 10

Comment:

яркий образец уморительного восточного юмора и изобретательной азиатской мудрости. Браво!

y_target: 9

y_predicted: 9

Comment:

Скажите нам ребята Скажите откровенно Было б скучно наверно На свете Без девчат !!!

y_target: 10

y_predicted: 10

Comment:

Всем кто любит историю -очень шикарный материал!

y_target: 10

y_predicted: 10

Comment:

Очень хорошо. Не могла отойти от компа пока не дочитала. Это моя первая книга авторов - буду с удовольствием читать дальше.

y_target: 10

y_predicted: 10

Comment:

фильм совершенно потрясающий!! великолепная актерская игра, герои - умопомрачительные и очень живые, действия развиваются динамично. смелись с самого начала и до конца!!! это именно то, чего так не хватает серой осенней порой советую всем ОБЯЗАТЕЛЬНО посмотреть :))

y_target: 10

y_predicted: 10

Or gives error in one point, which from my point of view is difficult to evaluate by human too.

Comment:

очень понравился!!! Смотрели с мужем в кинотеатре. Ухохотались!!!

y_target: 9

y_predicted: 10

Comment:

А мне понравилось. Если книга хорошо читается и нет у автора ляпов - это уже много. И одна из немногих книг с отрицательным (здесь - вроде как, но всё же!) главным героем.

y_target: 8

y_predicted: 7

Comment:

Обожаю этот фильм, хоть мне давно уже не страшно его смотреть, и все знаю наизусть, все равно люблю этот фильм! Смотрела я его наверное раз 30, чем-то он меня зацепил!)

y_target: 9

y_predicted: 10

Also there is the same situation in neutral comments

Comment:

«Беовульф» (VIII в.) — англосаксонская эпопея о герое того же имени. Б. побеждает Бреку в плавании, поражает морских чудовищ, освобождает датчан от великана-людоеда Гренделя (похищавшего воинов из палат короля Хродгара) и его свирепой матери. Затем он становится королем, обороняет свой народ от дракона, но сам погибает от яда убитого им чудовища... Наряду с песнями о Нибелунгах, Роланде, Сиде и тому подобным Б. является эпическим памятником мирового значения». Это «Литературная энциклопедия», том 1, 1930 года издания.

y_target: 5

y_predicted: 4

Comment:

я, видимо, смотрела какое-то другое аниме... не понравилось совсем, правда я осилила серий 5-7. впечатление одно - СКУЧНО. о чем это все? кто все эти люди? дальше выяснять не захотелось. прочитав еще раз отзывы, подумала, что может стоит досмотреть до конца, пока правда совсем не хочется

y_target: 4

y_predicted: 5

Comment:

Творчество Стивена Кинга значит для Фрэнка Дэрабонта невероятно много. Его дебютом стала короткометражка «Женщина в комнате» /1983/, основанная на рассказе знаменитого писателя. И наивысшее признание – как зрителей, так и критиков – пришло к Фрэнку именно благодаря двум замечательным экранизациям1 прозы беллетриста, хотя и не имеющим никакого касательства к фильмам ужасов. Может показаться, что в своей восьмой (в общем и целом) постановке Дэрабонт решил наверстать упущенное и, как говорится, напустить туману – и в прямом, и в переносном смысле.

y_target: 6

y_predicted: 5

Comment:

Ничего так фильма. В смысле совсем ничего. картинка – ничего плохого. Режиссура – никаких нареканий, впрочем как и прорывов тоже. Парк в порядок привели – приятно. Выглядит парк не как в девятнадцатом веке – неприятно. Актеры (старшие в особенности) играют прилично. Истории в фильме нет совсем. Мысли – идеи, нет совсем. То что Пушкина не выпячивают с криками о величии приятно. То что воспитательный процесс в результате которого Россия получила столько волнующихся за нее наивных Граждан не показан нормально – не приятно. Короче не рекомендую. Впустую потраченное время.

y_target: 5

y_predicted: 5

And for negative samples

Comment:

Не рекомендую читать тем, кто уважает русскую словесность. Можно наткнуться на такие перлы, как: "она задумчиво воскликнула", "покосилась на него затравленными глазами" и др. А в целом – полный бред, мура и примитив. Герои постоянно ведут какие-то бессмысленные разговоры, причем сами обрывают себя, признавая их бессмысленными :). А судя по отзывам здесь – такое впечатление, что я читала совсем другого какого-то Рудазова :). Да нет. Тот же самый :)

y_target: 2

y_predicted: 3

But also, sometimes model gives absolutely wrong results.

Comment:

В первый день февраля в российский прокат выходит "Вавилон". Это лучший фильм-драма по версии "Золотого глобуса", обладатель премии за режиссуру на фестивале в Каннах, и наверняка будущий лауреат премии "Оскар" хоть в какой-нибудь из основных номинаций. Если американские киноакадемики предпочтут эту картину остальным, то проявят завидную последовательность. Год назад, в 2006-м, на "Оскаре" уже победило "Столкновение" - произведение не столь нашумевшее, но посвященное все той же теме: непримиримым конфликтам людей и наций в современном мире. Однако скромное "Столкновение", демонстрируя ссоры и перестрелки белых, черных и цветных, ограничилось Лос-Анджелесом и буднями тамошней полиции. "Вавилон" же берет шире: действие происходит и в Штатах, и в Мексике, и в Марокко, и в Японии. Название фильма (в оригинале - "Babel") отсылает к библейскому сюжету о башне, которую человечество совместными усилиями хотело выстроить до небес. Всевышнему эта самонадеянная затея не понравилась, и он позаботился о разобщении народов, заставив их разговаривать на разных языках. И, несмотря на то, что многие сегодня подучили язык международного общения - английский, американцу до сих пор не понять араба, мексиканцу - американца, а японцы, судя по сценарию картины, даже друг дружку не понимают, если принадлежат к разным поколениям.

http://www.gzt.ru/culture/2007/01/31/220_030.html

y_target: 4

y_predicted: 9

Comment:

Ремейк знаменитого одноименного фильма ужасов 1974 года (судя по всему, пройдет совсем немного времени, и все без исключения хорроры прошлых лет обзаведутся новыми версиями) -- довольно предсказуемая, но жуткая история про маньяка, который в сочельник терроризирует учениц школы медсестер -- сначала с помощью тревожных телефонных звонков, а потом и по-настоящему, с помощью всевозможных колюще-режущих предметов. Даже в свое время подобные сюжеты не выглядели особенно свежо, а уж теперь и подавно. Но определенная категория зрителей -- немногочисленная, но преданная жанру -- в любом случае на праздниках предпочтет эту поделку танцующим и поющим пингвинам. Тем более что несколько впечатляющих сцен тут все-таки имеется. В ролях Кэйти Кэссиди, Мишель Трахтенберг, Мэри Элизабет Уинстед, Лейси Чеберт и Эрон Перл. <http://www.vremya.ru/2006/235/10/168130.html>

y_target: 6

y_predicted: 10

Comment:

Еще детский классик советовал: на ходите, дети, ни за что на свете в Африку гулять. Но персонажи нового американского фильма ужасов "Первобытное зло" (режиссер Майкл Кэтлман), похоже, не были знакомы с этой здоровой рекомендацией. В картине "Первобытное зло" группа любопытных тележурналистов отправилась в Бурунди, на озеро Танганьика. Тамошние жители утверждали, что в озере водится гигантский крокодил. Он поедает всех, кто попадется ему на зуб, счет его жертвам ведется уже на сотни. И что самое скверное, среди них числится сотрудница миссии ООН. Возможно, работники прессы в поисках горячего материала о мифическом чудовище предпочли бы совершить командировку в более приятные места, скажем, на озеро Лох-Несс. Но что поделаешь -- есть конкретное задание редакции телеканала! Бурунди так Бурунди. Журналистов сопровождает зоолог, который

страстно желает изучить диковинную рептилию. http://www.gzt.ru/culture/2007/04/11/220_006.html
y_target: 4
y_predicted: 8

Comment:

Отвратительная американская комедия про то, как парень поселился в доме с ОГРОМНЫМ количеством тараканов! А они ещё говорили человеческим голосом! А как они лазили по всему дому, по главному герою! Тьфу! В одной сцене таракан залез на героя и открыл ему глаз! Ужас! А концовка - это вообще полный ужас! Тот дом сожгли, герой сошёлся с любимой девушкой и поселился у неё. И угадайте, кто ещё там поселился? Правильно, ОНИ! И они сразу стали приводить дом в "порядок". Кошмар! В общем, ужасно грязный, тупой и бессмысленный фильм! P.S. Честно говоря, очень удивлена, что многим этот фильм нравится. Смешным его называют... Ну, я тоже в одном моменте улыбнулась, но в остальном это самая настоящая гадость!

y_target: 1
y_predicted: 7

Comment:

На российские экраны выходит новый фильм про Ганнибала Лектера. Одной из наиболее естественных человеческих потребностей, безусловно, является еда. Случается, что ее выбор и добыча перестает быть делом вкуса и меряется категориями морали. Известный по нескольким книгам и фильмам доктор Ганнибал Лектер не гнушался отведать человечинки, и многие из тех, кто становился для него изысканным блюдом, не выжили. За это, как мы помним, доктор был заключен в тюрьму строгого режима и хитроумно бежал из нее в "Молчании ягнят". Новая картина "Ганнибал: восхождение" проливает свет на то, как Лектер дошел до жизни такой, иначе говоря, взошел к высокой степени безумства. Ибо поедать себе подобных среди цивилизованных людей, конечно, не принято. Даже ради выживания, когда другой пищи нет. Лента по сценарию все того же Томаса Харриса, автора всей ганнибалоканнибальской эпопеи, не преподносит сюрпризов в плане поведения героя. Оказалось, что в детстве Ганнибал, литовец по происхождению, пережил страшную психологическую травму. Когда в конце Второй мировой войны близ его родного дома развернулись бои между советской армией и фашистами, родители погибли под обстрелом, а сам мальчонка вместе с младшей сестренкой оказался пленником местного подобию "лесных братьев". В фильме группа коллаборационистов, не успевших удрать с эсэсовцами, выглядит весьма жестоко - настоящая страшилка для детского ночного кошмара. Скрывавшимся от Советов амбалам вскоре стало нечего кушать, и тогда они, как и подобает сказочным злодеям, решили подкрепиться детишками. Сестренку съели, Ганнибала спасло наступление "красных". Юный Лектер оказался в детдоме и поклялся отомстить за пережитое. Он подрос, перебрался в Париж, отыскал своих врагов и стал хладнокровно исполнять задуманное. Не забывая проявлять те зачатки эстетства и тяги к красоте, которым суждено будет раскрыться позже, когда доктор Лектер достигнет возраста Энтони Хопкинса. http://www.gzt.ru/culture/2007/02/06/220_005.html

y_target: 4
y_predicted: 7

Usually it happens, when comment is long.

And sometimes without any justifiable reasons.

Comment:

Дрянь редкостная.

y_target: 3

y_predicted: 9

Comment:

Бойся исполнения своих желаний.

y_target: 5

y_predicted: 10

Comment:

Очень понравился. Читала захлеб. Из тех его книг, что я прочла - лучшее. А рассказ про минеральную воду - это сильно.

y_target: 9

y_predicted: 6

То, что легче представить самому Лукьяненко, мне, как читателю его дозоров представить не удалось. Серо, топорно, сумрак...точнее муть, без атмосферы, имхо само собой.

y_target: 5

y_predicted: 9

Comment:

Удивительный шедевр Голливуда, тонко сочетающий в себе правду (порой жестокую) жизни и возвышенную философию красоты, любви, доброты и всепрощения. Работы режиссера, актеров, оператора и композитора просто выше всяких похвал. Мой любимый фильм! Однозначно!

y_target: 10

y_predicted: 4

Comment:

Книжка по фильму, это всегда скучно! Те же диалоги, то же описание! Вообще на сценарий похоже было!

y_target: 5

y_predicted: 10

Comment:

На одном дыхании. Никогда не знала, что Айседора, оказывается, потеряла троих детей, рассталась с несколькими возлюбленными, ее мечта о своей школе так не осуществилась - но никакие трагедии ее не сломали, и она продолжала жить, любить, мечтать и строить планы. Невероятная сила духа у женщины!

y_target: 9

y_predicted: 6

Comment:

Честно, я ожидала лучшего, так как в списке авторов стоит НОРТОН. Читается плохо.

y_target: 5
y_predicted: 9

Comment:

Не понравился язык автора, до конца так и не дочитала.

y_target: 3
y_predicted: 10

Comment:

чего все повзбесились по этому фаренгейту? книга - тупая. персонажи - не прописаны. слог - топорный. идея - столетней давности. или Бэдбери жил в прошлом веке? тогда его можно было бы понять.

y_target: 1
y_predicted: 9

Comment:

ха-ха-ха, тупое книжко для необразованных бездарей! за основу взята (да что там взята - содрана) из старой хасидской сказки, которой двести лет отроду. с первого абзаца ясен конец. принципиально прочла всю. дерьмо для сопливых малолеток. ни пищи для ума, ни приятности сердцу

y_target: 1
y_predicted: 10

Comment:

не понравилось однозначно! Не люблю перебор мистики и фантастики.

y_target: 1
y_predicted: 9

Also I tried to solve it with (mean for all tensor) => input value is not a tensor, but one number, but accuracy was smaller (62%). So I don't use this way. But I can add proofs if it will be needed.

I faced with a problem, that preview of github does not show all rows of notebook (proofs in classification report), so I saved them to a csv file (Figure 4 –github).

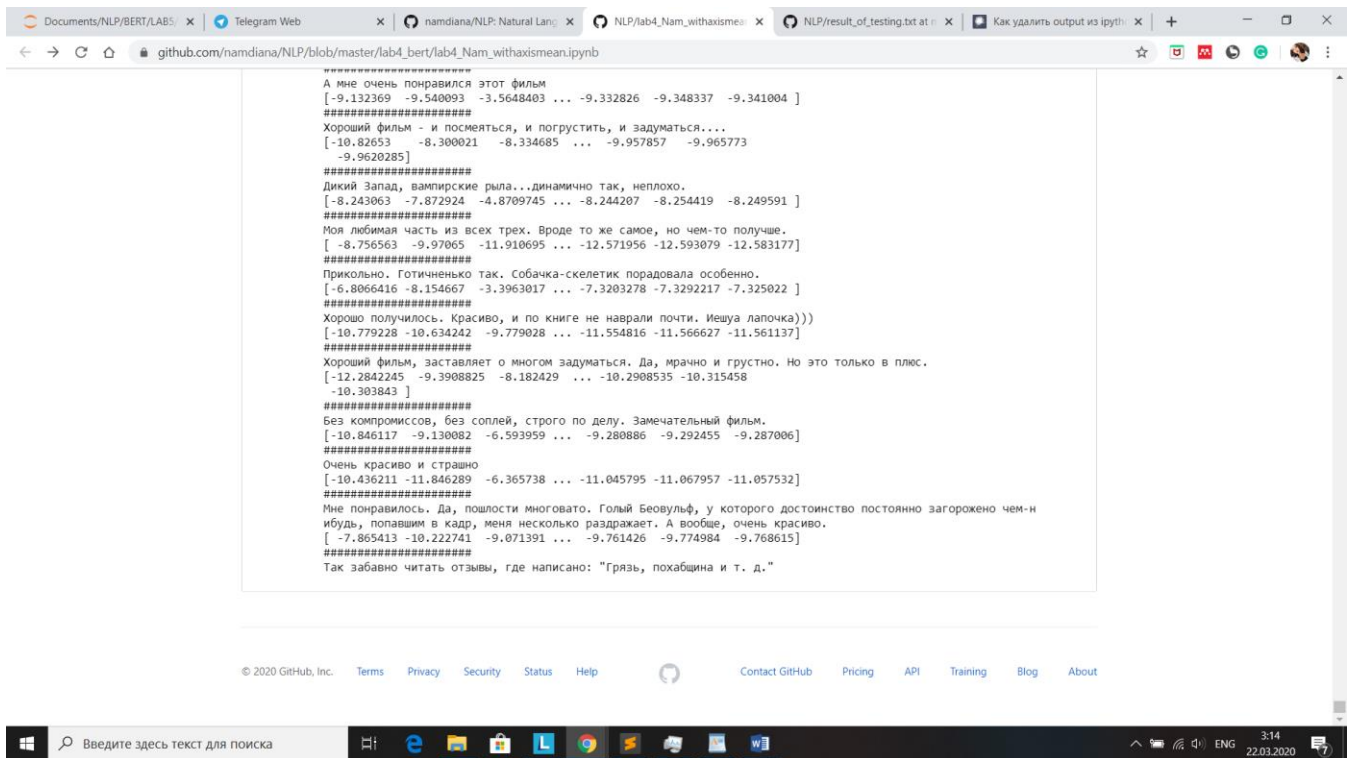
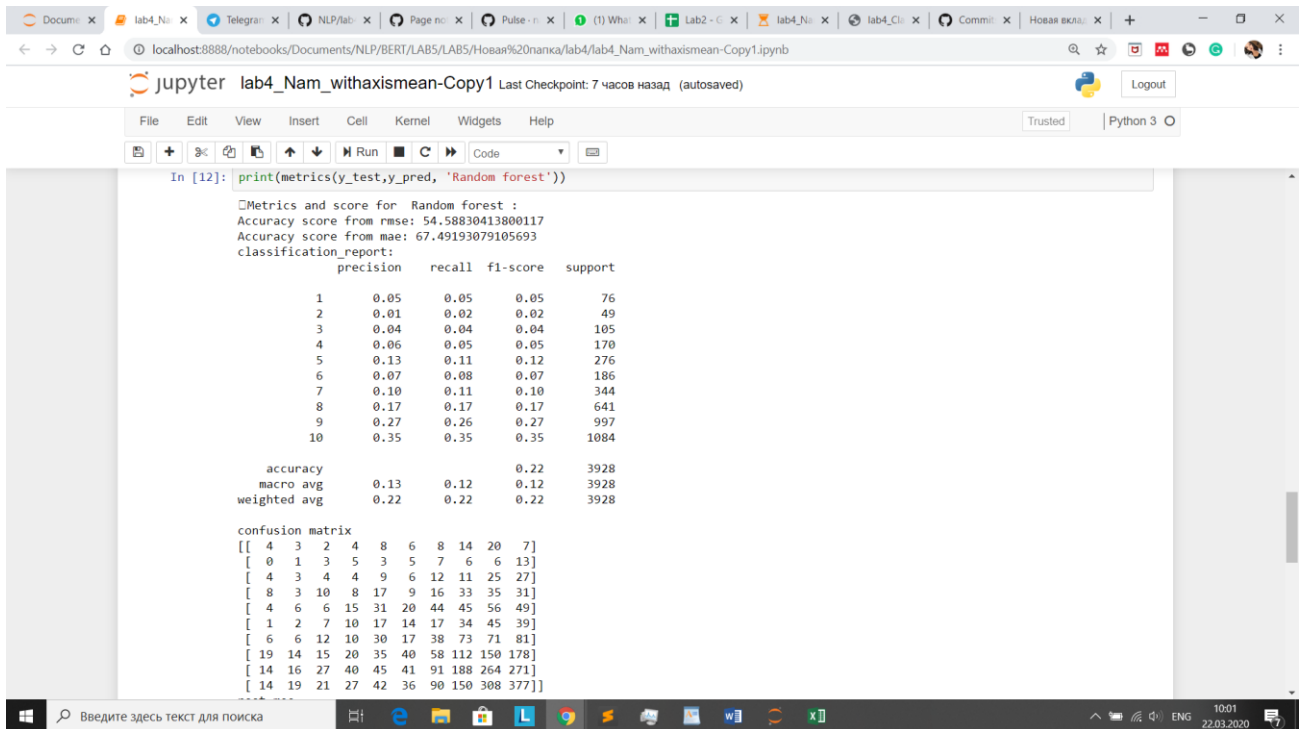


Figure4 – github.



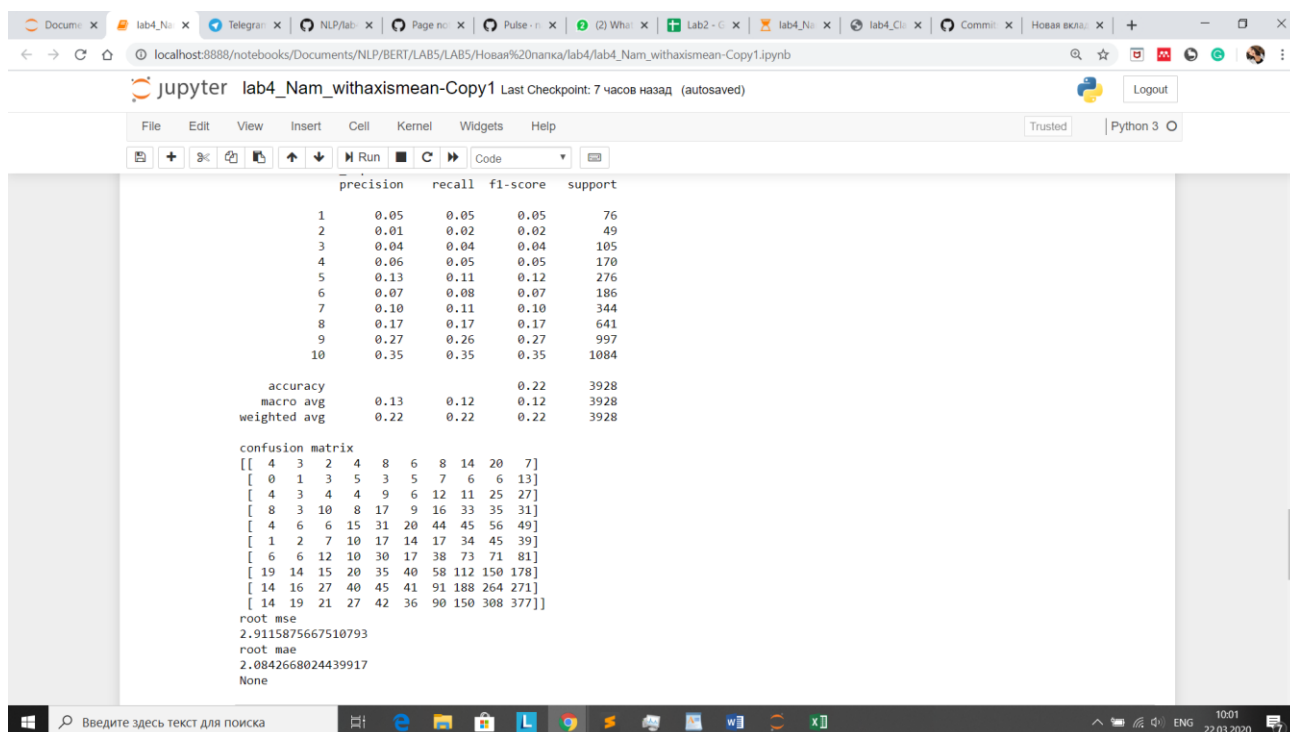


Figure 5 – Metrix