

## Разбор основных недочётов при написании курсовых / ВКР / диссертаций

для студентов Дьяконова А.Г. (dyakonov.org), версия 25.04.2022

Памятка посвящена стилю оформления, допустимым формулировкам, использованию формул, таблиц и т.п. Используются примеры из реальных курсовых и выпускных работ.

### Название, автор, научный руководитель.

Проверьте, что здесь нет опечаток. Название должно совпадать с названием в приказе о защите.

Название работы может упоминаться в официальных документах, например вкладыше к диплому, поэтому запрещается использовать нестандартные для официальных документов аббревиатуры, тем более на иностранных языках (SVM, DL, CNN и т.п.), их легко переформулировать (в тексте их использовать можно, при первом упоминании аббревиатуры расшифровываются).

Не используйте в названии фамилий ныне живущих учёных, даже научных руководителей.

плохо	хорошо
<i>Обобщение теоремы доцента ЛГУ Герасимова.</i>	<i>Обобщение теоремы Минского для многослойных трансформеров.</i>

### Структура работы

- введение – о чём работа, актуальность исследования,
- обзор предыдущих работ – кто и что делал, как это связано с настоящей работой,
- постановка задачи – должно быть кристально ясно, какая задача решается,
- предложенный метод – описание метода, отличия от предложенных ранее,
- эксперименты – проверка метода на реальных/модельных данных,
- заключение – за что давать диплом,
- список литературы – перечисляете всё что использовалось и первоисточники используемых концепций.

Часто забывают, но в научных работах очень полезно

- как был придуман предложенный подход,
- где и как применялись аналогичные подходы или идеи.

Также лучше честно описать недостатки предложенных решений и планы дальнейших исследований.

### Аннотация

Можно не делать. В аннотации, введении и заключении не должно быть повторяющихся предложений. В аннотации не должно быть ссылок на литературу. Надо кратко описать «о чём работа».

Избегайте повторений слов и фраз.

плохо	хорошо
<i>В работе рассмотрены методы построения рекомендательных систем. В данной работе решается задача ...</i>	<i>Рассмотрены методы построения рекомендательных систем. Решается задача ...</i>

## Введение

Объём от 1 страницы. Самое главное здесь – введение читателя в предмет исследования, актуальность исследования, возможен краткий обзор существующих подходов к решению, описание решаемой задачи («на пальцах»). Здесь должно быть много ссылок!

плохо	хорошо
<i>Значительных результатов в рассматриваемой области удалось добиться с помощью использования свёрточных нейронных сетей.</i>	<i>Значительных результатов в рассматриваемой области удалось добиться с помощью использования свёрточных нейронных сетей [1]-[9], [14], [17].</i>

Каждая Ваша фраза или следует из предыдущего изложения или сопровождается ссылкой – где это можно прочитать подробнее.

## Обзор литературы

В обзор помещайте только то, что имеет отношение к работе. Помните о целевой аудитории (это Ваши коллеги), не надо подробно расписывать то, что известно читателям (например, определение вектора или матрицы, не надо и ссылаться на работы, в которых они были предложены).

## Постановка задачи

Задача определяется своей ДНК:

- дано,
- найти,
- критерий.

Не забывайте использовать математический формализм при постановке задачи.

## Методы решения

Чётко разграничивайте свои и чужие методы. При описании чужих ставится ссылка на первоисточник. При описании своих – фразы типа «предложим метод, решающий описанную проблему».

## Графики и таблицы

Числа в таблицах указываются с необходимой точностью (не надо чисел типа 0.3816216), выделяйте жирным наиболее интересные значения (например, максимальную полученную точность решения).

Графики должны иметь названия, подписанные оси, контрастные цвета (различимые при печати на чёрно-белом принтере), шрифты должны соответствовать шрифтам в тексте.

Графики и таблицы – это всего лишь пояснения к тексту, поэтому «какой метод лучше» должно быть явно прописано в тексте.

К каждому графику и каждой таблице читатель должен быть хотя бы раз «отослан» в тексте.

Если используете чужие рисунки, схемы и т.п. – ссылайтесь на первоисточник! Но лучше минимизировать объём «чужого».

## Эксперименты

Описывайте с такой степенью подробности, чтобы не было дополнительных вопросов:

- какой был язык программирования, среда разработки, какие библиотеки использовались (почему?),
- как производилось обучение моделей (гиперпараметры, выбор темпа обучения и т.п.),
- сколько это всё (обучение, тестирование) занимало времени, на каких мощностях.

Если Вы что-то выбрали и зафиксировали (интервал изменения гиперпараметра, глубину деревьев и т.п.), надо объяснить, почему выбрано такое значение. Чем меньше к работе будет вопросов «почему», тем лучше.

Сравнивайтесь с существующими методами и/или естественными бенчмарками (например, прогнозом «завтра как вчера»).

## Благодарности

Не обязательно кого-то благодарить, но если Вы хотите это сделать – делайте в специальном разделе. Указывайте людей, коллективы, компании, фонды и т.п. Только Вы решаете кого и как благодарить, но не повторяйте благодарности несколько раз, также лучше не перечислять титулы и звания человека (Вы благодарите человека, а не должность).

плохо	хорошо
<i>Благодарю ректора Петрова К.В. за внимание к работе. Благодарю ректора Петрова К.В. и Вознесенского И.И. за обсуждение результатов.</i>	<i>Благодарю Петрова К.В. за внимание к работе и обсуждение результатов, Вознесенского И.И. за ряд ценных замечаний.</i>

## Заключение

Часто читают только введение и заключение, они должны быть выверены и идеальны.

Лучше в заключении написать фразу «на защиту выносятся» и описать «за что Вам надо выдать диплом». Это может быть:

- оригинальная постановка задачи,
- доказательство теоремы,
- реализация методов на ЭВМ и проведение экспериментов,
- сбор и разметка данных (с пояснением, почему это нетривиально).

Это НЕ может быть:

- знакомство с литературой,
- запуск готовых решений,
- аренда вычислительных мощностей,
- обзор чужих результатов.

### Заголовки

Заголовки должны быть краткими, но всё-таки полностью отражать содержание.

плохо	хорошо
<b>3.1. Некоторые подходы к построению</b>	<b>3.1. Некоторые подходы к построению покрывающих множеств</b>

### Термины

Пишите по-русски! Большинство объектов/предметов/явлений имеют нормальные русские названия.

плохо	хорошо
<i>фичи</i> <i>рандомный</i>	<i>признаки</i> <i>случайный</i>

Если общепринятого перевода нет, можно использовать «естественный», указав в скобках английский вариант (кстати, его всегда разумно указывать).

плохо	хорошо
<i>В работе [2] был предложен <b>BN</b>.</i>	<i>В работе [2] была предложена батч-нормализация (BatchNorm).</i>

Все аббревиатуры должны быть пояснены!

Многие термины в ML неудачны, хотя активно употребляются не очень грамотными людьми. Например, термин «метрика качества» – есть математический термин «метрика», а в том смысле, в котором он используется в ML, это не метрика, а показатель качества! Если Вы пишете для BI-аналитиков, то можно писать «метрика», если для математиков, то пишете «показатель» (в скобках может быть английский аналог).

плохо	Критика
<i>Гейты являются своеобразными <b>фильтрами</b> сигнала.</i>	В теории нейронных сетей термин «фильтр» (filter) уже, к сожалению, занят, поэтому лучше не использовать его при пояснении. Это может привести к путанице.

Лучше писать так, чтобы при удалении аббревиатур из текста, он оставался нормальным текстом на русском языке. Это решит проблему согласования.

плохо	хорошо
Классический SVM vs классическое SVM	Классический метод SVM.  А лучше написать по-русски: Классический метод опорных векторов (SVM).

## Определения и обозначения

Объекты должны вводиться до или во время первого упоминания.

Используйте общепринятые обозначения.

плохо	хорошо
Пусть $n$ – равномерно распределённая на отрезке $[d, z]$ случайная величина. Рассмотрим функцию $o(n)$ .	Пусть $\xi$ – равномерно распределённая на отрезке $[a, b]$ случайная величина. Рассмотрим функцию $f(\xi)$ .  Здесь ещё скобки стали прямыми!
Опишем определение используемой функции трансформации. Считаем, что $x > 0$ . Пусть функция трансформации $f(x) = \log(x+1)$ .	Используем функцию трансформации $f(x) = \log(x+1)$ , определённую при $x > 0$ .

В математике одно обозначение не используется дважды (для разных вещей), хотя иногда полиморфизм может быть удобен (но его отдельно надо объяснять читателю «если это не вызовет путаницу»).

Даже если обозначение интуитивно, его надо сначала ввести.

плохо	хорошо
Если выполняется условие (1), то полагаем $i \leftrightarrow j$ .	До использования написать: Запись « $i \leftrightarrow j$ » обозначение обмен значениями между переменными $i$ и $j$ .

## Стиль

Стиль соответствует написанию научных статей – обязательно посмотрите несколько! Предложения короткие и чёткие. Проанализируйте свой текст, многие цепочки предложений можно превратить в одно.

плохо	хорошо
Последнее время выходят статьи, посвященные как классификации одноканальных электрокардиограмм, так и методам классификации классических многоканальных стационарных электрокардиограмм. В данной главе будет дан их краткий обзор.	В данной главе представим обзор современных работ по классификации одноканальных электрокардиограмм и по классификации классических многоканальных стационарных электрокардиограмм.

<p>... где <math>\varphi</math> – сигмоидная функция активации. Данная функция активации задается как</p> $\varphi(z) = \frac{1}{1 + \exp(-z)}.$	<p>... где <math>\varphi</math> – сигмоидная функция активации:</p> $\varphi(z) = \frac{1}{1 + \exp(-z)}.$
--	--

Не допускайте просторечия («хорошенький», «получше», «малька» и т.п.).

плохо	хорошо
А сейчас рассмотрим несколько примеров, чтобы было всё понятно.	Далее приведём несколько поясняющих примеров.

При формулировании мысли не впадайте в крайности.

плохо	критика
Наиболее обще задачу можно поставить так...	Уверены ли Вы, что это наиболее общая постановка?

Текст состоит из предложений (в начале – заглавная буква, в конце – точка).

Каждое предложение должно быть согласовано.

Каждое предложение – законченная мысль, поэтому предложение не может начинаться с «то есть», «где» и т.п., т.к. это означает, что предыдущая мысль не закончена.

плохо	хорошо
Т.е. описанная модификация не улучшает качества.	..., т.е. описанная ...
Пусть $y=ax+b$ . Где $b=-c/d$ .	Пусть $y=ax+b$ , где $b=-c/d$ .
... является ансамблирование. Оно используется...	... является ансамблирование, которое используется...

Предложения формируют абзацы, каждый абзац начинается с красной строки. И только абзацы начинаются с отступа (часто в TeXе после формул идёт отступ, убирайте его директивой \noindent).

плохо	хорошо
$f(x, y) = \sin(xy / \tau)$ <p>Где <math>\tau</math> – температура (гиперпараметр метода).</p>	$f(x, y) = \sin(xy / \tau),$ <p>где <math>\tau</math> – температура (гиперпараметр метода).</p>

В научном тексте нет местоимения «я».

плохо	хорошо
Я реализовал метод.	<p>Автором был реализован метод.</p> <p>Нами был реализован метод.</p> <p>Опишем реализованный метод.</p>

Используйте меньше профессионального жаргона.

плохо	хорошо
вектора	векторы

### Формулировки

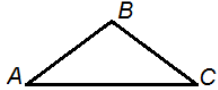
Следите за смыслом предложений.

плохо	замечание
Показатели <i>могут скрывать</i> недостатки процесса разработки модели.	Показатели не могут ничего скрывать! В них может быть отражено/заложено что-то или нет.
Градиенты могут <i>мерить</i> по переменной $x$ .	Градиенты вычисляют, а не «мерят».
Возьмём <i>значение</i> акции...	У акции есть цена, а не значение.
Обзор современных <i>подходов к задаче</i> диаризации.	Подходы бывают не к задаче, а к решению задачи.
Оба этих подхода <i>разделяют</i> один общий крупный недостаток, <i>основанный</i> на том ...	Недостаток нельзя разделять, лучше: «имеют общий недостаток», Недостаток не может быть на чём-то основан, лучше: «недостаток, состоящий в том».
Перестановка слов – метод аугментации, который часто используется для текстов [4].	Это не определение, лучше так: Одним из часто используемых методов аугментации текстов является перестановка слов [4].
Выберем <i>метод аугментации</i> , на котором сеть <i>работает</i> хуже всего.	Нейронная сеть не работает на методе аугментации! Часто жаргонно говорят, что она работает на данных, здесь на аугментированных данных (так написать уже допустимо).

Предложения не должны вызывать вопросов. Также не оставляйте читателю загадок.

плохо	Критика
Сравним графики (1) и (2). Теперь опишем модификацию предложенного алгоритма.	Стоп! А что получается в результате сравнения? И зачем мы сравнивали? Это должно быть чётко прописано:  <i>Сравнив графики (1) и (2), видим, что прямой зависимости между предсказаниями алгоритма (5) и результатами оценки методом (3) нет.</i>
Сначала с обучающей выборкой было проведено <i>несколько преобразований</i> . Затем решалась задача бинарной классификации...	Каких преобразований? Надо конкретизировать!
На данный момент в биологических исследованиях используется <i>ряд алгоритмов кластеризации</i> . Для обзора автор выбрал наиболее простой...	Неплохо бы перечислить алгоритмы из этого ряда.

Аналогично, ничего в тексте не должно вызывать простых вопросов.

плохо	Критика
<p>Рассмотрим треугольник ABC, см. рис.</p> 	<p>Почему на рис. изображён равносторонний треугольник? Это существенно или треугольник может быть произвольным?</p>

Предложения должны легко читаться, а смысл пониматься с первого прочтения.

плохо	хорошо
<p>Основной идеей является применение двунаправленной модели к сгенерированному тексту однонаправленной моделью.</p>	<p>Основной идеей <b>предложенного метода</b> является применение двунаправленной модели <b>к тексту</b>, сгенерированному <b>с помощью</b> однонаправленной модели.</p>

## Формулы

Формула – элемент предложения, поэтому, например, после неё обычно стоит знак препинания.

плохо	хорошо
<p>Введём нормировку следующим образом</p> $N f(x) = \frac{f(x + \varepsilon) - f(x - \varepsilon)}{\int_{-\varepsilon}^{\varepsilon} f(x + t) \partial t}$	<p>Введём нормировку следующим образом:</p> $N f(x) = \frac{f(x + \varepsilon) - f(x - \varepsilon)}{\int_{-\varepsilon}^{\varepsilon} f(x + t) \partial t}.$

Правильно ставьте знаки пунктуации внутри формулы.

плохо	хорошо
$x(t) = \begin{cases} 0 & t \leq 1 \\ 1 & t > 0, \end{cases}$	$x(t) = \begin{cases} 0, & t \leq 1, \\ 1, & t > 0, \end{cases}$

Обратите внимание, каким шрифтом в формулах пишутся буквы, векторы, функции и т.п. Некоторые функции в TeXе имеют стандартное обозначение, например `\sin`, некоторые Вы вводите сами – они должны быть написаны как стандартные. При включении в текст формул не забывайте оформлять их как формулы (в TeXе – помещать в знаки \$).

плохо	хорошо
$f(x, y) = \sin(x) + \text{cumsun}(y)$	$f(x, y) = \sin(x) + \text{cumsun}(y).$
<p>Пусть <math>x=1</math>.</p>	<p>Пусть <math>x = 1</math>.</p>

Если формула нумеруется, то где-то на неё ссылаются.



Все элементы формулы где-то определены.

Лучше писать так, чтобы при удалении формул предложение оставалось нормальным предложением русского языка.

плохо	хорошо
На выходе используем $\text{ReLU}$ .	На выходе используем <b>функцию активации</b> $\text{ReLU}$ .
Отметим, что $a, b, c$ попарно различны.	Отметим, что <b>значения</b> $a, b, c$ попарно различны.

Не начинайте предложения с формул и разделяйте формулы словами, кванторы желательно заменять словами (правило Д. Кнута). Во включённых формулах лучше не использовать некоторых кванторов, например существования.

плохо	хорошо
<b>Теорема.</b> $f(x) + g(x) = 0$ имеет $k$ корней.	<b>Теорема.</b> <b>Уравнение</b> $f(x) + g(x) = 0$ имеет $k$ корней.
Рассмотрим $f_k(x) = 0$ , $k \in \{1, 2, \dots, n\}$ .	Рассмотрим $f_k(x) = 0$ <b>при</b> $k \in \{1, 2, \dots, n\}$ .
Пусть задан граф $G = (V, E)$ , $\exists(i, j) \in E$ .	Пусть задан граф $G = (V, E)$ , одно из рёбер которого – $(i, j)$ .

Формулы должны иметь смысл.

плохо	критика
$(X, Y) = \{(x_i, y_i)\}_{i=1}^m$	Слева – набор (вектор) из двух элементов, справа – множество пар, они формально не могут быть равны!
Множество $R \in \mathbb{Z}^{m \times n}$ представляет собой матрицу.	Множество не может представлять собой матрицу.
Рассмотрим матрицу $M = (m_{ij}) = \begin{cases} 1, & i > j, \\ 0, & i \leq j. \end{cases}$	Слева – матрица, справа – значения одного элемента. Равенство не может быть выполнено.

Кстати, если идёт несколько равенств подряд, проверьте, точно ли равны левая и правая части.

«Высокие» формулы лучше делать выключенными (в TeXе есть специальные способы написания формул внутри текста, которые экономят высоту).

В математическом тексте многое можно выразить с помощью формул, причём очень коротко. Например, нам надо сказать, что три числа  $a, b, c$  попарно не равны:

обычный способ с помощью формул	экономный способ
$a \neq b, a \neq c, b \neq c$	$ \{a, b, c\}  = 3$

Очень часто в формулах индексы забывают заключить в фигурные скобки (в TeXе).

плохо	хорошо
$x_i j$	$x_{ij}$

## Теоремы

Утверждение теоремы лучше делать самостоятельным (не зависящим от предыдущего текста).

плохо	хорошо
<p>Выберем <math>k \in \{1, 2, \dots, n\}</math>, тогда справедлива следующая</p> <p><b>Теорема.</b> Уравнение <math>f_k(x) = 0</math> не имеет вещественных корней.</p>	<p><b>Теорема.</b> Уравнение <math>f_k(x) = 0</math> не имеет вещественных корней при <math>k \in \{1, 2, \dots, n\}</math>.</p>

## Списки

Списки тоже элементы основного текста, не забывайте пунктуацию.

плохо	хорошо
<p>В описанных случаях чаще всего используют</p> <ul style="list-style-type: none"> <li>• батч-нормализацию</li> <li>• регуляризацию</li> <li>• дропаут</li> </ul>	<p>В описанных случаях чаще всего используют:</p> <ul style="list-style-type: none"> <li>• батч-нормализацию,</li> <li>• регуляризацию,</li> <li>• дропаут.</li> </ul>

В приведённом примере если пункты нумерованы и номер заканчивается точкой, то каждый пункт пишется с заглавной буквы (но только в этом случае, лишний раз не используйте заглавные буквы):

хорошо
<p>В описанных случаях чаще всего используют:</p> <ol style="list-style-type: none"> <li>1. Батч-нормализацию.</li> <li>2. Регуляризацию.</li> <li>3. Дропаут.</li> </ol>

## Мелочи: пробелы, тире, сокращения и т.п.

После точки, как правило, идёт пробел. Часто ошибочно ставят две точки подряд.

Следите за написанием терминов.

При копировании формул часто забывают удалить/заменить индексы. Часто в тексте бывают дубликаты предложений и слов.

плохо	хорошо
<p>аудио сигнал</p> <p>аудио-сигнал</p>	<p>аудиосигнал</p>

Дефис и тире – разные, вот дефис – «-», а вот тире – «—».

Не забывайте про неразрывные символы (например, дефис и пробелы). Плохо, когда инициалы отрываются от фамилии.

плохо	хорошо
<i>В экспериментах <b>Петрова</b> <b>А.А.</b> было показано, что ...</i>	<i>В экспериментах Петрова А.А. было показано, что ...</i>

Всё должно быть однообразно, например, если букву «Ё» в слове всё Вы пишете с двумя точками, то надо писать «определённо», а не «определенно».

Есть ГОСТ регламентирующий сокращения (погуглите).

плохо	хорошо
<i>i-ый i-ого</i>	<i>i-й i-го</i>

## Ссылки

У всего чужого (цитаты, рисунки и т.п.) должен быть указан первоисточник.

При ссылке на статьи в русскоязычных работах чаще не упоминают авторов, особенно если это не важно с точки зрения повествования.

плохо	хорошо
<i>Капсульные сети были предложены Хинтоном [3]-[4].</i>	<i>Капсульные сети были предложены в [3]-[4].</i>

## Список литературы.

У каждой статьи должны быть указаны авторы, название, журнал и выходные данные (год, номер, страницы).


Оформляйте по ГОСТу, уже готовые оформления есть в гугл-академии (но там тоже бывают некорректные примеры).

Training **batchnorm** and only **batchnorm**: On the expressive power of random features in cnns

[PDF] [arxiv.org](#)

[J Frankle](#), [DJ Schwab](#), [AS Morcos](#) - arXiv preprint arXiv:2003.00152, 2020 - [arxiv.org](#)

Batch normalization (**BatchNorm**) has become an indispensable tool for training deep neural networks, yet it is still poorly understood. Although previous work has typically focused on its normalization component, **BatchNorm** also adds two per-feature trainable parameters: a ...

☆  Цитируется: 15 Похожие статьи Все версии статьи (3) »»

## пример правильной ссылки

*Frankle J., Schwab D. J., Morcos A. S. Training batchnorm and only batchnorm: On the expressive power of random features in cnns //arXiv preprint arXiv:2003.00152. – 2020.*

## Код, данные

Для воспроизводимости научных исследований необходимо выложить код / данные / значения гиперпараметров в открытый доступ. Ссылка на соответствующий репозиторий гитхаба существенно повышает доверие к работе.

## Про себя

Не забывайте указать все свои публикации, которые сделаны по теме работы. Лучше упомянуть выступления на конференции, полученные патенты, свидетельства о регистрации программ и другие вещи, если они связаны с работой.

## Не описывайте

- свои личные проблемы («я болел и поэтому не доделал эксперименты»),
- шутки, вопросы к читателю.

плохо	хорошо
<i>Какой же метод лучше работает? Давайте разбираться.</i>	<i>Проведём эксперименты для определения оптимального метода решения описанной задачи.</i>

Вообще, в Вашем тексте вряд ли будут знак вопроса, восклицательный знак и многоточие.

плохо	хорошо
<i>Что будет, если изменить гиперпараметры метода? Мы этого не знаем...</i>	<i>Эксперименты проводились только при указанных значениях гиперпараметров.</i>

## Перечитывание

Перечитывайте и редактируйте текст много раз.

Удаляйте ненужные слова.

Упрощайте сложные предложения.

## Орфография

Напоследок обязательно проверьте, благо есть автоматические средства проверки.