

Министерство общего и профессионального образования
Российской Федерации
УФИМСКИЙ ГОСУДАРСТВЕННЫЙ АВИАЦИОННЫЙ
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

УДК № 378.14

№ госрегистрации 01970006723

Инв. № _____

УТВЕРЖДАЮ

Проректор университета
по научной работе

_____ Н.С. Жернаков

«_____» _____ 1999 г.

ОТЧЁТ
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

Социально-экономические проблемы подготовки военных специалистов
в гражданских вузах России

по теме:

ФЕМИНИЗАЦИЯ АРМИИ КАК СОЦИАЛЬНЫЙ ПРОЦЕСС
(промежуточный)

Зам. проректора по научной работе

_____ Р.А. Бадамшин

Руководитель темы

_____ Г.А. Кабакович

Уфа 1999

СПИСОК ИСПОЛНИТЕЛЕЙ

Первый исполнитель _____ ФИО

Второй исполнитель _____ ФИО

РЕФЕРАТ

Отчет содержит 24 стр., 3 рис., 2 табл., 3 источника, 2 прил.

Это пример каркаса расчётно-пояснительной записки, желательный к использованию в РПЗ проекта по курсу РСОИ .

Данный опус, как и более новые версии этого документа, можно взять по адресу (<https://github.com/latex-g7-32/latex-g7-32>).

Текст в документе носит совершенно абстрактный характер.

СОДЕРЖАНИЕ

Введение	6
1 Аналитический раздел	7
1.1 Цель и задачи работы	7
1.2 Обнаружение аномалий	7
1.2.1 Классификация методов обнаружений аномалий	8
1.3 Постановка задачи	9
1.4 Существующие подходы к созданию всячины	9
2 Конструкторский раздел	13
2.1 Архитектура всячины	13
2.1.1 Протестируем подпункт	13
2.1.1.1 А теперь подподпункт	13
2.2 Подсистема всякой ерунды	13
2.2.1 Блок-схема всякой ерунды	14
3 Технологический раздел	15
4 Экспериментальный раздел	17
5 Организационно-экономический раздел	18
5.1 Протестируем специальные символы.	18
6 Промышленная экология и безопасность	19
Заключение	21
Список использованных источников	22
Приложение А Картинки	23
Приложение Б Еще картинки	24

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

Выборка/выборка данных — конечный набор прецедентов (объектов, случаев, событий, испытуемых, образцов, и т.п.), некоторым способом выбранных из множества всех возможных прецедентов, называемого генеральной совокупностью[2].

ВВЕДЕНИЕ

Задача поиска аномалий относится к одному из популярных способов машинного обучения - обучению без учителя. В настоящее время задачу поиска аномалий активно решают во многих областях жизнедеятельности:

- а) Защита информации и безопасность
- б) Социальная сфера и медицина
- в) Банковская и финансовая отрасль
- г) Распознавание и обработка текста, изображений, речи
- д) Другие сферы деятельности (например, мониторинг неисправностей механизмов)

Задачей поиска выбросов, как частный случай задачи поиска аномалий так же занимаются во вышеперечисленных отраслях.

Количество данных в мире удваивается примерно каждые два года. Поэтому актуальной задачей является разработка новых методов и усовершенствования старых методов поиска выбросов.

В данной работе предлагается новый метод, позволяющий найти аномалии в выборках данных.

1 Аналитический раздел

Цель и задачи работы

Целью данной работы является создание программного комплекта для обнаружения выбросов временных рядов в собираемых данных. Для достижения данной цели необходимо решить следующие задачи:

- ПЕРЕПИСАТЬ ИЗ ПРЕЗЕНТАЦИИ
- проанализировать предметную область и существующие методы обнаружения выбросов
- разработать метод обнаружения выбросов
- smth
- создать ПО, реализующего разработанный метод обнаружения выбросов
- провести вычислительный эксперименты с использованием разработанного метода

Обнаружение аномалий

В машинном обучении обнаружение "ненормальных" экземпляров в наборах данных всегда представляло большой интерес. Этот процесс широко известен как обнаружение аномалий или обнаружение выбросов. Вероятно, первое определение было дано Граббсом[3] в 1969 году: "Относительное наблюдение или выброс - это элемент выборки, который, заметно отличается от других членов выборки, в которых он встречается". Хотя это определение по-прежнему актуально и сегодня, мотивация для обнаружения этих выбросов сейчас совсем другая. Тогда основная причина обнаружения заключалась в том, чтобы удалить выбросы из данных обучения, поскольку алгоритмы распознавания образов были весьма чувствительны к выбросам в данных. Эта процедура также называется очищением данных. После разработки более надежных классификаторов интерес к обнаружению аномалий значительно снизился. Однако в 2000

году произошел поворотный момент, когда исследователи стали больше интересоваться самими аномалиями, поскольку они часто связаны с особенно интересными событиями или подозрительными данными. С тех пор было разработано много новых алгоритмов, которые оцениваются в этой статье. В этом контексте определение Граббса также было расширено, так что сегодня аномалии, как известно, имеют две важные характеристики:

Классификация методов обнаружений аномалий

В анализе данных есть два основных направления, которые занимаются поиском аномалий - это детектирование новизны и обнаружение выбросов. "Новый объект это так же объект, который отличается по своим свойствам от объектов выборки. Однако, в отличие от выброса, он его ещё нет в самой выборке и задача анализа сводится к его обнаружению при появлении. Например, если вы анализируете замеры уровня шума и отбрасываете слишком высокие или слишком низкие значения, то вы боретесь с выбросом. А если Вы создаёте алгоритм, который для каждого нового замера оценивает, насколько он похож на прошлые, и выбрасывает аномальные — вы "боретесь с новизной"[1]. Выбросы являются следствием:

- а) ошибок в данных
- б) неверно классифицированных объектов
- в) присутствием объектов других выборок
- г) намеренным искажением данных

На рисунке 1.1 можно увидеть желтые точки - выброс "слабом смысле". Они незначительно отклоняются от основных данных (зеленые точки). Красные же точки являются аномальными - выбросами "в сильном смысле они значительно отклоняются от основных данных. В данной работе будет изучаться вопрос нахождения "сильных выбросов" и критериев отличия сильного выброса от основных данных. В дальнейшем под словом "выброс" будет подразумеваться "сильный выброс а под

аномалией - в выброс (выброс является частным случаем аномалии). Понятие аномалии зачастую интерпретируют по-разному в зависимости от характера данных. Обычно аномалией называют некоторое отклонение от нормы. Это определение нуждается в формальном уточнении.

Постановка задачи

Кстати, про картинки. Во-первых, для фигур следует использовать [ht]. Если и после этого картинки вставляются «не по ГОСТ», т.е. слишком далеко от места ссылки, — значит у вас в РПЗ слишком мало текста! Хотя и ужасный параметр !ht у окружения figure тоже никто не отменял, только при его использовании документ получается страшный, как в ворде, поэтому просьба так не делать по возможности.

Существующие подходы к созданию всячины

Известны следующие подходы...

а) Перечисление с номерами.

б) Номера первого уровня. Да, ГОСТ требует именно так — сначала буквы, на втором уровне — цифры. Чуть ниже будет вариант «нормальной» нумерации и советы по её изменению. Да, мне так нравится: на

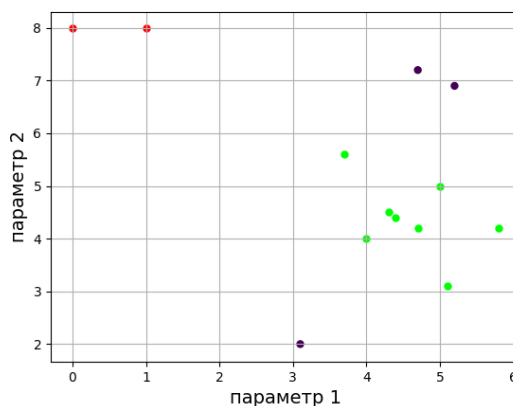


Рисунок 1.1 — Пример размеченного набора данных

первом уровне выравнивание элементов как у обычных абзацев. Проверим теперь вложенные списки.

1) Номера второго уровня.

2) Номера второго уровня. Проверяем на длииииной-предлиииии-ииииинной строке, что получается.... Сойдёт.

в) По мнению Лукьяненко, человеческий мозг старается подвести любую проблему к выбору из трех вариантов.

г) Четвёртый (и последний) элемент списка.

Теперь мы покажем, как изменить нумерацию на «нормальную», если вам этого захочется. Пара команд в начале документа поможет нам.

1) Изменим нумерацию на более привычную...

2) ... нарушим этим гост.

а) Но, пожалуй, так лучше.

В заключение покажем произвольные маркеры в списках. Для них нужен пакет `enumerate`.

1. Маркер с арабской цифрой и с точкой.

2. Маркер с арабской цифрой и с точкой.

I. Римская цифра с точкой.

II. Римская цифра с точкой.

В отчётах могут быть и таблицы — см. табл. 1.1 и 1.2. Небольшая таблица делается при помощи `tabular` внутри `table` (последний полностью аналогичен `figure`, но добавляет другую подпись).

Для больших таблиц следует использовать пакет `longtable`, позволяющий создавать таблицы на несколько страниц по ГОСТ.

Для того, чтобы длинный текст разбивался на много строк в пределах одной ячейки, надо в качестве ее формата задавать `p` и указывать явно ширину: в мм/дюймах (`110mm`), относительно ширины страницы (`0.22\textwidth`) и т.п.

Таблица 1.1 — Пример короткой таблицы с коротким названием

Тело	F	V	E	$F + V - E - 2$
Тетраэдр	4	4	6	0
Куб	6	8	12	0
Октаэдр	8	6	12	0
Додекаэдр	20	12	30	0
Икосаэдр	12	20	30	0
Эйлер	666	9000	42	$+\infty$

Можно также использовать уменьшенный шрифт — но, пожалуйста, тогда уж во всей таблице сразу.

Таблица 1.2 — Пример длинной таблицы с длинным названием на много длинных-длинных строк

Вид шума	Громкость, дБ	Комментарий
Порог слышимости	0	
Шепот в тихой библиотеке	30	Конечно, это было до эпохи мобильных (внутри машины)
Обычный разговор	60-70	
Звонок телефона	80	
Уличный шум	85	
Гудок поезда	90	
Шум электрички	95	
Порог здоровой нормы	90-95	Длительное пребывание на более громком шуме может привести к ухудшению слуха
Мотоцикл	100	(модель бензокосилки) (Doom в целом вреден для здоровья)
Power Mower	107	
Бензопила	110	
Рок-концерт	115	
Порог боли	125	feel the pain

Продолжение на след. стр.

Продолжение таблицы 1.2

Клепальный молоток	125	(автор сам не знает, что это)
Порог опасности	140	Даже кратковременное пребывание на шуме большего уровня может привести к необратимым последствиям
Реактивный двигатель	140 180	Необратимое полное повреждение слуховых органов
Самый громкий возможный звук	194	Интересно, почему?..

2 Конструкторский раздел

В данном разделе проектируется новая всячина.

Архитектура всячины

Протестируем подпункт

А теперь подподпункт

Проверка параграфа. Вроде работает.

Вторая проверка параграфа. Опять работает.

Вот.

— Это список с «палочками».

— Хотя он и по ГОСТ, но...

1) Для списка, начинающегося с заглавной буквы, лучше список с цифрами.

Формула 2.1 совершенно бессмысленна.

$$a = cb \tag{2.1}$$

Окружение cases опять работает (см. 2.2), спасибо И. Короткову за исправления..

$$a = \begin{cases} 3x + 5y + z, & \text{если хорошо} \\ 7x - 2y + 4z, & \text{если плохо} \\ -6x + 3y + 2z, & \text{если совсем плохо} \end{cases} \tag{2.2}$$

Подсистема всякой ерунды

Культурная вставка dot-файлов через утилиту dot2tex (рис. ??).

Блок-схема всякой ерунды

Кстати о заголовках

У нас есть и subsubsection. Только лучше её не нумеровать.

3 Технологический раздел

В данном разделе описано изготовление и требование всячины. Кстати, в Latex нужно эскейпить подчёркивание (писать «some_function» для some_function).

Для вставки кода есть пакет listings. К сожалению, пакет listings всё ещё работает криво при появлении в листинге русских букв и кодировке исходников utf-8. В данном примере он (увы) на лету конвертируется в koi-8 в ходе сборки pdf.

Есть альтернатива listingsutf8, однако она работает лишь с \lstinputlisting, но не с окружением \lstlisting

Вот так можно вставлять псевдокод (питоноподобный язык определен в listings.inc.tex):

Листинг 3.1 — Алгоритм оценки дипломных работ

```
1 def EvaluateDiplomas():
2     for each student in Masters:
3         student.Mark ← 5
4     for each student in Engineers:
5         if Good(student):
6             student.Mark ← 5
7         else:
8             student.Mark ← 4
```

Еще в шаблоне определен псевдоязык для BNF:

Листинг 3.2 — Грамматика

```
1 ifstmt → "if" "(" expression ")" stmt |
2         "if" "(" expression ")" stmt1 "else" stmt2
3 number → digit digit *
```

В листинге ?? работают русские буквы. Сильная магия. Однако, работает только во включаемых файлах, прямо в TeX нельзя.

Можно также использовать окружение verbatim, если listings чем-то не устраивает. Только следует помнить, что табы в нём «съедаются». Существует так же команда \verbatiminput для вставки файла.

```
a_b = a + b; // русский комментарий  
if (a_b > 0)  
    a_b = 0;
```


4 Экспериментальный раздел

В данном разделе проводятся вычислительные эксперименты.

5 Организационно-экономический раздел

Протестируем специальные символы.

И заодно переключение шрифтов.

"--* Прямая речь "--- <<после „тире“ неразрывный пробел>>

\cyrillicfonttt — Прямая речь — «после „тире“ неразрывный пробел».

\cyrillicfontsf — Прямая речь — «после „тире“ неразрывный пробел».

\cyrillicfont — Прямая речь — «после „тире“ неразрывный пробел».

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

1) First itemtext

a) First itemtext

- First itemtext
- Second itemtext
- Last itemtext
- First itemtext
- Second itemtext

б) Last itemtext

в) First itemtext

г) Second itemtext

д) Last itemtext

2) First itemtext

3) Second itemtext

4) Last itemtext

5) First itemtext

ЗАКЛЮЧЕНИЕ

В результате проделанной работы стало ясно, что ничего не ясно...

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Дьяконов Александр. Поиск аномалий (Anomaly Detection). — 2017. — URL: <https://alexanderdyakonov.wordpress.com/2017/04/19/%D0%BF%D0%BE%D0%B8%D1%81%D0%BA-%D0%B0%D0%BD%D0%BE%D0%BC%D0%B0%D0%BB%D0%B8%D0%B9-anomaly-detection/>.
2. machinelearning.ru/. Выборка. — URL: <http://www.machinelearning.ru/wiki/index.php?title=%D0%92%D1%8B%D0%B1%D0%BE%D1%80%D0%BA%D0%B0>.
3. F.E. Grubbs. Procedures for Detecting Outlying Observations in Samples. Technometrics). — 1969.

ПРИЛОЖЕНИЕ А

КАРТИНКИ

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Рисунок А.1 — Картинка в приложении. Страшная и ужасная.

ПРИЛОЖЕНИЕ Б

ЕЩЕ КАРТИНКИ

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Рисунок Б.1 — Еще одна картинка, ничем не лучше предыдущей. Но надо же как-то заполнить место.