РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ

Факультет физико-математических и естественных наук

Кафедра прикладной информатики и теории вероятностей

Утверждаю

Заведующий кафедрой   
прикладной информатики   
и теории вероятностей  
д.т.н., профессор  
\_\_\_\_\_\_\_\_\_\_\_\_ К.Е. Самуйлов

« » 20 г.

КУРСОВАЯ РАБОТА

на тему

**«Статистический анализ выборок малого объема»**

по дисциплине «Компьютерный практикум по математическому моделированию»

шифр и наименование учебной дисциплины

Выполнил

Студент группы НФИбд-01-17

Студенческий билет №:

И.О. Фамилия

« » 20 г.

Руководитель

доцент кафедры   
прикладной информатики и теории вероятностей, к.ф.-м.н.

И.О. Фамилия

Москва 2020

Оглавление

[Список используемых сокращений 3](#_Toc90295793)

[Введение 4](#_Toc90295794)

[Кластеризация малой выборки 5](#_Toc90295795)

[Иерархическая кластеризация 6](#_Toc90295796)

[Кластеризация методом К-средних 7](#_Toc90295797)

[Классификация малой выборки 8](#_Toc90295798)

[Определение лучшего метода классификации для выборки малого объема 8](#_Toc90295799)

[Экстраполяция на основе малой выборки 9](#_Toc90295800)

[Экстраполяция методом наименьших квадратов 10](#_Toc90295801)

[Экстраполяция методом скользящей средней 11](#_Toc90295802)

[Заключение 12](#_Toc90295803)

[Список литературы по теме 13](#_Toc90295804)

# 

# Список используемых сокращений

# Введение

Актуальность работы

Актуальность данной работы обусловлена ….

Цель работы

Целью моей курсовой работы является …...

Задачи работы

Основными задачами мой работы являются:

...

...

Методы исследования

Методом исследования является ...

Структура работы

Курсовая работа состоит из введения, трех разделов, заключения и списка используемой литературы. Во введении ….

В первом разделе ....

Во втором разделе ...

Во третьем разделе ...

В заключении подведены общие итоги курсовой работы, изложены основные выводы.

# 

# Кластеризация малой выборки

Кластеризация (или кластерный анализ) — это задача разбиения множества объектов на группы, называемые кластерами. Внутри каждой группы должны оказаться «похожие» объекты, а объекты разных группы должны быть как можно более отличны. Главное отличие кластеризации от классификации состоит в том, что перечень групп четко не задан и определяется в процессе работы алгоритма.  
  
Применение кластерного анализа в общем виде сводится к следующим этапам:

* Отбор выборки объектов для кластеризации.
* Определение множества переменных, по которым будут оцениваться объекты в выборке. При необходимости – нормализация значений переменных.
* Вычисление значений меры сходства между объектами.
* Применение метода кластерного анализа для создания групп сходных объектов (кластеров).
* Представление результатов анализа.

## Иерархическая кластеризация

Среди алгоритмов иерархической кластеризации выделяются два основных типа: восходящие и нисходящие алгоритмы. Нисходящие алгоритмы работают по принципу «сверху-вниз»: в начале все объекты помещаются в один кластер, который затем разбивается на все более мелкие кластеры. Более распространены восходящие алгоритмы, которые в начале работы помещают каждый объект в отдельный кластер, а затем объединяют кластеры во все более крупные, пока все объекты выборки не будут содержаться в одном кластере. Таким образом строится система вложенных разбиений. Результаты таких алгоритмов обычно представляют в виде дерева – дендрограммы. Классический пример такого дерева – классификация животных и растений.

## Кластеризация методом К-средних

Метод К-средних обычно применяется, когда не известны ни количество кластеров, ни присущие им качества. Таким образом, специалисты по машинному обучению используют данный алгоритм, чтобы обнаружить закономерности в датасете, о которых они еще ничего не знают.

Задачу кластеризации можно рассматривать как построение оптимального разбиения объектов на группы. При этом оптимальность может быть определена как требование минимизации среднеквадратической ошибки разбиения:  
  
Изображение выглядит как текст, часы

Автоматически созданное описание  
  
где cj — «центр масс» кластера j (точка со средними значениями характеристик для данного кластера).  
  
Алгоритмы квадратичной ошибки относятся к типу плоских алгоритмов. Самым распространенным алгоритмом этой категории является метод k-средних. Этот алгоритм строит заданное число кластеров, расположенных как можно дальше друг от друга. Работа алгоритма делится на несколько этапов:

Случайно выбрать k точек, являющихся начальными «центрами масс» кластеров.

Отнести каждый объект к кластеру с ближайшим «центром масс».

Пересчитать «центры масс» кластеров согласно их текущему составу.

Если критерий остановки алгоритма не удовлетворен, вернуться к п. 2.

В качестве критерия остановки работы алгоритма обычно выбирают минимальное изменение среднеквадратической ошибки. Так же возможно останавливать работу алгоритма, если на шаге 2 не было объектов, переместившихся из кластера в кластер.

# Классификация малой выборки

Классификация — один из разделов машинного обучения, посвященный решению следующей задачи. Имеется множество объектов (ситуаций), разделённых некоторым образом на классы. Задано конечное множество объектов, для которых известно, к каким классам они относятся. Это множество называется обучающей выборкой. Классовая принадлежность остальных объектов не известна. Требуется построить алгоритм, способный классифицировать произвольный объект из исходного множества.

Классифицировать объект — значит, указать номер (или наименование класса), к которому относится данный объект.

Классификация объекта — номер или наименование класса, выдаваемый алгоритмом классификации в результате его применения к данному конкретному объекту.

## Определение лучшего метода классификации для выборки малого объема

# 

# Экстраполяция на основе малой выборки

Экстраполяция – это метод научного исследования, который основан на распространении прошлых и настоящих тенденций, закономерностей, связей на будущее развитие объекта прогнозирования.

## Экстраполяция методом наименьших квадратов

## Экстраполяция методом скользящей средней

# Заключение

Текст

# Список литературы по теме

Фамилия И.О. Название книги. — Место издательства: Название издательства, год.

Фамилия И.О. Название статьи // Название журнала. — год. — Т. , № . — С.  - . — DOI:

Фамилия И.О. Название публикации в сборнике материалов конференции // Название сборника материалов конференции. — Место издательства: Название издательства, год. — С.  - . —URL: