|  |  |
| --- | --- |
|  | **Министерство науки и высшего образования Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение**  **высшего образования**  **«Московский государственный технический университет**  **имени Н.Э. Баумана**  **(национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

ФАКУЛЬТЕТ \_\_\_ИУ\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

КАФЕДРА \_\_\_\_\_ИУ-5\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА**

***К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ***

***НА ТЕМУ:***

***Информационная Система\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

***\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_***

Студент \_\_\_\_ИУ5-\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_  **\_\_\_**

(Группа) (Подпись, дата) (И.О.Фамилия)

Руководитель **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_**

(Подпись, дата) (И.О.Фамилия)

Консультант **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_М.В.Черненький\_**

(Подпись, дата) (И.О.Фамилия)

*Оценка –*

*2021 г.*

**Министерство науки и высшего образования Российской Федерации**

**Федеральное государственное бюджетное образовательное учреждение**

**высшего образования**

**«Московский государственный технический университет имени Н.Э. Баумана**

**(национальный исследовательский университет)»**

**(МГТУ им. Н.Э. Баумана)**

УТВЕРЖДАЮ

Заведующий кафедрой \_\_\_\_\_\_\_\_\_\_

(Индекс)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(И.О.Фамилия)

« \_\_\_\_\_ » \_\_\_\_\_\_\_\_\_\_\_\_ 20 \_\_\_\_ г.

**ЗАДАНИЕ**

**на выполнение научно-исследовательской работы**

по теме \_\_Система \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Студент группы \_\_\_ИУ5-\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Фамилия, имя, отчество)

Направленность НИР (учебная, исследовательская, практическая, производственная, др.)

\_\_\_\_\_\_\_\_\_\_учебная\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Источник тематики (кафедра, предприятие, НИР) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

График выполнения НИР: 25% к \_\_\_ нед., 50% к \_\_\_ нед., 75% к \_\_ нед., 100% к \_\_\_ нед.

Техническое задание \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_выполнить обзор существующих методов

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***Оформление научно-исследовательской работы:***

Расчетно-пояснительная записка на \_12\_ листах формата А4.

Перечень графического (иллюстративного) материала (чертежи, плакаты, слайды и т.п.)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Дата выдачи задания « \_21\_ » \_сентября\_ 2021 г.

**Руководитель НИР**  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_

(Подпись, дата) (И.О.Фамилия)

**Студент \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_**

(Подпись, дата) (И.О.Фамилия)

Примечание: Задание оформляется в двух экземплярах: один выдается студенту, второй хранится на кафедре.

Содержание

[Цель работы 4](#_Toc89898962)

[Задачи 4](#_Toc89898963)

[Классификация 4](#_Toc89898964)

[Классификаторы 4](#_Toc89898965)

[Обзор статей по классификации 5](#_Toc89898966)

[Анализ … 6](#_Toc89898967)

[Классификация … 6](#_Toc89898968)

[Выводы 7](#_Toc89898969)

[Модель … 7](#_Toc89898970)

[Список использованных источников 8](#_Toc89898971)

# Цель работы

Провести обзор существующих методов …

# Задачи

В ходе выполнения работы были выделены следующие подзадачи:

1. Рассмотреть классификацию …
2. Провести обзор публикаций по …
3. Исследовать …
4. Рассмотреть модель …

# Классификация

Задача классификации может быть формализована как задача …. [1].

Принцип классификации может быть определен как множество категорий и набор документов. Категории состоят из положительных или отрицательных примеров, членов множества. Процесс классификации состоит из нескольких этапов, на каждом из которых используются разные методы, зависящие от набора текстов и заданной цели.

Далее представлены наиболее общие этапы:

1. Классификация …

2. Преобразование …

3. Индексация …

4. Выбор подмножества классифицирующих признаков и связывающей их целевой функции...

5. Использование одного из алгоритмов (классификаторов).

6. На последнем этапе проводится оценка качества проведенной классификации. Если получен неудовлетворительный результат, проводится корректировка признаков или даже смена классификатора.

# Классификаторы

В качестве классификаторов используют различные алгоритмические подходы [2]. В отчете будут рассмотрены наиболее популярные классификаторы.

1. ….
2. ….
3. ….
4. …

Описанные здесь классификаторы часто используются на практике и хорошо зарекомендовали себя при решении задач классификации, но ни один из них не может считаться лучшим и претендовать на роль стандарта при классификации текста. Вероятно, это связано с тем, что каждый классификатор ориентирован на конкретное приложение или определенный тип данных.

# Обзор статей по классификации

Последние исследования в области классификации реализуют идею объединения разных классификаторов, т.е. создание гибридного метода. Этот подход фокусируется только на индивидуальных преимуществах классификаторов, отбрасывая их недостатки и тестируя возможность совместного использования. Результаты исследований доказывают перспективность сочетания классификаторов, за счет снижения индивидуальных ограничений каждого из них. Ниже приведены примеры подобных работ.

В работе … для сравнения классификаторов измерялась их производительность при ранжировании ...

Вероятностный метод объединения классификаторов предложен в работе … [3].

В работе … [4] представлено подробное исследование эффективности объединения классификаторов при классификации …

В двух работах …..[5] и …. [6] приведены результаты исследований совместимости ….

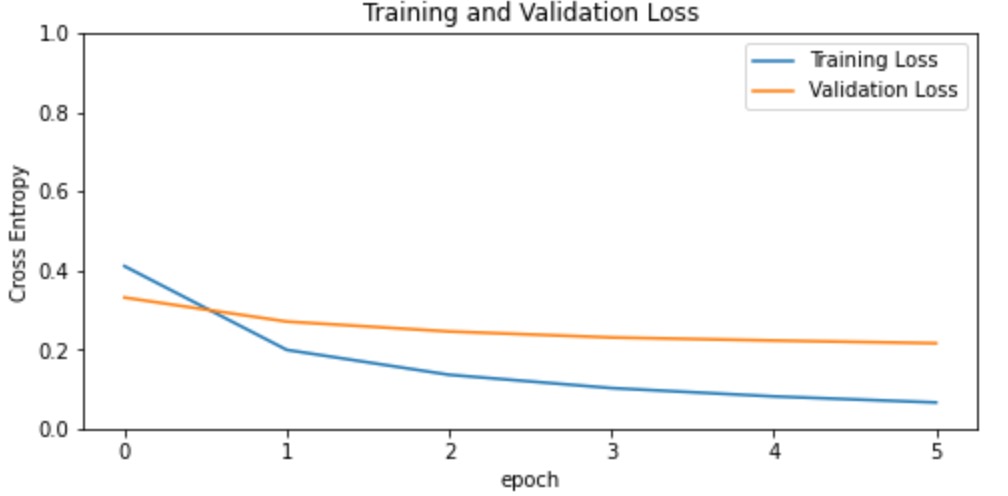
Результаты анализа совместной работы методов … [7]. Первичный набор используется для формирования нижних и верхних границ функций для каждого класса, классифицируемого ….

Метод улучшения алгоритма классификации …. [8]. Гибридный классификатор назван авторами ...

Приведенные примеры подтверждают, что сочетание классификаторов стало популярной областью исследований. Интерес исследователей объясняется тем, что результаты, полученные …

# Анализ …

Визуализировать его результаты удобно в виде …. Эта диаграмма содержит … [9]



# 

# Классификация …

В общем случае …. Это повышает точность классификации.

Важной особенностью … является их устойчивость к …. Кроме того, в ...

Рассмотрим отдельные этапы классификации. Поскольку процесс классификации основан на сравнении ...

Построение профиля включает следующие процедуры:

1. …;
2. …;
3. …;
4. ….

В результате получается файл … Заключительной процедурой классификации является сравнение ...

# Выводы

Задачи анализа ...

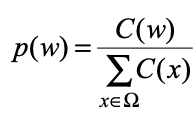
Интерес к новым средствам ...

Существует большое количество подходов к решению задач ...

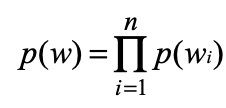
Несмотря на общность задачи ...

# Модель …

Пусть задан некоторый …



Подобно определяют вероятность p…



Таким образом, …

Тогда вероятность задается в виде …



где …

# Список использованных источников

1. Vandana Korde, C Namrata Mahender, “Text Classification and Classifiers: A Survey”, International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.2, March 2012.
2. L.S. Larkey. and W. B. Croft, “Combining classifiers in text categorization”,Proc. SIGIR-96, 19th ACM International Conference on Research and Development in Information Retrieval (Zurich, CH, 1996), pp. 289–297 1996.
3. Paul N. Bennett, Susan T. Dumais, Eric Horvitz, “Probabilistic Combination of Text Classifiers Using Reliability Indicators: Models and Results”, Proceedings of 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Tampere, Finland, August 2002. ACM Press.
4. B. Grilheres, S. Brunessaux, and P. Leray, “Combining classifiers for harmful document filtering”, RIAO '04 Coupling approaches, coupling media and coupling languages for information retrieval, Pages 173-185, 2004.
5. Dino Isa, Lam Hong lee, V. P Kallimani, and R. Raj Kumar, “Text Documents Preprocessing with the Bayes Formula for Classification using the Support vector machine”, IEEE Transactions of Knowledge and Data Engineering, vol.20, no. 9, pp.1264-1272, September 2008.
6. Dino Isa, V. P Kallimani and Lam Hong lee, “Using Self Organizing Map for Clustering of Text Documents”, Expert System with Applications, vol. 36, no. 5, pp. 9584-9591, July, 2009.
7. Duoqian Miao , QiguoDuan, Hongyun Zhang, and Na Jiao, “Rough set based hybrid algorithm for text classification”, Journal of Expert Systems with Applications, vol. 36, no. 5, pp. 9168-9174, July 2009
8. S.Ramasundaram, “NGramsSA Algorithm for Text Categorization”, International Journal of Information Technology & Computer Science (IJITCS ), Volume 13, Issue No : 1, pp.36-44, 2014.