МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

им. Н.Э. Баумана

Кафедра «Системы обработки информации и управления»

|  |  |  |  |
| --- | --- | --- | --- |
|  | Утверждаю  Заведующий кафедрой ИУ-5 |  | Согласовано  научный руководитель |
|  | \_\_\_\_\_\_\_\_\_\_\_ В.И.Терехов  "\_\_"\_\_\_\_\_\_\_\_\_\_2024 г. |  | \_\_\_\_\_\_\_\_\_\_ Ю.Е. Гапанюк  "\_\_"\_\_\_\_\_\_\_\_\_\_2024 г. |

**Метрика скорости реакции поисковых систем**

Техническое задание

(вид документа)

писчая бумага

(вид носителя)

5

(количество листов)

ИСПОЛНИТЕЛЬ: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Сафонов Федор Алексеевич

"\_\_"\_\_\_\_\_\_\_\_\_\_\_2024 г.

Москва - 2024

1. **Введение**

Поисковая система  — это специальная программа, которая в ответ на запрос пользователя через веб-интерфейс (сайт) выдает список ресурсов, отсортированных по релевантности этому запросу.

Чтобы отслеживать качество работы различных аспектов поисковой системы необходимы соответствующие метрики. С помощью метрик разработчики понимают, насколько хорошо работают их алгоритмы и модели машинного обучения. Метрики являются основным критерием качества уже выполненной работы. Благодаря метрикам разработчики могут быстро заметить какую-то неисправность в алгоритмах ранжирования и немедленно исправить ее.

Большая доля запросов в Яндекс Поиске касаются новостей, происшествий и других внезапных событий, произошедших в мире. Такие запросы обладают особой спецификой, т.к. на них крайне важно показать недавно созданные страницы в интернете, ведь именно на них отражены последние актуальные новости, которые нужны пользователю.

Метрика скорости реакции показывает, насколько хорошо Яндекс Поиск сработал на запросах про внезапные новостные события. Имея такую метрику, можно сделать вывод о текущем качестве ранжирования и спланировать дальнейшую работу по улучшению данной метрики.

1. **Основания для разработки**

Основанием для разработки является рабочая задача в компании Яндекс, отделе Поисковых систем. Задание было поставлено руководителем группы отдела ранжирования.

**3. Назначение разработки**

Разрабатываемая метрика предназначена для внутренних разработчиков Яндекс Поиска, для отслеживания и улучшения качества работы алгоритмов ранжирования на пользовательских новостных запросах.

**4. Требования к программе или программному изделию**

4.1 Требования к функциональным характеристикам

Программа должна выполнять следующие функции:

4.1.1 Загрузка запросов и логов за соответствующий день

4.1.2 Выделение набора признаков для классификации

4.1.3 Классификация запроса на новостной, касательно какого-то внезапного события или происшествия.

4.1.4 Определение времени происшествия новостного события по данному запросу

4.1.5 Оценка качества ранжирования выдачи по запросам, заданным после момента происшествия новостного события

4.1.6 Итоговое вычисление метрики, определяющее качество ранжирования за текущий день.

**5. Требования к программной документации**

Для представления заказчику разрабатываются следующие документы:

1. Описание решения подзадач в рамках построение процесса для вычисления метрик.

2. Описание итогового процесса

3. Сам дашборд, необходимый для отслеживания метрики скорости реакции.

**6. Технико-экономические показатели**

Процесс должен потреблять допустимое количество CPU и GPU на кластере для ежедневного построения точек на графике метрики скорости реакции. Также в процессе должна быть задействована асессорская разметка, стоимость которой будет укладываться в общий бюджет компании.

**7. Стадии и этапы разработки**

График выполнения отдельных этапов работ приведен в соответствии с приказом об организации учебного процесса в 2020/2021 учебном году.

Таблица 1: Этапы разработки

| № п/п | Наименование этапа и содержание работ | Сроки исполнения |
| --- | --- | --- |
| 1 | Разработка и утверждение ТЗ | Март 2024г. |
| 2 | Исследование предметной области | Март-Апрель 2024г. |
| 3 | Разработка архитектуры программного обеспечения | Апрель 2024 г. |
| 4 | Создание программ | Апрель — Май 2024 г. |
| 5 | Тестирование и отладка | Май 2024 г. |
| 6 | Оформление документации | Май — Июнь 2024 г. |
| 7 | Защита работы | Июнь 2021 г. |

**8. Порядок контроля и приёмки**

Приём и контроль программного изделия осуществляется в соответствие с методикой испытаний (см. документ «Программа и методика испытаний»).

Приём программного изделия в виде … испытаний осуществляется в ходе «Защиты макетов программ – предварительной защиты ВКРБ» в период с 15 по 24 мая 2024 года в соответствие с разработанной программой и методикой испытаний.